JMIR Dermatology

All topics related to diseases of the skin, hair, and nails, with special emphasis on technologies for information exchange, education, and clinical care

Volume 9 (2026) ISSN 2562-0959 Editor in Chief: Robert Dellavalle, MD, PhD, MSPH

Contents

Original Papers

Accuracy, Reliability, and Quality of YouTube and TikTok Videos as Sources of Information on Urticaria: Cross-Sectional Study (e88827)

Abrar Bukhari, Saad Altalhab, Suliman AlFaraj, Huda Alrwebah, Mishari Alrubaiaan, Abdulrahman Alfawzan. 6

Assessment of Quality and Utility of Patient-Taken Smartphone Photographs of Atopic Dermatitis: Clinical Survey Study (e72916)

Zarqa Ali, Kenneth Thomsen, Christian Vestergaard, Simon Thomsen. 18

Comments on Dermatologists' TikTok Videos on Atopic Dermatitis: Content Analysis (e90649)

Sahithi Gangavarapu, Tazeena Khan, Morgan McCarthy, Molly Hales, Stephanie Rangel. 28

Enhanced Diagnosis of Generalized Pustular Psoriasis With the Legit.Health Device as a Diagnosis Support Tool: Multireader Multicase Study (e82030)

Alfonso Medela, Ignacio Hernández Montilla, Alberto Sabater, Andy Aguilar, Taig Mac Carthy, Gurpreet Singh Chowdhry, Juan Semeco, Antonio Martorell. 43

Association of Skin Cancer With Clinical Depression and Poor Mental Health Days: Cross-Sectional Analysis (e80710)

Riona Ray, Mytien Nguyen. 57

Perceptions of a Culturally Tailored Narrative Video for Skin Cancer Prevention Among Spanish-Speaking Hispanic Outdoor Workers: Exploratory Study (e98130)

David Perez, Valeria Gómez, Dariana Sedeño-Delgado, Carlos Orellana García, Melida Chacon, Miguel Gonzalez, Chiranjeev Dash, Roxanne Mirabal-Beltran, Alejandra Hurtado-de-Mendoza. 69

Pre-Post Evaluation of a Sun Safety Social Media Campaign for Young Adults to Reduce Skin Cancer Risk (e90595)

Janessa Mendoza, David Perez, Muriel Statman, Mia Price, Madeleine Brown, Omar Anwar, Marcelo Sleiman Jr, Kenneth Tercyak. 86

ChatGPT-Generated Advice on Sun Protection and Skin Cancer Prevention Compared to American Academy of Dermatology Guidelines: Cross-Sectional Content Analysis (e93839)

Hanadi Qeyam, Ahmed Al-Rusan. 100

Remote Monitoring of Cryosurgery Response Using a Smartphone App: Prospective Study (e63467)

Vanessa Weir, Emily Cowen, Trina Salvador, Mary Sun, Lilly Gu, Maura Gillis, Nicholas Kurtansky, Veronica Rotemberg, Allan Halpern. 111

XSL•FO
RenderX

Patient Satisfaction, Side Effects, and Other Reactions Reported by Adult Men Prescribed Compounded Topical Finasteride via a National Telehealth Platform: Retrospective Analysis of Real-World Data (e84676)
Jessica Yu, Sachie Mochida, Michele Emery, Patrick Carroll, Justin Ko, Arash Mostaghimi. 121

Community-Based Teledermatology for Urgent Suspected Skin Cancer: Health Economic Cost-Comparison and Discrete Event Simulation Study (e86402)
Tim Hoogenboom, Pablo Martínez, Piyush Mahapatra, Nurul Nizar. 128

Public Interest in Stress-Induced Dermatologic Conditions and Psychodermatologic Disorders Before, During, and After the COVID-19 Pandemic: Segmented Regression Analysis (e93458)
Marisa Fat, Hayden Johnson, Michael Bernas, Cheryl Hurd, Robert Smith. 139

Identifying Over- and Underfunded Diseases by Comparing National Institutes of Health Funding for Skin Disease Research With US Skin Disease Burden According to 2021 Global Burden of Disease Data: Cross-Sectional Analysis (e71468)
Aileen Park, Emily Woolhiser, Hannah Riva, Leo Wan, Haaris Kadri, Elizabeth Lamberty, Parker Juels, Sandra Jaroonwanichkul, Madison Reed, Catherine Hegedus, Dana Chen, Danielle Duffle, Jessica Kirk, Sydney Christensen, Emma Shelby, Robert Dellavalle. 150

Dermatologic Conditions and Incident Anxiety in Young Adults: Propensity Score–Matched Retrospective Cohort Study (e90820)
Isabella Zai, Adrian Zai. 156

Leveraging AI Large Language Models for Writing Clinical Trial Proposals in Dermatology: Instrument Validation Study (e76674)
Megan Hauptman, Daniel Copley, Kelly Young, Tran Do, Joseph Durgin, Albert Yang, Jungsoo Chang, Allison Billi, Mio Nakamura, Trilokraj Tejasvi. 165

Evaluation of a Self-Test Technology for Chlamydia and Gonorrhea Testing: Observational Before-and-After Study (e99694)
Bettina Trettin, Ida Josefsson, Nadja Munk, Tine Vestergaard, Tue Kjølhede, Mette Skjøth. 173

The Effects of Outreach Initiatives on Equitable Access to Dermatology Services: Single-Center Retrospective Analysis (e94872)
J Albrecht, Justin Lyon, Michael Birdsall, Eliza Broadbent, Nicole Ufkes, Stephanie Zone, Anneli Bowen, Christopher Hull, Lauren Madigan, Ashley Snyder, Chaorong Wu, Bethany Lewis. 183

Efficiency and Sustainability in the Skin Cancer Excision Pathway: Life Cycle Assessment and Process Map (e90087)
Zahra Ahmed, Alexander Zargaran, Sara Moeini, Sara Sousi, David Zargaran, Ann-Marie Quigley, Norbert Kang, Afshin Mosahebi. 198

Quality and Readability of ChatGPT and Gemini Responses to Patient Questions About Systemic Isotretinoin: Comparative Evaluation Study (e99780)
Elif Bal Avcı, Bu ra Da ta , Dilara lhan Erdil, Ozan Erdem, Andaç Salman, Selçuk Korkmaz, Sezgi Sarıkaya Solak. 208

Exploring the Lived Experience of Acne in the United States and the United Kingdom: Social Media Analysis (e91126)
John S Barbieri, Valentine Delore, Isabelle Vieillard, Sandrine Samson, Esteban Puentes, Joelle Malaab, Melanie Sambrès, Manissa Talmatkadi, Stéphane Schück. 217

Promotion of Tanning Behaviors on Instagram: Content Analysis (e110157)
Tailisha Gonzalez, Grace Hillyer, Bailey Gold, Corey Basch. 229

Skin Health Information Seeking on Short Video Platforms in Indonesia: Mixed Methods Approach (e93461)
Fathi Ahimsa, Kelvin Saputra, Michelle Hutasoit, Putu Handayani, Hana Fitriani. 240

XSL•FO
RenderX

Harmonized Dual Deep Learning Architectures for Image-Based Diagnostics of Skin Neglected Tropical Diseases: Benchmark Study via Novel Funnel Framework (e91544)
Yohannes Minyilu, Mohammed Yimer, Million Meshesha. . . . 264

Deep Learning–Based Classification of Melanoma Using Dermoscopic Images: Comparative Evaluation of Convolutional Neural Network and DenseNet Models (e104086)
Harshal Sanghvi, Argyrios Agritis, Isabella Alvarez, Mitsu Patel, Shuwen Wang, Riki Patel, Raphael Banoub, Timonthy Makkar. . . . 288

AI-Based Grading of Facial Aging Signs and Prediction of Visible Age From Standardized Facial Images: Algorithm Development and Internal Evaluation Study (e93236)
Frederic Flament, Panagiotis-Alexandros Bokaris, Tao Li, Aurelie Abric, Sileye Ba, Julien Despois, Hussein Jouni, Laudine Bertrand, Paul Michel, Thomas Dauxerre, Diala Haykal, Hughes Cartier. . . . 403

Online Discussions About Alopecia Areata in 5 Arabian Gulf Countries: Social Media Listening Study (e78262)
Ashraf Reda, Ghadear AL Alawi, Ameen Al Awadhi, Amina Al Obaidli, Fadwa Ebeid, Haytham Ahmed. . . . 461

Assessing Alopecia Areata Misinformation: Social Media Analysis (e90088)
Shanika Francis, Ank Agarwal, Aaron Fanous, Dawn Siegel, Roxana Daneshjou. . . . 471

Climate, Humidity, and Population-Level Interest in Dry Skin: Infodemiology Analysis Using Google Trends Across the United States (e93639)
Kimiya Aframian, Shaya Naimi, Joy Xu, Gordon Bae. . . . 477

Corrigenda and Addendas

Corrigendum: Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study (e106381)
JMIR Editorial Office. . . . 171

Correction: Informed Consent Practices for Publication of Patient Images in Dermatology Journals (e94194)
Toluwani Taiwo, Bianca Obiakor, Sarah McClung, Kanade Shinkai. . . . 397

Research Letters

The Role of TikTok in Education on Hidradenitis Suppurativa in Skin of Color: Cross-Sectional Analysis (e71566)
Arsema Zadu, Jordan Young, Janyla Seltzer, Angel Byrd, Cheri Frey. . . . 237

Patient Perceptions of Climate Change Impacts on Atopic Dermatitis: Cross-Sectional Survey Study (e80679)
Gunnar Mattson, Sarah Coates, Amanda Twigg. . . . 338

Shifting Perceptions of Cosmetic Procedures: Rise of Medical Spas in Online Search Trends (e83509)
Hong Ye, Sophia Zhu, Britta Verhey, Joy Xu, Gordon Bae. . . . 342

Public Interest in Sun Protection and Misinformation: A Google Trends and TikTok Analysis (e99635)
Rajyk Bhala, Sriram Ambadi, Reid Oldenburg. . . . 346

Differences in Electronic Consultation Conversion Rates Between Advanced Practice Providers and Board-Certified Dermatologists (e83922)
Dakota Hitchcock, Sabrina Newman. . . . 350

XSL•FO
RenderX

Analysis of Melanoma Educational Content and Skin of Color Representation on TikTok: Cross-Sectional Study (e87064)
Daniella Jaguan, Zoya Siddiqui, Jennifer Choi. . . . 354

Public Interest in Janus Kinase (JAK) Inhibitors for Alopecia Areata: A Google Trend Analysis (e75119)
Jade Howard, Nourine Kamili, Hala Idris, Loren Krueger. . . . 359

What's Trending: Cross-Sectional Analysis of the Content and Quality of Popular Melanoma Videos on TikTok (e79250)
Nazaneen Saleh, Alexa Rahaman, Emily Siegel. . . . 362

Reviews

The Rise in Homemade Sunscreen Trends and Future Impacts on Skin Cancer Risk: Systematic Review (e84694)
Gabi Kaftan, Mia Panillo, Lydia Maxon, Emily Esposito. . . . 308

Navigating the Intersection of Radiofrequency Microneedling and Surgical Facelifts: Scoping Review (e78385)
Mia Panlilio, Rebecca Bolen, Olnita Martini, Alexa Bonk, John Tedesco. . . . 315

Informatics-Based Psychotherapeutic and Psychiatric Interventions in Dermatology: Scoping Review of Impacts on Skin Disease Severity and Mental Health Outcomes (e82096)
Caroline Lamarre, Jeffrey Chivinski, Alexandre Hudon. . . . 323

Case Reports

Asymptomatic Tattoo-Associated Cutaneous Nontuberculous Mycobacterial Infection: Case Report (e108587)
Chris Alihosseini, Milana Stein, Ana Madiedo, Roksana Hesari, Matthew Kauffman, Martin Zaiac. . . . 365

Atypical Fibroxanthoma Presenting as a Rapidly Enlarging Tender Papule: Case Report (e84550)
Elizabeth Tchernogorova, Nicole Skalka, Mary Tidwell, Binh Nguyen, Julien Lanoue, Robin Burger, McKenna Abercrombie. . . . 370

Chronic Facial Abscess Mimicking Cervicofacial Actinomyces From Dermal Filler Migration: Case Report (e80278)
Monika Ziogaite, Sarah Mannlein, Nicole Bender, Scott Mahlberg. . . . 378

Anterior Cervical Hypertrichosis (Hairy Throat Syndrome): Pediatric Case Report and Brief Literature Review (e95391)
Nancy Shehata, Husna Thalib, Heba Alahwal. . . . 384

Topical Recombinant Collagen Following Fractional Radio Frequency Microneedling for Nonsurgical Facelift: Case Report (e86892)
Thien-Chong Wong, Yizhi Ong. . . . 392

XSL•FO
RenderX

Editorial

Understanding Gender Wage Disparities in Dermatology (e65003)
Danielle Duffle, Ramiro Rodriguez, Benjamin Brooks, Robert Dellavalle. 399

Viewpoints

Local Deployment of Open-Weight Language Models in Dermatology: Viewpoint on Privacy, Equity, and Practical Implementation (e94764)
William Nahm, Emily Yin, Emily Milam, Jason Weed. 422

Agentic AI in Dermatology: A Call to Action (e94909)
Brian Chu, Heather Shen, Ivy Lee, Jules Lipoff. 434

Beta-Alanine and Aquagenic Pruritus: Proposed Neuroimmune Mechanism (e90737)
Natalie Piserchio, Bailey Baratta, Benjamin Brooks, Brandon Muse, Katelin Ball. 438

Current Management and Unmet Needs in Alopecia Areata: Viewpoint From a Malaysian Perspective (e89583)
Azura Mohd Affandi, Adawiyah Jamil, Noor Azizan, Felix Yap, Kin Leong, Agnes Heng, Siti Nuraihan, Wooi Tan, Latha Selvarajah, Teeba Raja, Evelyn Yap, Peter Ch'ng, Tze Teoh, Mei Tey. 441

AI and Digital Tools in Dermatology: Addressing Access and Misinformation (e79044)
Dominique du Crest, Monisha Madhumita, Wendemagegn Enbiale, Jose Ruiz Postigo, Josep Malvehy, Shannon Wongvibulsin, Somesh Gupta, Harald Kittler, Charbel Skayem, Anjali Mahto, Adewole Adamson, Jules Lipoff, Art Papier, Hugues Cartier, Sébastien Garson, Esther Freeman.
4 5 1

Retraction

Retraction: Assessment and Evaluation of Social Engagement in Dermatology Residency Programs on Instagram: Cross-sectional Study (e106666)
JMIR Publications Editorial Office. 484

XSL•FO
RenderX

Accuracy, Reliability, and Quality of YouTube and TikTok Videos as Sources of Information on Urticaria: Cross-Sectional Study

Abrar Bukhari[1], MD; Saad Altalhab[1], MD; Suliman AlFaraj[2], MD; Huda Alrwebah[3], MD; Mishari Alrubaiaan[4], MD; Abdulrahman Alfawzan[3], MD

[1]Department of Dermatology, College of Medicine, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

[2]Department of Dermatology and Dermatologic Surgery, Prince Sultan Military Medical City, Riyadh, Saudi Arabia

[3]Department of Pediatric Dermatology, King Abdullah Specialized Children's Hospital, National Guard Health Affairs (NGHA), Riyadh, Saudi Arabia

[4]Department of Dermatology, King Abdulaziz Medical City, National Guard Health Affairs (NGHA), King Abdullah District, 4551, Riyadh, Saudi Arabia

Corresponding Author:
Mishari Alrubaiaan, MD
Department of Dermatology, King Abdulaziz Medical City, National Guard Health Affairs (NGHA), King Abdullah District, 4551, Riyadh, Saudi Arabia

Abstract

Background: The era of social media platforms has transformed the dissemination of health information, raising concerns about the quality and accuracy of content. Urticaria, a common skin condition, necessitates reliable information for effective management. Unreliable information on the internet can lead to misconceptions about urticaria, which could contribute to overemphasis on triggers or underrepresentation of guideline-based interventions, delaying effective management approaches.

Objective: This study aimed to systematically evaluate urticaria-related content on 2 prominent social media platforms, YouTube and TikTok, using validated tools to assess information reliability and explore the characteristics of videos and their uploaders.

Methods: On February 16, 2023, a comprehensive search on YouTube and TikTok was conducted for videos related to urticaria. Two independent dermatologists rated 92 eligible videos using DISCERN, the Global Quality Score (GQS), and the Accuracy in Digital-Health Instrument (ANDI). Statistical analyses, including mean values, SDs, percentages, and chi-square tests, were used to interpret the data.

Results: TikTok videos were predominantly uploaded by patients (29/46, 63%), whereas YouTube exhibited a higher engagement of health care professionals (31/46, 67.4%). YouTube videos exhibited higher mean scores for the GQS (2.9, SD 0.7), ANDI (2.3, SD 0.8), and DISCERN (40.2, SD 7.6) compared to TikTok (GQS: mean 2, SD 0.8; ANDI: mean 1.4, SD 1; DISCERN: mean 32.9, SD 7.9). Videos uploaded by professionals on TikTok had significantly higher mean GQS (2.8, SD 0.7), ANDI (2.1, SD 0.8), and DISCERN (40, SD 6.3) scores compared to those uploaded by nonprofessionals (P=.001). Similarly, videos from professional uploaders on YouTube exhibited significantly higher mean GQS (3, SD 0.6), ANDI (2.6, SD 0.6), and DISCERN (42.2, SD 7.1) scores compared to nonprofessionals (P=.01 for GQS, P=.001 for ANDI, and P=.003 for DISCERN).

Conclusions: Urticaria videos on TikTok showed a significantly higher engagement ratio and less accuracy and quality of health-related information than YouTube videos. These insights can guide content creators, platform administrators, and health care professionals in collaboratively fostering a trustworthy and accurate health information environment on social media.

(JMIR Dermatol 2026;9:e88827) doi:10.2196/88827

KEYWORDS

urticaria; social media; YouTube; TikTok; DISCERN; Global Quality Score; GQS; Accuracy in Digital-Health Instrument; ANDI; content quality; health care professionals; patient uploaders

Introduction

Urticaria, commonly known as hives, is a prevalent dermatological condition characterized by the sudden appearance of red, itchy edematous skin lesions. The condition has a lifetime prevalence of nearly 20% [1]. These lesions vary in size and shape and often emerge as a result of an allergic reaction or other triggers, such as immunoglobulin E (IgE)–mediated hypersensitivity to drugs or insect venoms and non–IgE-mediated mechanisms, that prompt the release of histamine from mast cells [2-4]. Non–IgE-mediated mechanisms in chronic spontaneous urticaria are triggered by immune complexes that bind immunoglobulin M and immunoglobulin G [5,6], whereas chronic inducible urticaria is triggered by physical stimuli, vaccines, or infections (viruses, parasites, and bacteria) [2]. The condition's transient nature and the potential

XSL•FO
RenderX

for severe discomfort make it a significant health concern affecting individuals of all ages [7]. Urticaria management is based on established clinical recommendations, which require medical practitioners to follow the clinical history (to make the initial diagnosis and examine the etiological aspects involved), carry out diagnoses to rule out other dermatological disorders, and examine the skin to detect the lesions [1].

Following examination, medical practitioners can recommend a standard dose of second-generation H1 receptor antihistamines as the first-line treatment [8]. Among individuals with persistent symptoms, the guidelines recommend increasing the dose of second-generation antihistamines up to four-fold as the second-line treatment [9,10]. In refractory cases, biologic treatment can be used to attain spontaneous resolution [11]. According to the American Academy of Allergy, Asthma & Immunology, there are various misconceptions concerning urticaria etiology, especially over attribution to dietary, stress, or pet causes [12]. Such misconceptions mostly contribute to patients seeking alternative interventions and explanations through digital platforms, smart devices, and computers outside formal medical guidance [13]. Additionally, while medical interventions are available for managing urticaria, the accessibility of information through digital platforms has led many individuals to seek insights and guidance from online sources [14].

The advent of social media has transformed the landscape of health communication, providing a platform for individuals to share personal experiences, seek advice, and access information [15]. User engagement in social media platforms has witnessed a substantial rise over the past years. This is evidenced by the fact that over half of the population in the world has an online social media account [16]. Such a rise was further intensified due to the global COVID-19 pandemic [17]. YouTube and TikTok, among other applications, have repeatedly proven to be the platforms with the most traffic from online consumers [18]. With millions of health-related videos circulating on these platforms and their broad reach and impactful potential [19,20], social media can act as a catalyst to propel valid and reliable educational materials forward [20].

The democratization of content creation allows patients, health care professionals, and various entities apart from health care organizations to contribute to the discourse surrounding health conditions. However, the unregulated nature of these platforms raises questions about the accuracy and quality of health-related information, particularly concerning dermatological conditions such as urticaria [21]. Generally, social media ranking algorithms optimized for users have revealed preferences, including user engagement such as comments, shares, likes, and watch times, to determine which content is ranked or pushed to wider audiences [22]. By concentrating on user engagement, the algorithms might amplify content that is popular but unaligned with scientific information or evidence-based guidelines. Consequently, low-quality or misleading medical-related content might gain traction to wider coverage, potentially shaping public perceptions, impacting decisions, and amplifying misinformation [23]. Such concerns are compounded by social media's widespread amplification of unverified information, which underscores the significance of systematically evaluating the reliability, quality, and accuracy of health-related content available on such platforms.

Health information on social media is often characterized by its dynamic and diverse nature. Patients share personal anecdotes, health care professionals offer expert advice, and entities ranging from medical groups to news companies contribute to the content pool [15]. Understanding the characteristics of urticaria-related content on platforms such as YouTube and TikTok is crucial considering the potential impact on public knowledge, perceptions, and health care–seeking behaviors [24].

The influence of social media on health-related decisions is a complex interplay of factors [19,25]. Patients often turn to these platforms seeking information that complements or even substitutes traditional health care channels [25]. The allure of relatable stories and visual content can provide a sense of connection and understanding, fostering a community of individuals navigating similar health challenges [24,25]. However, the lack of regulation and oversight raises concerns about the accuracy and reliability of the information disseminated, especially when it comes to conditions such as urticaria, where misinformation can impact treatment choices and overall well-being [21,26,27].

The concept of health literacy in the digital age has garnered attention, emphasizing individuals' ability to access, understand, and apply health information from online sources [24,28]. For dermatological conditions such as urticaria, where self-management and lifestyle choices play a role, the impact of online information becomes even more pronounced. As individuals navigate the plethora of content available, it becomes imperative to assess the quality, accuracy, and reliability of health-related information on these platforms [27].

Existing literature highlights the challenges posed by health misinformation on social media and its potential consequences for public health [25]. While studies have explored the landscape of health-related content on platforms such as YouTube, TikTok's role in shaping health perceptions is relatively uncharted territory. TikTok's format, characterized by short videos and a younger user demographic, introduces unique considerations in understanding the dynamics of health communication [29,30]. Nevertheless, the lack of content regulation on social media platforms raises concerns about the reliability and accuracy of the information disseminated. In the context of urticaria, misinformation might promote misconceptions on the etiology (eg, diet, treatments, causes, and supplementation), potentially contributing to delayed management [31].

As the prevalence of urticaria persists and individuals increasingly turn to online platforms for health information, there is a pressing need to critically examine the quality of content available. This study endeavored to fill a gap in the literature by systematically evaluating urticaria-related content on YouTube and TikTok. We aimed to assess and compare the quality, accuracy, and reliability of urticaria-related information available on YouTube and TikTok using validated assessment tools. Given the variations in the presentation, content creation features, and format of the videos between the 2 platforms, we

XSL•FO
RenderX

hypothesized that the reliability and accuracy of the information would vary. Through the application of validated scoring criteria and a rigorous methodology, we aimed to provide insights into the characteristics of videos, the role of different uploader types, and the overall quality of the health information available on these platforms. The findings of this study hold implications for health care professionals, content creators, and platform administrators in shaping a reliable and informative health communication landscape in the digital age.

Methods

Study Design

This study used a cross-sectional observational design to assess the quality of urticaria-related content on YouTube and TikTok.

Search Strategy

Search Terms

The search for relevant content was conducted on February 16, 2023. To ensure a comprehensive representation of urticaria-related videos, we used clinically and standardized search terms: “urticaria” and “hives.”

Search Procedure and Sampling Strategy

On TikTok, searches were performed in both incognito mode and through an author’s account. The results from both approaches were merged before assessing duplication and screening according to eligibility criteria. The first or top 50 videos for each search term were selected, sorted by views, and included in the dataset to ensure relevant sorting. Duplicate videos were defined as identical videos uploaded several times (including by different users). We only relied on unique videos (1 instance) to ensure that observation of each one represented unique content while avoiding overrepresentation of duplicated material. We retained only 1 version of the content when it was uploaded multiple times. We identified duplicates through comparison of the upload metadata, audiovisual content, titles, and platform-specific video identifiers (URLs).

A parallel process was applied to YouTube. Duplicate videos were removed to ensure the uniqueness of the dataset.

Eligibility Criteria

We included publicly accessible English-language videos directly related to urticaria and excluded videos exceeding 20 minutes in duration.

Video Classification and Characteristics

The final dataset consisted of 92 videos, reflecting a diverse representation of urticaria-related content. The compiled dataset underwent thorough characterization, extracting key variables for analysis. Video characteristics that we included were views; likes; shares; followers; verification status; gender; and presence of animation, subtitles, music, written material, educational content, anecdotal information, home remedies, advertisements, and other elements. We categorized the videos based on the uploader type, specifically, health care professionals or nonprofessionals. Furthermore, we classified video content types as animation, music, educational, anecdotal, home remedy, advertisement, written material, and subtitles.

Quality Assessment

Two independent expert dermatologists were involved in the video assessment process, ensuring a comprehensive and professional assessment. To evaluate the quality and accuracy of urticaria-related content, 3 established tools were used: DISCERN [32], the Global Quality Score (GQS) [33], and the Accuracy in Digital-Health Instrument (ANDI) [34]. DISCERN is a tool comprising 16 questions rated from 1 to 5, where 1 means “no” and 5 means “yes,” while 3 is “partial” [32]. It measures 3 aspects, including reliability, quality of information, and overall rating of the videos. The GQS and ANDI’s primary aim is to provide a standardized method for assessing the overall quality of digital health content. All tools used are provided in Multimedia Appendix 1.

Platform Comparisons

Comparisons were performed between TikTok and YouTube to assess differences in video characteristics and quality assessment scores. Within each platform, videos were also compared according to uploader type (professional vs nonprofessional) with respect to views; likes; shares; followers; daily views; engagement percentage; video duration; and DISCERN, GQS, and ANDI scores. DISCERN score categories were also compared between professional and nonprofessional uploaders within each platform.

Statistical Analysis

We summarized video characteristics and selection using means and SDs provided they were continuous variables, while we used frequencies and percentages for categorical variables. Similarly, we used means and SDs to summarize quality assessment of TikTok and YouTube urticaria videos and platform comparison. The distribution of continuous variables was assessed using the Shapiro-Wilk test. Normally distributed continuous variables were summarized as means and SDs and compared using parametric tests, whereas nonnormally distributed variables were summarized as medians and IQRs and compared using nonparametric tests. For comparisons between TikTok and YouTube, the 2-tailed independent-samples *t* test was used for normally distributed continuous variables, and the Mann-Whitney *U* test was used for nonnormally distributed continuous variables. Categorical variables were compared using the chi-square test. The level of significance was set at .05.

Results

Search Results

On TikTok, 63 videos were identified through the incognito search, and 76 were identified through the author’s account. Across the 2 search terms, “urticaria” retrieved 75 videos, whereas “hives” retrieved 64 videos. Following the removal of duplicates and irrelevant videos based on the inclusion criteria, 46 unique TikTok videos were included in the final analysis. On YouTube, the search yielded 85 videos, of which 51 were retrieved using “urticaria” and 34 were retrieved using “hives.”

XSL•FO
RenderX

Following deduplication and application of the inclusion criteria, 46 unique YouTube videos were included in the final analysis. Overall, 92 videos, comprising 46 videos from each platform, were included in the final analysis.

Video Classification and Characteristics

Table 1 shows the characteristics of the uploading channels and uploaded videos on TikTok and YouTube, providing insights into the diverse landscape of content creators and their respective platforms. All TikTok videos had humans, while 80.4% (37/46) of the videos on YouTube had humans. In addition, there was a predominant presence of patients on TikTok (29/46, 63%) compared to YouTube (5/46, 10.9%).

Table . Characteristics of uploading channels and uploaded videos on TikTok and YouTube (N=46).

Parameters	TikTok	YouTube
Uploader, n (%)		
Not professional	31 (67.4)	15 (32.6)
Professional	15 (32.6)	31 (67.4)
Views		
Mean (SD; range)	631,513 (1,040,097; 63,300-4,800,000)	171,130 (336,435; 11,000-1,800,000)
Median (IQR)	330,050 (102,625-526,150)	62,000 (35,000-179,000)
Likes		
Mean (SD; range)	29,688 (66,353; 175-324,200)	1761 (3046; 0-17,000)
Median (IQR)	3907 (621-26,050)	595 (308-1425)
Shares		
Mean (SD; range)	410 (772; 11-4259)	419 (614; 0-2916)
Median (IQR)	104 (40-407)	189 (50-529)
Verified, n (%)	3 (6.5)	21 (45.7)
Sex, n (%)		
No human in the video	0 (0)	9 (19.6)
Male	8 (17.4)	13 (28.3)
Female	38 (82.6)	18 (39.1)
Both	0 (0)	6 (13)
Animation, n (%)	1 (2.2)	11 (23.9)
Subtitles, n (%)	11 (23.9)	6 (13)
Music, n (%)	26 (56.5)	23 (50)
Written material, n (%)	44 (95.7)	42 (91.3)
Educational, n (%)	19 (41.3)	39 (84.8)
Anecdotal, n (%)	27 (58.7)	9 (19.6)
Home remedy, n (%)	4 (8.7)	13 (28.3)
Advertisement, n (%)	3 (6.5)	12 (26.1)

Engagement metrics varied across the 2 platforms, as evidenced by the differences in the mean views, likes, shares, and followers between TikTok and YouTube. TikTok videos exhibited a higher mean number of views (631,513, SD 1,040,097; median 330,050, IQR 102,625-526,150) compared to YouTube (mean 171,130, SD 336,435; median 62,000, IQR 35,000-179,000). Additionally, TikTok videos garnered substantially more likes (mean 29,688, SD 66,353; median 3907, IQR 621-26,050) and shares (mean 410, SD 772; median 104, IQR 40-407) than YouTube videos (mean 1761, SD 3046 and median 595, IQR 308-1425 likes; mean 419, SD 614 and median 189, IQR 50-529 shares). In addition, YouTube exhibited slightly comparable mean numbers of shares (419, SD 614) to those of TikTok (410, SD 772). Moreover, a noteworthy proportion of YouTube videos were verified (21/46, 45.7%), indicating potential credibility, while most of the videos on TikTok (43/46, 93.5%) were not.

TikTok videos were less likely to include animation (1/46, 2.2%) and more likely to include subtitles (11/46, 23.9%) compared to YouTube (11/46, 23.9% and 6/46, 13%, respectively). Health professionals, including physicians, nonphysicians, and medical groups or foundations, were more active on YouTube (31/46, 67.4%) than on TikTok (15/46, 32.6%). Written material content comprised most videos across TikTok (44/46, 95.7%) and

XSL•FO
RenderX

YouTube (42/46, 91.3%), followed by educational and musical content.

Quality Assessment

Regarding the evaluation of videos using DISCERN, GQS, and ANDI scores by 2 independent expert dermatologists, the DISCERN scores, indicative of the quality of information, demonstrated a moderate level, with rater 1 assigning a mean score of 37 (SD 11) and rater 2 assigning a mean score of 36 (SD 13). The interrater reliability, as measured using the Cronbach α, was low (0.102), suggesting variations in scoring between the raters. The GQS and ANDI scores exhibited higher levels of agreement between raters, with mean scores of 3 (SD 1) and 3 (SD 1), respectively, and higher interclass correlation coefficients (0.619 and 0.766, respectively).

Table 2 delves into the comparison of average DISCERN, GQS, and ANDI scores across different platforms. Notably, YouTube videos demonstrated higher mean scores for the GQS (2.9, SD 0.7), ANDI (2.3, SD 0.8), and DISCERN (40.2, SD 7.6) compared to TikTok (GQS: mean 2, SD 0.8; ANDI: mean 1.4, SD 1; DISCERN: mean 32.9, SD 7.9), indicating a potential disparity in the quality and accuracy of the information between the 2 social media platforms (*P*=.001).

Table . Average DISCERN, Global Quality Score (GQS), and Accuracy in Digital-Health Instrument (ANDI) scores among different platforms.

Items	Platform, mean (SD)		Mean difference (95% CI)	*P* value[a]
	TikTok	YouTube		
GQS (1-5)	2 (0.8)	2.9 (0.7)	−0.82 (–1.5 to –0.59)	.001
ANDI (0-4)	1.4 (1)	2.3 (0.8)	−0.95 (–1.5 to –0.59)	.001
DISCERN (16-80)	32.9 (7.9)	40.2 (7.6)	−7.28 (–1.4 to –0.51)	.001

[a]The 2-tailed independent-samples *t* test was used to calculate the *P* values.

Platform Comparisons

Table 3 provides a comprehensive analysis of TikTok videos, categorizing them based on the uploader type and comparing various metrics such as views, likes, shares, followers, daily views, engagement percentage, and video duration. Interestingly, TikTok videos from nonprofessionals exhibited higher mean views (670,129, SD 981,086) than those from professionals (551,707, SD 1,185,129), although the difference was not statistically significant (*P*=.72). However, nonprofessionals had nonsignificantly higher shares (mean 397, SD 778) compared to professionals (mean 437, SD 786; *P*=.87). The daily views (*P*=.84) and engagement percentages (*P*=.08) did not show significant differences between the two 2 groups. In terms of video quality assessment, videos from professionals on TikTok had significantly higher mean GQS (2.8, SD 0.7), ANDI (2.1, SD 0.8), and DISCERN (40, SD 6.3) scores compared to nonprofessionals (*P*=.001 for all three comparisons), indicating a potential association between professional status and the quality of urticaria-related content on TikTok (Figure 1).

XSL•FO
RenderX

Table . Video characteristics and average DISCERN, Global Quality Score (GQS), and Accuracy in Digital-Health Instrument (ANDI) scores among TikTok videos compared by uploader type.

Items	Nonprofessionals	Professionals	*P* value[a]	Mean difference (95% CI)
Views			.72	—[b]
Mean (SD)	670,129 (981,086)	551,707 (1,185,129)		
Median (IQR)	61,000 (43,000-171,000)	69,000 (34,500-195,500)		
Likes			.60	—
Mean (SD)	33,337 (64,594)	22,145 (71,556)		
Median (IQR)	552 (292-1900)	631 (321-1450)		
Shares			.87	—
Mean (SD)	397 (778)	437 (786)		
Median (IQR)	196 (24-667)	181 (61-485)		
Followers			.09	—
Mean (SD)	87,335 (151,167)	891,109 (2,598,803)		
Median (IQR)	189,000 (6120-1,532,500)	148,000 (22,800-1,655,000)		
Daily views, mean (SD)	2798.1 (4318)	3089.9 (4644.1)	.84	—
Engagement (%), mean (SD)	5.84 (7.73)	2.15 (2.68)	.08	—
Duration (s), mean (SD)	46 (45)	58 (46)	.44	—
GQS (1-5), mean (SD)	1.7 (0.7)	2.8 (0.7)	.001	0.61 (0.20-1.53)
ANDI (0-4), mean (SD)	1 (0.9)	2.1 (0.8)	.001	1.01 (0.76-2.19)
DISCERN (16-80), mean (SD)	29.5 (6.2)	40 (6.3)	.001	7.14 (0.35-1.70)

[a]*P* values are unadjusted. The 2-tailed independent-samples *t* test was used to calculate them.

[b]Not applicable.

Figure 1. DISCERN score categories across different platforms.

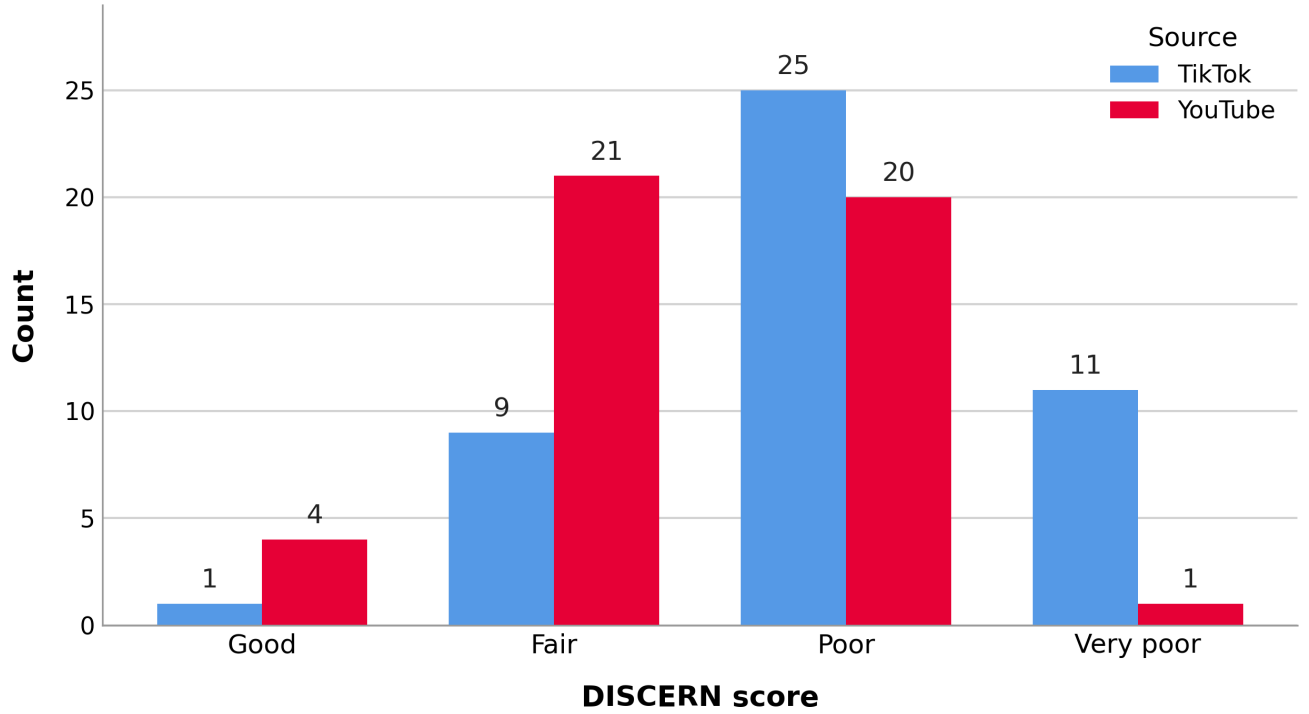


Table 4 extends a similar comparison to YouTube videos based on uploader types. Unlike on TikTok, YouTube videos from professionals had higher mean views (194,545, SD 391,001) compared to nonprofessionals (111,692, SD 106,941), although the difference was not statistically significant (*P*=.46). No significant differences were observed in likes (*P*=.56), shares

XSL•FO
RenderX

(*P*=.99), followers (*P*=.41), daily views (*P*=.40), engagement percentage (P=.61), and video duration (P=.30) between nonprofessional and professional YouTube uploaders. However, videos from professional uploaders on YouTube exhibited significantly higher mean GQS (3, SD 0.6), ANDI (2.6, SD 0.6), and DISCERN (42.2, SD 7.1) scores compared to nonprofessionals (*P*=.01 for GQS, *P*=.001 for ANDI, and *P*=.003 for DISCERN), suggesting a potential association between professional status and the quality of urticaria-related content on YouTube.

Table . Video characteristics and average DISCERN, Global Quality Score (GQS), and Accuracy in Digital-Health Instrument (ANDI) scores among YouTube videos compared by uploader type.

Items	Nonprofessionals	Professionals	*P* value[a]	Mean difference (95% CI)
Views			.46	—[b]
Mean (SD)	111,692 (106,941)	194,545 (391,001)		
Median (IQR)	375,700 (83,300 - 757,600)	237,200 (134,300 - 368,200)		
Likes			.56	—
Mean (SD)	1342 (1480)	1927 (3482)		
Median (IQR)	5389 (784 - 46,900)	138 (331-7659)		
Shares			.99	—
Mean (SD)	421 (625)	418 (619)		
Median (IQR)	149 (30-498)	74 (43-298)		
Followers			.41	—
Mean (SD)	1,641,722 (4,006,427)	930,805 (1,834,514)		
Median (IQR)	14,100 (2894 - 78,000)	61,700 (23,400 - 547,500)		
Daily views, mean (SD)	101.5 (162.3)	190.1 (355.6)	.40	—
Engagement (%), mean (SD)	1.21 (0.54)	1.1 (0.67)	.61	—
Duration (s), mean (SD)	401 (325)	305 (259)	.30	—
GQS (1-5), mean (SD)	2.4 (0.9)	3 (0.6)	.01	1.07 (0.86-2.25)
ANDI (0-4), mean (SD)	1.6 (0.9)	2.6 (0.6)	.001	1.10 (0.58-1.92)
DISCERN (16-80), mean (SD)	35.1 (6.6)	42.2 (7.1)	.003	8.16 (0.98-2.41)

[a]The 2-tailed independent-samples *t* test was used to calculate the *P* values.

[b]Not applicable.

Table 5 delves into the distribution of DISCERN score categories among professional and nonprofessional uploaders on TikTok and YouTube. On TikTok, among videos categorized as “fair,” 77.8% (7/9) were uploaded by professionals. Among videos categorized as “poor,” 72% (18/25) were uploaded by nonprofessionals, while all videos categorized as “very poor” (11/11, 100%) were uploaded by nonprofessionals. This discrepancy was statistically significant (χ^{2}_{3} =16.0; *P*=.001). On YouTube, the distribution of DISCERN score categories did not reach statistical significance (χ^{2}_{3}=6.4; *P*=.10), although videos from professionals tended to fall into the higher categories (“good” and “fair”) compared to nonprofessionals.

XSL•FO
RenderX

Table . DISCERN score categories among professional vs nonprofessional uploaders on TikTok and YouTube.

DISCERN score (16-80)	Uploaders, n/N (%)		Chi-square (*df*)	*P* value[a]
	Nonprofessional	Professional		
TikTok			16.0 (3)	.001
Very good	0 (0)	0 (0)		
Good	0 (0)	1/1 (100)		
Fair	2/9 (22.2)	7/9 (77.8)		
Poor	18/25 (72)	7/25 (28)		
Very poor	11/11 (100)	0 (0)		
YouTube			6.4 (3)	.10
Very good	0 (0)	0 (0)		
Good	0 (0)	4/4 (100)		
Fair	4/21 (19)	17/21 (81)		
Poor	8/20 (40)	12/20 (60)		
Very poor	1/1 (100)	0 (0)		

[a]The chi-square test was used to calculate the *P* values.

Figure 1 shows the DISCERN categories among TikTok and YouTube videos. Under the "good" and "fair" categories, YouTube videos were more predominant than TikTok videos. On the other hand, TikTok videos were predominant under the "poor" and "very poor" categories compared to YouTube videos.

Discussion

Video Characteristics

Urticaria, characterized by the sudden onset of raised, itchy welts on the skin, poses a significant burden on individuals and necessitates reliable information sources for effective management [4]. With the rise of the social media era, platforms such as YouTube and TikTok have become accessible venues for disseminating health-related content [24]. However, the quality and accuracy of health information on these platforms remain a concern. In this context, our study aimed to systematically evaluate urticaria-related content on YouTube and TikTok using validated tools to assess the information's reliability and the characteristics of the videos and their uploaders.

Our comprehensive analysis of 92 videos, spanning diverse characteristics and platforms, revealed intriguing patterns in content quality and uploader types. TikTok videos had almost 4 times the views of YouTube videos and more than 20 times the daily views, with a much higher engagement ratio. Notably, TikTok videos demonstrated a predominant presence of patient uploaders (29/46, 63%), while YouTube exhibited a higher engagement of health care professionals (31/46, 67.4%). The observed significant difference in viewer engagement and content characteristics between the platforms underlines the need for changing the distribution of health-related content across diverse social media spaces [28] as health care professionals are more likely to use YouTube for their content, which has significantly lower views and engagement ratios.

The variation in video characteristics can also be attributed to the short form of TikTok videos. In particular, educational content was slightly higher on YouTube than on TikTok, which is advantageous in urticaria, a condition that necessitates broader understanding. Specifically, YouTube is designed to hold long-form videos, allowing content for over 1 or 2 hours [35]. In such cases, long-form videos support nuanced comprehension of initiators, stepwise treatment escalation, and chronicity. The step-by-step explanations on YouTube might lead to comprehensive content, reducing common misconceptions. In contrast, the brevity of content on TikTok creates oversimplification of information, limiting discussion of guideline-based handling of urticaria.

DISCERN, GQS, and ANDI Scores

Our study found that TikTok videos from professionals exhibited significantly higher mean scores across the DISCERN, GQS, and ANDI measures, suggesting a positive association between professional status and the reliability of urticaria-related information. This aligns with the notion that health care professionals bring a level of expertise and accuracy to content creation, potentially enhancing the overall quality of health-related information on social media platforms [36].

Even though condensed and short-form TikTok videos might limit the comprehensive coverage of educational content, online information is likely influenced by content creator attributes. Hill and colleagues [37] note that the quality of TikTok information consumed can be impacted by multiple factors, including content creator characteristics such as their professional medical training, reliance on evidence-based sources, and motivation behind their videos. A similar observation was made by Molitierno et al [38], who noted that videos from health care professionals demonstrated a higher actionability, accuracy, and understandability compared to patient-generated content regarding induction of labor. Consequently, the length of the videos is not the sole determiner

XSL•FO
RenderX

of information quality as the characteristics of the creators play a crucial role across social media platforms.

On YouTube, a similar trend was observed, with videos from professional uploaders surpassing those of their nonprofessional counterparts in mean GQS, ANDI, and DISCERN scores. High-quality videos from professionals reinforce the significance of expert-driven communications in diseases such as urticaria [39]. Putting stress on high-quality information among such conditions is vital as management decisions are based on evolving clinical evidence and necessitate individualized assessment [39]. This consistency across platforms underscores the potential influence of the uploader's background on the quality of urticaria-related content [31]. The higher scores among professional uploaders may be attributed to their expertise, commitment to evidence-based information, and adherence to professional standards [36].

For both platforms, the variations in the quality observed have vital implications for urticaria management. As stated by Zhang et al [40], physicians and patients should be aware of the internet being a source of health-related misinformation even though it is a valuable information source. Similarly, given that chronic urticaria is frequently idiopathic, inaccurate or incomplete information may lead patients to pursue unnecessary trigger elimination strategies, particularly restrictive diets, while delaying escalation to evidence-based therapies such as high-dose antihistamines or biologics. Consequently, it is vital for patients and physicians to critically evaluate YouTube and TikTok information on urticaria as misconceptions might lead to suboptimal treatment, discord between patients and physicians, and delayed diagnosis [41].

This study further explored the impact of various video characteristics on the DISCERN score categories. Notably, nonprofessional uploaders exhibited a higher prevalence of videos in the "poor" and "very poor" categories, emphasizing potential concerns about the quality of the information disseminated by nonprofessionals [42].

Our findings align with existing literature emphasizing the importance of health care professional involvement in health-related content creation. The positive correlation between professional status and content quality is consistent with studies indicating that expert-driven content tends to be more accurate, reliable, and aligned with evidence-based practices [43,44]. The prevalence of verified accounts contributing to higher-quality content corroborates prior research highlighting the role of platform verification in ensuring reliability [43].

Implications and Recommendations

Urticaria videos on TikTok showed a significantly higher engagement ratio and less accuracy and quality of health-related information than YouTube videos. To enhance the quality of urticaria-related content, collaborations between health care professionals and social media platforms should be encouraged to understand the preferences and expectations of viewers, guiding creators in developing engaging yet accurate health-related content. In addition, initiatives promoting verification and professional engagement can contribute to a more reliable health information environment.

Limitations

The cross-sectional nature of the analysis provides a snapshot of the current state of urticaria-related content on YouTube and TikTok. Furthermore, the evaluation was based on a specific set of criteria, and the subjective nature of video evaluation may introduce inherent biases. Third, we only included English-language videos in the analysis, which could introduce some limitations. In particular, the findings might not be generalizable to urticaria-related content in populations speaking languages other than English, where the accuracy, quality, or reliability of the information might differ. In addition, our study reflects a cross-sectional evaluation of digital information available at the time of data collection in 2023. Given the rapidly evolving nature of social media platforms such as TikTok and YouTube, the quality and characteristics of available content may have changed since then. However, such snapshot analyses remain important for establishing baseline trends and determining persistent gaps in the quality of online health information. Future studies with larger sample sizes, longitudinal designs, and more diverse evaluation criteria could further refine our comprehension of health-related content on various popular social media platforms and examine how content evolves over time.

Conclusions

In conclusion, our study offers an exploration of urticaria-related content on YouTube and TikTok, highlighting the influence of uploader types and video characteristics on the perceived quality and accuracy of the information. The findings underscore the importance of health care professional and professional societies' involvement in content creation and suggest avenues for improving the reliability of health-related information on social media platforms. Involvement in creating accessible, reliable, and evidence-based educational content could enhance the availability of accurate and guideline-based data that may help mitigate the spread of misleading information, support solid health care decisions, and elevate public understanding of urticaria. As these platforms continue to evolve, collaborative efforts among health care professionals, content creators, and platform administrators are essential to cultivate a trustworthy and accurate health information environment.

Acknowledgments

The authors would like to express their sincere appreciation to Dr Razan AlSubaie for her valuable contribution to this manuscript. Her thorough proofreading and thoughtful editorial support enhanced the clarity and quality of the final work.

XSL•FO
RenderX

Generative AI Use

The authors declare that no AI, machine learning, or generative AI tools were used in the conduct of this study or in the preparation of the original manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon request.

Authors' Contributions

All authors contributed to the study and manuscript preparation. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tools.
[PDF File, 40 KB - derma_v9i1e88827_app1.pdf]

References

1. Ryan D, Tanno LK, Angier E, et al. Clinical review: the suggested management pathway for urticaria in primary care. Clin Transl Allergy 2022 Oct;12(10):e12195. [doi: 10.1002/clt2.12195] [Medline: 36225262]
2. Ensina LF, Min TK, Félix MM, de Alcântara CT, Costa C. Acute urticaria and anaphylaxis: differences and similarities in clinical management. Front Allergy 2022;3:840999. [doi: 10.3389/falgy.2022.840999] [Medline: 35958944]
3. Antia C, Baquerizo K, Korman A, Alikhan A, Bernstein JA. Urticaria: a comprehensive review: treatment of chronic urticaria, special populations, and disease outcomes. J Am Acad Dermatol 2018 Oct;79(4):617-633. [doi: 10.1016/j.jaad.2018.01.023] [Medline: 30241624]
4. Beck LA, Bernstein JA, Maurer M. A review of international recommendations for the diagnosis and management of chronic urticaria. Acta Derm Venereol 2017 Feb 8;97(2):149-158. [doi: 10.2340/00015555-2496] [Medline: 27349620]
5. Church MK, Kolkhir P, Metz M, Maurer M. The role and relevance of mast cells in urticaria. Immunol Rev 2018 Mar;282(1):232-247. [doi: 10.1111/imr.12632] [Medline: 29431202]
6. Nguyen SM, Rupprecht CP, Haque A, Pattanaik D, Yusin J, Krishnaswamy G. Mechanisms governing anaphylaxis: inflammatory cells, mediators, endothelial gap junctions and beyond. Int J Mol Sci 2021 Jul 21;22(15):7785. [doi: 10.3390/ijms22157785] [Medline: 34360549]
7. Kolkhir P, Giménez-Arnau AM, Kulthanan K, Peter J, Metz M, Maurer M. Urticaria. Nat Rev Dis Primers 2022 Sep 15;8(1):61. [doi: 10.1038/s41572-022-00389-z] [Medline: 36109590]
8. Zuberbier T, Aberer W, Asero R, et al. The EAACI/GA(2) LEN/EDF/WAO guideline for the definition, classification, diagnosis, and management of urticaria: the 2013 revision and update. Allergy 2014 Jul;69(7):868-887. [doi: 10.1111/all.12313] [Medline: 24785199]
9. Guillén-Aguinaga S, Jáuregui Presa I, Aguinaga-Ontoso E, Guillén-Grima F, Ferrer M. Updosing nonsedating antihistamines in patients with chronic spontaneous urticaria: a systematic review and meta-analysis. Br J Dermatol 2016 Dec;175(6):1153-1165. [doi: 10.1111/bjd.14768] [Medline: 27237730]
10. Sussman G, Hébert J, Gulliver W, et al. Insights and advances in chronic urticaria: a Canadian perspective. Allergy Asthma Clin Immunol 2015;11(1):7. [doi: 10.1186/s13223-015-0072-2] [Medline: 25705232]
11. Ferrer M, Giménez-Arnau A, Saldana D, et al. Predicting chronic spontaneous urticaria symptom return after omalizumab treatment discontinuation: exploratory analysis. J Allergy Clin Immunol Pract 2018;6(4):1191-1197.e5. [doi: 10.1016/j.jaip.2018.04.003] [Medline: 29655772]
12. Facts and myths about urticaria (hives). American Academy of Allergy, Asthma & Immunology. URL: https://www.aaaai.org/tools-for-the-public/conditions-library/allergies/myths-hives [accessed 2025-08-07]
13. Cherrez-Ojeda I, Vanegas E, Cherrez A, et al. How are patients with chronic urticaria interested in using information and communication technologies to guide their healthcare? A UCARE study. World Allergy Organ J 2021 Jun;14(6):100542. [doi: 10.1016/j.waojou.2021.100542] [Medline: 34141048]
14. Maurer M, Weller K, Magerl M, et al. The usage, quality and relevance of information and communications technologies in patients with chronic urticaria: a UCARE study. World Allergy Organ J 2020 Nov;13(11):100475. [doi: 10.1016/j.waojou.2020.100475] [Medline: 33204387]

XSL•FO
RenderX

15. Sumayyia MD, Al-Madaney MM, Almousawi FH. Health information on social media. Perceptions, attitudes, and practices of patients and their companions. Saudi Med J 2019 Dec;40(12):1294-1298. [doi: 10.15537/smj.2019.12.24682] [Medline: 31828284]
16. Geist R, Militello M, Albrecht JM, et al. Social media and clinical research in dermatology. Curr Dermatol Rep 2021;10(4):105-111. [doi: 10.1007/s13671-021-00350-5] [Medline: 34840860]
17. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
18. Haenlein M, Anadol E, Farnsworth T, Hugo H, Hunichen J, Welte D. Navigating the new era of influencer marketing: how to be successful on Instagram, TikTok, & Co. Calif Manage Rev 2020;63(1):5-25. [doi: 10.1177/0008125620958166]
19. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. Annu Rev Public Health 2020 Apr 2;41:101-118. [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
20. Sobaih AE, Hasanein AM, Abu Elnasr AE. Responses to COVID-19 in higher education: social media usage for sustaining formal academic communication in developing countries. Sustainability 2020;12(16):6520. [doi: 10.3390/su12166520]
21. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
22. Milli S, Carroll M, Wang Y, Pandey S, Zhao S, Dragan AD. Engagement, user satisfaction, and the amplification of divisive content on social media. PNAS Nexus 2025 Mar;4(3):pgaf062. [doi: 10.1093/pnasnexus/pgaf062] [Medline: 40070432]
23. Trethewey SP. Medical misinformation on social media: cognitive bias, pseudo-peer review, and the good intentions hypothesis. Circulation 2019 Oct;140(14):1131-1133. [doi: 10.1161/CIRCULATIONAHA.119.041719] [Medline: 31567022]
24. DeBord LC, Patel V, Braun TL, Dao H Jr. Social media in dermatology: clinical relevance, academic value, and trends across platforms. J Dermatolog Treat 2019 Aug;30(5):511-518. [doi: 10.1080/09546634.2018.1530444] [Medline: 30265614]
25. Madke B, Gardner JM. Enhanced worldwide dermatology-pathology interaction via Facebook, Twitter, and other social media platforms. Am J Dermatopathol 2018 Mar;40(3):168-172. [doi: 10.1097/DAD.0000000000000963] [Medline: 28796693]
26. Mueller SM, Jungo P, Cajacob L, Schwegler S, Itin P, Brandt O. The absence of evidence is evidence of non-sense: cross-sectional study on the quality of psoriasis-related videos on YouTube and their reception by health seekers. J Med Internet Res 2019 Jan 16;21(1):e11935. [doi: 10.2196/11935] [Medline: 30664460]
27. Ranpariya V, Chu B, Fathy R, Lipoff JB. Dermatology without dermatologists? Analyzing Instagram influencers with dermatology-related hashtags. J Am Acad Dermatol 2020 Dec;83(6):1840-1842. [doi: 10.1016/j.jaad.2020.05.039] [Medline: 32416205]
28. Sierro TJ, Young PM, Kassabian SK, Wu KK, Armstrong AW. Dermatologists in social media: a study on top influencers, posts, and user engagement. J Am Acad Dermatol 2020 Nov;83(5):1452-1455. [doi: 10.1016/j.jaad.2020.03.001] [Medline: 32151630]
29. Szeto MD, Mamo A, Afrin A, Militello M, Barber C. Social media in dermatology and an overview of popular social media platforms. Curr Dermatol Rep 2021;10(4):97-104. [doi: 10.1007/s13671-021-00343-4] [Medline: 34692234]
30. Cooper BR, Concilla A, Albrecht JM, et al. Social media as a medium for dermatologic education. Curr Dermatol Rep 2022;11(2):103-109. [doi: 10.1007/s13671-022-00359-4] [Medline: 35493063]
31. Finnegan P, Murphy M, O'Connor C. Reinventing the wheal: a review of online misinformation and conspiracy theories in urticaria. Clin Exp Allergy 2023 Jan;53(1):118-120. [doi: 10.1111/cea.14246] [Medline: 36263623]
32. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
33. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007 Sep;102(9):2070-2077. [doi: 10.1111/j.1572-0241.2007.01325.x] [Medline: 17511753]
34. Salah LA, AlTalhab S, Omair A, AlJasser M. Accuracy and quality of YouTube videos as a source of information on vitiligo. Clin Cosmet Investig Dermatol 2022;15:21-25. [doi: 10.2147/CCID.S330015] [Medline: 35027834]
35. Seidel N. Short, long, and segmented learning videos: from YouTube practice to enhanced video players. Tech Know Learn 2024 Dec;29(4):1965-1991. [doi: 10.1007/s10758-024-09745-2]
36. Cinquantasei M, Albanesi M. Allergy and immunology diffusion via social media: performance of video and audio contents on Instagram. Postepy Dermatol Alergol 2023 Dec;40(6):817-819. [doi: 10.5114/ada.2023.133607] [Medline: 38282881]
37. Hill MD, Fox KM, Gardiner OY, Prempeh AG, Laws-Barker AM. Quality of preeclampsia videos on TikTok: a social media analysis. J Natl Med Assoc 2025 Sep;117(1):26-27. [doi: 10.1016/j.jnma.2025.08.052]
38. Molitierno R, Cerillo A, Fordellone M, et al. Labor induction and social media influence: an analytical video-based cross-sectional study to assess quality, reliability, and accuracy of related content on TikTok. J Gynecol Obstet Hum Reprod 2026 Sep;55(7):103205. [doi: 10.1016/j.jogoh.2026.103205] [Medline: 42082105]

XSL•FO
RenderX

39. Nettis E, Rossi O, Diluvio L, et al. Patient-reported outcome measures for chronic spontaneous urticaria: a strategy centred on optimizing patient management. Clin Exp Dermatol 2025 Jul 24;50(8):1522-1529. [doi: 10.1093/ced/llaf110] [Medline: 40172064]
40. Zhang DG, Sørensen JA, Pedersen NH, et al. Online depiction of urticaria is often flawed and does not reflect the spectrum of clinical manifestation. Dermatology 2024;240(3):507-513. [doi: 10.1159/000535932] [Medline: 38142684]
41. Murray E, Lo B, Pollack L, et al. The impact of health information on the internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. J Med Internet Res 2003;5(3):e17. [doi: 10.2196/jmir.5.3.e17] [Medline: 14517108]
42. Heague M, Ray C, Bowers J, Guckian J, Arents BW, Layton A. Patient and public involvement in dermatology research: a review. Am J Clin Dermatol 2022 May;23(3):319-329. [doi: 10.1007/s40257-022-00680-5] [Medline: 35349092]
43. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85. [doi: 10.2196/jmir.1933] [Medline: 23615206]
44. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021 Apr;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]

Abbreviations

ANDI: Accuracy in Digital-Health Instrument
GQS: Global Quality Score
IgE: immunoglobulin E

Edited by J Meisenheimer; submitted 02.Dec.2025; peer-reviewed by A Adil, S Bestavros; revised version received 15.Sep.2026; accepted 14.Oct.2026; published 09.Nov.2026.

Please cite as:
Bukhari A, Altalhab S, AlFaraj S, Alrwebah H, Alrubaiaan M, Alfawzan A
Accuracy, Reliability, and Quality of YouTube and TikTok Videos as Sources of Information on Urticaria: Cross-Sectional Study
JMIR Dermatol 2026;9:e88827
URL: https://derma.jmir.org/2026/1/e88827
doi:10.2196/88827

© Abrar Bukhari, Saad Altalhab, Suliman AlFaraj, Huda Alrwebah, Mishari Alrubaiaan, Abdulrahman Alfawzan. Originally published in JMIR Dermatology (http://derma.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Quality and Utility of Patient-Taken Smartphone Photographs of Atopic Dermatitis: Clinical Survey Study

Zarqa Ali[1], MD, PhD; Kenneth Thomsen[2], MD, PhD; Christian Vestergaard[2], MD, PhD; Simon Francis Thomsen[1], MD, PhD

[1]Department of Dermatology, Bispebjerg Hospital, Bispebjerg Bakke 23, Copenhagen, Denmark

[2]Department of Dermatology and Venereology, Aarhus University Hospital, Aarhus, Denmark

Corresponding Author:
Zarqa Ali, MD, PhD
Department of Dermatology, Bispebjerg Hospital, Bispebjerg Bakke 23, Copenhagen, Denmark

Abstract

Background: Atopic dermatitis (AD) has a relapsing and remitting nature, and scheduled clinic visits only provide a snapshot of the skin condition at the moment.

Objective: This study aimed to investigate the quality of patient-taken smartphone photographs of AD skin lesions and characterize patients using smartphone photographs as a tool to assist the physician to show disease activity in between consultations.

Methods: Patients from 2 university outpatient clinics specialized in AD were surveyed. A questionnaire regarding digital readiness was completed, and a previously taken skin lesion photograph on the patients' own smartphone was evaluated.

Results: Between February 2024 and September 2024, a total of 100 questionnaires were completed, 60 (60%) by participants from the capital region of Denmark and 40 (40%) by participants from an urban area, including 62 (62%) men and 38 (38%) women. The mean age of the recruited patients was 33.9 (SD 19.9) years. A total of 78% (78/100) of the patients used a desktop computer, laptop, or tablet often or always, and 86% (86/100) corresponded with the health care system using technology (eg, via email to the general practitioner or contact with hospitals via apps). More than 50% (52/100, 52%) strongly agreed or agreed with the statement that they would prefer a remote online visit with, for example, upload of skin lesion photographs over a routine in-person office visit. Almost 3 out of 4 patients had a photograph of their AD skin lesion on their smartphone, most (38/71, 54%) with the sole intention of presenting it to a physician. The photographs were of good quality in 85% (60/71) of the cases, and most (61/71, 86%) of the smartphone photographs were assessed to be useful for diagnostic and clinical evaluation. Receiving topical monotherapy was significantly associated with increased risk of having taken a skin lesion smartphone photograph (P=.006).

Conclusions: Patients with AD followed up on in an outpatient clinic often took good-quality photographs of their skin lesions before consultations with the intention of presenting them to the physician.

(JMIR Dermatol 2026;9:e72916) doi:10.2196/72916

KEYWORDS

atopic dermatitis; photograph; telemedicine; teledermatology; outpatient clinic; personalized; follow-up

Introduction

Atopic dermatitis (AD) has a fluctuating nature, including unpredictable flares [1], which is why scheduled visits to an outpatient clinic only provide a momentary snapshot of the disease course. Patients' perception of the use of photographs of skin lesions in clinical settings to improve medical care is overall positive [2]. A qualitative study highlighted that patients often feel unheard when consulting their physicians in times of disease remission. It also demonstrated an unfulfilled desire to be able to show a flair either by writing down symptoms or photographing lesions during flairs. Patients also indicated that the ability to evaluate the skin in between consultations provides increased autonomy and ownership [3]. A study conducted in an urticaria outpatient clinic showed that patients often took photographs of their skin lesions with their own smartphones before their first consultation, providing the physician with an insight into their disease severity at times of flare [4]. It has also been confirmed that the use of smartphones to take photographs of skin lesions is growing rapidly, a trend that might reduce the need for referrals to face-to-face visits [5] and thereby mitigate the growing shortage of dermatologists [6]. Furthermore, the severity of AD can be reliably assessed using photographs taken using smartphones as there is a high agreement between assessments conducted in the clinic directly looking at the skin and assessments conducted based on photographs [7,8].

Due to the clearly visible morphology of AD and the growing use of photographs taken using smartphones for medical

XSL•FO
RenderX

documentation purposes, we aimed to investigate the quality of patient-taken photographs of AD skin lesions using smartphones. Second, we aimed to characterize the group of patients who take smartphone photographs as a tool to assist the physician's evaluation of disease activity in between consultations. This knowledge might help understand patient preferences and tailor an individualized plan for follow-up either face-to-face or remotely based on photographs, thereby reducing health care costs while increasing patient autonomy.

Methods

Overview

Patients were consecutively recruited from 2 university outpatient clinics specialized in AD; one clinic in the capital region of Copenhagen and one from the second-largest urban area in Denmark, Aarhus. From February 2024 to September 2024, patients with a consultation in one of the outpatient clinics were asked to complete a questionnaire and select a possible previously taken smartphone photograph of their own AD lesions for severity assessment and quality evaluation by the physician. For pediatric patients, the questionnaire was completed by the parents.

To measure the perception of the impact of AD on quality of life, the Skindex-Mini, a 3-item questionnaire assessing 3 domains (symptoms, emotions, and function) graded on a Likert scale from 0 to 6, was used [9]. The Skindex-Mini total score was used to stratify impact of skin conditions on patient's quality of life as follows: a score of 0 to 1 indicated no impact, a score of 2 to 5 indicated low impact, a score of 6 to 10 indicated moderate impact, a score of 11 to 14 indicated high impact, and a score of 15 to 18 indicated very high impact on quality of life. The questionnaire has also been validated in pediatric patients with AD [10]. Questions related to use of technology in general and for communication with health care professionals were also included [11].

On the basis of the selected photograph of an AD lesion taken by the patient on their smartphone, a questionnaire regarding the quality and utility of smartphone photographs of AD skin lesions was completed by the attending physician. The quality assessment was based on focus of the photograph, resolution, lighting, and blurriness [12,13]. The utility of smartphone photographs for diagnostic use was based on the overall assessment of the treating physician (ie, whether the treating physician felt confident when using the photograph to establish diagnosis and for clinical evaluation and severity assessment). The clinical signs from the Eczema Area and Severity Index (EASI) and Scoring Atopic Dermatitis (SCORAD), including erythema, edema or papulation, excoriation, lichenification, oozing or crusts, and dryness, were assessed for each photograph along with their intensity (0-3 for the EASI and 0-4 for SCORAD) [14,15].

Statistical Analysis

Chi-square and independent-sample 2-tailed *t* tests were used to characterize patients who took smartphone photographs of skin lesions, and 95% CIs were provided when applicable. The Fisher exact test was used when one or more of the cells had an expected frequency of 5 or less. Multiple logistic regression was used to identify the variables best related to the likelihood of patients having a smartphone photograph of a skin lesion, including age (<30 years vs >30 years), sex, capital or urban area residence, AD onset (<2 years of age vs >2 years of age), quality of life (Skindex-Mini total score), systemic treatment vs topical treatment, daily use of technology, digital contact with the health care system, and whether they preferred a remote visit. A *P* value of <.05 was considered statistically significant. All tests were carried out using the SPSS software (version 25.0; IBM Corp) [16].

Ethical Considerations

As this was a questionnaire study, there was no requirement of governmental approval or written informed consent according to Danish guidelines [17]. All study participants gave oral consent to be included in the study and have their data stored. All data used in this study were fully anonymized. No personally identifiable information was collected, stored, or processed, ensuring the privacy and confidentiality of all participants. Participants did not receive any financial or nonfinancial compensation for their participation in this study.

Results

Cohort Description

A total of 100 questionnaires were completed, 60 (60%) by participants from the capital region and 40 (40%) by participants from the urban area, including 62 (62%) men and 38 (38%) women. The median age of the recruited patients was 28.0 (IQR 20.25-48.75; mean age 33.9, SD 19.9) years. Most (n=53, 53%) had an AD onset before the age of 2 years, 25% (n=25) had an AD onset between the ages of 2 and 6 years, 10% (n=10) had an AD onset between the ages of 6 and 18 years, and the remaining 12% (n=12) had an AD onset after the age of 18 years. A total of 37% (n=37) of the patients were treated with topical corticosteroids in monotherapy at the time of consultation, 36% (n=36) were treated with dupilumab, 12% (n=12) were treated with methotrexate, 3% (n=3) were treated with tralokinumab, and 3% (n=3) were treated with baricitinib or abrocitinib. Most patients (n=64, 64%) estimated AD to have none or a small impact on quality of life, 19% (n=19) estimated AD to have a moderate impact, 10% (n=10) estimated AD to have a large impact, and 7% (n=7) estimated AD to have a very large impact based on the Skindex-Mini questionnaire (Table 1).

XSL•FO
RenderX

Table . Characteristics of the included patients from 2 atopic dermatitis outpatient clinics (N=100).

Characteristic	Values
Sex, n (%)	
Male	62 (62)
Female	38 (38)
Age (years), mean (SD)	33.9 (19.9)
Current treatment, n (%)	
Topical treatment[a] only	38 (38)
UVB	1 (1)
Traditional immunosuppressants[b]	16 (16)
Prednisolone	1 (1)
JAK[c] inhibitors	3 (3)
Biologics[d]	39 (39)
None	2 (2)
Skindex-Mini score (0-18), mean (SD)	
Symptoms	2.25 (1.88)
Emotions	1.51 (1.81)
Function	1.30 (1.77)
Total	5.02 (5.03)
Impact on quality of life, n (%)	
None	33 (33)
Small	31 (31)
Moderate	19 (19)
Large	10 (10)
Very large	7 (7)

[a]Topical corticosteroids and topical calcineurin inhibitors.

[b]Azathioprine, methotrexate, and mycophenolate mofetil.

[c]JAK: Janus kinase; inhibitors included abrocitinib and baricitinib.

[d]Dupilumab and tralokinumab.

Digital Readiness

In total, 78% (78/100) of the patients used a computer, laptop, or tablet often or always; 18% (18/100) used them seldom or once in a while; and 4% (4/100) never used them. A vast majority (86/100, 86%) corresponded with the health care system using technology (eg, via email to the general practitioner or contact with hospitals via apps). More than 50% (52/100, 52%) strongly agreed or agreed with the statement that they would prefer a remote online visit with, for example, upload of skin lesion photographs over a routine in-person office visit. Table 2 provides further details.

XSL•FO
RenderX

Table . Items related to attitudes toward digital solutions (N=100).

Digital readiness	Participants, n (%)
Daily use of a computer, laptop, or tablet	
Often or always	78 (78)
Seldom or once in a while	18 (18)
Never	4 (4)
Digital correspondence with the health care system	
Often or always	56 (56)
Seldom or once in a while	30 (30)
Never	14 (14)
Digital access to blood samples or medical records	
Often or always	53 (53)
Seldom or once in a while	34 (34)
Never	13 (13)
Search for information related to morbidity on the internet	
Often or always	42 (42)
Seldom or once in a while	35 (35)
Never	23 (23)
"I would like to replace a physical in-office visit with a remote visit."	
Strongly agree	19 (19)
Agree	33 (33)
Neutral	27 (27)
Disagree	12 (12)
Strongly disagree	9 (9)

Smartphone Photographs

Almost 3 out of 4 patients (71/100, 71%) had a photograph of their AD skin lesion on their smartphone. Of the remaining 29% (29/100) who did not have any photographs of their AD lesions on their smartphones, most (15/29, 52%) indicated that the reason was a well-controlled disease for a longer period without experiencing any flair or worsening of AD, only 3% (1/29) did not have a smartphone, 7% (2/29) used another smartphone to take photographs, and the remaining 38% (11/29) did not give a reason. The number of smartphone photographs of AD lesions taken in the previous year varied from 1 to 100, the mean number of photographs taken was 21.4 (SD 22.7), and the median number of photographs was 15 (IQR 5-25). Most of those who took photographs did so with the sole intention of presenting them to a physician (38/71, 54%), only 8% (6/71) took the photographs for their own use, and 38% (27/71) took the photographs both for their own use and for the physician. Most of the photographs were of upper limbs (26/71, 37%) or the head and neck (23/71, 32%). Of all evaluated photographs, 85% (60/71) were of good quality, 7% (5/71) were of acceptable quality, and 9% (6/71) were of bad quality based on lighting, resolution, clarity, and focus. In total, 89% (63/71) of the smartphone photographs had the skin lesion in focus, of which 92% (65/71) were sharp and 9% (6/71) were blurred. Most of the smartphone photographs (61/71, 86%) were assessed to be useful for diagnostic and clinical evaluation (Table 3).

XSL•FO
RenderX

Table . Smartphone photographs taken by the patients coming to consultation in outpatient clinics (n=71).

	Photographs, n (%)
Body region	
Head and neck	23 (32)
Chest and stomach	6 (8)
Back	11 (15)
Upper limb	26 (37)
Lower limb	4 (6)
Missing	1 (1)
Lesion in focus	
Agree	63 (89)
Disagree	8 (11)
Sharp photograph	
Agree	65 (92)
Disagree	6 (9)
Useful in diagnostic evaluation	
Agree	61 (86)
Disagree	10 (14)
Useful in severity assessment	
Agree	59 (83)
Disagree	12 (17)
Resolution	
Good	63 (89)
Acceptable	8 (11)
Bad	0 (0)
Lighting	
Good	61 (86)
Acceptable	4 (6)
Bad	6 (8)
Photo quality	
Good	60 (85)
Acceptable	5 (7)
Bad	6 (8)

For EASI items, induration (14/71, 20%) and lichenification (10/71, 14%) were often difficult to assess (Table 4), and for SCORAD items, lichenification (11/71, 16%) and dryness (13/71, 18%) proved the biggest challenge (Table 5).

Table . Severity assessment of atopic dermatitis lesion photographs based on Eczema Area and Severity Index (EASI) (n=71).

	EASI score, n (%)				
	None	Mild	Moderate	Severe	Difficult to assess
Erythema	1 (1)	20 (28)	24 (34)	22 (31)	4 (6)
Induration	16 (23)	13 (18)	21 (30)	7 (10)	14 (20)
Excoriation	27 (38)	16 (23)	13 (18)	9 (13)	6 (8)
Lichenification	26 (37)	20 (28)	7 (10)	8 (11)	10 (14)

XSL•FO
RenderX

Table . Severity assessment of atopic dermatitis lesion photographs based on Scoring Atopic Dermatitis (SCORAD) tool (n=71).

	SCORAD score, n (%)					
	None	Mild	Moderate	Severe	Very severe	Difficult to assess
Erythema	2 (3)	19 (27)	20 (28)	14 (20)	12 (17)	4 (6)
Edema	19 (27)	16 (23)	12 (17)	9 (13)	7 (10)	8 (11)
Oozing	41 (58)	15 (21)	6 (9)	2 (3)	1 (1)	6 (9)
Excoriation	31 (44)	12 (17)	12 (17)	8 (11)	2 (3)	6 (9)
Lichenification	28 (39)	12 (17)	7 (10)	10 (14)	3 (4)	11 (16)
Dryness	15 (21)	19 (27)	11 (16)	9 (13)	4 (6)	13 (18)

Characteristics of Patients Who Took Smartphone Photographs of Skin Lesions

We found a significant difference in mean age between patients who took photographs and those who did not of 16.3 years (95% CI 8.15-24.46; *P*<.001). The mean age of patients who took smartphone photographs was 29.2 (SD 18.9) years, and that of patients who did not take smartphone photographs was 45.5 (SD 17.8) years. Previous digital contact with the health care system was associated with an increased odds ratio (OR) of 7.19 (95% CI 1.31-39.51; *P*=.01) of taking a skin lesion smartphone photograph. Patients receiving topical monotherapy had a higher chance of taking a skin lesion photograph (OR 4.17, 95% CI 1.42-12.16; *P*=.006), and patients receiving systemic treatment had a lower risk of taking a skin lesion photograph (OR 0.20, 95% CI 0.07-0.59; *P*=.002; Table 6). In logistic regression analysis, use of topical treatment was a statistically significant predictor for the probability of taking a photograph of a skin lesion (OR 5.67, 95% CI 1.20-26.77; β=1.74; SE 0.79; *P*=.03).

XSL•FO
RenderX

Table . Comparison between patients who took at least 1 smartphone photograph of their skin lesions and those who did not.

Characteristic	Photograph (n=71), n (%)	No photograph (n=29), n (%)	OR[a] (95% CI)	*P* value
Sex			1.24 (0.50 - 3.05)	.64
Male	43 (61)	19 (66)		
Female	28 (39)	10 (34)		
Age (years)			0.18 (0.07-0.49)	<.001
<30	45 (63)	7 (24)		
>30	26 (37)	22 (76)		
Residence			2.03 (0.79-5.21)	.14
Capital region	40 (56)	21 (72)		
Urban area	31 (44)	8 (28)		
Age at disease onset (years)			1.08 (0.45-2.55)	.87
<2	38 (54)	15 (52)		
>2	33 (46)	14 (48)		
Topical treatment only			4.17 (1.42-12.16)	.006
Yes	33 (46)	5 (17)		
No	38 (54)	24 (83)		
Traditional immunosuppressants			0.88 (0.28-2.80)	.83
Yes	11 (15)	5 (17)		
No	60 (85)	24 (83)		
Systemic treatment[b]			0.20 (0.07-0.59)	.002
Yes	35 (49)	24 (83)		
No	36 (51)	5 (17)		
Biologics or JAK[c] inhibitors			0.25 (0.10-0.63)	.002
Yes	23 (32)	19 (66)		
No	48 (68)	10 (34)		
Preferred remote visit[d]			1.23 (0.42-3.66)	.71
Yes	37 (52)	15 (52)		
No	14 (20)	7 (24)		
Daily use of technology			8.07 (0.80-81.17)	.07
Yes	70 (99)	26 (90)		
No	1 (1)	3 (10)		
Digital contact with the health care system[e]			7.19 (1.31-39.51)	.01
Yes	69 (97)	24 (83)		
No	2 (3)	5 (17)		
Impact of disease on quality of life			0.64 (0.42-0.97)	.04
None	17 (24)	16 (55)		
Small	26 (37)	5 (17)		
Moderate	15 (21)	4 (14)		
Large	6 (8)	4 (14)		
Very large	7 (10)	0 (0)		

[a]OR: odds ratio.

XSL•FO
RenderX

[b]Systemic treatment included dupilumab, tralokinumab, baricitinib, abrocitinib, methotrexate, azathioprine, and mycophenolate mofetil.

[c]JAK: Janus kinase.

[d]Includes “strongly agree” or “agree” vs “strongly disagree” or “disagree.”

[e]Includes both digital correspondence with the health care system and digital access to blood samples or medical records.

Discussion

Hospital outpatients with AD had high digital readiness, with 78% (78/100) using a computer, laptop, or tablet often or always. Almost 3 out of 4 had taken a photograph of their AD skin lesion on their smartphone, mostly with the intention of presenting it to a physician. Furthermore, 85% (60/71) of the photographs were of good quality; however, induration, lichenification, and dryness were often difficult to assess. Receiving topical monotherapy was associated with a higher chance of taking a skin lesion photograph, supporting the demand for tailored monitoring depending on patients’ preferences and risk of flair. AD is very heterogeneous in terms of symptoms, skin manifestations, body area involved, extent, course, and comorbidities. Therefore, it is very unlikely that all patients with AD will respond equally well to treatments. Biomarkers will lead to better identification of patients who will benefit from immunomodulatory treatments, leading to more individualized management [18]. Traditionally, patients on immunosuppressive drugs have often planned consultations in the clinic at certain intervals. Due to better disease control with targeted therapies, these patients only need to be followed up on, for example, once every year; however, due to the expenses related to the treatments, close monitoring will be beneficial for timely drug dose tapering to reduce unnecessary health care expenditures. On the other hand, many patients with mild to moderate disease will still be on traditional immunosuppressive drugs, not meeting the criteria for expensive biological treatments. These patients will often experience flairs in between scheduled consultations. Our study showed that more than half of patients with AD followed up on in an outpatient clinic preferred a remote or online visit instead of an in-person visit at the clinic. Furthermore, there is increasing evidence that patients with skin diseases often take good-quality photographs of their skin lesions with their smartphones [4] and that photographs have high validity and reliability [7,8,19]. This is supported by our findings. Tailored monitoring considering the age, digital readiness, type of treatment, and preferences of the patients may lead to a reduction in health care costs and help pivot consultations toward focused care based on individual needs.

Smartphones are easily accessible and extensively used to take photographs. Many photographs are taken on a daily basis, and more than 90% of all photographs are taken in 2020 using smartphones [20]. Many people find it natural to take photographs for memory or documentation [20]; hence, taking photographs of skin lesions is widely practiced [4]. There is a demand for integrating smartphone photographs into clinical practice to assess disease fluctuation in between physical examinations. Educating patients in how to take a good clinical photograph of AD skin lesions may improve the quality and utility of the photographs in a clinical setting. Information regarding distance between the camera and the skin lesion (approximately 20 cm), using a uniform background, and taking the photograph in good natural lighting is especially important. Furthermore, using photographs in a clinical setting through a remote visit to replace a physical consultation requires thorough patient education in the assessment of body surface area and selection of representative lesions in each anatomical area included in the EASI or SCORAD.

Even though the task of evaluating the quality of photographs was clearly defined to create consistency in evaluations, this study was limited by a lack of multiple raters to evaluate the same photograph due to logistical challenges in a clinical survey.

In conclusion, patients with AD followed up on in an outpatient clinic often took high-quality photographs of their skin lesions before consultations with the intention of presenting them to the physicians. More evidence for tailored or personalized monitoring through remote visits using photographs of skin lesions and its effect on health care costs is warranted.

Conflicts of Interest

CV has received grants from Pfizer, LEO Pharma, Almirall, and Sanofi and has been a speaker or served on advisory boards for Pfizer, Almirall, LEO Pharma, AbbVie, Sanofi, Galderma, Pierre Fabre, AstraZeneca, and Novartis. All other authors declare no other conflicts of interest.

References

1. Thomsen SF. Atopic dermatitis: natural history, diagnosis, and treatment. ISRN Allergy 2014 Apr 2;2014:354250. [doi: 10.1155/2014/354250] [Medline: 25006501]
2. Kim W, Sivesind T. Patient perceptions of dermatologic photography: scoping review. JMIR Dermatol 2022 Jan 26;5(1):e33361. [doi: 10.2196/33361] [Medline: 37632867]
3. de Wijs LE, van Egmond S, Devillers AC, Nijsten T, Hijnen D, Lugtenberg M. Needs and preferences of patients regarding atopic dermatitis care in the era of new therapeutic options: a qualitative study. Arch Dermatol Res 2023 Jan;315(1):75-83. [doi: 10.1007/s00403-021-02321-z] [Medline: 35112162]
4. Ali Z, Sørensen JA, Zhang DG, et al. Smartphone photographs of chronic urticaria taken by patients are of good quality and useful in the clinic. Dermatology 2024;240(3):357-361. [doi: 10.1159/000535027] [Medline: 37926073]

XSL•FO
RenderX

5. Marin-Gomez FX, Vidal-Alaball J, Poch PR, Sariola CJ, Ferrer RT, Peña JM. Diagnosis of skin lesions using photographs taken with a mobile phone: an online survey of primary care physicians. J Prim Care Community Health 2020;11:2150132720937831. [doi: 10.1177/2150132720937831] [Medline: 32590923]
6. Kimball AB, Resneck JS Jr. The US dermatology workforce: a specialty remains in shortage. J Am Acad Dermatol 2008 Nov;59(5):741-745. [doi: 10.1016/j.jaad.2008.06.037] [Medline: 18723242]
7. Ali Z, Chiriac A, Bjerre-Christensen T, et al. Mild to moderate atopic dermatitis severity can be reliably assessed using smartphone-photographs taken by the patient at home: a validation study. Skin Res Technol 2022 Mar;28(2):336-341. [doi: 10.1111/srt.13136] [Medline: 35020960]
8. Ali Z, Joergensen KM, Andersen AD, et al. Remote rating of atopic dermatitis severity using photo-based assessments: proof-of-concept and reliability evaluation. JMIR Form Res 2021 May 25;5(5):e24766. [doi: 10.2196/24766] [Medline: 34032580]
9. Swerlick RA, Zhang C, Patel A, Chren MM, Chen S. The Skindex-Mini: a streamlined quality of life measurement tool suitable for routine use in clinic. J Am Acad Dermatol 2021 Aug;85(2):510-512. [doi: 10.1016/j.jaad.2018.12.035] [Medline: 30590072]
10. Sheth AP, Blumstein AJ, Rangel SM, Jackson KL, Lai JS, Paller AS. Three-question Skindex-Mini measures quality of life in children with atopic dermatitis. J Am Acad Dermatol 2023 Feb;88(2):493-495. [doi: 10.1016/j.jaad.2022.06.1167] [Medline: 35753552]
11. Holgersen N, Nielsen VW, Ali Z, et al. Attitudes towards clinical research in adult patients with hidradenitis suppurativa during the COVID-19 pandemic: insights from a survey. J Eur Acad Dermatol Venereol 2024 May;38(5):e371-e373. [doi: 10.1111/jdv.19758] [Medline: 38108515]
12. Kashetsky N, Mar K, Liu C, Rivers JK, Mukovozov I. Photography and image acquisition in dermatology a scoping review: techniques for high-quality photography. Dermatol Surg 2023 Sep 1;49(9):877-884. [doi: 10.1097/DSS.0000000000003864] [Medline: 37389542]
13. Zoltie T, Blome-Eberwein S, Forbes S, Theaker M, Hussain W. Medical photography using mobile devices. BMJ 2022 Aug 26;378:e067663. [doi: 10.1136/bmj-2021-067663] [Medline: 36028231]
14. Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. Exp Dermatol 2001 Feb;10(1):11-18. [doi: 10.1034/j.1600-0625.2001.100102.x] [Medline: 11168575]
15. Severity scoring of atopic dermatitis: the SCORAD index. Consensus report of the European Task Force on Atopic Dermatitis. Dermatology 1993;186(1):23-31. [doi: 10.1159/000247298] [Medline: 8435513]
16. IBM support. IBM. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-25 [accessed 2025-11-12]
17. Hvad skal IKKE anmeldes? [Web page in Danish]. Videnskabsetiske. URL: https://www.videnskabsetik.dk/ansoegning-til-etisk-komite/overblik-over-anmeldelsespligten/hvad-skal-ikke-anmeldes [accessed 2026-01-22]
18. van der Schaft J, Thijs JL, Garritsen FM, Balak D, de Bruin-Weller MS. Towards personalized treatment in atopic dermatitis. Expert Opin Biol Ther 2019 May;19(5):469-476. [doi: 10.1080/14712598.2019.1583204] [Medline: 30768375]
19. Ali Z, Robert Zibert J, Dahiya P, et al. Mild-to-moderate severity of psoriasis may be assessed remotely based on photographs and self-reported extent of skin involvement. JAAD Int 2023 Feb 22;11:129-136. [doi: 10.1016/j.jdin.2023.02.004] [Medline: 37128265]
20. Kislinger L, Kotrschal K. Hunters and gatherers of pictures: why photography has become a human universal. Front Psychol 2021;12:654474. [doi: 10.3389/fpsyg.2021.654474] [Medline: 34168589]

Abbreviations

AD: atopic dermatitis
EASI: Eczema Area and Severity Index
OR: odds ratio
SCORAD: Scoring Atopic Dermatitis

Edited by R Dellavalle; submitted 20.Feb.2025; peer-reviewed by H Wang, OS Kucuk; revised version received 13.Nov.2025; accepted 14.Nov.2025; published 27.Jan.2026.

Please cite as:
Ali Z, Thomsen K, Vestergaard C, Thomsen SF
Assessment of Quality and Utility of Patient-Taken Smartphone Photographs of Atopic Dermatitis: Clinical Survey Study
JMIR Dermatol 2026;9:e72916
URL: https://derma.jmir.org/2026/1/e72916
doi:10.2196/72916

XSL•FO
RenderX

© Zarqa Ali, Kenneth Thomsen, Christian Vestergaard, Simon Francis Thomsen. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comments on Dermatologists' TikTok Videos on Atopic Dermatitis: Content Analysis

Sahithi Gangavarapu[1*], BA; Tazeena Khan[2*], BS; Morgan McCarthy[3], BS; Molly Hales[4], MD, PhD; Stephanie M Rangel[1], PhD

[1]Department of Dermatology, Feinberg School of Medicine, Northwestern University, 676 N. St. Clair Street, Suite 1600, Chicago, IL, United States

[2]College of Medicine, University of Illinois Chicago, Chicago, IL, United States

[3]Chicago College of Osteopathic Medicine, Midwestern University, Downers Grove, IL, United States

[4]Department of Dermatology, UChicago Medicine, University of Chicago, Chicago, IL, United States

[*]these authors contributed equally

Corresponding Author:
Stephanie M Rangel, PhD
Department of Dermatology, Feinberg School of Medicine, Northwestern University, 676 N. St. Clair Street, Suite 1600, Chicago, IL, United States

Abstract

Background: TikTok is one of the fastest-growing social media platforms in the world. It has become an important space for sharing information on a wide range of topics, including medical conditions such as atopic dermatitis (AD). Advice on skin conditions has become popular on TikTok, and most previous research in this area focuses on the credibility of the information being shared. However, little research has focused specifically on physician-created videos and their audience engagement and interaction.

Objective: Our study aimed to (1) characterize the audience's online response to board-certified physicians' TikTok content related to AD according to established protocols and (2) better understand the interactions that happen among members of the audience in the comment section of this content.

Methods: In December 2023, searches were conducted for the terms "atopic dermatitis" and "eczema" on 3 unique TikTok accounts to identify videos created about AD by board-certified dermatologists. A total of 28 final videos were analyzed and classified into the following categories: (1) explanation of disease, (2) recommendation, (3) debunking misinformation, and (4) informal or anecdotal. The top 50 original comments on each of the 28 videos were collected and classified into one of the following categories: (1) "positive personal experience," (2) "negative personal experience," (3) "neutral personal experience," (4) "requesting advice," (5) "learning," (6) "appreciative reaction," (7) "critical reaction," (8) "giving advice," (9) "humor," (10) "tagging another user," and (11) "off-topic." Replies to comments were also analyzed and grouped into similar categories.

Results: Video category did not have a significant impact on engagement rate (P>.99). Across all video categories, comments that involved personal experiences or sharing information made up a larger percentage than those that were critical or off-topic (P<.001). Of the comments related to personal experience, the percentage of negative personal experience comments was significantly higher than that of positive personal experience comments (P=.001). Among replies to comments, "recommendation" and "emotional support" replies were significantly more common than other types of replies (P<.001).

Conclusions: Our study suggests that videos created by dermatologists on TikTok are generally well received regardless of video style or category. The comment sections appear to provide transient supportive environments where users connect over shared challenges and exchange personal experiences and recommendations. There is a gap in dermatologist-produced TikTok content involving darker skin tones.

(JMIR Dermatol 2026;9:e90649) doi:10.2196/90649

KEYWORDS

atopic dermatitis; eczema; TikTok; social media; dermatology; peer support; online community; online engagement

XSL•FO
RenderX

Introduction

Background

Social media platforms have become increasingly popular forums for people who share similar experiences to connect. TikTok has gained traction since its creation in 2016 and has become one of the most popular platforms used by people to share their stories, request advice, and connect with others [1].

Medical content, including videos that discuss symptoms, diagnoses, and treatment of various conditions, is widespread on TikTok, including content pertaining to skin health and diseases.

Atopic Dermatitis on Social Media

Atopic dermatitis (AD) is a common chronic skin condition with a strong patient presence in social media communities on various platforms, where patients often post about their experiences and concerns related to the disease [2]. AD affects 15% to 30% of children and 2% to 10% of adults in industrialized countries [3,4]. Due to its itchy, painful, and visible nature, AD can often impact self-image and quality of life. Patients with AD have reported lower satisfaction with life scores and lower mental and physical health scores [5]. Previous studies have found that online posts from patients with dermatoses, including AD, often discuss issues such as self-image, self-confidence, and therapeutic options [6].

There are many systemic barriers to care that patients with AD may experience, including financial and geographic barriers as well as patient knowledge and beliefs [7]. In light of these potential access issues, the public nature of videos makes TikTok an easy and accessible way for patients to find information related to their disease. However, misinformation can reach patients just as easily. Studies have found that TikTok content produced by board-certified physicians is more accurate and reliable than content produced by others [8,9], but they do not produce much of the content that reaches patients. One study found that most skin care–related TikTok content was produced by laypeople, and of the most liked and most viewed skin care posts, only 2.5% were produced by board-certified dermatologists [10].

By combining accessibility and accuracy, videos produced by board-certified physicians may be a helpful avenue for patients to obtain accurate information regarding AD. Understanding how people respond to these videos by analyzing the comments they leave may be helpful to determine the utility of such content and assess whether patients are able to find supportive communities on TikTok.

Patient Experience With Illness-Related Social Media Content

Prior studies have analyzed illness-related content on other diseases on social media to better understand its purpose and utility for patients. Many of these studies have found that most illness-related content is produced by patients or loved ones, and producing content on social media can serve as an outlet for these individuals to discuss their experiences with a disease or treatment [11].

Some studies have attempted to characterize patients' perceptions of such illness-related content and have found mixed reactions. Concern about misinformation has been cited as a primary reason for disliking health-related social media content, whereas observing empathy and shared experiences has been reported as a primary benefit [12]. This suggests that while patients have varying responses to illness-related content on social media, many appreciate the feelings of relatability and support that this content can produce.

Prior studies have found that public reactions on social media can be reasonably representative of public opinion gathered via conventional surveys [13]. Additionally, multiple studies have used comments on social media to assess and analyze public audience reactions in the context of illness-related content (eg, videos; [14,15]). Naslund et al [16] found that interactions among commenters on social media conveyed themes of hopefulness, minimizing isolation, peer support, and learning from shared experiences. This suggests that commenters may feel support from the initial video or post itself but also from interactions with other users or the video creator in the comment section.

Our study aimed to (1) characterize the audience's online response to board-certified physicians' TikTok content related to AD according to established protocols [14,15,17] and (2) better understand the interactions that happen among members of the audience in the comment section of this content. In doing this, we aimed to understand whether the supportive environments that have been observed as a result of patient-driven illness-related content on social media are still observed with physician-driven content related to AD.

Methods

Video Identification

Three unique TikTok accounts were used to search 2 separate terms ("atopic dermatitis" and "eczema") in December 2023. Two accounts were previously established personal accounts, whereas 1 account was newly created for the purpose of the study in an attempt to mitigate the effects of a personalized algorithm. Using the "Top" video categorization on TikTok, the first 30 videos to appear under each term for each account (for a total of 180 videos) were screened for inclusion. From this initial list of 180 videos, duplicate videos (75/180, 41.7%), videos that were not in English (12/180, 6.7%), and videos that did not appear to be posted by a board-certified physician (57/180, 31.7%) were excluded, resulting in 36 videos. To confirm board certification status, we did a preliminary search of the creator's publicly available full name to find an associated hospital, university, or clinic. On the basis of the location of their associated institution, we cross-checked their certification status with national online databases. This resulted in the use of the American Board of Medical Specialties (United States), General Medical Council (United Kingdom), or Royal College of Physicians and Surgeons of Canada (Canada) databases. Of the remaining 36 videos, 7 were further excluded through this process, and 1 additional video was excluded as the content was not related to AD, leading to 28 final videos to be analyzed. All

XSL•FO
RenderX

analyzed videos were posted to the social media platform between 2020 and 2023.

Categorization

Thematic Approach

To analyze these qualitative data, a thematic analysis method was used. An inductive approach was used to identify themes in the content by following the guidelines by Braun and Clarke [17]. After collecting the final videos that met the inclusion criteria, 3 separate reviewers (TK, SG, and MM) watched all videos to familiarize themselves with the data and perform initial coding. Next, after discussion among the reviewers, 4 categories reflecting video themes were generated. A rubric was created with category definitions and examples to aid in the categorization and analysis process. A similar methodology was followed with the comments for categorization and analysis.

Video Categorization

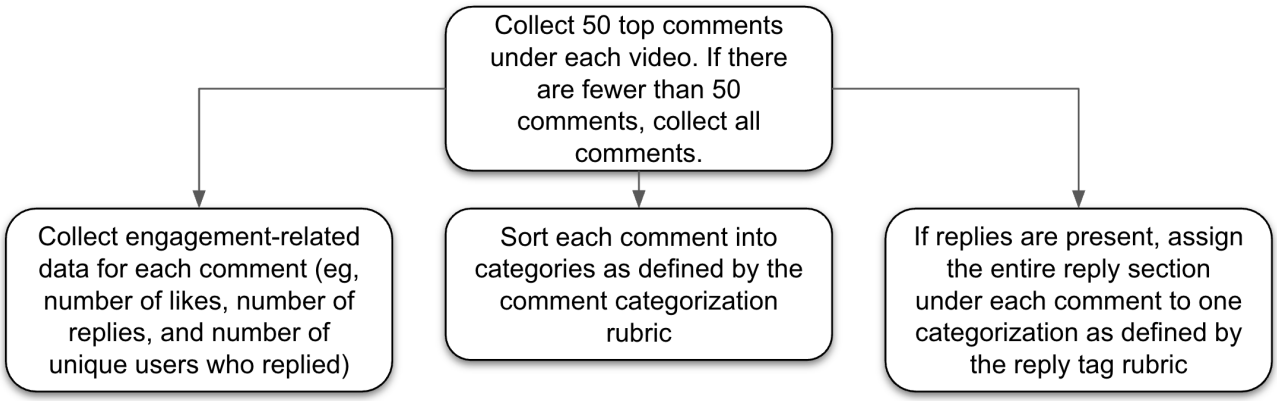
From each video, the following information was collected: creator handle; number of creator followers; date of posting; video duration; and number of views, likes, saves, shares, and total comments. Engagement rate was calculated as the sum of the number of video likes, comments, shares, and saves divided by the number of video views. This formula was derived from engagement rate calculations commonly used in the social media industry and in prior literature [18,19].

Through consensus discussion among 3 raters, each of the 28 videos was classified into one or more of the following categories so that videos could be counted in more than one category if they met the respective inclusion criteria: (1) explanation of disease, (2) recommendation, (3) debunking misinformation, and (4) informal or anecdotal. Videos were also tagged if they (1) used the TikTok "stitch" feature (ie, the video creator adds their content to the end of a different video, typically as a response), (2) were sponsored, (3) used a trending sound vs trending song vs original audio, and (4) were inclusivity focused (ie, if the video focused on presentation of AD in darker skin tones) to consider additional factors that may impact engagement.

Comment Categorization

For each video, the top 50 original comments (ie, not replies to other comments) were collected. If the video had less than 50 original comments at the time of analysis, all available original comments were collected (Figure 1).

Figure 1. Video comment identification framework for analysis.



Original comments and secondary comments were divided among 3 individual raters. To confirm interrater reliability, raters first coded a sample of 1.7% (20/1149) of the total comments analyzed and compared their categorizations. We found 80% perfect agreement and 90% agreement among at least 2 raters (Krippendorff α=0.72).

For each original comment, the following information was collected: date, number of likes, whether the creator liked the comment, number of secondary comments (ie, replies to an original comment), number of unique secondary commenters, number of secondary comments from the creator, number of likes on the creator's secondary comments (if applicable), and number of replies from another public board-certified physician or dermatologist. The original comments were classified into one of the following categories: (1) "positive personal experience," (2) "negative personal experience," (3) "neutral personal experience," (4) "requesting advice," (5) "learning," (6) "appreciative reaction," (7) "critical reaction," (8) "giving advice," and (9) "humor." We also had two categories for unrelated comments: (1) "tagging another user" and (2) "off-topic." Comments in which users self-identified as patients with AD or relatives or loved ones of patients with AD were tagged.

Additionally, for each original comment with secondary comments, the secondary comment section was scanned and classified into one of the following categories: (1) "recommendation" for discussion focused on requesting and sharing recommendations; (2) "emotional support/shared experience" for discussion focused on relating to, showing support for, or agreeing with the original comment; (3) "argumentative" for discussion that was marked with hostility or nonrespectful disagreement; (4) "educational" for discussion that focused on providing explanations or answers without any recommendations given; and (5) "other" for any secondary comment sections that did not fit into the aforementioned categories.

Statistical Analysis

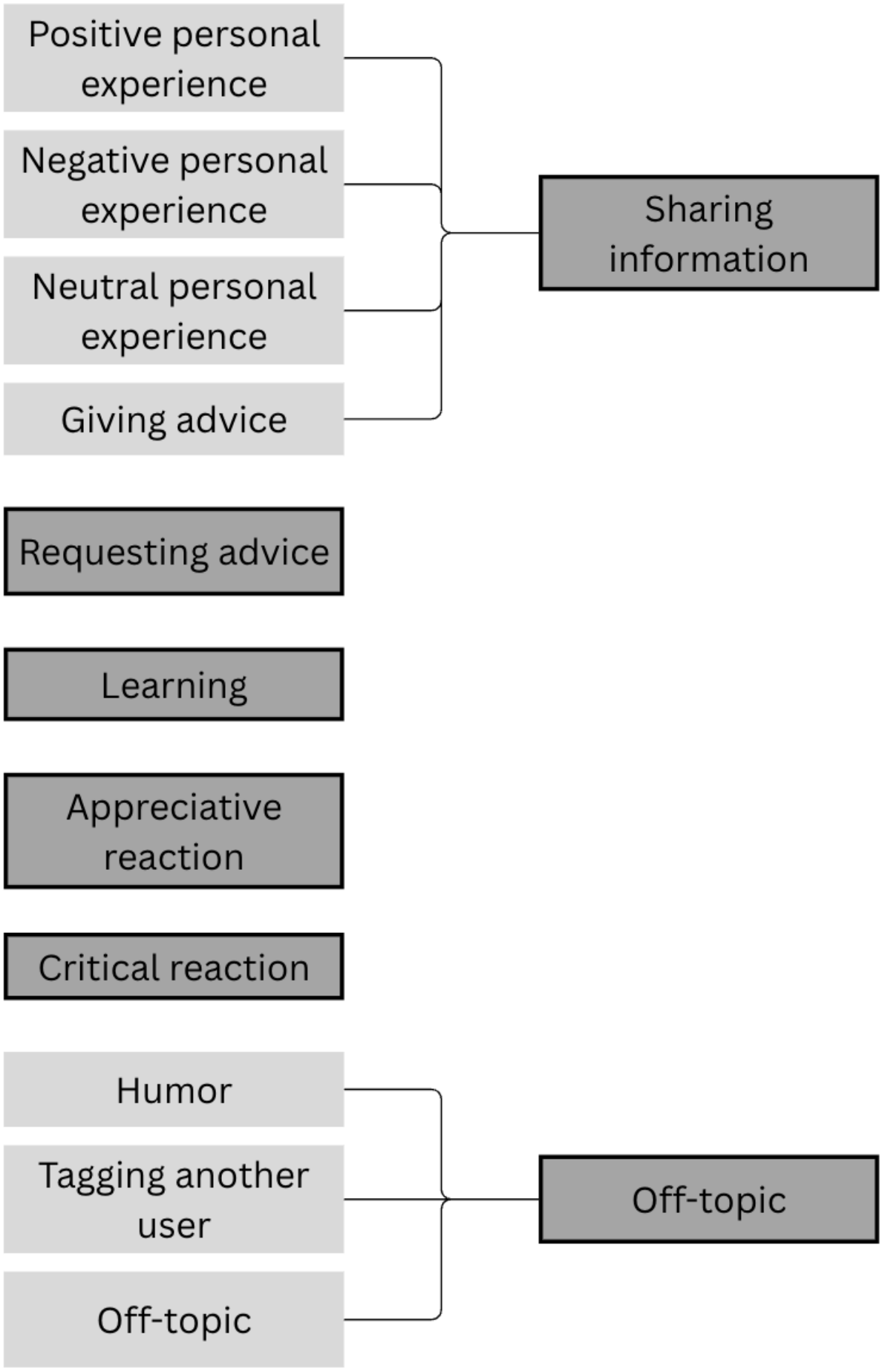
All analyses were performed in Prism (version 10.2.3; GraphPad Software), and *P* values of less than .05 were considered statistically significant. Frequency distributions were performed

XSL•FO
RenderX

on categorical video, original comment, and secondary comment data separately. Engagement rate analyses were performed using Kruskal-Wallis one-way ANOVAs. For ease of analysis, certain related comment categories were consolidated (Figure 2). "Personal experiences" and "giving advice" were combined into "sharing information." "Tagging another user," "off-topic," and "humor" were combined into a single "off-topic" category.

XSL•FO
RenderX

Figure 2. Consolidation of comment categories into broader thematic groups.



XSL•FO
RenderX

Ethical Considerations

This study was exempt from institutional review board approval as there was no involvement with human participants and only publicly available data were used.

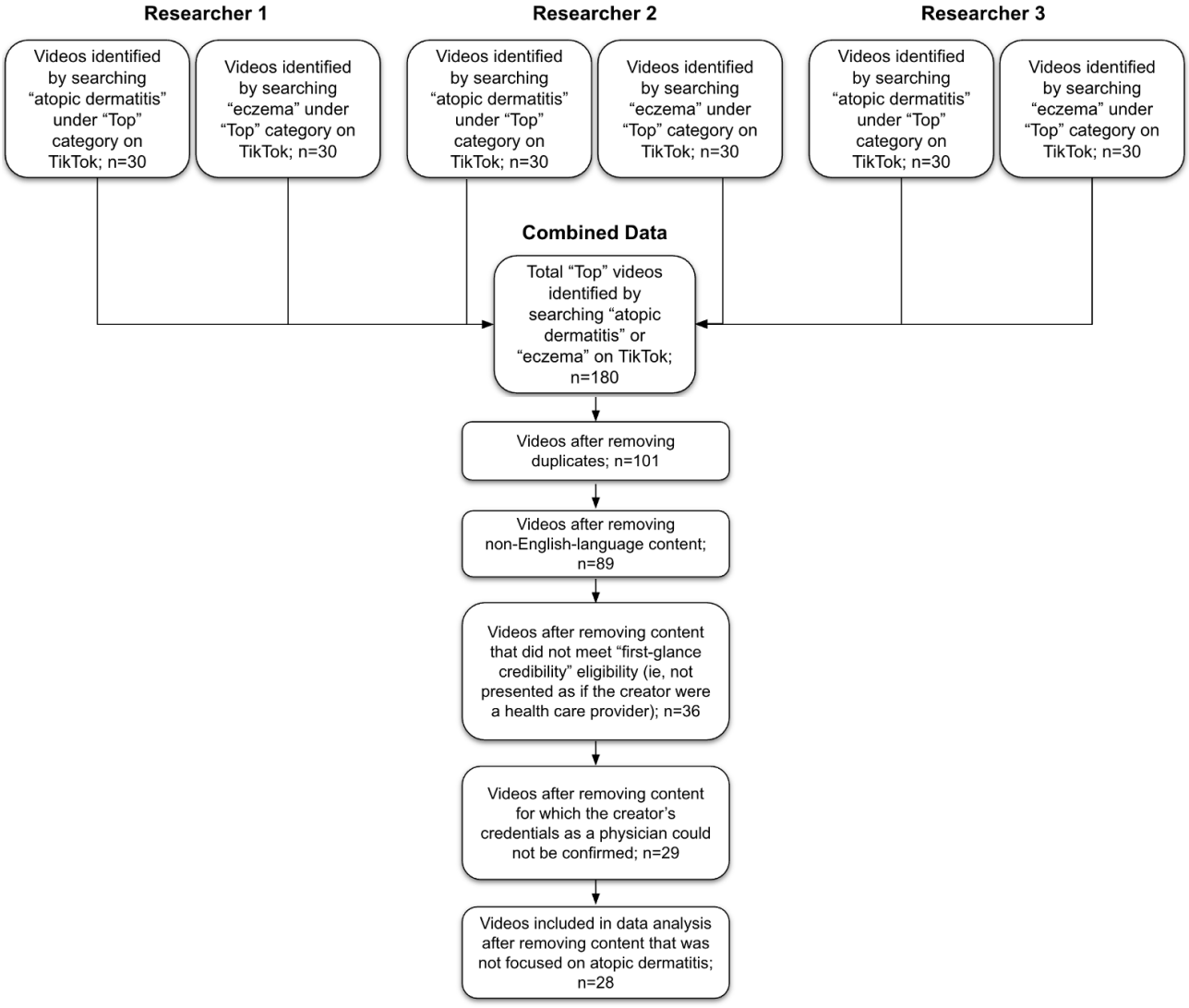
Results

Video Metrics

The 28 videos analyzed (Multimedia Appendix 1) had a total view count of 98,900,200, with an average of 3,532,150 views per video (SD 9,756,600; range 21,200-51,400,000). Video length ranged from 4 to 89 seconds, with an average length of video of 40.25 seconds. Among the total videos, 7.1% (n=2) were tagged for being inclusivity focused, 7.1% (n=2) were tagged for containing sponsored content, 35.7% (n=10) were tagged for having a "stitched" format, and 53.6% (n=15) were tagged for use of a trending sound or song (Figure 3).

Most videos categorized fell into the "recommendation" category (21/28, 75%), followed by "explanation of disease" (16/28, 57.1%), "debunking misinformation" (5/28, 17.9%), and "informal or anecdotal" (5/28, 17.9%). Most videos fell under 2 categories (Table 1).

Figure 3. TikTok video identification flowchart for inclusion in the analysis.



XSL•FO
RenderX

Table . Excerpted transcripts of sample videos for each video category.

Video category	Sample transcript
Recommendation: provides recommendations for people with atopic dermatitis	• "One of the first line treatments are topical corticosteroids like triamcinolone which is prescribed frequently. These are safe to use when you follow the directions of your dermatologist, but if you overuse them they can cause thinning of the skin and stretch marks." [Dr. Portela, 2020] • "I was asked to comment on this recent post of diluted bleach baths. Would you let your child swim in a chlorinated swimming pool? Well, diluted bleach baths are really no different, they're safe and effective at reducing bacteria on infection prone skin and like in children with eczema. Many dermatologists and pediatric dermatologists recommend them routinely. They can be quite helpful." [Dr. Sheilagh, 2021]
Explanation of disease: includes attempts to explain the presentation, diagnosis, or pathology of atopic dermatitis	• "In atopic dermatitis we have an upregulation of this Th2 pathway which causes the physical manifestations and itch. Interestingly, this same pathway is responsible for fighting parasites such as worms." [Dr. Morris, 2022] • "Eczema is also known as atopic dermatitis. When you have this condition you have dry, itchy skin which is really common in kids but can also present for the first time as an adult." [Dr. Yadav, 2022]
Debunking misinformation: focuses on dispelling myths and misconceptions regarding the disease or treatment	• ["Stitched" video of someone saying they found out babies have eczema because they are allergic to dairy so they switched their child's formula to nutramigin] "This is for educational purposes only, not specific medical advice but I would be careful when considering eliminating a food in a baby's diet to help treat eczema. The reason being is that many babies who have this—it may not actually be beneficial and could be potentially harmful." [Dr. Rubin, 2023]
Informal or anecdotal: focuses on the creator's specific experiences or attempts to be humorous	• Text: "When my patient with eczema takes lukewarm showers, trims their nails short, and dunks their body in Vaseline after showers." Trending audio: "Someone cooked here." [Dr. Park, 2023]

Engagement Rate

Video engagement rate ranged from 0.38 to 7.54 (Multimedia Appendix 2). Video category did not have a significant impact on engagement rate (Figure 4).

XSL•FO
RenderX

Figure 4. TikTok video engagement rate across all video categories. ns: not significant.

XSL•FO
RenderX

Comment Categorization Results

The 28 videos analyzed had a total of 30,019 comments, with comment counts on individual videos ranging from 10 to 14,700. For videos with more than 50 comments, the top 50 comments were analyzed, resulting in a total of 1149 comments included in the analysis. The length of time that a video continued to generate top replies varied substantially, and our analysis was limited by the fact that some videos were posted only weeks before the date of analysis. Nevertheless, among the videos we analyzed, the median interval between the earliest and most recent top comment was 268 (IQR 113-477) days.

Of note, 37.2% (427/1149) of the analyzed comments were made by users who self-identified as patients with AD in their comments, and 6.8% (78/1149) were made by users who self-identified as relatives or loved ones of patients with AD (Table 2).

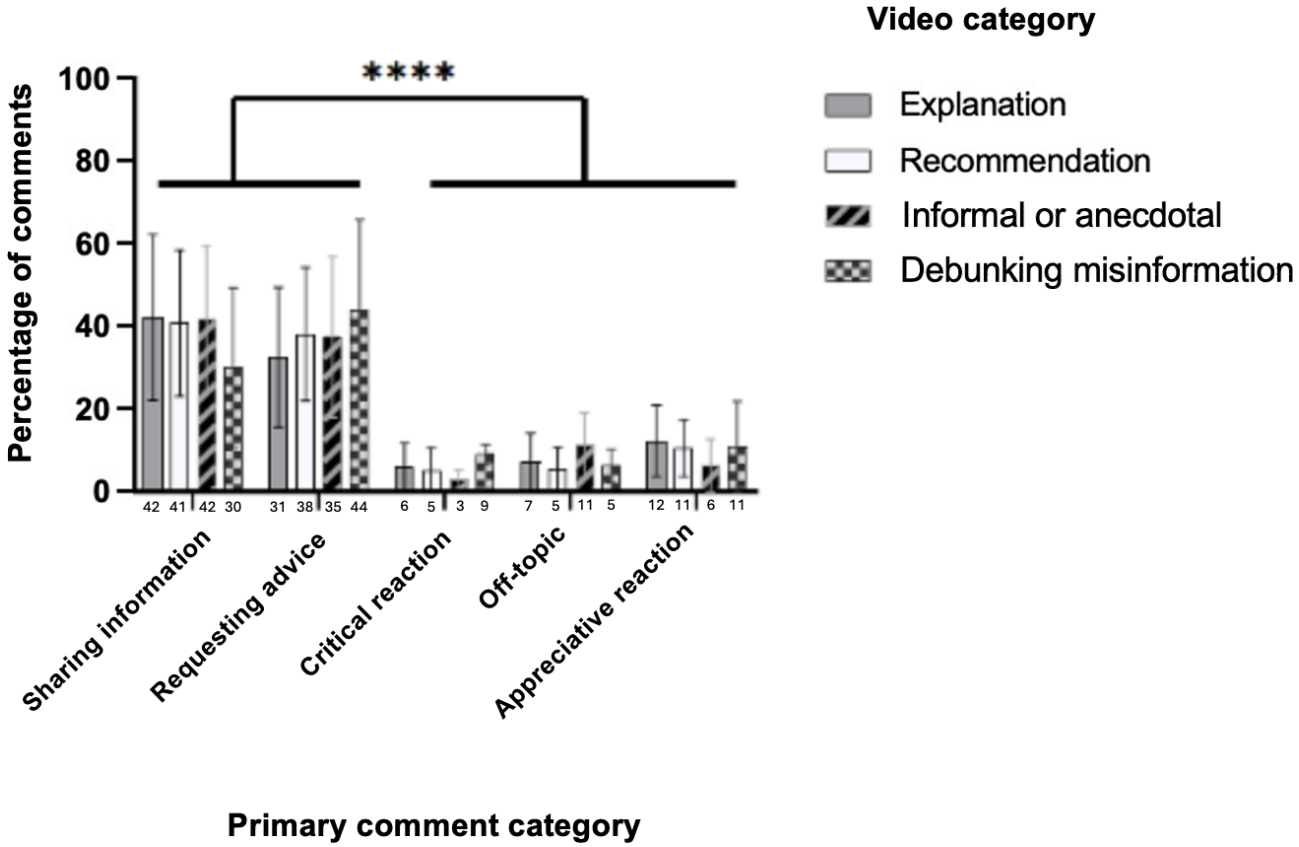
Across all video categories, "sharing information" and "requesting advice" comments made up a larger percentage than those that were critical, off-topic, or appreciative (P<.001; Figure 5). There was no significant difference between the percentage of comments involving sharing information and requesting advice (P=.95). Comments that were categorized as "learning" made up only 4.3% (49/1149) of the total.

Table . Excerpted comments for each comment category.

Comment category	Sample comment
Sharing information	
Positive personal experience: discusses the user's experience with the disease or product in a positive or praising tone	• "This has seriously been a game changer for my daughter. When hers is bad she'll ask for a bleach bath" [Video 2; comment 2] • "4 weeks and my skin is clear! No side effects" [Video 8; comment 24]
Negative personal experience: discusses the user's experience with the disease or product in a negative or complaining tone	• "Ugh it sucks and looks ugly. It's on my chest, stomach and back mostly... ugh" [Video 14; comment 30]
Neutral personal experience: discusses the user's experience with the disease or product without praising or complaining	• "I notice that I get eczema on my chin /forehead 2 weeks before my period" [Video 12; comment 27]
Giving advice: attempts to provide their own recommendations or resources for other users	• "guys use advantan cream it worked so well for me" [Video 17; comment 14]
Requesting advice: asks the creator or other users for more information or specific advice	• "How can you differentiate between eczema and allergic reaction" [Video 14; comment 15] • "Do all eczema itch?" [Video 16; comment 5]
Learning: expresses that they learned something new	• "This explains why my scalp is so clear after being in the pool for a long time!" [Video 2; comment 12] • "... THIS HAS A NAME!!!???" [Video 24; comment 22]
Appreciative reaction: shows appreciation to the creator or their overall account, not necessarily for the information presented in the video	• "THANK YOU! Pls dont stop sharing!" [Video 25; comment 49] • "Thank god I found this video! I have the same thing" [Video 26; comment 23]
Critical reaction: expresses a direct negative reaction or criticizes the creator or their overall account, not necessarily against the information presented in the video	• "THIS IS NOT ECZEMA, THIS IS TOPICAL STEROID WITHDRAWAL" [Video 15; comment 42] • "Girl you have to explain more because everyone gunna just be taken a bath tub full of bleach" [Video 3; comment 27]
Off-topic	
Humor: did not fit under other categories but was attempting to be humorous	• "Honestly, if my arms looked and felt like that, I'd take worms." [Video 22; comment 37]
Tagging another user: the comment's only content was tagging someone	• @[username]
Off-topic: any other comments, such as spam comments	• "First!!" [Video 26; comment 37] • "Why do some Americans say 'egg-zema' THERES NO G and it's not an egg" [Video 23; comment 31]

XSL•FO
RenderX

Figure 5. Video comment topics related to sharing information and experiences constituted a greater percentage compared to critical or off-topic comments. ****P<.0001

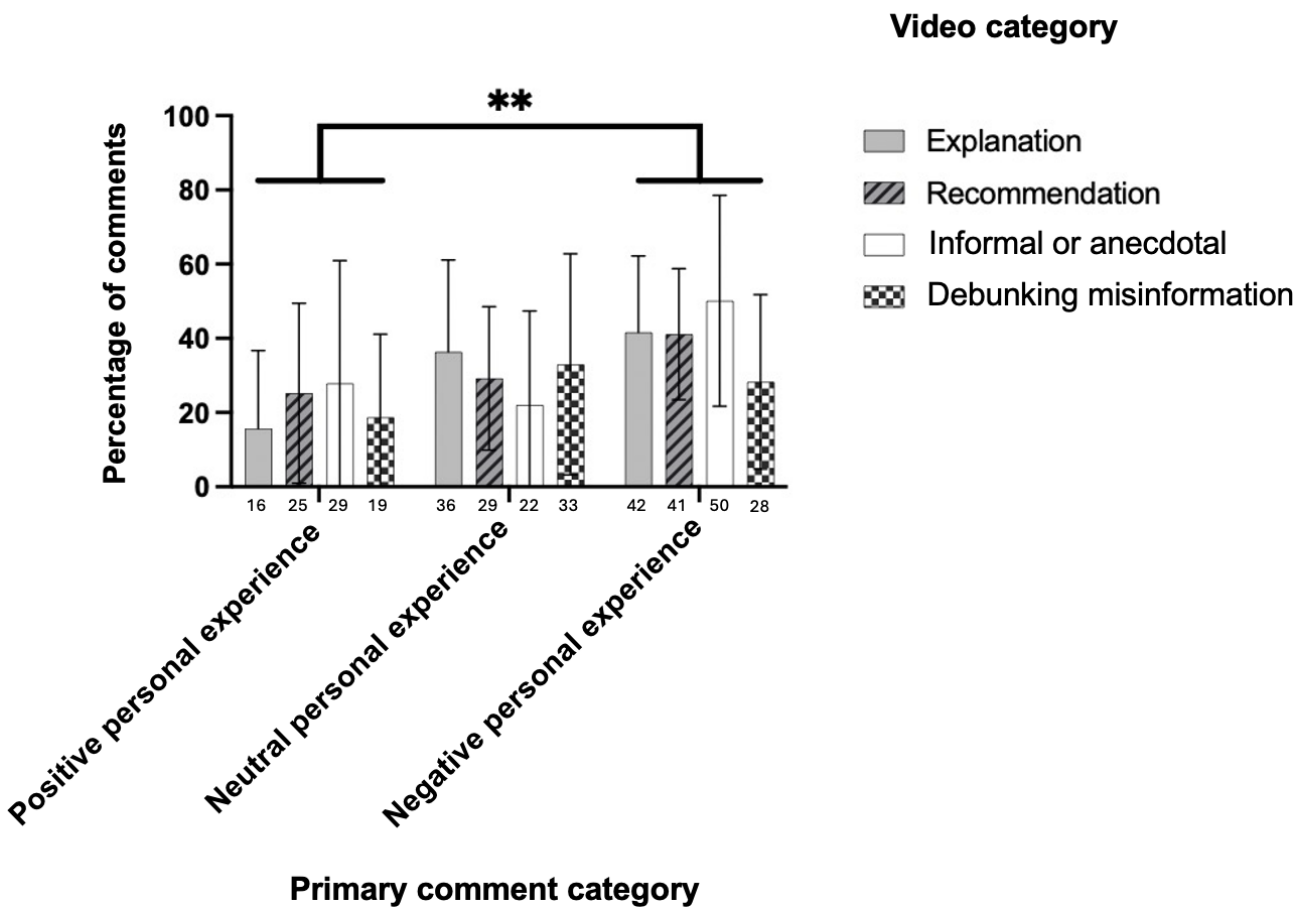


Personal Experience Comments

Of the comments related to personal experience, the percentage of negative personal experience comments trended higher than that of neutral or positive personal experience comments across all video categories. The percentage of negative personal experience comments was significantly higher than that of positive personal experience comments (P=.001; Figure 6); however, there was no significant difference between the percentage of neutral and negative personal experience comments across video categories or between the percentage of positive and neutral personal experience comments.

XSL•FO
RenderX

Figure 6. Across all video categories, comments focused on sharing negative experiences with atopic dermatitis vs neutral or positive experiences. **P<.01

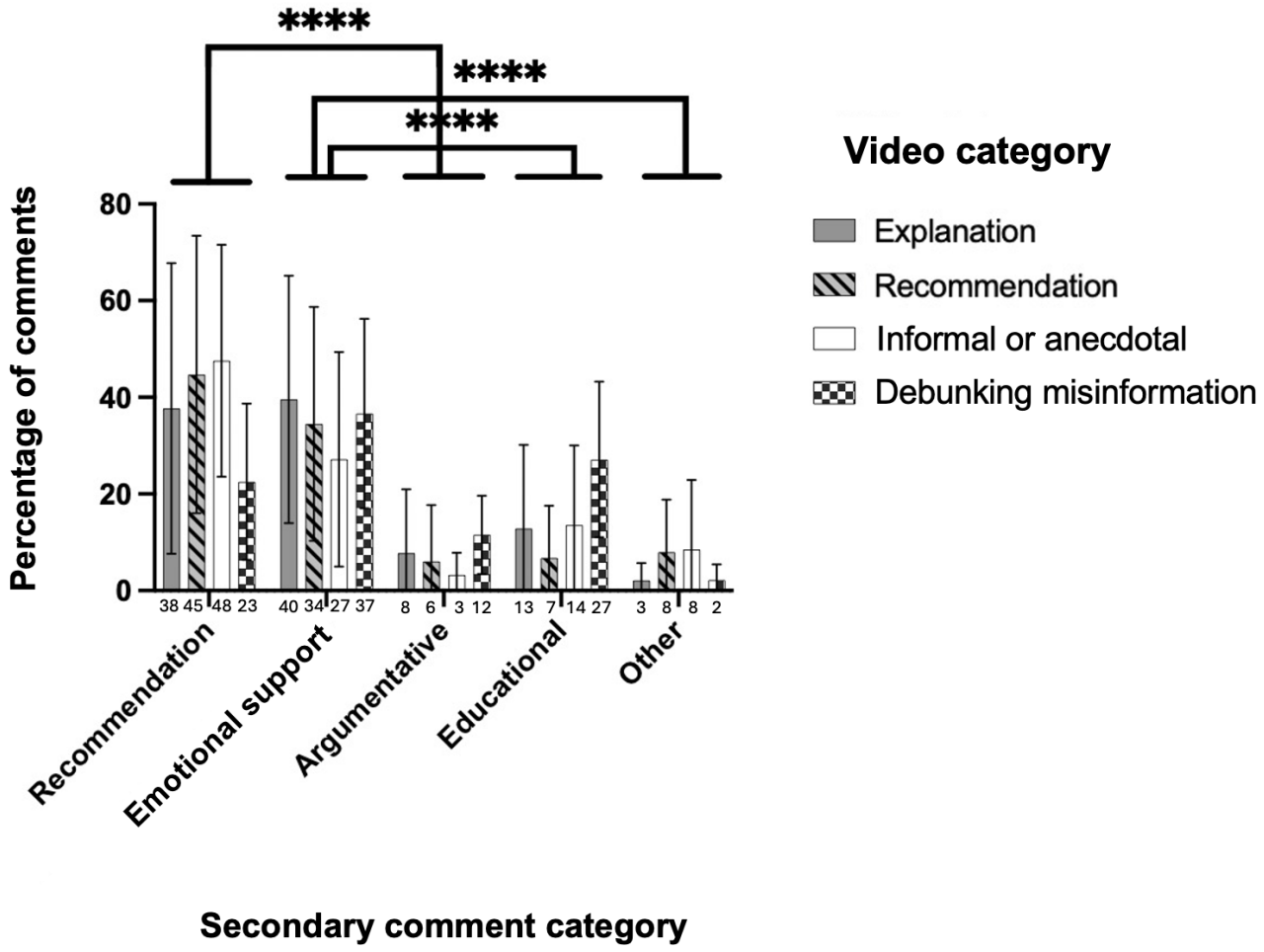


Secondary Comment Results

Video category did not have a significant impact on the distribution of secondary comment categories. Across video categories, the percentage of "recommendation" secondary comments was significantly higher than the percentage of "argumentative" (P<.001), "educational" (P<.001), and "other" (P<.001) secondary comments (Figure 7). Across video categories, the percentage of "emotional support" secondary comments was also significantly higher than the percentage of "argumentative" (P<.001), "educational" (P<.001), and "other" (P<.001) secondary comments (Figure 7). There was no significant difference between the percentage of "recommendation" and "emotional support" secondary comments (P=.96). There were also no significant differences between the percentage of "argumentative," "educational," and "other" secondary comments (Figure 7). The number of creator replies did not have a significant impact on the distribution of secondary comment categories (data not shown).

XSL•FO
RenderX

Figure 7. "Recommendation," "personal experience," and "emotional support" secondary comments trended higher than argumentative or educational comments. ****P<.0001



Discussion

Principal Results

Video Content

Research has shown that most videos related to dermatology on TikTok are created by laypeople rather than board-certified dermatologists [10], and few previous studies have specifically analyzed TikTok content produced by physicians. Among the 28 physician-produced videos we analyzed, format, content, and tone varied widely even within a single category. For instance, in the "recommendation" video category, some videos adopted a formal, explanatory, and detailed approach that was typically longer, whereas others opted for shorter and more casual formats. In the "explanation of disease" category, some videos formally explained the scientific pathogenesis of AD, whereas others offered brief, visual representations of the condition.

Our study found that video category did not affect engagement rate (Figure 4), suggesting that physicians can be versatile with their content and still receive engagement. Only 7.1% (2/28) of the top-appearing videos contained sponsored content, suggesting that sponsorship is not an important component of a successful TikTok video. In total, 28.6% (8/28) of the videos contained a trending sound, whereas 46.4% (13/28) used an original audio, showing that following specific audio formats or trends may not be necessary either. Additionally, the video with the greatest number of views or highest engagement rate was not created by the physician with the greatest number of followers. These findings suggest that physicians can be flexible with the format, content, and length of their videos and produce active engagement regardless of social media experience or an existing large follower base.

We noted that none of the videos analyzed showed practical demonstrations of the techniques being recommended (eg, preparation of bleach bath or application of emollients), instead relying on verbal discussion and occasionally photos. We believe that physician creators could use the video format supported by TikTok and similar platforms to provide such demonstrations, making their recommendations easier for patients to understand.

Virtual Supportive Environment

Access to physicians continues to be difficult for many people, with 9.5% of Americans being uninsured and 23% being underinsured in 2023 [20]. Access to dermatologists also varies by location as 88% of US counties have 0 dermatologists [21]. Although these videos are not and do not claim to be a replacement for regular care from a board-certified dermatologist, it is important to recognize that, in the context of barriers to care, they are still a helpful way for users to obtain accurate and reliable general information about their disease.

XSL•FO
RenderX

Additionally, the comment sections under these videos may provide much-needed validation and support.

Across all video categories, comments discussing users' personal experiences or giving and receiving advice were significantly more common than comments consisting of critiques, jokes, or off-topic messages (Figure 5). In other words, most comments were "productive" as personal experiences and interuser advice can be helpful for both patients and medical professionals to read and share. This is encouraging as it suggests that the discussions prompted by physician-produced content are likely to remain beneficial and unlikely to become negative, hateful, or unproductive.

Additionally, the fact that sharing personal experiences and advice was more prominent than other types of comments suggests that patients with AD may be using comment sections to seek support. Dermatological conditions, including AD, are often highly visible and impact the quality of day-to-day life for patients. For patients with AD, these videos and comment sections may provide emotionally supportive content by reminding them that others have similar experiences (eg, "Thank you for telling me all this information makes me happy to see other people have eczema. I'm not the only one. Thank you."), allowing them to vent (eg, "The itching is so horrible but the peeling that happens afterwards is very painful."), or providing them with nonmedical recommendations that can ease the disease's visibility or impact on quality of life for the patient (eg, "Also with eczema avoid fragrance in your skin products and avoid anything with alcohol it will dry your skin or ask your dermatologist for the dupixent shot which I'm on and has helped significantly life changing").

The idea that users may be seeking supportive environments is reinforced by the fact that, among "personal experience" comments, negative personal experiences were more commonly shared than positive or neutral experiences (Figure 6). This suggests that people may have been engaging with the video and potentially finding it in the first place via search terms due to frustration or stress about their disease or in an attempt to find validation as they navigated their AD.

Additionally, among secondary comments, recommendations and emotional support comments trended higher than other types of comments (Figure 7). This shows that users were not only sharing their own experiences and frustrations but also responding to and relating to each other in a transient way. For instance, in response to one comment saying, "I hate having eczema especially when I have it behind my knee it's so hard to reach," another user replied, "It really is i feel your pain same here."

Future studies could further evaluate the quality, reciprocity, and time frame of interactions among commenters to better understand the role that comment sections play in providing a supportive environment. Additionally, it would be helpful to understand whether the same users typically engage with multiple AD-related videos and, thus, form a more connected network.

Darker Skin Tones

Notably, we found that only 7.1% (2/28) of the videos had content involving darker skin tones. A previous study found that, when looking specifically for Black skin on TikTok, posts were even less likely to be created by dermatologists compared to general dermatology-related content [22]. Our findings reinforce this content gap and highlight the need for greater inclusivity of darker skin tones in physician-created dermatology content on TikTok. This is especially important when viewing the online spaces created by this content as supportive environments as exclusion of darker skin tones in content may translate to people of color being denied online support and validation.

Limitations

Our study analyzed TikTok videos retrieved through searches conducted in December 2023. Given the constantly evolving nature of social media, and TikTok in particular, our findings are limited to the period in which the videos were posted (June 2020 to December 2023) and may not reflect content from other time frames. Additionally, studies have critiqued systematic analysis of TikTok content as limited by variability in the content presented to each user based on the app's algorithm [23]. We attempted to improve replicability by using 3 separate accounts, including 1 newly created account with no established user preferences, to search for AD-related videos. Among the "Top" videos collected by the new account, 60% of the videos (18/30) and 75% (21/28) of the video creators overlapped with the "Top" videos collected by the other 2 accounts, suggesting some similarity among video presentation across users with different personalized algorithms. However, a larger number of varying accounts would need to be studied to confirm this. It is also important to note that the "Top" videos under a search term may not accurately reflect the videos that are most likely to appear in a user's "For You Page" or personalized feed. The "Top" videos may also not be representative of all AD content on TikTok.

Another limitation is that comment sections represent only a subset of all viewers and may not accurately reflect the general population's perceptions of these videos. The categorizations in this study were based on interpretations of online comments, and because we did not directly interview individuals, some comments may have been misinterpreted. This is particularly challenging due to the text form nature of comments, which makes tone and intention difficult to judge. Therefore, future research could incorporate qualitative interviews with patients with AD and their loved ones to gain a deeper understanding of how they are affected by these videos and what value they derive from them.

It is also important to note that our methodology excluded all creators whose board certification status was not publicly verifiable, resulting in mostly US-based creators, which may have disproportionately excluded international creators. Additionally, the formula used to calculate the engagement rate may have resulted in inflated engagement rates, particularly for videos with many spam comments or creator replies. Finally, the subset used for interrater reliability only represented 1.7%

XSL•FO
RenderX

(20/1149) of the total comment dataset and may constitute a limitation in the reliability of ratings across all samples.

Conclusions

Our study suggests that videos created by dermatologists on TikTok are generally well received. The comment sections appear to provide transient supportive environments where users connect over shared challenges and exchange personal experiences and recommendations. While previous studies have identified virtual support spaces, they have primarily focused on patient-generated content. Our finding that supportive interactions also occur under physician-produced videos suggests that increasing physician-created content on social media could help expand access to accurate, evidence-based information while maintaining supportive communities for those affected by an emotionally and socially distressing illness such as AD.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Video metric information ordered by engagement rate.
[PNG File, 467 KB - derma_v9i1e90649_app1.png]

Multimedia Appendix 2
Video categorization ordered by engagement rate.
[PNG File, 437 KB - derma_v9i1e90649_app2.png]

References

1. Harries B, Zenone M, Hartwell G. Feeling lonely in the online crowd: what TikTok tells us about young people and loneliness. Health Promot Int 2025 Jul 1;40(4):daaf094. [doi: 10.1093/heapro/daaf094] [Medline: 40621771]
2. Silverberg JI, Feldman SR, Smith Begolka W, et al. Patient perspectives of atopic dermatitis: comparative analysis of terminology in social media and scientific literature, identified by a systematic literature review. J Eur Acad Dermatol Venereol 2022 Nov;36(11):1980-1990. [doi: 10.1111/jdv.18442]
3. Silverberg JI. Public health burden and epidemiology of atopic dermatitis. Dermatol Clin 2017 Jul;35(3):283-289. [doi: 10.1016/j.det.2017.02.002] [Medline: 28577797]
4. Margolis JS, Abuabara K, Bilker W, Hoffstad O, Margolis DJ. Persistence of mild to moderate atopic dermatitis. JAMA Dermatol 2014 Jun;150(6):593-600. [doi: 10.1001/jamadermatol.2013.10271] [Medline: 24696036]
5. Silverberg JI, Gelfand JM, Margolis DJ, et al. Patient burden and quality of life in atopic dermatitis in US adults: a population-based cross-sectional study. Ann Allergy Asthma Immunol 2018 Sep;121(3):340-347. [doi: 10.1016/j.anai.2018.07.006] [Medline: 30025911]
6. Marty T, Khadar M, Autelitano L, et al. Patients' testimonies, feelings, complaints and emotional experiences with dermatoses on open social media: the French infodemiologic patient's free speech study. J Eur Acad Dermatol Venereol 2024 Jul;38(7):1373-1382. [doi: 10.1111/jdv.19781] [Medline: 38314863]
7. Toy J, Gregory A, Rehmus W. Barriers to healthcare access in pediatric dermatology: a systematic review. Pediatr Dermatol 2021 Nov;38 Suppl 2:13-19. [doi: 10.1111/pde.14748] [Medline: 34338358]
8. Borkhetaria RK, Devireddy N, Cannon N, Soni A, Ely AL. An analysis of pediatric ophthalmology content on TikTok. J Pediatr Ophthalmol Strabismus 2024;61(2):90-97. [doi: 10.3928/01913913-20230627-01] [Medline: 37489626]
9. Mansour MR, Abushukur Y, Potts GA. Keratosis pilaris on TikTok: a cross-sectional analysis of trending content. JAAD Int 2022 Jun;8:116-117. [doi: 10.1016/j.jdin.2022.06.015] [Medline: 35875398]
10. Nguyen M, Youssef R, Kwon A, Chen R, Park JH. Dermatology on TikTok: analysis of content and creators. Int J Womens Dermatol 2021 May;7(4):488-489. [doi: 10.1016/j.ijwd.2021.04.006] [Medline: 34621965]
11. Zehrung RF, Chen Y. Self-expression and sharing around chronic illness on TikTok. AMIA Annu Symp Proc 2024;2023:1334-1343. [Medline: 38222376]
12. Kirubarajan A, Han S, Patel P, Black A, Sierra S. TikTok as a resource for fertility information and support: a patient survey. J Family Reprod Health 2024 Dec;18(4):240-245. [doi: 10.18502/jfrh.v18i4.17424] [Medline: 40342653]
13. Davis MA, Zheng K, Liu Y, Levy H. Public response to Obamacare on Twitter. J Med Internet Res 2017 May 26;19(5):e167. [doi: 10.2196/jmir.6946] [Medline: 28550002]

XSL•FO
RenderX

14. Boatman D, Jarrett Z, Starkey A, Conn ME, Kennedy-Rea S. HPV vaccine misinformation on social media: a multi-method qualitative analysis of comments across three platforms. PEC Innov 2024 Aug;5:100329. [doi: 10.1016/j.pecinn.2024.100329] [Medline: 39206222]
15. Berdida DJ, Franco FM, Santos XA, et al. Filipinos' COVID-19 vaccine hesitancy comments in TikTok videos: a manifest content analysis. Public Health Nurs 2023 Jan;40(1):135-143. [doi: 10.1111/phn.13143] [Medline: 36300833]
16. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS One 2014;9(10):e110171. [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Wilkens FM, Ganter C, Kriegsmann K, et al. YouTube-videos for patient education in lymphangioleiomyomatosis? Respir Res 2022 Apr 27;23(1):103. [doi: 10.1186/s12931-022-02022-9] [Medline: 35477513]
19. Heales LJ, Gunasinghe B, Melrose E, Lorraway J, Obst SJ. A systematic review of the quality and engagement of TikTok videos of resistance exercises for the management of lateral elbow tendinopathy. J Hand Ther 2025;38(4):834-840. [doi: 10.1016/j.jht.2025.02.016] [Medline: 40274443]
20. Collins SR, Gupta A. The state of health insurance coverage in the U.S.: findings from the Commonwealth Fund 2024 Biennial Health Insurance Survey. The Commonwealth Fund. 2024. URL: https://doi.org/10.26099/byce-qc28 [accessed 2026-07-01]
21. Vaidya T, Zubritsky L, Alikhan A, Housholder A. Socioeconomic and geographic barriers to dermatology care in urban and rural US populations. J Am Acad Dermatol 2018 Feb;78(2):406-408. [doi: 10.1016/j.jaad.2017.07.050] [Medline: 29332710]
22. Campbell JR, Woolery-Lloyd H, Williams K. TikTok and Black skin: is this a missed opportunity for dermatologists? J Drugs Dermatol 2023 Jul 1;22(7):698-700. [doi: 10.36849/JDD.7061] [Medline: 37410044]
23. Kanthawala S, Cotter K, Foyle K, Decook JR. It's the methodology for me: a systematic review of early approaches to studying TikTok. Presented at: Hawaii International Conference on System Sciences; Jan 4-7, 2022. [doi: 10.24251/HICSS.2022.382]

Abbreviations

AD: atopic dermatitis

Edited by R Dellavalle; submitted 31.Dec.2025; peer-reviewed by N Hu, T Schopf; revised version received 15.Jun.2026; accepted 17.Jun.2026; published 16.Jul.2026.

Please cite as:
Gangavarapu S, Khan T, McCarthy M, Hales M, Rangel SM
Comments on Dermatologists' TikTok Videos on Atopic Dermatitis: Content Analysis
JMIR Dermatol 2026;9:e90649
URL: https://derma.jmir.org/2026/1/e90649
doi:10.2196/90649

© Sahithi Gangavarapu, Tazeena Khan, Morgan McCarthy, Molly Hales, Stephanie M Rangel. Originally published in JMIR Dermatology (http://derma.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhanced Diagnosis of Generalized Pustular Psoriasis With the Legit.Health Device as a Diagnosis Support Tool: Multireader Multicase Study

Alfonso Medela[1], MSc; Ignacio Hernández Montilla[1], MSc; Alberto Sabater[1], PhD; Andy Aguilar[2], BA; Taig Mac Carthy[2], BA; Gurpreet Singh Chowdhry[3], MSc, MBA; Juan Semeco[4], MD, MBA; Antonio Martorell[2,5], MD, PhD

[1]Department of Medical Data Science, Legit.Health, Gran Via 1 BAT Tower, Bilbao, Spain

[2]Department of Clinical Endpoint Innovation, Legit.Health, Bilbao, Spain

[3]TA Inflammation, Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany

[4]TA Inflammation Medicine, Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany

[5]Dermatology Department, Hospital de Manises, Valencia, Spain

Corresponding Author:
Alfonso Medela, MSc
Department of Medical Data Science, Legit.Health, Gran Via 1 BAT Tower, Bilbao, Spain

Abstract

Background: Generalized pustular psoriasis (GPP) is a rare, chronic, systemic inflammatory disease with an unpredictable and heterogeneous clinical course characterized by chronic symptoms and periods of flaring. GPP presents diagnostic challenges due to its rarity and high similarity to other dermatologic diseases.

Objective: This study aimed to assess the performance of Legit.Health, a medical device powered by artificial intelligence software, in assisting health care practitioners (HCPs) in identifying GPP.

Methods: The medical device used in this study includes a deep neural network for skin disease recognition trained on thousands of images of over 200 skin conditions (classified based on the *International Classification of Diseases, 11th Revision*). The sensitivity and specificity of the algorithm for the differential diagnosis of GPP were assessed. Due to the scarcity of GPP-related images, the medical device was fine-tuned using a dataset that included 4397 new GPP images. Thereafter, 15 HCPs (n=11, 73.3% primary care practitioners and n=4, 26.7% dermatologists) prospectively reviewed a total of 100 images of 15 visually similar skin conditions virtually in a clinical setting. After their diagnostic prediction, Legit.Health provided a prompt giving them the top 5 possible skin conditions to assist them with their choice.

Results: Legit.Health demonstrated high accuracy in identifying GPP, with top-1, top-3, and top-5 sensitivity and specificity of 0.80, 0.86, and 0.90 and 0.99, 0.99, and 0.96, respectively. Results showed a notable increase in the diagnostic accuracy of the HCPs with assistance from Legit.Health (P<.001), with an increase in GPP diagnostic accuracy of 22.97% overall, 24.24% for primary care practitioners, and 19.45% for dermatologists.

Conclusions: This improvement highlights the potential of Legit.Health in assisting HCPs in diagnosing rare diseases such as GPP, particularly in primary care settings where expertise may be limited, thereby improving patient outcomes.

Trial Registration: ClinicalTrials.gov NCT03782792; https://clinicaltrials.gov/study/NCT03782792 and NCT04399837; https://clinicaltrials.gov/study/NCT04399837

(JMIR Dermatol 2026;9:e82030) doi:10.2196/82030

KEYWORDS

generalized pustular psoriasis; psoriasis; hidradenitis suppurativa; artificial intelligence; AI; rare diseases; diagnostic accuracy

Introduction

Generalized pustular psoriasis (GPP) is a rare, systemic neutrophilic inflammatory disease with an unpredictable and heterogeneous clinical course [1-5], which is characterized by chronic symptoms such as pustules, erythema, and periods of flaring [5,6]. GPP is associated with a considerable clinical burden, which can greatly impact patients' lives both physically and emotionally. Patients with GPP may experience multiple flares per year that can be triggered by infections, stress, medication withdrawal (such as from corticosteroids and biologics), menstruation, and pregnancy [5,7-12]. GPP flares and symptoms are often unpredictable and highly heterogeneous due to the presence of cyclical variations in disease severity that occur over long periods [13,14]. GPP manifests as painful, itchy, and visible pustules and also induces systemic symptoms,

XSL•FO
RenderX

including fever and fatigue [5]. Without timely and appropriate intervention, GPP flares can lead to life-threatening complications that require emergency treatment, such as multisystem organ failure and sepsis [5,6]; therefore, there is a critical need for prompt diagnosis and effective management of symptoms.

A study by Reisner et al [15] of 66 patients living with GPP in the United States showed that 86% experienced 2 or more flares annually. Despite the profound impact of GPP on patients, the diagnostic process can be challenging. Insights from the study by Reisner et al [15] highlighted significant delays and misdiagnoses experienced by patients. Approximately 40% of surveyed patients reported that it took years for them to receive an accurate diagnosis of GPP, with over 50% being consulted by multiple health care practitioners (HCPs) before the correct diagnosis was made [15]. A common reason for delays in diagnosis is the frequent misdiagnosis of GPP as an infection, leading to incorrect treatments and delayed specialist referrals [16]. This misdiagnosis is attributed to the rarity of GPP and the consequent lack of experience and knowledge of the presentation and pathogenesis of GPP among many HCPs [16].

The impact of GPP extends beyond physical symptoms as GPP can have a profound effect on patients' emotional health and quality of life, influencing daily activities, work, and social relationships [15]. The patients in the study by Reisner et al [15] reported a significant emotional impact, with 71% reporting that they lived in fear of flares and two-thirds reporting that they experienced anxiety related to their GPP. Emotional stress, which is itself a major trigger for flares, affected 83% of those surveyed [15].

In recent years, artificial intelligence (AI) has been increasingly implemented in the medical field, offering promising advancements in disease diagnosis and management. However, despite the increased number of AI algorithms, their implementation in clinical practice remains challenging. Many AI tools lack the appropriate medical device certification and fail to prove their usefulness in real-world settings. Therefore, there is a pressing need for AI solutions that are technically proficient and beneficial in real-world clinical practice.

Legit.Health, a medical device powered by AI, uses computer vision algorithms (deep neural networks) that have been shown to be effective in many clinical settings, such as in assessing the severity of hidradenitis suppurativa (HS) [17], urticaria [18], and atopic dermatitis [19]. In this study, we aimed to evaluate the diagnostic accuracy of HCPs for GPP before and after the use of the Legit.Health device and its skin disease recognition algorithm. This multireader, multicase study is the first to evaluate the real-world clinical utility of an AI-powered medical device in improving diagnostic accuracy for GPP, addressing a critical gap in dermatologic care.

Methods

The study methodology adheres to the CLEAR Derm (Checklist for Evaluation of Image-Based AI Reports in Dermatology) guidelines [20].

Device Fine-Tuning and Validation

Dataset

Legit.Health, a class IIb, CE-marked medical device (International Organization for Standardization 13485; approval number: MD792784; certificate number: MDR 792790; single registration number: ES-MF-000025345), among other features, incorporates a neural network for skin disease recognition trained on an extensive dataset, which comprises hundreds of thousands of skin images and represents over 200 categories from the *International Classification of Diseases, 11th Revision* (*ICD-11*). Images from the Legit.Health dataset were acquired using a combination of digital cameras, smartphone cameras, and dermatoscopes. Although this dataset already included a small subset of images of patients with GPP, 2 additional batches of images (Figure 1) were incorporated to enhance the differential diagnosis capabilities for GPP. These 2 batches of patient images were supplied by Boehringer Ingelheim and acquired using digital cameras. The first batch comprised 101 images, most of which focused on close-up views of small and discrete body sites (Figure 1A-C). These images provided detailed views of the pustules, which are characteristic of GPP, facilitating precise lesion recognition and analysis. The second batch comprised 4296 high-resolution images depicting larger body sites (ie, images of the trunk and legs) viewed from both the front and back (Figure 1D-F). This batch of images provided a comprehensive visual context, capturing the widespread distribution of pustules that is typical in GPP. All images were sourced from patients with confirmed GPP diagnoses (see Table 1 for demographic information), ensuring the reliability and accuracy of the dataset. In both batches, images depicting a patient with GPP with no visible signs at the time of the picture were also identified; therefore, no additional review or diagnosis of each image was required. These patients were not actively involved in the testing of the Legit.Health device, and no patient images were presented to the public. Once the new batches of GPP images were provided, they were properly stratified into training, validation, and test sets to prevent data leakage. The training and validation images were then combined with the Legit.Health dataset to refine the algorithm's response for GPP.

XSL•FO
RenderX

Figure 1. Examples of generalized pustular psoriasis images from the dataset. Images A, B, and C correspond to samples from the first batch, while D, E, and F correspond to samples from the second.



Table . Summary of patient demographics for the dataset of generalized pustular psoriasis images (N=4397).

Sex	Number of patients	Samples, n (%)
Female	36	3024 (68.8)
Male	21	1373 (31.2)

For this study, we prepared a subset of 100 images representing a range of dermatologic conditions, including GPP and various conditions commonly encountered in primary and secondary care that may resemble GPP and which, thereby, complicate the diagnostic process. The list of conditions included in this study was defined by one of the authors (A Martorell) and aimed to simulate real-world clinical scenarios and test the device's ability to accurately differentiate GPP from other similar skin conditions. To create this subset, we picked images from the test split of the Legit.Health dataset and the test split of the new GPP image batches. No formal sample size calculation was carried out. This initial dataset was deemed sufficient to answer the research question due to the number of images within the existing neural network for skin disease recognition incorporated into the Legit.Health medical device.

Fine-Tuning

As indicated above, Legit.Health already included GPP within its diagnostic outputs, classified as EA90.40 according to the *ICD-11* [21]. However, due to the scarcity of GPP cases and related GPP images compared with other *ICD-11* categories, the medical device's algorithm was fine-tuned using an updated dataset that included the 2 new batches of GPP images. We began by detecting the images that did not present visual signs of GPP (6/101, 5.9% of the images from the first batch and 688/4296, 16.0% of the images from the second batch) and removing them from the dataset's GPP category. The remaining images (n=3608) were split into training (n=2563), validation (n=546), and test (n=499) sets. As the images included metadata such as patient identification codes, we were able to stratify the dataset and avoid data leakage at every stage. The device's skin disease recognition algorithm was retrained using the original setup and procedures established by the manufacturer,

XSL•FO
RenderX

integrating the new GPP images with the existing validated dataset.

To enhance the performance and generalizability of the medical device, specific data augmentation techniques such as random color jittering, histogram equalization, motion blur, rotation, and resizing were applied to the GPP images. These techniques accounted for variations in image quality, lighting conditions, and background features, thereby increasing the robustness of the medical device. The augmented dataset helped in simulating a diverse range of clinical scenarios, which is crucial for the accurate recognition of GPP in real-world settings.

The training process used superconvergence methods, significantly expediting the learning phase and optimizing the device's ability to handle the increased volume of data. Superconvergence leverages a cyclic learning rate schedule, which allows the device to converge rapidly while maintaining high accuracy even with the substantial addition of new images [22].

Validation

The performance of Legit.Health was evaluated using the top-1, top-3, and top-5 sensitivity and specificity metrics for the GPP category. While top-1 sensitivity and specificity were equivalent to the common sensitivity and specificity metrics, the top-3 and top-5 variants accounted for cases in which the correct GPP category was not exactly in the first position of the device's output but appeared among the top 3 and top 5 predicted categories. These variants provided a more realistic understanding of the device's performance; moreover, they provided a comprehensive understanding of the medical device's accuracy and its ability to correctly identify GPP among other dermatologic conditions.

Clinical Study

A multireader, multicase clinical study was conducted to assess the effectiveness of Legit.Health in helping enhance diagnostic accuracy for GPP among HCPs. This evaluation involved 15 HCPs (female: n=8, 53.3%) comprising 11 (73.3%) primary care practitioners (PCPs; female: n=6, 54.5%) and 4 (26.7%) dermatologists (female: n=2, 50%), thereby providing insights on the usefulness of the medical device across different levels of dermatologic expertise.

Participant Selection and Training

The HCPs (all based in Spain) who participated in this study were selected based on their clinical experience and willingness to commit to the full evaluation process. Prior to the evaluation, the HCPs received training on how to use Legit.Health. This included an overview of the device's diagnostic capabilities, guidance on interpreting its output, and an understanding of the confidence metrics provided by the system. They were also given a brief introduction to the evaluation experiment without any hints of which conditions they would be presented with.

Design

The multireader, multicase clinical evaluation, which took place virtually, consisted of reviewing 100 high-resolution images of 15 various skin conditions, including GPP and other visually similar dermatologic conditions. All participants were presented with the same dataset from the combined image pool. Each image was accompanied by a detailed patient anamnesis, including medical history, systemic symptoms, and any other relevant laboratory findings. This information was intended to mimic real-world clinical scenarios, providing a holistic context for the diagnosis. Each case was structured as a 2-stage sequence (Figure 2) to measure the impact of the Legit.Health medical device on the accuracy of HCPs in diagnosing GPP and other conditions.

In the first stage, HCPs were presented with the image and clinical information and recorded their initial diagnosis without assistance from the medical device. This first stage aimed to capture the baseline diagnostic accuracy of the HCP.

The second stage aimed to evaluate how the integration of Legit.Health aided the diagnostic process and whether it improved the identification of GPP by HCPs. The image was supplemented with the medical device's top 5 diagnostic predictions, along with the corresponding confidence values for each prediction. These predictions were intended to serve as a second opinion, aiding the HCPs in refining their initial diagnoses. The HCPs were encouraged to consider the medical device's output during their final diagnostic decisions, using its insights to enhance the accuracy and confidence of their initial clinical assessments.

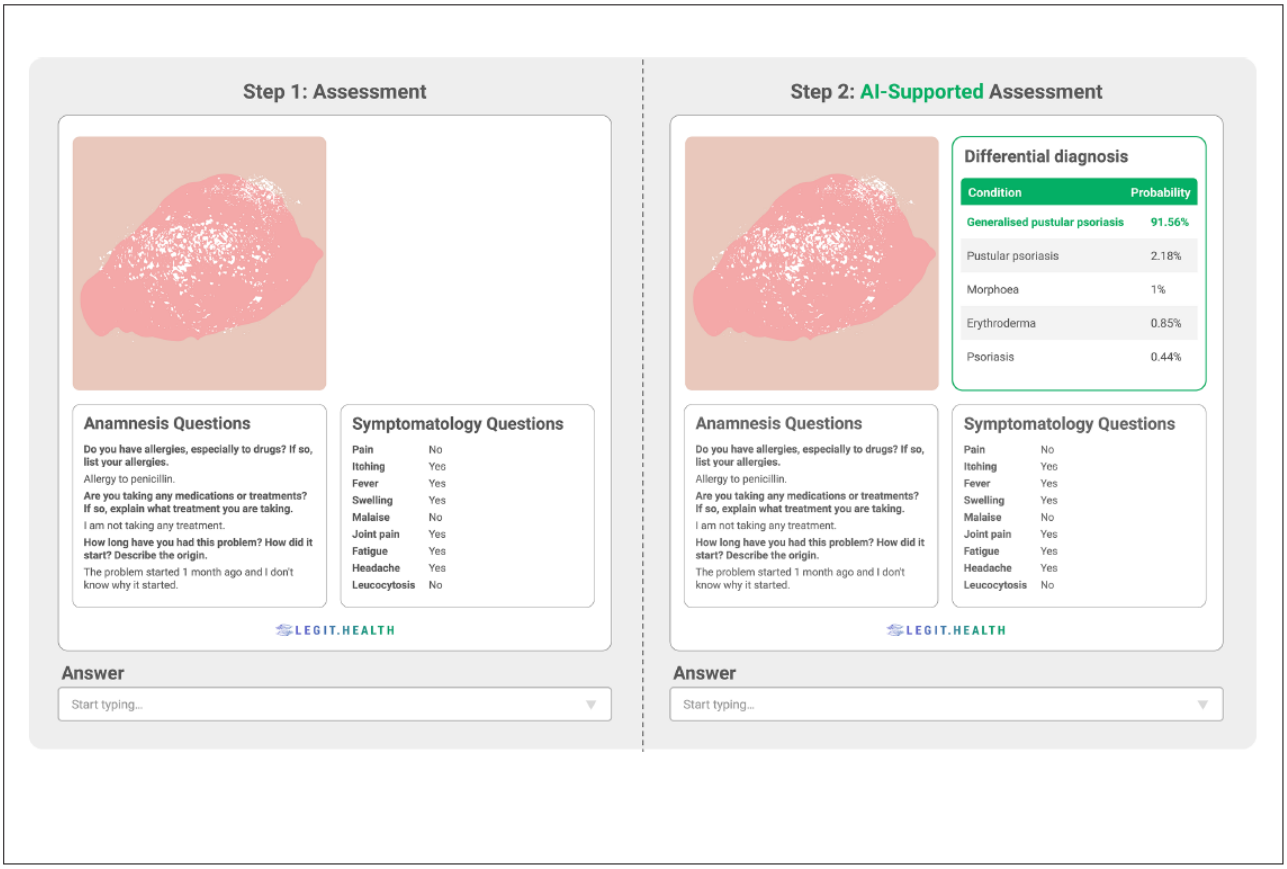
In both stages, participants had to make their choices from a list of 406 conditions, which could be easily explored by typing in either English or Spanish. This list, which was curated from Legit.Health's internal dataset, covered all pathologies encompassed by the device at different levels of detail, as well as those used in this study. Once all responses were collected, they were normalized to avoid redundant or synonymous terms.

Image quality may impact the effectiveness of the Legit.Health medical device and is a major concern both in clinical practice and in research. If the user inputs an image into the device that does not capture the region of interest correctly or that does not provide enough resolution to correctly identify the condition based on the pixels in the image, the reliability of the output may decrease. To overcome this limitation in this study, we used a specialized neural network for Dermatology Image Quality Assessment (DIQA), which was provided by Legit.Health [23]. All images included in this clinical evaluation were subjected to DIQA and successfully met the image quality requirements.

The participants first accessed the platform on June 24, 2024, on their computers. The first answer was recorded on June 25, 2024, and the last answer was recorded on July 6, 2024.

XSL•FO
RenderX

Figure 2. Comparison between the first and second stages of clinical evaluation. AI: artificial intelligence.



Ethical Considerations

This study analyzed patient images from the EFFISAYIL 1 (NCT03782792) and EFFISAYIL 2 (NCT04399837) clinical trials, which were conducted in accordance with the trial protocols, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use good clinical practice guidelines, Regulation 536/2014 (European Union), the Japanese good clinical practice regulations, and applicable local regulations. The studies were approved by the ethics committees of participating institutions and countries. The study protocols were approved at each study site and/or country.

All patients from the EFFISAYIL 1 and EFFISAYIL 2 trials provided written informed consent, and confidentiality agreements were in place between the authors and Boehringer Ingelheim. Patient consent was obtained for the inclusion of patient photographs in this publication. More detailed information regarding data compilation, sharing policies, review board approvals, and ethical standards can be found in the original source publications for these trials [24,25].

The rest of the images were extracted from public and open dermatology atlases and open web images. Due to the retrospective use of anonymized, open-source data, no separate approval was required from an ethics committee. This study did not require ethics committee approval because it is observational and noninterventional.

The original protocols for the Effisayil 1 (NCT03782792) and Effisayil 2 (NCT04399837) trials explicitly allow for the sharing of deidentified participant-level data and clinical study documents for secondary analyses. The data-sharing process ensures that all shared materials are deidentified to protect participant privacy and remain within the scope of the original informed consent. Access to the clinical study data and images was obtained in accordance with the sponsor's (Boehringer Ingelheim) Transparency and Publication Policy. This access is granted to external authors to fulfill their roles under ICMJE criteria and is governed by a formal Legal Agreement that authorizes the secondary use of the data for this manuscript.

Results

Device Performance

The Legit.Health medical device achieved top-1, top-3, and top-5 sensitivity values of 0.80, 0.86, and 0.90, respectively, with specificity values exceeding 0.90 (Table 2), indicating robust performance across various diagnostic thresholds.

XSL•FO
RenderX

Table . Generalized pustular psoriasis (GPP) detection performance of Legit.Health before and after fine-tuning with additional GPP images.

Metric	Before fine-tuning	After fine-tuning
Top-1 sensitivity	0.4000	0.8026
Top-1 specificity	1.0000	0.9983
Top-3 sensitivity	0.8000	0.8633
Top-3 specificity	0.9994	0.9957
Top-5 sensitivity	0.8000	0.9002
Top-5 specificity	0.9986	0.9611

To better understand the medical device's performance, we inspected the GPP images from the test set when another condition was predicted (top 1 prediction). The top 5 conditions with which GPP was most confused are shown in Table 3.

Table . Summary of the classes predicted in the top-1 position instead of generalized pustular psoriasis (GPP) assessed by Legit.Health**.**

Top 1 predicted condition	*ICD-11*[a] code [21]	Number of misclassified GPP images
Generalized eczematous dermatitis of unspecified type	EA89	16
Pustular psoriasis[b]	EA90.4	15
Nonspecific lesion (no condition)	—[c]	12
Plaque psoriasis	EA90.0	12
Morphea	EB61	5

[a]*ICD-11*: *International Classification of Diseases, 11th Revision*.

[b]Pustular psoriasis (EA90.4) encompasses a set of images that include localized pustular psoriasis subcategories such as palmoplantar pustulosis (EA90.42); other specified pustular psoriasis (EA90.4Y); and pustular psoriasis, unspecified (EA90.4Z). GPP (EA90.40) was excluded from the pustular psoriasis group and included as an independent category.

[c]Not applicable.

Clinical Evaluation

The overall diagnostic accuracy of HCPs across several different dermatologic conditions improved significantly with the use of Legit.Health. PCPs exhibited an increase in diagnostic accuracy of 17.74%, whereas dermatologists exhibited an improvement of 8.4%. Dermatologists changed their diagnoses more often than PCPs; on average, the switch ratio (ie, how many times a participant changed their responses) for dermatologists was 0.87 (SD 0.04), whereas for PCPs, it was 0.67 (SD 0.13). The overall switch ratio considering all HCPs was 0.72 (SD 0.14). The overall impact of the device was assessed using a McNemar test (Table 4), with the conclusion that the device had an impact on the improvement in diagnostic capabilities ($P<.001$).

Table . Results of the McNemar tests for all groups. The resulting contingency tables from these McNemar tests have been reorganized to represent positive rate (correct answer after device use), neutral positive rate (correct answer before and after use of the device), neutral negative rate (incorrect answer before and after use of the device), and negative rate (incorrect answer after use of the device).

Group	Positive rate, % (n/N)	Neutral positive rate, % (n/N)	Neutral negative rate, % (n/N)	Negative rate, % (n/N)
HCPs[a]	16.76 (236/1408)	46.31 (652/1408)	35.3 (497/1408)	1.63 (23/1408)
Primary care practitioners	19.61 (199/1015)	42.46 (431/1015)	36.06 (366/1015)	1.87 (19/1015)
Dermatologists	9.41 (37/393)	56.23 (221/393)	33.33 (131/393)	1.02 (4/393)

[a]HCP: health care practitioner.

To better understand the influence of the device outputs on the participants' responses, we counted the number of diagnosis changes across several top-1 probability ranges and the position of the correct diagnosis. The motivation behind using only the top-1 probability is that it serves as a simple proxy for the system's overall certainty for a given image. The results, presented in Table 5 & Table 6, show that participants who changed their diagnoses did so regardless of the device's confidence or the position of the confirmed, ground-truth diagnosis.

XSL•FO
RenderX

Table . Influence of the device top-1 output and the ratio of changed diagnoses.

Device top-1 confidence	Number of images in range	Switch ratio
0.0 - 0.2	5	0.63
0.2 - 0.4	5	0.60
0.4-0-6	4	0.80
0.6 - 0.8	17	0.71
0.8 - 1.0	69	0.75

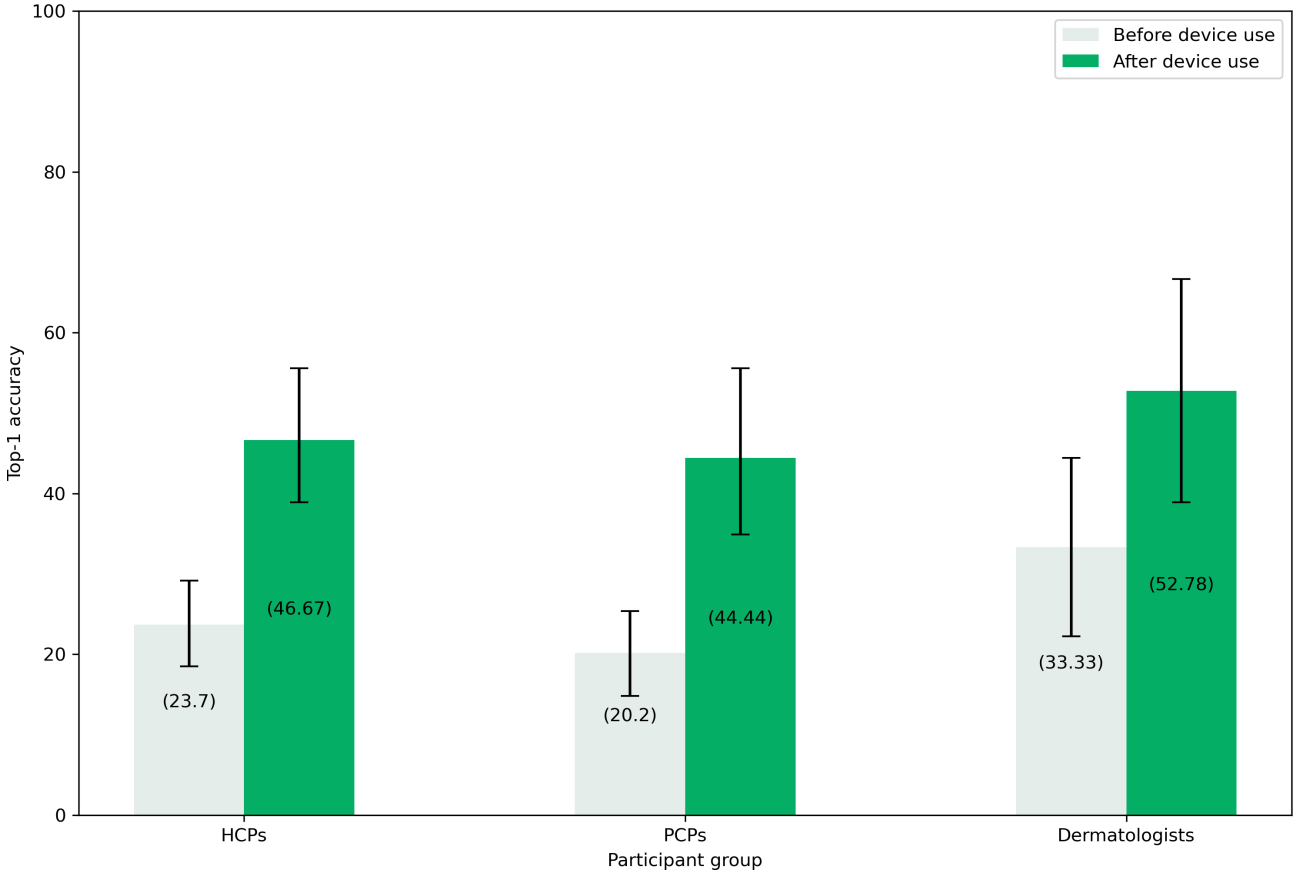
Table . Influence of the position of the correct condition on the ratio of changed diagnoses.

Position of correct condition	Number of images	Switch ratio
Top 1	76	0.67
Top 2	8	0.74
Top 3	2	0.71
Top 4	2	0.70
Top 5	1	0.77
Not in top 5	11	0.87

The impact that Legit.Health had on diagnostic performance varied across conditions. For GPP (Figure 3), the diagnostic accuracy of HCPs increased from 23.7% to 46.7%, representing a 22.97% increase (Table 7). PCPs experienced a 24.24% increase in diagnostic accuracy (Table 8), whereas dermatologists experienced a 19.45% increase (Table 9). For HS, PCPs showed a 10.11% increase in diagnostic accuracy, with no substantial change observed for dermatologists (2.85%). One of the most notable improvements was observed in the diagnosis of palmoplantar pustulosis, where PCPs exhibited a 47.82% increase in diagnostic accuracy. Overall, Legit.Health enhanced diagnostic performance in 16.76% of cases, demonstrating significant diagnostic improvements (P<.001) for dermatologic conditions that are frequently misdiagnosed due to similar clinical presentations.

XSL•FO
RenderX

Figure 3. Top-1 accuracy of participants for generalized pustular psoriasis diagnosis before and after use of the device (with 95% CIs). HCP: health care practitioner; PCP: primary care practitioner.



XSL•FO
RenderX

Table . Top-1 accuracy of all participants.

Condition	*ICD-11*[a] code	Accuracy before use of the device (%; 95% CI)	Accuracy after use of the device (%; 95% CI)	Difference (%; 95% CI)
Acne	ED80	37.50 (29.27-46.15)	56.25 (47.06-66.67)	18.75 (8.50-29.41)
Acne conglobata	ED80.41	18.40 (12.31-23.81)	38.40 (27.03-50.00)	20.00 (13.52-27.42)
Drug-induced acute generalized exanthematous pustulosis	EH67.0	5.00 (0.00-9.38)	5.00 (0.00-9.38)	0.00 (0.00-0.00)
Generalized eczematous dermatitis of unspecified type	EA89	71.34 (66.66-76.14)	73.17 (68.33-78.41)	1.83 (0.00-3.90)
Generalized pustular psoriasis	EA90.40	23.70 (18.52-29.17)	46.67 (38.89-55.56)	22.97 (18.52-28.39)
Hidradenitis suppurativa	ED92.0	85.48 (79.71-92.00)	93.55 (89.87-98.44)	8.07 (3.79-12.65)
Impetigo	1B72	57.43 (47.96-67.36)	76.35 (66.25-88.78)	18.92 (12.86-25.00)
Palmoplantar pustulosis	EA90.42	45.31 (30.56-56.76)	79.69 (71.43-88.89)	34.38 (22.58-47.37)
Pemphigus vulgaris	EB40.0	28.77 (18.17-40.00)	56.16 (40.00-74.42)	27.39 (16.00-39.40)
Plaque psoriasis	EA90.0	91.89 (87.18-97.50)	97.30 (94.87-100.00)	5.41 (2.04-9.09)
Seborrheic dermatitis and related conditions	EA81.Z	75.34 (68.42-82.86)	90.41 (82.86-100.00)	15.07 (9.99-20.93)
Seborrheic keratosis	2F21.0	94.67 (91.11-100.00)	97.33 (95.00-100.00)	2.66 (0.00-5.00)
Severe inflammatory acne	ED80.4Z	10.61 (4.88-16.13)	43.94 (33.33-53.85)	33.33 (26.19-41.38)
Subcorneal pustular dermatosis	EB2Y	2.67 (0.00-5.00)	2.67 (0.00-5.00)	0.00 (0.00-0.00)
Tinea corporis	1F28.Y	36.36 (28.85-42.60)	62.50 (55.00-71.15)	26.14 (17.28-34.80)

[a]*ICD-11*: *International Classification of Diseases, 11th Revision.*

XSL•FO
RenderX

Table . Top-1 accuracy of primary care practitioners.

Condition	*ICD-11*[a] code	Accuracy before use of the device (%; 95% CI)	Accuracy after use of the device (%; 95% CI)	Difference (%; 95% CI)
Acne	ED80	36.36 (25.93-46.89)	61.36 (50.00-73.91)	25.00 (10.81-40.93)
Acne conglobata	ED80.41	21.18 (13.46-28.89)	48.24 (35.19-62.97)	27.06 (20.00-35.71)
Drug-induced acute generalized exanthematous pustulosis	EH67.0	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Generalized eczematous dermatitis of unspecified type	EA89	68.33 (63.08-73.85)	70.83 (64.47-77.27)	2.50 (0.00-5.26)
Generalized pustular psoriasis	EA90.40	20.20 (14.81-25.40)	44.44 (34.92-55.56)	24.24 (18.50-31.48)
Hidradenitis suppurativa	ED92.0	82.02 (74.55-90.74)	92.13 (87.04-100.00)	10.11 (5.00-15.38)
Impetigo	1B72	50.00 (39.71-60.29)	72.22 (58.57-87.94)	22.22 (15.00-30.89)
Palmoplantar pustulosis	EA90.42	32.61 (23.08-41.94)	80.43 (74.07-86.21)	47.82 (37.93-58.35)
Pemphigus vulgaris	EB40.0	22.64 (10.53-36.36)	43.40 (22.86-66.67)	20.76 (8.58-33.38)
Plaque psoriasis	EA90.0	88.89 (82.76-94.87)	96.30 (92.00-100.00)	7.41 (2.86-11.77)
Seborrheic dermatitis and related conditions	EA81.Z	69.81 (60.71-78.79)	88.68 (78.57-100.00)	18.87 (12.12-26.09)
Seborrheic keratosis	2F21.0	92.73 (86.67-100.00)	96.36 (92.00-100.00)	3.63 (0.00-6.67)
Severe inflammatory acne	ED80.4Z	10.87 (3.85-17.24)	50.00 (37.92-62.96)	39.13 (30.77-48.39)
Subcorneal pustular dermatosis	EB2Y	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Tinea corporis	1F28.Y	29.69 (23.53-35.29)	59.38 (50.00-70.59)	29.69 (19.04-42.50)

[a]*ICD-11*: *International Classification of Diseases, 11th Revision.*

XSL•FO
RenderX

Table . Top-1 accuracy of dermatologists.

Condition	*ICD-11*[a] code	Accuracy before use of the device (%; 95% CI)	Accuracy after use of the device (%; 95% CI)	Difference (%; 95% CI)
Acne	ED80	40.00 (30.00-50.00)	45.00 (30.00-60.00)	5.00 (0.00-10.00)
Acne conglobata	ED80.41	12.50 (5.00-20.00)	17.50 (10.00-25.00)	5.00 (0.00-10.00)
Drug-induced acute generalized exanthematous pustulosis	EH67.0	18.75 (0.00-37.50)	18.75 (0.00-37.50)	0.00 (0.00-0.00)
Generalized eczematous dermatitis of unspecified type	EA89	79.55 (72.73-86.36)	79.55 (72.73-86.36)	0.00 (0.00-0.00)
Generalized pustular psoriasis	EA90.40	33.33 (22.22-44.44)	52.78 (38.89-66.67)	19.45 (16.67-22.23)
Hidradenitis suppurativa	ED92.0	94.29 (86.67-100.00)	97.14 (93.33-100.00)	2.85 (0.00-6.67)
Impetigo	1B72	77.50 (65.00-90.00)	87.50 (75.00-100.00)	10.00 (5.00-15.00)
Palmoplantar pustulosis	EA90.42	77.78 (50.00-100.00)	77.78 (50.00-100.00)	0.00 (0.00-12.50)
Pemphigus vulgaris	EB40.0	45.00 (30.00-60.00)	90.00 (80.00-100.00)	45.00 (30.00-60.00)
Plaque psoriasis	EA90.0	100.00 (100.00-100.00)	100.00 (100.00-100.00)	0.00 (0.00-0.00)
Seborrheic dermatitis and related conditions	EA81.Z	90.00 (80.00-100.00)	95.00 (90.00-100.00)	5.00 (0.00-10.00)
Seborrheic keratosis	2F21.0	100.00 (100.00-100.00)	100.00 (100.00-100.00)	0.00 (0.00-0.00)
Severe inflammatory acne	ED80.4Z	10.00 (0.00-20.00)	30.00 (20.00-40.00)	20.00 (10.00-30.00)
Subcorneal pustular dermatosis	EB2Y	10.00 (0.00-20.00)	10.00 (0.00-20.00)	0.00 (0.00-0.00)
Tinea corporis	1F28.Y	54.17 (41.67-66.67)	70.83 (58.33-83.33)	16.66 (8.33-25.00)

[a]*ICD-11*: *International Classification of Diseases, 11th Revision*.

Discussion

Principal Findings

In this study, the Legit.Health medical device enhanced HCPs' diagnostic accuracy for GPP, especially among PCPs with limited specialized dermatologic expertise. Legit.Health demonstrated its potential in helping HCPs distinguish GPP from other dermatologic conditions, which could fulfill a critical need for timely and precise diagnosis of this rare and complex disease. By enhancing diagnostic accuracy, AI-powered tools such as Legit.Health have the potential to reduce health care disparities, particularly in underserved areas with limited access to dermatologic specialists.

While individual neural networks for disease recognition have previously been shown to be useful for the automatic assessment of the severity of HS [17], urticaria [18], or atopic dermatitis [19] cases, with results comparable with those of human experts, this clinical evaluation is the first to reveal substantial improvements in diagnostic accuracy across various skin conditions, highlighting the potential of this medical device in reducing diagnostic errors; streamlining clinical decision-making; and, ultimately, improving patient outcomes.

These findings highlight the critical role that AI can play in bridging the expertise gap in dermatology, particularly for rare and complex diseases such as GPP. By aiding HCPs in differentiating GPP from other skin conditions, the Legit.Health medical device could facilitate timely and accurate diagnoses, which are crucial for prompt and effective treatment. Improved diagnostic accuracy can, in turn, reduce the risk of complications and improve quality of life for patients with GPP. This capability could be especially valuable in primary care environments, where immediate access to dermatologic specialists may not be possible.

Moreover, the performance of Legit.Health across a range of conditions highlights its versatility, further validating its utility and potential to improve diagnostic accuracy in the broad field of dermatology. Thanks to the extensive dataset used to train the algorithm, the device can provide the practitioner with a broader scope of pathologies, making them aware of less common diseases. We believe this could have occurred to some extent in this study: the participants may be biased toward common diseases (ie, the ones they are more likely to encounter in real life), and the output of the device changed their final assessments to more specific diagnoses (such as GPP or HS). Our results support the hypothesis that the device outputs served as a trigger to reconsider the diagnosis rather than an authoritative answer: the mere presence of an alternative diagnosis was enough to make the participants switch regardless of the confidence metric attached to it.

Limitations

However, this study has several limitations. First, this study only analyzed the impact of 1 medical device. However,

XSL•FO
RenderX

Legit.Health was deemed appropriate for this study due to its approved regulatory status and previous published evidence supporting its use in clinical assessments [17-19,26]. Furthermore, to the best of our knowledge, there is no other equivalent device with GPP diagnosis capabilities. For that reason, the results of this study may not generalize to other similar technologies developed for the assessment of different conditions. Further studies may focus on performing a similar assessment using different devices. Another limitation of this study is that, due to the rarity of GPP, assembling a diverse dataset is particularly challenging. Higher skin phototypes (ie, darker skin) [27] were underrepresented in the GPP images used to train the algorithm; this limitation is planned to be addressed in future iterations by supplementing the dataset with additional images with darker skin tones.

This study has demonstrated that the Legit.Health device had a positive impact on participants' performance. However, our analysis of users' dynamics before and after use of the device was limited by the amount of data collected during the experiment. In future work, additional variables such as seniority and the participants' perceived confidence in their initial responses should also be taken into account.

It is also worth noting that image-based diagnostic tools face limitations beyond image quality alone. The use of DIQA assured that there were no low-quality images, but there can be other issues that may impact the performance of the medical device. During the study, an interesting pattern of confusion between GPP and "nonspecific lesions" was observed, often in images with subtle or low-intensity visible signs. This highlights the crucial role of users in capturing high-quality images with a clear focus on the region of interest.

Furthermore, a noted limitation of this study is that it only considered teledermatology assessment based on images and information transmitted electronically. This context may not be fully replicable in in-person clinical environments where HCPs have the advantage of direct patient examination, additional questioning, and follow-up assessments. Nevertheless, this study incorporated detailed anamnesis and other relevant clinical information alongside the images to provide a comprehensive evaluation and equate the remote consultation with an in-person consultation.

Overall, this study shows that the use of an AI-powered medical device such as Legit.Health represents a potential advancement in dermatologic diagnostics. Legit.Health enhanced the accuracy of GPP diagnoses and could provide valuable support to HCPs, which in turn may improve patient care and outcomes.

Conclusions

Use of Legit.Health was associated with an improvement in the accuracy of HCPs in diagnosing GPP and other dermatologic conditions such as HS. This AI-powered medical device could be particularly beneficial in primary care settings, where specialized dermatologic expertise may not be available. This study revealed that the implementation of Legit.Health enhanced the overall diagnostic accuracy of HCPs and reduced the rate of misdiagnoses, potentially streamlining the decision-making process.

Although this study only considered the medical device Legit.Health, these results highlight the promise of the growing field of AI technologies as a beneficial tool for the future of dermatologic diagnostics and patient management.

Funding

This study was supported and funded by Boehringer Ingelheim. Eleanor James, BSc, of Ashfield MedComms (United Kingdom), an Inizio company, provided editorial and formatting support, which was funded by Boehringer Ingelheim.

Data Availability

The data generated or analyzed during this study are available in zenodo [28]. The analytical code is available at GitHub [29]. The image datasets used to train and validate the Legit.Health device are private and cannot be shared publicly.

Authors' Contributions

The authors meet the criteria for authorship as recommended by the International Committee of Medical Journal Editors.
Conceptualization: A Medela, TMC, A Martorell, IHM, AS
Data curation: IHM, AS, GSC
Formal analysis: IHM, A Martorell
Funding acquisition: AA, GSC
Investigation: AS, A Martorell
Methodology: A Medela, IHM, TMC, GSC
Project administration: A Medela, IHM, AA
Software: IHM, AS
Supervision: IHM, AS, A Medela, AA, A Martorell
Validation: IHM, AS, TMC
Visualization: IHM, TMC
Writing—original draft: A Medela
Writing—review and editing: A Medela, IHM, AS, TMC, GSC, JS, A Martorell

XSL•FO
RenderX

Conflicts of Interest

AA, A Medela, A Martorell, AS, IHM, and TMC are employees of Legit.Health and have collaborated with Boehringer Ingelheim, which funded this study. GSC is an employee of Boehringer Ingelheim. JS was an employee of Boehringer Ingelheim at the time of this study and is currently an employee of Almirall, S.A., Barcelona, Spain. A Martorell has received honoraria and/or travel grants from and/or has acted as an advisory board member for AbbVie, Almirall, Amgen, Boehringer Ingelheim, Bristol Myers Squibb, Cilag AG, LEO Pharma, Lilly, L'Oréal, Novartis, Sanofi, and UCB; has worked as a principal investigator in clinical trials supported by AbbVie, Bristol Myers Squibb, Galderma, Cilag AG, Lilly, Novartis, Sanofi, and UCB; and receives fees for his position as chief medical lead in Legit.Health.

References

1. Armstrong AW, Elston CA, Elewski BE, et al. Generalized pustular psoriasis: a consensus statement from the National Psoriasis Foundation. J Am Acad Dermatol 2024 Apr;90(4):727-730. [doi: 10.1016/j.jaad.2023.09.080] [Medline: 37838256]
2. Fujita H, Terui T, Hayama K, et al. Japanese guidelines for the management and treatment of generalized pustular psoriasis: the new pathogenesis and treatment of GPP. J Dermatol 2018 Nov;45(11):1235-1270. [doi: 10.1111/1346-8138.14523] [Medline: 30230572]
3. Navarini AA, Burden AD, Capon F, et al. European consensus statement on phenotypes of pustular psoriasis. J Eur Acad Dermatol Venereol 2017 Nov;31(11):1792-1799. [doi: 10.1111/jdv.14386] [Medline: 28585342]
4. Choon SE, van de Kerkhof P, Gudjonsson JE, et al. International consensus definition and diagnostic criteria for generalized pustular psoriasis from the International Psoriasis Council. JAMA Dermatol 2024 Jul 1;160(7):758-768. [doi: 10.1001/jamadermatol.2024.0915] [Medline: 38691347]
5. Choon SE, Navarini AA, Pinter A. Clinical course and characteristics of generalized pustular psoriasis. Am J Clin Dermatol 2022 Jan;23(Suppl 1):21-29. [doi: 10.1007/s40257-021-00654-z] [Medline: 35061227]
6. Gooderham MJ, Van Voorhees AS, Lebwohl MG. An update on generalized pustular psoriasis. Expert Rev Clin Immunol 2019 Sep;15(9):907-919. [doi: 10.1080/1744666X.2019.1648209] [Medline: 31486687]
7. Bachelez H. Pustular psoriasis and related pustular skin diseases. Br J Dermatol 2018 Mar;178(3):614-618. [doi: 10.1111/bjd.16232] [Medline: 29333670]
8. Baker H, Ryan TJ. Generalized pustular psoriasis. A clinical and epidemiological study of 104 cases. Br J Dermatol 1968 Dec;80(12):771-793. [doi: 10.1111/j.1365-2133.1968.tb11947.x] [Medline: 4236712]
9. Choon SE, Lai NM, Mohammad NA, Nanu NM, Tey KE, Chew SF. Clinical profile, morbidity, and outcome of adult-onset generalized pustular psoriasis: analysis of 102 cases seen in a tertiary hospital in Johor, Malaysia. Int J Dermatol 2014 Jun;53(6):676-684. [doi: 10.1111/ijd.12070] [Medline: 23967807]
10. Fujita H, Gooderham M, Romiti R. Diagnosis of generalized pustular psoriasis. Am J Clin Dermatol 2022 Jan;23(Suppl 1):31-38. [doi: 10.1007/s40257-021-00652-1] [Medline: 35061226]
11. Krueger J, Puig L, Thaçi D. Treatment options and goals for patients with generalized pustular psoriasis. Am J Clin Dermatol 2022 Jan;23(Suppl 1):51-64. [doi: 10.1007/s40257-021-00658-9] [Medline: 35061230]
12. Ohata C, Tsuruta N, Yonekura K, et al. Clinical characteristics of Japanese pustular psoriasis: a multicenter observational study. J Dermatol 2022 Jan;49(1):142-150. [doi: 10.1111/1346-8138.16217] [Medline: 34723399]
13. Kharawala S, Golembesky AK, Bohn RL, Esser D. The clinical, humanistic, and economic burden of generalized pustular psoriasis: a structured review. Expert Rev Clin Immunol 2020 Mar;16(3):239-252. [doi: 10.1080/1744666X.2019.1708193] [Medline: 32073341]
14. Puig L, Choon SE, Gottlieb AB, et al. Generalized pustular psoriasis: a global Delphi consensus on clinical course, diagnosis, treatment goals and disease management. J Eur Acad Dermatol Venereol 2023 Apr;37(4):737-752. [doi: 10.1111/jdv.18851] [Medline: 36606566]
15. Reisner DV, Johnsson FD, Kotowsky N, Brunette S, Valdecantos W, Eyerich K. Impact of generalized pustular psoriasis from the perspective of people living with the condition: results of an online survey. Am J Clin Dermatol 2022 Jan;23(Suppl 1):65-71. [doi: 10.1007/s40257-021-00663-y] [Medline: 35061229]
16. Strober B, Leman J, Mockenhaupt M, et al. Unmet educational needs and clinical practice gaps in the management of generalized pustular psoriasis: global perspectives from the front line. Dermatol Ther (Heidelb) 2022 Feb;12(2):381-393. [doi: 10.1007/s13555-021-00661-2] [Medline: 34904208]
17. Hernández Montilla I, Medela A, Mac Carthy T, et al. Automatic International Hidradenitis Suppurativa Severity Score System (AIHS4): a novel tool to assess the severity of hidradenitis suppurativa using artificial intelligence. Skin Res Technol 2023 Jun;29(6):e13357. [doi: 10.1111/srt.13357] [Medline: 37357665]
18. Mac Carthy T, Hernández Montilla I, Aguilar A, et al. Automatic urticaria activity score: deep learning-based automatic hive counting for urticaria severity assessment. JID Innov 2023 Jul 12;4(1):100218. [doi: 10.1016/j.xjidi.2023.100218] [Medline: 38075673]
19. Medela A, Mac Carthy T, Aguilar Robles SA, Chiesa-Estomba CM, Grimalt R. Automatic SCOring of Atopic Dermatitis using deep learning: a pilot study. JID Innov 2022 Feb 11;2(3):100107. [doi: 10.1016/j.xjidi.2022.100107] [Medline: 35990535]

XSL•FO
RenderX

20. Daneshjou R, Barata C, Betz-Stablein B, et al. Checklist for evaluation of image-based artificial intelligence reports in dermatology: CLEAR derm consensus guidelines from the International Skin Imaging Collaboration Artificial Intelligence Working Group. JAMA Dermatol 2022;158(1):90-96. [doi: 10.1001/jamadermatol.2021.4915] [Medline: 34851366]
21. ICD-11 coding tool. International Classification of Diseases. URL: https://icd.who.int/ct/icd11_mms/en/release [accessed 2024-11-15]
22. Smith LN, Topin N. Super-convergence: very fast training of neural networks using large learning rates. arXiv. Preprint posted online on Aug 23, 2017. [doi: 10.48550/arXiv.1708.07120]
23. Hernández Montilla I, Mac Carthy T, Aguilar A, Medela A. Dermatology Image Quality Assessment (DIQA): artificial intelligence to ensure the clinical utility of images for remote consultations and clinical trials. J Am Acad Dermatol 2023 Apr;88(4):927-928. [doi: 10.1016/j.jaad.2022.11.002] [Medline: 36526082]
24. Bachelez H, Choon SE, Marrakchi S, et al. Trial of spesolimab for generalized pustular psoriasis. N Engl J Med 2021 Dec 23;385(26):2431-2440. [doi: 10.1056/NEJMoa2111563] [Medline: 34936739]
25. Morita A, Strober B, Burden AD, et al. Efficacy and safety of subcutaneous spesolimab for the prevention of generalised pustular psoriasis flares (Effisayil 2): an international, multicentre, randomised, placebo-controlled trial. The Lancet 2023 Oct;402(10412):1541-1551. [doi: 10.1016/S0140-6736(23)01378-8]
26. Medela A, Sabater A, Montilla IH, MacCarthy T, Aguilar A, Chiesa-Estomba CM. The utility and reliability of a deep learning algorithm as a diagnosis support tool in head & neck non-melanoma skin malignancies. Eur Arch Otorhinolaryngol 2025 Mar;282(3):1585-1592. [doi: 10.1007/s00405-024-08951-z] [Medline: 39242415]
27. Gupta V, Sharma VK. Skin typing: Fitzpatrick grading and others. Clin Dermatol 2019;37(5):430-436. [doi: 10.1016/j.clindermatol.2019.07.010] [Medline: 31896400]
28. Medela A, Hernández Montilla I, Sabater A, Aguilar SA, MacCarthy TD, Martorell A. Participant responses from the study “evaluating diagnostic accuracy of generalized pustular psoriasis with the legit. health device”. Zenodo. 2025. URL: https://doi.org/10.5281/zenodo.15704954 [accessed 2025-05-25]
29. Legit-health/gpp-diagnosis-support. GitHub. URL: https://github.com/Legit-Health/gpp-diagnosis-support [accessed 2025-05-25]

Abbreviations

AI: artificial intelligence
CLEAR Derm: Checklist for Evaluation of Image-Based Artificial Intelligence Reports in Dermatology
DIQA: Dermatology Image Quality Assessment
GPP: generalized pustular psoriasis
HCP: health care practitioner
HS: hidradenitis suppurativa
***ICD-11*:** *International Classification of Diseases, 11th Revision*
PCP: primary care practitioner

Edited by R Dellavalle; submitted 07.Aug.2025; peer-reviewed by E Karampinis, G Lim; revised version received 02.Mar.2026; accepted 03.Mar.2026; published 10.Jun.2026.

Please cite as:
Medela A, Hernández Montilla I, Sabater A, Aguilar A, Mac Carthy T, Singh Chowdhry G, Semeco J, Martorell A
Enhanced Diagnosis of Generalized Pustular Psoriasis With the Legit.Health Device as a Diagnosis Support Tool: Multireader Multicase Study
JMIR Dermatol 2026;9:e82030
URL: https://derma.jmir.org/2026/1/e82030
doi:10.2196/82030

© Alfonso Medela, Ignacio Hernández Montilla, Alberto Sabater, Andy Aguilar, Taig Mac Carthy, Gurpreet Singh Chowdhry, Juan Semeco, Antonio Martorell. Originally published in JMIR Dermatology (http://derma.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Skin Cancer With Clinical Depression and Poor Mental Health Days: Cross-Sectional Analysis

Riona Ray[1]; Mytien Nguyen[2], MS, PhD

[1]Marvin Ridge High School, 7117 Stonehaven Drive, Waxhaw, NC, United States

[2]Yale School of Medicine, New Haven, CT, United States

Corresponding Author:
Riona Ray
Marvin Ridge High School, 7117 Stonehaven Drive, Waxhaw, NC, United States

Abstract

Background: Mental health is becoming increasingly recognized as an important part of overall health, especially for patients with cancer. However, the relationship between nonmelanoma skin cancer and mental health has not been widely studied.

Objective: The aim of this study was to examine the association between nonmelanoma skin cancer diagnosis and 2 key mental health outcomes (ie, clinical depression and the number of poor mental health days).

Methods: This study used the 2023 Behavioral Risk Factor Surveillance System, a nationally representative survey of adults in the United States, which included 312,317 participants. Nonmelanoma skin cancer diagnosis, depression, and self-reported mental health days were analyzed. Logistic regression was used to evaluate the association between nonmelanoma skin cancer and depression, whereas Poisson regression was used to model the number of poor mental health days, adjusting for age, sex, race and ethnicity, education, BMI, income, and major comorbid conditions (other cancers, heart disease, lung disease, and kidney disease).

Results: Individuals with nonmelanoma skin cancer (5086/26,552, 19.15%) reported a lower overall rate of depression compared to those without nonmelanoma skin cancer (61,438/285,765, 21.50%; *P*<.001) but reported more poor mental health days on average (4.54, SD 8.37 d vs 3.20, SD 7.37 d; *P*<.001). After adjustment, nonmelanoma skin cancer diagnosis was not significantly associated with depression (adjusted odds ratio 1.01, 95% CI 0.98 - 1.05) and was associated with a slightly lower number of poor mental health days (adjusted rate ratio 0.94, 95% CI 0.91 - 0.97).

Conclusions: Adults with nonmelanoma skin cancer experienced a meaningful mental health burden, and unadjusted analyses suggested greater day-to-day distress than among adults without nonmelanoma skin cancer. However, these differences were reduced and no longer significant for depression after adjusting for sociodemographic factors and comorbid chronic illnesses. These findings support the need for mental health screenings and support services in dermatologic and oncologic care.

(JMIR Dermatol 2026;9:e80710) doi:10.2196/80710

KEYWORDS

mental health; nonmelanoma skin cancer; depression; sociodemographic variables; analysis

Introduction

Background

In recent years, public health conversations have continued to emphasize the importance of mental health. Mental health is increasingly viewed not as a stand-alone issue, but as a factor that deeply interacts with physical illness, such as cancer [1].

While nonmelanoma skin cancer has obvious physical consequences, it poses serious complications regarding mental health. This aspect has not received sufficient attention in the health care field [2]. Dermatologists and oncologists are facing a dramatic rise in cases of nonmelanoma skin cancer, with melanoma rates doubling over the past two decades [3]. Nonmelanoma skin cancer is one of the most diagnosed malignancies in the world today, with rates for both melanoma and nonmelanoma types on the rise [3]. The relationship between mental state and nonmelanoma skin cancer is a complex feat. While the stress of diagnosis and treatment can create or worsen mental health conditions [4,5], existing mental health conditions can also increase the chances of developing nonmelanoma skin cancer through behavioral, immunological, and systemic mechanisms [6]. Recent evidence shows that approximately 30% of patients with melanoma experience anxiety, and nearly 20% experience depression. The highest risk has been observed among women and young adults [7]. Other studies have confirmed similar trends, showing that psychological distress and the fear of recurrence remain substantial even in patients with early-stage melanoma.

XSL•FO
RenderX

Research conducted recently has started to uncover the complex relationships between mental health and nonmelanoma skin cancer. A 2016 cross-sectional study using Behavioral Risk Factor Surveillance System data found that individuals who had frequent poor mental health days had a significantly higher chance of being diagnosed with the disease of nonmelanoma skin cancer [8].

These data were confirmed even after using the multivariate logistic regression analyses. These analyses suggest a possible link between poor mental health and keratinocyte carcinoma. This could be possible through factors such as dysregulated immune responses [9]. Other studies have shown similar results, mostly highlighting a high dose of psychological distress among patients with cancer [5,8]. Additionally, approximately one-third of patients with melanoma skin cancer require professional mental health care but are not receiving that treatment [1,2].

Further literature reviews on neuroendocrine-immune interactions support the biological plausibility of this connection. Chronic mental distress is a well-known contributor to the disruption of skin immunity, wound healing, and active inflammatory mediators, which can all contribute to the progression of cancer [6]. Additionally, factors including hostility and depression have been connected to melanoma and its treatment outcomes [4]. There is an extremely minimal amount of information regarding the demographic or socioeconomic factors that shape the outcomes of mental health across nonmelanoma skin cancer subtypes [3,10].

This study aimed to address these gaps in knowledge by analyzing the association between mental health disorders and the rate of nonmelanoma skin cancer diagnosis by using the information provided by the Behavioral Risk Factor Surveillance System (BRFSS). Focusing on nonmelanoma skin cancers, assessing the link between nonmelanoma skin cancer and mental health status by sociodemographic factors, such as age, sex, race, income, BMI, and education, will provide critical insights into how mental health influences the risk and experience of nonmelanoma skin cancer.

Literature Review

Recent studies document consistent associations between multiple indicators of psychological distress and nonmelanoma skin cancer. A proportional meta-analysis of patients with melanoma reported prevalence estimates of 30% for anxiety and 20% for depression, with higher odds observed among women, younger adults, and individuals with lower education levels [7]. Similar findings have been reported in earlier clinical and observational studies, showing elevated levels of psychological symptoms across different stages of melanoma, including treatment and posttreatment phases [5,11].

Beyond symptom prevalence, multiple studies have examined behavioral and biological pathways linking mental health to nonmelanoma skin cancer. Young adults with mental health problems demonstrate higher rates of cancer-related risk behaviors, such as smoking, alcohol use, sleep disturbances, and inactivity, which may contribute to disease development or worse outcomes [10]. Experimental research has also shown that chronic psychological stress alters neuroendocrine and immune signaling, increasing inflammatory activity and harming skin repair processes [6]. Additional studies have reported relationships between melanoma severity and personality traits, such as hostility and depressive tendencies.

More recent research has shifted attention to survivorship and early-stage disease. Patients diagnosed with melanoma have reported reduced emotional well-being. They have also stated persistent uncertainty despite a favorable clinical prognosis. This suggests that psychological effects extend further than cancer itself. Moreover, fear of recurrence has been identified as a primary contributor to ongoing mental distress following the completion of treatment [12]. These findings indicate that mental health challenges in nonmelanoma skin cancer populations can include forms of distress that may not be clinically diagnosed.

Population-based research has identified variation in mental health outcomes among patients with nonmelanoma skin cancer across demographic and socioeconomic subgroups. Studies have shown that mental health service use remains limited, with unmet psychological needs concentrated among older adults and lower-income populations [1,2]. Global assessments have revealed a lower quality of life in regions with lower access to supportive care resources [3].

Globally, the burden of skin disease is high in many regions, especially Asia, and is linked to socioeconomic status and inflammatory conditions [3]. Tools such as the Skin Cancer Index have been developed to measure the quality of life in patients with nonmelanoma skin cancer [13].

Methods

Participants

This study used data from the BRFSS, a nationally representative survey conducted by the Centers for Disease Control and Prevention [9]. The data used were from the year 2023. The BRFSS surveys US adults aged 18 years or older, collecting data on health conditions, behaviors, and preventive health practices [9]. This dataset included responses to questions related to nonmelanoma skin cancer, mental health, and sociodemographic characteristics. Participants with missing, refused, or “don’t know” responses were excluded from the analyses to ensure the high quality and reliability of the study.

Exposure

The independent variable was a self-reported diagnosis of nonmelanoma skin cancer. These individuals did not have to have a current diagnosis; the diagnosis could be from any time in the past. Respondents were asked whether a health professional had ever told them they had skin cancer, including melanoma and nonmelanoma types. Individuals who answered “yes” were categorized as having a nonmelanoma skin cancer diagnosis. Those who answered “no” were the comparison group. Those with missing or ambiguous responses were excluded from the analysis to maintain the integrity of the data.

Outcomes

The 2 primary mental health–related outcomes that were examined were depression and the number of poor mental health

XSL•FO
RenderX

days an individual had. Depression was defined as being diagnosed with a depressive disorder by a health care professional [9,10]. Poor mental health days were based on the number of days during the past 30 days that an individual reported that their mental health was "not good," including stress, depression, and other emotional issues [1,3]. Respondents with invalid responses were excluded from the analysis.

Covariates

The sociodemographic variables that were included in the analysis were age (18 - 64 and ≥65 years), sex (male or female), race or ethnicity (White only, Black only, Asian only, American Indian or Alaskan Native only, Native Hawaiian or other Pacific Islander only, multiracial, and other), education (did not graduate high school, graduated high school, attended college or technical school, and graduated from college or technical school), and BMI (underweight, normal weight, overweight, and obese). Additional health-related covariates included self-reported diagnoses of other (non–skin) cancer, heart disease, chronic lung disease, and kidney disease. These covariates were specifically selected based on the evidence linking them to mental health and cancer-related outcomes [3,10].

Statistical Analysis

Descriptive statistics were first used to summarize the distribution of depression status and the number of poor mental health days by nonmelanoma skin cancer diagnosis and sociodemographic variables, including age, sex, race and ethnicity, education level, and BMI. All statistical models were run on the entire BRFSS sample, and individuals without a history of nonmelanoma skin cancer served as the reference group. This allowed a direct comparison between those with and those without nonmelanoma skin cancer. Categorical variables were summarized using frequencies and percentages, while continuous variables were described using means and SDs. Group differences in categorical variables were assessed using Pearson χ^2 tests, and differences in continuous outcomes were interpreted using independent samples *t* tests. These tests described unadjusted differences between adults with and without a history of nonmelanoma skin cancer.

To examine the association between nonmelanoma skin cancer diagnosis and depression, a multivariable logistic regression model was used [9]. Depression was treated as a yes or no outcome, and a nonmelanoma skin cancer diagnosis (yes or no) was the main comparison of interest. All statistical models were run on the full BRFSS sample, and individuals without a history of nonmelanoma skin cancer served as the reference group in all analyses. This approach allowed direct comparison of depression prevalence and poor mental health days between respondents with and without nonmelanoma skin cancer, instead of limited analyses to only the skin cancer subgroup. Additionally, adjusted odds ratios (aORs) and corresponding 95% CIs were reported.

For the continuous outcome of mental health days, a multivariable Poisson regression model with standard errors to account for potential overdispersion was used. The results were expressed as adjusted rate ratios (aRRs) with 95% CIs, which allowed for the calculation of the relative increase or decrease in the expected number of poor mental health days among individuals with nonmelanoma skin cancer compared to those without, after accounting for sociodemographic factors. Both regression models adjusted for age, sex, race and ethnicity, education level, BMI, household income, and comorbid conditions (other cancer, heart disease, lung disease, and kidney disease). These sociodemographic and health-related variables are independently associated with both mental health outcomes and cancer risk in prior studies. Logistic regression was used for the binary depression outcome, whereas a multivariable Poisson regression model was used for the count-based outcome of poor mental health days. Poisson regression was selected because the outcome represents a count of days within a fixed 30-day period and was not normally distributed, making linear regression inappropriate. Standard errors were adjusted to account for overdispersion. The distribution of days with poor mental health was examined. It was discovered that, although the data showed variability, it did not exhibit sufficient overdispersion to warrant switching to an alternative model. Therefore, the Poisson model was the best option.

All statistical tests were 2 sided. Analyses were conducted using JASP, ensuring appropriate complex survey weighting to reflect the nationally representative design of the BRFSS dataset [9].

Ethical Considerations

This study involved secondary analysis of publicly available, deidentified data from the BRFSS, administered by the US Centers for Disease Control and Prevention. As the dataset contains no identifiable private information, this study did not constitute human subjects research and was therefore exempt from institutional review board review in accordance with US federal regulations. The BRFSS protocol is reviewed and approved annually by the US Centers for Disease Control and Prevention Institutional Review Board, and informed consent is obtained from all participants at the time of data collection.

Results

Overview

Among 433,323 participants in the 2023 BRFSS questionnaire, 312,317 (72.07%) had complete demographic and disease information and were included in the analysis (Table 1). Among the analytical cohort, 154,230 (49.38%) were men, and 158,087 (50.62%) were women. Additionally, 253,634 (81.21%) identified as White participants only. The remaining racial and ethnic distribution included 26,936 (8.62%) Asian only, 6551 (2.10%) Black only, 8865 (2.84%) American Indian or Alaska Native only, 2041 (0.65%) Native Hawaiian or Pacific Islander, 5742 (1.84%) multiracial, and 8548 (2.74%) identifying as other race and ethnicity.

XSL•FO
RenderX

Table . Characteristics of the study cohort [8].

Characteristics	Values
Skin cancer diagnosis, n (%)	
No	285,765 (91.50)
Yes	26,552 (8.50)
Depression, n (%)	
No	245,793 (78.70)
Yes	66,524 (21.30)
Mental health days, mean (SD)	4.42 (8.29)
Race and ethnicity, n (%)	
White only	253,634 (81.21)
Asian only	26,936 (8.62)
Black only	6551 (2.10)
American Indian or Alaskan Native only	8865 (2.84)
Native Hawaiian or other Pacific Islander only	2041 (0.65)
Multiracial	5742 (1.84)
Other race only	8548 (2.74)
Sex, n (%)	
Male	154,230 (49.38)
Female	158,087 (50.62)
Age (y), n (%)	
18-64	198,394 (63.52)
≥65	113,923 (36.48)
BMI, n (%)	
Underweight	4802 (1.54)
Normal weight	89,431 (28.63)
Overweight	111,680 (35.76)
Obese	106,404 (34.07)
Education, n (%)	
Did not graduate high school	14,184 (4.54)
Graduated high school	73,285 (23.46)
Attended college or technical school	83,761 (26.82)
Graduated from college or technical	141,087 (45.17)
Other cancer, n (%)	
No	275,645 (88.26)
Yes	36,672 (11.74)
Heart disease, n (%)	
No	277,856 (88.97)
Yes	34,461 (11.03)
Lung disease, n (%)	
No	251,263 (80.45)
Yes	61,054 (19.55)
Kidney disease, n (%)	
No	297,584 (95.28)

XSL•FO
RenderX

Characteristics	Values
Yes	14,733 (4.72)

Most of the 312,317 respondents were aged between 18 and 64 years (n=198,394, 63.52%), with 113,923 (36.48%) aged 65 years or older. BMI classifications showed that 4802 (1.54%) were underweight, 89,431 (28.63%) had a normal BMI, 111,680 (35.76%) were overweight, and 106,404 (34.07%) were obese. Educational attainment also varied, with 14,184 (4.54%) not graduating from high school, 73,285 (23.46%) graduating from high school, 83,761 (26.82%) attending some college or technical school, and 141,087 (45.17%) graduating from a college or technical program.

Most respondents did not report a nonmelanoma skin cancer diagnosis, with 285,765 (91.50%) of 312,317 indicating no history of nonmelanoma skin cancer and 26,552 (8.50%) reporting a diagnosis. Additionally, 245,793 (78.70%) participants did not report depression, whereas 66,524 (21.30%) reported having been diagnosed with depression by a health care professional. The high average number of mental health days was consistent with high fluctuations in mental health experiences across many individuals.

Comorbid health conditions were also reported. A total of 36,672 (11.74%) participants reported another form of cancer, 34,461 (11.03%) reported heart disease, 61,054 (19.55%) reported lung disease, and 14,733 (4.72%) reported kidney disease.

The average number of poor mental health days in the past 30 days was 4.42 (SD 8.29). This was consistent with substantial variation in mental health experiences across the population.

Depression

Of the entire sample, 88,524 (21.31%) of 312,317 participants reported experiencing depression. Of those without a nonmelanoma skin cancer diagnosis, 61,428 (21.50%) of 285,765 reported depression. However, of those with a nonmelanoma skin cancer diagnosis, 5086 (19.15%) of 26,552 individuals reported depression. After the analysis was adjusted for the included covariates, nonmelanoma skin cancer diagnosis was not significantly associated with depression (aOR 1.01, 95% CI 0.98 - 1.05; P<.001; Table 2).

XSL•FO
RenderX

Table . Association between nonmelanoma skin cancer and depression [8].

Characteristics	Depression			
	No, n (%)	Yes, n (%)	*P* value	aOR[a] (95% CI)
Nonmelanoma skin cancer diagnosis			<.001	
No	224,327 (78.5)	61,438 (21.4)	<.001	Ref[b]
Yes	21,466 (80.8)	5086 (19.1)	<.001	1.01 (0.98 - 1.05)
Race and ethnicity			<.001	
White only	1,97,650 (77.9)	55,984 (22)	<.001	Ref
Asian only	22,476 (83.4)	4460 (16.5)	<.001	0.50 (0.48 - 0.52)
Black only	5135 (78.3)	1416 (21.6)	<.001	0.70 (0.66 - 0.75)
American Indian or Alaskan Native only	7863 (88.6%)	1002 (11.3%)	<.001	0.47 (0.440.50)
Native Hawaiian or other Pacific Islander only	1721 (84.3)	320 (15.6)	<.001	0.51 (0.45 - 0.58)
Multiracial	4778 (83.2)	964 (16.7)	<.001	0.60 (0.56 - 0.64)
Other race only	6170 (72.1)	2378 (27.8)	<.001	1.08 (1.04 - 1.15)
Sex			<.001	
Male	130,827 (84.8)	23,403 (15.1)	<.001	Ref
Female	114,966 (72.7)	43,121 (27.2)	<.001	1.98 (1.94 - 2.02)
Age (years)			<.001	
18-64	150,115 (75.6)	48,279 (24.3)	<.001	Ref
≥65	95,678 (83.9)	18,245 (16)	<.001	0.44 (0.440.45)
BMI			<.001	
Underweight	3581 (74.5)	1221 (25.4)	<.001	Ref
Normal weight	72,440 (81)	16,991 (18.9)	<.001	0.83 (0.77 - 0.89)
Overweight	91,376 (81.8)	20,304 (18.1)	<.001	0.87 (0.81 - 0.94)
Obese	78,396 (73.6)	28,008 (26.3)	<.001	1.18 (1.10 - 1.27)
Education			<.001	
Did not graduate high school	10,816 (76.2)	3368 (23.7)	<.001	Ref
Graduated high school	58,033 (79.1)	15,252 (20.8)	.27	0.98 (0.94 - 1.02)
Attended college or technical school	63,782 (76.1)	19,979 (23.8)	<.001	1.02 (0.98 - 1.07)
Graduated from college or technical	113,162 (80.2)	27,925 (19.7)	<.001	0.91 (0.87 - 0.95)
Other cancer			<.001	1.12 (1.08 - 1.15)
No	217,198 (78.7)	58,447 (21.2)	<.001	Ref
Yes	28,595 (77.9)	8077 (22)	<.001	1.12 (1.08 - 1.15)
Heart disease			<.001	
No	220,365 (79.3)	57,491 (20.6)	<.001	Ref
Yes	25,428 (73.7)	9033 (26.2)	<.001	1.41 (1.37 - 1.45)
Lung disease			<.001	
No	206,137 (82)	45,126 (17.9)	<.001	Ref
Yes	39,656 (64.9)	21,398 (35)	<.001	2.10 (2.05 - 2.14)
Kidney disease			<.001	

XSL•FO
RenderX

Characteristics	Depression			
	No, n (%)	Yes, n (%)	*P* value	aOR[a] (95% CI)
No	235,503 (79.1)	62,081 (20.8)	<.001	Ref
Yes	10,290 (69.8)	4443 (30.1)	<.001	1.47 (1.41 - 1.53)

[a]aOR: adjusted odds ratio.

[b]Ref: reference.

When analyzing all racial and ethnic groups, there were many considerable differences in the prevalence of depression. White respondents were used as the reference group. Using the reference group, Asian (aOR 0.50, 95% CI 0.48 - 0.52), Black (aOR 0.70, 95% CI 0.66 - 0.75), American Indian or Alaska Native (aOR 0.47, 95% CI 0.44 - 0.50), Native Hawaiian or other Pacific Islander (aOR 0.51, 95% CI 0.45 - 0.58), and multiracial respondents (aOR 0.60, 95% CI 0.56 - 0.64) all had lower adjusted odds of depression. Participants in the "other" category had slightly higher odds of depression compared to White respondents (aOR 1.08, 95% CI 1.04 - 1.15).

Women (43,121/158,087, 27.28%) reported significantly higher rates of depression compared to men (23,403/154,230, 15.19%). After adjustment, women had almost double the odds of depression when compared to men (aOR 1.98, 95% CI 1.94 - 2.02). Participants (18,245/113,923, 16.0%) aged 65 years or older had significantly lower rates of depression compared to adults (48,279/198,394, 24.3%) aged 18 to 64 years. BMI also played a substantial role. With underweight individuals as the reference group, obese individuals experienced higher odds of depression (aOR 1.18, 95% CI 1.10 - 1.27), while those who were underweight or of normal weight had lower odds compared to those who were considered overweight or obese.

After adjusting for covariates, high school graduates had similar odds of depression to the reference group (aOR 0.98, 95% CI 0.94 - 1.02). Participants who had reached the college level of education had slightly different odds (aOR 1.02, 95% CI 0.98 - 1.07), and college graduates had lower odds (aOR 0.91, 95% CI 0.87 - 0.95).

Poor Mental Health Days

Respondents with a history of nonmelanoma skin cancer reported a higher average number of poor mental health days (mean 4.54, SD 8.37) compared to those without a nonmelanoma skin cancer diagnosis (mean 3.20, SD 7.37). However, after adjustment, individuals with nonmelanoma skin cancer experienced a slight decrease in poor mental health days compared to those without (aRR 0.94, 95% CI 0.91 - 0.97; Table 3).

XSL•FO
RenderX

Table . Association between nonmelanoma skin cancer and poor mental health days [8].

Characteristics	Mental health days		
	Mean (SD)	*P* value	aRR[a] (95% CI)
Skin cancer diagnosis		<.001	
No	3.19 (7.36)	<.001	Ref[b]
Yes	4.54 (8.36)	<.001	0.94 (0.91 - 0.97)
Race and ethnicity		<.001	
White only	4.30 (8.17)	<.001	Ref
Asian only	4.80 (8.65)	<.001	0.93 (0.90 - 0.95)
Black only	5.86 (9.52)	<.001	1.07 (1.01 - 1.13)
American Indian or Alaskan Native only	3.36 (6.85)	<.001	0.82 (0.78 - 0.86)
Native Hawaiian or other Pacific Islander only	5.42 (9.34)	<.001	1.11 (1.00 - 1.24)
Multiracial	4.65 (8.70)	<.001	0.93 (0.88 - 1.00)
Other race only	6.45 (9.74)	<.001	1.25 (1.19 - 1.32)
Sex		<.001	
Male	3.64 (7.71)	<.001	Ref
Female	5.18 (8.74)	<.001	1.36 (1.34 - 1.39)
Age (y)		<.001	
18-64	5.43 (8.85)	<.001	Ref
≥65	2.66 (6.85)	<.001	0.40 (0.40 - 0.41)
BMI		<.001	
Underweight	6.37 (9.81)	<.001	Ref
Normal weight	4.18 (7.95)	<.001	0.76 (0.71 - 0.82)
Overweight	3.78 (7.70)	<.001	0.71 (0.67 - 0.76)
Obese	5.20 (8.98)	<.001	0.83 (0.78 - 0.89)
Education		<.001	
Did not graduate high school	6.13 (10.21)	<.001	Ref
Graduated high school	5.01 (9.06)	<.001	1.01 (1.01 - 1.01)
Attended college or technical school	5.01 (8.84)	<.001	1.03 (1.03 - 1.03)
Graduated from college or technical	3.59 (7.15)	<.001	0.88 (0.88 - 0.88)
Other cancer diagnosis		<.001	
No	4.47 (8.28)	<.001	Ref
Yes	4.09 (8.35)	<.001	1.12 (1.08 - 1.14)
Heart disease		<.001	
No	4.31 (8.11)	<.001	Ref
Yes	5.30 (9.55)	<.001	1.33 (1.30 - 1.37)
Lung disease		<.001	
No	3.87 (7.73)	<.001	Ref
Yes	6.67 (9.95)	<.001	1.50 (1.47 - 1.53)
Kidney disease		<.001	
No	4.36 (8.21)	<.001	Ref

XSL•FO
RenderX

Characteristics	Mental health days		
	Mean (SD)	*P* value	aRR[a] (95% CI)
Yes	5.60 (9.63)	<.001	1.28 (1.23 - 1.33)

[a]aOR: adjusted odds ratio.

[b]Ref: reference.

Significant differences in mental health days were observed by race and ethnicity. Black individuals reported the highest average (5.85 d). This group had significantly increased rates of mental health issues compared to White individuals (aRR 1.07, 95% CI 1.01 - 1.13). In contrast, American Indian or Alaska Native participants (aRR 0.82, 95% CI 0.78 - 0.86), multiracial individuals (aRR 0.93, 95% CI 0.88 - 1.00), and Asian respondents (aRR 0.93, 95% CI 0.90 - 0.95) reported fewer mental health days compared to the White reference group.

Women had significantly more poor mental health days (mean 5.18, SD 8.75) compared to men (mean 3.65, SD 7.72). After adjustment, women had substantially higher rates of mental health distress (aRR 1.36, 95% CI 1.34 - 1.39). Respondents aged 65 years and older reported fewer mental health days than those in lower age groups (aRR 0.40, 95% CI 0.40 - 0.41).

BMI was strongly associated with mental health outcomes. Underweight individuals experienced the highest number of poor mental health days (mean 6.40, SD 9.85) and served as the reference group. Compared to them, respondents of normal weight (aRR 0.76, 95% CI 0.71 - 0.82), overweight individuals (aRR 0.71, 95% CI 0.67 - 0.76), and those with obesity (aRR 0.83, 95% CI 0.78 - 0.89) all had significantly lower rates of poor mental health days.

Individuals who did not graduate high school reported the highest average number of poor mental health days (6.13 d), while college graduates reported the fewest number (3.60 d). After adjustment, graduating from college or technical school was associated with significantly fewer mental health days (aRR 0.88, 95% CI 0.88 - 0.88) compared to individuals with less education.

Several comorbid health conditions were also associated with increased mental distress. Individuals with another cancer diagnosis had more poor mental health days (mean 5.01, SD 7.56) and higher adjusted rates compared to those without other cancers (aRR 1.12, 95% CI 1.08 - 1.14). Lung disease was associated with the strongest increase in mental health burden (mean 6.77, SD 7.81; aRR 1.50, 95% CI 1.47 - 1.53). Respondents with kidney disease (aRR 1.28, 95% CI 1.23 - 1.33) and heart disease (aRR 1.33, 95% CI 1.30 - 1.37) also reported significantly higher adjusted rates of poor mental health days compared to their respective reference groups.

Discussion

Principal Findings

In this nationally representative sample, study findings reveal a subtle relationship between nonmelanoma skin cancer and mental health: while individuals with a history of nonmelanoma skin cancer were slightly less likely to report a formal diagnosis of depression in unadjusted comparisons, nonmelanoma skin cancer was not significantly associated with depression after adjusting for demographics, other cancers, and chronic diseases. However, individuals with nonmelanoma skin cancer reported a higher number of poor mental health days before adjustment but slightly fewer poor mental health days after adjustment. These findings suggest that the differences in mental health burden are largely explained by sociodemographic and comorbid factors instead of the nonmelanoma skin cancer itself.

Prior research has suggested that the association between nonmelanoma skin cancer and mental health may operate in both biological and psychological directions. Chronic psychological stress has been shown to alter neuroendocrine-immune pathways, increasing inflammatory activity, impairing wound repair, and weakening immune surveillance, which may elevate susceptibility to certain nonmelanoma skin cancers [6]. However, a nonmelanoma skin cancer diagnosis may contribute to psychological distress through concerns about recurrence, uncertainty during long-term surveillance, scarring, and changes in visible appearance. These have all been documented as drivers of anxiety and depressive symptoms in melanoma and nonmelanoma patient populations [12,13].

A cancer diagnosis itself is often associated with increased stress. Prior research has shown that uncertainty about outcomes and concerns about physical appearance can elevate psychological stress, particularly in patients with visible scars [5,11]. Although stress was not directly measured in this study, the higher number of poor mental health days reported by individuals with nonmelanoma skin cancer may reflect this psychological impact. These findings support the notion that cancer-related stress can appear in daily tasks, even when it does not meet clinical criteria for depression [2,5].

Interestingly, in adjusted models, individuals with a history of nonmelanoma skin cancer reported fewer poor mental health days compared with those without nonmelanoma skin cancer, while no association was observed with depression. Several potential mechanisms may help explain this counterintuitive pattern. Nonmelanoma skin cancer is typically detected early, treated effectively, and associated with excellent long-term outcomes, which may mitigate sustained psychological distress. Successful removal of visible lesions can also create a sense of resolution or restored control, potentially improving daily emotional well-being. In addition, patients with nonmelanoma skin cancer may often engage in regular dermatologic care, providing frequent health care touchpoints that may reduce uncertainty, reinforce preventive health behaviors, and reflect a population with generally higher health literacy or

XSL•FO
RenderX

wellness-oriented behaviors, factors that are linked to more favorable mental health profiles.

The sociodemographic differences observed in this study are consistent with broader public health literature, showing that mental health outcomes are shaped by structural, cultural, and economic factors. Higher rates of poor mental health days among women and younger adults may reflect increased stress, body image concerns, or work-related pressures. Racial variation may be influenced by differences in health care access. Educational and income-related disparities may also reflect gaps in early detection resources. These findings underscore the importance of tailoring mental health support within dermatologic and oncologic care to the needs of various groups rather than applying a uniform approach.

The nature of being diagnosed with nonmelanoma skin cancer itself may contribute significantly to this distress. Patients often experience fear of imperfections due to visible scarring from surgery, concerns about cancer recurrence, or anxiety over potential mortality, especially with melanoma [5,11]. The continuation of dermatological watch and uncertainty with treatments can further elevate emotional strain for individuals. This specifically takes place when the cancer affects visible areas, such as the face or neck [2]. These stressors may not meet the clinical definition of depression but can still influence day-to-day mental well-being [5].

These results align with previous studies that highlight psychological distress among patients with nonmelanoma skin cancer. However, some research has found higher rates of depression, suggesting variability based on sample demographics or methods of measurement [5,11]. This study adds to the conversation by emphasizing subjective mental distress, which may not always manifest as a clinical diagnosis, while also showing that much of the observed association may be explained by comorbid illness and sociodemographic factors.

We also observed key sociodemographic differences. Women, younger adults, individuals with higher BMI, and those with lower levels of education reported a higher number of poor mental health days and higher levels of depression. These outcomes are consistent with a large amount of public health literature and suggest that mental health improvements should be tailored to the vulnerabilities of different subgroups [3,10].

Following these results, a consistent routine of mental health screenings for those diagnosed with nonmelanoma skin cancer is recommended to help relieve mental distress. This may include screening tools such as the Patient Health Questionnaire-9 during dermatology or oncology visits. This incorporates automatic referral pathways to licensed mental health providers with outstanding scores. Integrated care models may also involve co-located behavioral health specialists (eg, psychologists, social workers, or psychiatric nurse practitioners) within dermatology or oncology clinics. Incorporating this may help address psychological needs associated with the diagnosis and its follow-up care. Moreover, support groups, cognitive behavioral therapy, or survivorship counseling should be offered as part of a thorough treatment plan, helping patients manage stressors, such as body image changes, fear, and long-term mental challenges [1,2].

This study has several limitations. As the BRFSS dataset is cross-sectional, the direction of the relationship between nonmelanoma skin cancer and mental health outcomes cannot be established. It is not possible to determine whether poor mental health causes the development of nonmelanoma skin cancer or arises because of diagnosis, treatment, and other factors of nonmelanoma skin cancer. Poor mental health days rely on self-report and capture broad, nonspecific distress, which may not align with clinical diagnoses. Reverse causality is possible if individuals with mental health issues are more likely to seek evaluation for skin changes, leading to higher rates of nonmelanoma skin cancer detection. Additionally, several confounding variables, such as family history of cancer, medication use, and factors such as sun exposure or smoking, were not used in the dataset and may partially explain the observed associations. Although this analysis adjusted for several major chronic illnesses (other cancers, lung disease, heart disease, and kidney disease), many clinically important conditions remain unmeasured. For instance, a participant may have both nonmelanoma skin cancer and a more psychologically burdensome condition, such as lung cancer or severe cardiac disease, which could influence their mental health outcomes. The inability to differentiate whether mental health symptoms stem from nonmelanoma skin cancer itself or from co-occurring illnesses limits the precision of our findings. Additionally, our analyses do not capture illness perceptions, cosmetic concerns, or treatment experiences that may influence psychological outcomes. Future work using datasets with richer clinical detail or linked cancer registry data may help more accurately isolate the independent effect of skin cancer on mental health. These findings should be interpreted as a correlation, and future research is needed to clarify the direction of this relationship. Finally, the BRFSS survey may not capture more nuanced mental health challenges such as anxiety or posttraumatic stress disorder, limiting the depth of insight into the psychological experiences of patients with nonmelanoma skin cancer [5,9]. All variables were self-reported, which may introduce misclassification of both exposures and outcomes.

It is also crucial to recognize that depression is frequently underdiagnosed in community populations, particularly among older adults, men, and individuals with limited access to health care. The BRFSS depression variable relies on self-reported clinical diagnosis, which does not capture unreported cases. More sensitive mental health assessments, such as the Patient Health Questionnaire-9 or validated cancer-specific screening tools, may better capture psychological distress in future studies.

Conclusions

This study highlights a major association between mental health challenges, particularly

depression and poor mental health days, and the presence of nonmelanoma skin cancer among US adults using nationally representative data from the 2023 BRFSS [9]. Adults with a history of skin cancer reported higher unadjusted levels of day-to-day mental distress than those without skin cancer, but analyses adjusted for covariates showed no significant association with depression and a slight decrease in poor mental health days. Moreover, sociodemographic factors play a

XSL•FO
RenderX

substantial role in shaping mental health, with certain groups showing greater vulnerability [3,10].

These results emphasize the importance of integrated care models that address both physical and mental health outcomes in patients with nonmelanoma skin cancer [1,2]. Public health initiatives should prioritize mental health screening and support within dermatologic and oncologic care, especially for disproportionately affected populations. The favorable mental health profile observed among individuals with nonmelanoma skin cancer may also highlight opportunities to leverage routine dermatologic care as a platform for promoting mental well-being and early identification of psychosocial needs. Future research should investigate longitudinal patterns, causal mechanisms, and the effectiveness of mental health interventions in improving quality of life and potentially clinical outcomes among patients with nonmelanoma skin cancer, and whether resilience, health care engagement, or other unmeasured attributes mediate these associations, and whether similar patterns emerge across diverse populations and cancer types [5,11].

Ultimately, recognizing and addressing the mental health burden associated with nonmelanoma skin cancer can lead to more holistic, equitable, and patient-centered care strategies.

Acknowledgments

The authors would like to recognize Mytien Nguyen from the Yale School of Medicine for her help throughout the process of evaluating data and constructing this study. Her guidance and mentorship are greatly appreciated.

Data Availability

The datasets generated or analyzed during this study are available in the Behavioral Risk Factor Surveillance System repository [8].

Conflicts of Interest

None declared.

References

1. Hewitt M, Rowland JH. Mental health service use among adult cancer survivors: analyses of the National Health Interview Survey. J Clin Oncol 2002 Dec 1;20(23):4581-4590. [doi: 10.1200/JCO.2002.03.077] [Medline: 12454116]
2. Ernstmann N, Neumann M, Ommen O, et al. Determinants and implications of cancer patients' psychosocial needs. Support Care Cancer 2009 Nov;17(11):1417-1423. [doi: 10.1007/s00520-009-0605-7] [Medline: 19283411]
3. Urban K, Chu S, Giesey RL, et al. Burden of skin disease and associated socioeconomic status in Asia: a cross-sectional analysis from the Global Burden of Disease Study 1990-2017. JAAD Int 2021 Mar;2:40-50. [doi: 10.1016/j.jdin.2020.10.006] [Medline: 34409353]
4. Kontoangelos K, Nikolaou V, Syrgianni A, Tsiori S, Papageorgiou C, Stratigos A. Correlation of psychosomatic factors and personality traits with the severity of melanoma. In Vivo 2024;38(6):2844-2852. [doi: 10.21873/invivo.13765] [Medline: 39477397]
5. Kungwengwe G, Gowthorpe C, Ali SR, et al. Prevalence and odds of anxiety and depression in cutaneous malignant melanoma: a proportional meta-analysis and regression. Br J Dermatol 2024 Jun 20;191(1):24-35. [doi: 10.1093/bjd/ljae011] [Medline: 38197404]
6. Wang J, Chen X, Li X, Li W. Role of stress in skin diseases. Front Immunol 2024;14:1295986. [doi: 10.3389/fimmu.2023.1295986] [Medline: 38223260]
7. Zheng L, Susanto AK, Soyer HP, Janda M, de Camargo Catapan S. Psychosocial wellbeing in people with melanoma in-situ: a systematic review. Melanoma Res 2025 Feb 1;35(1):11-23. [doi: 10.1097/CMR.0000000000001007] [Medline: 39526661]
8. Behavioral Risk Factor Surveillance System (BRFSS) 2023 survey data and documentation. : Centers for Disease Control and Prevention; 2024 URL: https://www.cdc.gov/brfss/annual_data/annual_2023.html [accessed 2026-01-02]
9. Dinh TT, Maibach HI, Huen AO. Association of poor mental health and skin cancer development. J Dermatol Treat 2020;31:810-814. [doi: 10.1080/09546634.2019.1708111] [Medline: 31846266]
10. Massetti GM, Thomas CC, King J, Ragan K, Buchanan Lunsford N. Mental health problems and cancer risk factors among young adults. Am J Prev Med 2017 Sep;53(3S1):S30-S39. [doi: 10.1016/j.amepre.2017.04.023] [Medline: 28818243]
11. Toma AO, Boeriu E, Decean L, et al. The effects of lack of awareness in age-related quality of life, coping with stress, and depression among patients with malignant melanoma. Curr Oncol 2023 Jan 23;30(2):1516-1528. [doi: 10.3390/curroncol30020116] [Medline: 36826077]
12. Mahama AN, Haller CN, Labrada J, et al. Lived experiences and fear of cancer recurrence among survivors of localized cutaneous melanoma. JAMA Dermatol 2024 May 1;160(5):495-501. [doi: 10.1001/jamadermatol.2023.6158] [Medline: 38353983]

XSL•FO
RenderX

13. Rhee JS, Matthews BA, Neuburg M, Logan BR, Burzynski M, Nattinger AB. Validation of a quality-of-life instrument for patients with nonmelanoma skin cancer. Arch Facial Plast Surg 2006;8(5):314-318. [doi: 10.1001/archfaci.8.5.314] [Medline: 16982987]

Abbreviations

aOR: adjusted odds ratio
aRR: adjusted rate ratio
BRFSS: Behavioral Risk Factor Surveillance System

Edited by R Dellavalle; submitted 17.Jul.2025; peer-reviewed by D Hitchcock, J Guo; accepted 17.Dec.2025; published 22.Jan.2026.

Please cite as:
Ray R, Nguyen M
Association of Skin Cancer With Clinical Depression and Poor Mental Health Days: Cross-Sectional Analysis
JMIR Dermatol 2026;9:e80710
URL: https://derma.jmir.org/2026/1/e80710
doi:10.2196/80710

© Riona Ray, Mytien Nguyen. Originally published in JMIR Dermatology (http://derma.jmir.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceptions of a Culturally Tailored Narrative Video for Skin Cancer Prevention Among Spanish-Speaking Hispanic Outdoor Workers: Exploratory Study

David Perez[1], BA; Valeria Gómez[1], BS; Dariana Sedeño-Delgado[1], MS; Carlos Orellana García[2], MPH; Melida Chacon[3]; Miguel Gonzalez[4]; Chiranjeev Dash[1], PhD, MD; Roxanne Mirabal-Beltran[1], PhD, RN; Alejandra Hurtado-de-Mendoza[1], PhD

[1]Department of Oncology, Georgetown University Medical Center, 2115 Wisconsin Avenue, Suite 300, Washington, DC, United States

[2]Epidemiology & Population Science, Baylor College of Medicine, Houston, TX, United States

[3]Catholic Charities, Washington, DC, United States

[4]We Are CASA, Wheaton, MD, United States

Corresponding Author:
Alejandra Hurtado-de-Mendoza, PhD
Department of Oncology, Georgetown University Medical Center, 2115 Wisconsin Avenue, Suite 300, Washington, DC, United States

Abstract

Background: Outdoor workers receive significantly more UV radiation exposure than indoor workers, increasing their skin cancer risk. Hispanic individuals comprise a large proportion of the outdoor workforce in the United States, but have limited access to culturally tailored prevention resources.

Objective: This exploratory study assessed the perceptions of a culturally tailored narrative video to promote skin cancer prevention among Spanish-speaking Hispanic outdoor workers.

Methods: A qualitative study was conducted using two focus groups with Hispanic outdoor workers. Prior to viewing the video, participants completed surveys assessing sociodemographic factors, sun exposure, and protective behaviors. Focus group discussions explored video comprehension, cultural relevance, emotional engagement, and dissemination strategies. Quantitative data were summarized descriptively, and qualitative data were analyzed using rapid qualitative analysis.

Results: Among 19 participants, the mean age was 46.5 (SD 15.5) years, and 94.7% (18/19) identified as male. While 47.4% (9/19) reported 6 - 9 hours of outdoor work per day, more than half (10/19, 52.6%) reported not engaging in sun protection. Participants described the video as culturally resonant, realistic, and engaging. Identification with characters and family-centered themes increased perceived relevance and message credibility. Many participants reported increased awareness of skin cancer risk and greater confidence in adopting sun-protective behaviors. Dissemination recommendations included workplaces, social media, schools, daycares, clinics, and other community settings.

Conclusions: Our video shows promise in improving skin cancer awareness and prevention, reflecting strong acceptability among participants. Future randomized controlled studies are needed to assess its effectiveness in enhancing long-term sun-protective behaviors.

(JMIR Dermatol 2026;9:e98130) doi:10.2196/98130

KEYWORDS

Hispanic outdoor workers; skin cancer prevention; health communication; narrative education; disparities

Introduction

Skin cancer continues to be the leading cancer diagnosis in the United States [1,2]. Basal cell carcinoma (BCC) is the most common form of skin cancer, followed by squamous cell carcinoma (SCC), while invasive melanoma is the deadliest form of skin cancer [3,4]. Each year, there are approximately 3.6 million cases of BCC diagnosed in the United States, and an estimated 1.8 million cases of SCC [4]. Within the past decade, from 2016 to 2026, the incidence of invasive melanoma has increased by 46.6% [2,5,6]. In 2026 alone, it is estimated that 234,680 melanoma cases will be diagnosed in the United States, of which 112,000 will be invasive [2,5]. These projections represent a 10.6% increase in cases and a 1% increase in mortality compared with 2025 [2,5]. UV radiation (UVR) from the sun is widely recognized as the primary risk factor for the development of skin cancer [2,7].

XSL•FO
RenderX

Sunburns, which result from prolonged and intense UVR exposure without sun protection, are reported to be directly associated with an increased risk of both nonmelanoma (BCC and SCC) and melanoma skin cancer [8]. Severe sunburns are among the strongest predictors of melanoma, with evidence showing that even 1 sunburn every 2 years can triple the risk [9,10]. Outdoor workers receive up to 8 times more UVR exposure than indoor workers and have a 60% greater risk of developing skin cancer [11].

Construction and landscaping are among the occupations with the highest proportion of outdoor work time [12]. Despite this, prior interventions targeting these outdoor sectors in the United States remain scarce [13-15].

Interventions for US outdoor workers in the construction and landscape industries are critically needed, especially for Hispanic workers, who represent 35% and 47% of the construction and landscape workforce in the United States, respectively [16]. Although Hispanic individuals have lower melanoma incidence rates than non-Hispanic White individuals, incidence among Hispanic individuals in the United States has increased by 20% in the past 2 decades [17]. Hispanic individuals also experience poorer melanoma outcomes as they are more likely to be diagnosed with thicker melanomas at more advanced stages and have lower melanoma-specific survival rates than non-Hispanic White individuals [17]. These disparities may be particularly relevant among Hispanic outdoor workers due to their increased occupational UVR exposure. Prior research among US Hispanic outdoor laborers (N=175) found that more than half reported experiencing at least 1 sunburn during the summer preceding the study [18]. Hispanic individuals also report low levels of sun-protective behaviors and self-skin examinations [18-20]. Cultural beliefs and social norms play an important role in shaping perceptions of sun exposure risk among Hispanic populations [20,21]. These cultural factors, along with others, such as limited knowledge and awareness, low health literacy and numeracy, limited access to health care, language barriers, and competing priorities, further complicate efforts to promote effective sun protection within this population [20,21].

While recent evidence suggests educational and multicomponent interventions are promising strategies to enhance sun-protective behaviors [14,22], past studies have mostly included workers with fair skin types and were often conducted outside the United States [14,23]. Narrative approaches have emerged as a promising strategy in health communication given their ability to increase engagement and their adaptability across diverse populations [24,25]. Narrative health education materials integrate health information into character-driven stories, unlike traditional didactic materials, which are driven by facts and statistics [26,27]. Narrative communication is hypothesized to influence behavior through multiple psychological processes, including narrative transportation and engagement [28]. Narrative transportation refers to the degree to which individuals are cognitively and emotionally immersed in a story [28]. In addition to decreasing resistance and increasing receptivity to the health message, experiencing “transportation” into the story can support engagement and increase understanding as well as interest in preventive behavior [28,29]. Narrative interventions have also been shown to influence behavioral determinants like self-efficacy by depicting protective behaviors in relatable characters and contexts [30,31]. This may help bridge the gap between knowledge and action by making recommended behaviors clear and feasible [31]. Narratives can also elicit emotional responses that may enhance the impact of the message and motivate behavioral change [32]. Emotion-focused models of health communication suggest that these emotional reactions affect how people pay attention to, understand, and seek information, making it feel more relevant and supporting intentions to change behavior [32].

Evidence indicates that narratives are effective cancer communication approaches among Hispanic individuals [33,34]. Additionally, research has shown that Hispanic outdoor workers prefer videos over printed brochures or fact sheets [35-37]. Most psychoeducation interventions, however, have used a traditional didactic approach, and only a limited number have been developed for Hispanic populations, with even fewer specifically tailored for Hispanic outdoor workers [14,22]. In response to these gaps, we developed a theoretically-informed character-driven video in an effort to increase sun-protective behaviors and promote skin cancer prevention among Hispanic-identifying outdoor workers [38]. To the best of our knowledge, our video is among the first Spanish-language narrative videos tailored for and developed in collaboration with Hispanic outdoor workers living in the United States. This exploratory study aimed to evaluate the perceptions of our culturally tailored narrative education video among Spanish-speaking Hispanic outdoor workers for skin cancer prevention using a qualitative approach.

Methods

Overview

To develop the video, we collaborated with a diverse multistakeholder advisory board (MAB) and a community advisory board that included Hispanic outdoor workers, patients, clinicians, and representatives from local and national advocacy organizations. The initial draft of the script was informed by formative focus group findings [20] and developed by the National Conservatory of Dramatic Arts in Washington, DC. Two rounds of script review were conducted, and community advisory board meetings preceded MAB meetings to support Spanish-language discussion and ensure community perspectives informed subsequent MAB deliberations. The finalized script was then translated into Spanish and reviewed by a Spanish-speaking dermatologist on the MAB for production by the National Conservatory of Dramatic Arts. A detailed description of this process is reported elsewhere [38].

Consistent with Kreuter et al [39] framework for culturally tailored health communication, the video included both superficial and deep structure cultural elements. Superficial tailoring was achieved with Hispanic actors, Spanish language dialogue, and culturally familiar elements such as traditional foods and soccer [39]. Deep tailoring targeted cultural values and beliefs identified as salient within Hispanic communities like familismo by depicting family relationships and support as important motivators to engage in sun-protective behaviors and

XSL•FO
RenderX

early detection [39,40]. To reflect these tailoring strategies, the video portrays a multigenerational family who owns a landscaping company. The main character, Miguel, is diagnosed with BCC as he is transitioning into taking the lead of his father's business. The diagnosis propels him and his wife, Sofia, to take an active role in skin cancer prevention at home and in the workplace. Throughout the story, the video provides information about skin cancer, its signs, protective behaviors, and the importance of early detection. The video also models sun-protective behaviors and highlights their significance in all generations. A scene-by-scene description of the video and their associated theoretical constructs are reported elsewhere [38].

To evaluate perceptions of the video, 2 in-person focus groups were conducted between October and November 2025. We recruited participants through trusted community organizations serving Hispanic communities, including Catholic charities and We Are CASA, both of which have established relationships with outdoor workers employed in construction, landscaping, and agricultural occupations. Participants were eligible for this study if they self-identified as Hispanic, were aged ≥18 years, and had current or prior outdoor work experience. Given the focused scope of the study, consisting of focus groups with a targeted sample, data collection was considered complete once data sufficiency was achieved across the focus groups. This was reflected in the convergence of participant responses across sessions with limited emergence of new viewpoints in the second group. Additionally, as this was an exploratory study designed to guide the development and improvement of a future randomized trial, the primary objective was to assess participants' initial perceptions and potential areas for intervention optimization instead of achieving broad generalizability. Practical factors related to budget deadlines also played a role in the decision to limit the number of focus groups.

Measures

Sociodemographics

Participants were asked to report on their age, gender, race, country of birth, number of years residing in the United States, marital status, educational attainment, health insurance status, health insurance type, annual income, and employment status.

Occupational Sun Exposure and Protection Behaviors

Occupational sun exposure and protection behaviors are listed as follows:

1. History of outdoor work: Participants indicated the outdoor occupational sectors they have experience in, including construction, landscaping, agriculture, and/or other.
2. Years of outdoor work experience: Participants recorded their years of outdoor work experience: less than 1 year, 1 - 5 years, 6 - 10 years, 11 - 15 years, 16 - 20 years, or more than 20 years.
3. Occupational sun exposure: Participants reported the typical number of hours they work outdoors per day. Responses were categorized as: 1 - 3 hours, 3 - 6 hours, 6 - 9 hours, and 10 or more hours.
4. Sun protection use**:** Participants answered a binary (yes or no) item asking if they engaged in any sun-protective behaviors.
5. Sun protection practices: participants indicated which sun protection strategies they used, if any. Behaviors included wearing sunscreen, a hat, sunglasses, long-sleeved shirts, and long pants.
6. Protection frequency: Participants reported on how often they protect their skin from the sun. Response options included: never, rarely, sometimes, often, or always.
7. Shade time: Participants indicated the amount of time they typically spend in the shade during work. Answer choices included never, rarely, sometimes, often, or always.

Focus Group Guide

The focus groups aimed to gather feedback from Hispanic outdoor workers on our culturally tailored video designed to increase skin cancer awareness and prevention strategies. Discussions began with introductions and ice-breakers to create a comfortable environment and establish focus group norms, emphasizing confidentiality and respectful dialogue.

Guided by the Learner Verification and Revision Framework [41], participants were then asked about their perceptions on comprehension, self-efficacy, attraction, and cultural-linguistic acceptability. To further capture narrative-specific mechanisms of engagement, per narrative transportation frameworks [42], participants also reported on emotional engagement, character identification, and transportation. The final section of the guide focused on feasible ways to disseminate the video within their community, including preferred platforms and settings.

Quantitative Data Analysis

Descriptive analysis was conducted to summarize sociodemographic and occupational sun exposure and protection behaviors using Statistical Product and Service Solutions (version 31; IBMorp).

Qualitative Data Analysis

We analyzed focus group responses using rapid qualitative data analysis, as this method aligns with standard qualitative techniques and is optimized for short, structured studies with defined goals [43]. Using recommended rapid qualitative data analysis practices [43], the analysis was performed in six steps. In step 1, two members of the research team independently analyzed a single focus group transcript and created a neutral domain summary template based on the focus group topics and questions. In step 2, the analysts met to discuss their summary templates and created a consolidated version of the summary template. In step 3, each analyst independently summarized both focus group transcripts using the summary template. In this summarization phase, the analysts developed descriptive summaries consisting of bullet statements organized by domains and identified key quotes, with the goal of providing a concise summary of the data while avoiding interpretation. In step 4, the analysts compared their focus group summaries and refined them through group discussion. In step 5, the analysts compiled the focus group summaries into a matrix organized by domains. In step 6, the team used the matrix to discuss patterns, connect domains, and develop themes.

XSL•FO
RenderX

Ethical Considerations

This study was reviewed by the Georgetown University Institutional Review Board and granted an exemption determination (IRB #STUDY00007990; exempt under category (2) (ii): tests, surveys, interviews, or observation, low risk). Trained research staff reviewed the informed consent form and study details with participants prior to each focus group, and verbal consent was documented by the study team before any data collection began. Participation was voluntary, and participants were free to decline or withdraw at any time. Data are coded rather than fully deidentified. Focus group audio recordings were transcribed and deidentified for analysis, while any files containing identifying information are stored separately in Georgetown University's secure, password-protected box repository, accessible only to authorized study team members. All participants received a US $50 gift card as compensation for completing the focus group.

Results

Sociodemographics

Of the 20 participants recruited, 19 met our inclusion criteria and were included in the final sample. The mean age was 46.5 (SD 15.5) years. Most (18/19, 94.7%) participants were male, and almost half (9/19, 47.7%) identified as “Other” for race. More than half (10/19, 52.6%) reported an education of middle school or less, 57.9% (11/19) indicated an annual income of US $10,000-$39,999, and 63.2% (12/19) did not have health insurance (Table 1).

XSL•FO
RenderX

Table . Sociodemographics (n=19).

Variable	Values
Age (years), mean (SD)	46.5 (15.5)
Gender, n (%)	
Men	18 (94.7)
Women	1 (5.3)
Race, n (%)	
White	2 (10.5)
Multiracial	1 (5.3)
Other	9 (47.4)
Prefer not to respond	4 (21.1)
Unknown	3 (15.8)
Country of birth, n (%)	
El Salvador	6 (31.6)
Peru	5 (26.3)
Colombia	1 (5.3)
Guatemala	1 (5.3)
Honduras	1 (5.3)
Costa Rica	1 (5.3)
Nicaragua	1 (5.3)
Venezuela	1 (5.3)
Other	1 (5.3)
Unknown	1 (5.3)
Number of years residing in the United States, mean (SD)	10.8 (8.4)
Marital status, n (%)	
Married/living with domestic partner	6 (31.6)
Never married	5 (26.3)
Divorced	1 (5.3)
Other	5 (2.3)
Prefer not to respond	2 (10.5)
Educational attainment	
Elementary school (through 5th grade)	5 (26.3)
Middle school (6th-8th grade)	5 (26.3)
High school diploma/GED[a]	4 (21.1)
Some college	3 (15.8)
Bachelor’s degree	1 (5.3)
Postgraduate degree	1 (5.3)
Annual income, n (%)	
Less than US $10,000	3 (15.8)
US $10,000-$39,999	11 (57.9)
US $40,000-$79,999	2 (10.5)
Prefer not to respond	2 (10.5)
Unknown	1 (5.3)
Health insurance, n (%)	

XSL•FO
RenderX

Variable	Values
Yes	7 (36.8)
No	12 (63.2)
Health insurance type, n (%)	
Private (individual)	3 (15.8)
Medicare	2 (10.5)
Medicaid	1 (5.3)
Employer	1 (5.3)
Unknown	12 (63.2)
Employment status, n (%)	
Full-time	6 (31.6)
Part-time	8 (42.1)
Unemployed	4 (21.1)
Student	1 (5.3)

[a]GED: General Educational Development.

Occupational Sun Exposure and Protection Behaviors

The majority (8/19, 42.1%) of participants reported 6 - 10 years of outdoor work experience. Nearly half (9/19, 47.4%) work 6 - 9 hours per day in the sun, and 52.6% (10/19) do not engage in sun-protective behaviors (Table 2). Wearing long-sleeved shirts (13/19, 76.5%) and long pants (12/19, 70.6%) were the most reported sun-protective behaviors, while sunscreen use (4/19, 23.5%) was the least reported (Table 2).

XSL•FO
RenderX

Table . Occupational sun exposure and protection behaviors.

Variable	Participants, n (%)
History of outdoor work[a]	
Construction	16 (84.2)
Landscaping	9 (47.4)
Agriculture	2 (10.5)
Other	3 (15.8)
Years of outdoor work experience	
Less than 1 year	2 (10.5)
1 - 5 years	2 (10.5)
6 - 10 years	8 (42.1)
11 - 15 years	6 (31.6)
16 - 20 years	0 (0.0)
Over 20 years	1 (5.3)
Daily hours working under the sun	
1 - 3 hours	3 (15.8)
3 - 6 hours	4 (21.1)
6 - 9 hours	9 (47.4)
10 hours or more	2 (10.5)
Unknown	1 (5.3)
Use of sun protection	
No	10 (52.6)
Yes	9 (47.4)
Frequency of sun protection	
Never	3 (15.8)
Rarely	2 (10.5)
Sometimes	7 (36.8)
Often	1 (5.3)
Always	6 (31.6)
Amount of time spent in shade during work	
Never	2 (10.5)
Rarely	5 (26.3)
Sometimes	11 (57.9)
Often	1 (5.3)
Always	0 (0.0)
Types of sun protection used[a]	
Long-sleeved shirt	13 (76.5)
Pants	12 (70.6)
Hat	10 (58.8)
Sunglasses	6 (35.3)
Sunscreen	4 (23.5)

[a]Could select more than one response.

XSL•FO
RenderX

Focus Group Findings

Satisfaction

Participants appreciated the video's focus on skin cancer, a topic they knew little of and were rarely taught about at work, despite their increased UV exposure:

> *I really like it because I was in construction for more than 20 years, and I don't remember–never–that anyone gave me a talk about the damage that can come from so much sun exposure…and very little is said about skin cancer.* [Focus group (FG) 1]

Others noted that it completely altered their viewpoint:

> *It changed mine 100%.* [FG 1]

Several described it as credible and trustworthy, stating that the information appeared up to date and was professionally presented.

Comprehension and Increased Knowledge

After watching the video, participants reported feeling significantly more knowledgeable regarding skin cancer. Although the video presented substantial information, many participants described it as well-organized and easy to understand:

> *I feel that, for me, the message is quite clear. Very clear. Why? Because it talks about what skin cancer is, what the causes are, and what possible actions need to be taken to prevent it* [FG 1]

Participants were particularly surprised that skin cancer could develop decades after unprotected sun exposure, believing that it happened as late as 1 or 2 years of being under the sun. Participants were also shocked to learn that skin cancer could develop in young and middle-aged adults as a result of early-life UV exposure. As one participant stated:

> *To me it seemed very important since I thought that I did not have the capacity to suffer from cancer at an early age. I used to say that only happened to old people, but the information you are giving us seemed very interesting to me, as it explains to us that we are exposed to radiation like ultraviolet rays at all times and we need to protect ourselves.* [FG 1]

Self-Efficacy

Participants especially valued that the content was actionable. When asked how confident they felt in implementing the strategies presented after watching the video, participants expressed strong confidence:

> *Compared to before, I'm very confident now.* [FG 1]

In addition to increasing their confidence and motivation to practice sun-protective behaviors, including using sunscreen, seeking shade, and wearing protective clothing, participants also expressed a strong desire to share this knowledge and empower their loved ones:

> *The most important thing is that, after watching this video, I now have more information for myself to take better care of myself. Even though this isn't currently mandatory at my company, I know I will do it. I will try to take better care of myself, and of course, care for those closest to me as well. I am very grateful for this.* [FG 1]

Attraction

Most participants expressed a clear preference for the narrative story format over didactic explanations delivered by a physician. They reported that the video was engaging and entertaining, capturing their attention without feeling "boring" or overly formal. Many noted that the pacing and length were appropriate, covering all necessary topics without feeling rushed or incomplete:

> *I found the video very informative, and I would neither remove nor add anything* [FG 2]

Cultural Relevance and Identification With Characters

Participants also expressed strong identification with the characters. Many commended the video's portrayal of the close-knit family dynamic representative of Hispanic culture:

> *The family support, being with family, home, and all of that is very much directed toward our community.* [FG 1]

Several particularly valued the story's themes of friendship, trust, and support for one another, highlighting the trust between Miguel and Sofia as partners and the supportive relationship between Miguel and his father. Participants additionally appreciated the video's attention to how crucial it is for coworkers to share information in order to protect one another:

> *I really liked the support as well, that among the workers there is support, and the knowledge is shared with others so they don't fall into the same mistake as someone else.* [FG 2]

The video was described as very relevant to their daily work and life experiences. Many emphasized how realistic and representative the characters' work experiences felt compared to their own lives. In addition to capturing their daily realities, several expressed that the video effectively clarified widespread myths among Hispanic individuals regarding sun protection and the risk of skin cancer. Participants specifically noted that it competently addressed machismo, a key cultural barrier to sun-protective behaviors among Hispanic men characterized by traditional masculinity, by alluding to the cultural influence of soccer players in practicing sun-protective behaviors (eg, sunscreen use). Many also appreciated that the video was produced in Spanish as it not only felt like it spoke to them directly but also addressed another key barrier for those with low literacy. According to one participant:

> *I think it is very focused on the position of the Hispanic worker. Obviously the language, the physical appearance, and all those qualities that represent us are very evident.* [FG 2]

Emotions

Participants further described feeling engaged and immersed in the storyline from start to finish. Several mentioned that the video evoked strong emotions like sadness and concern,

XSL•FO
RenderX

particularly when thinking about how a cancer diagnosis could affect a person's family. One participant shared the emotional impact:

> *I was about to cry. [Laughs] I looked down twice so they [other focus group participants] wouldn't see me.* [FG 1]

At the same time, participants liked that the video ended on a hopeful note, focusing on self-agency and prevention:

> *I'm thankful, as she [other participant] said, to let family know that there is a way to treat cancer. There are those of us who say, 'We go through it, and at that point, I'm going to die, I'm going to die.' But always, at the end of the road, there is always a way out. Thank you.* [FG 1]

This transition allowed participants to acknowledge the risks of sun exposure while feeling encouraged to take actionable steps.

Suggestions for Improvement

To enhance the effectiveness of the video, participants suggested having the video presented by a professional available to answer questions and pairing it with a pamphlet to include (1) listing providers they could contact and locations where they could seek treatment if needed, (2) additional information on sunscreen types and application, and (3) providing illustrations of how skin cancer can appear across different skin tones and ages. As one participant expressed:

> *The information I got from watching the whole video about the doctor—that he [Miguel] has been well cared for—was before things progressed further. But I think that, like all of us Latinos [used interchangeably with the term "Hispanic" at a similar frequency], there's something here in this country that we fear even more than cancer itself. I mean, going to a doctor. We think, if we go to a doctor, how much is it going to cost me? What will the bill be? And because of that fear, we don't get treated; we just let it go. So, I think that including a reference, that is, a phone number where you can contact this doctor, or this person, someone who can see you, or something along those lines would be helpful.* [FG 2]

Implementation and Dissemination Strategies

Barriers to Implementation

Participants identified several barriers related to implementing sun-protective behaviors and sharing the video. Personal factors, like forgetting sunscreen, past allergic reactions, and occasional inconvenience, were mentioned as potentially hindering protective behaviors. Occupational barriers included lack of company support for sun protection, absence of formal orientation or training, and work uniforms that do not provide enough coverage. As expressed by one of the participants:

> *There's definitely a lack of education in this area, right? A huge lack of education when it comes to sun protection, and clearly companies don't care because they're focused on production. I mean, they need production, so it's not going to matter to them. Like, 'look, let's stop the equipment—right now the sun is so hot, we're going to shut it down for half an hour or an hour so the workers can rest and hydrate a bit.' That generally doesn't happen, and it doesn't happen either because the necessary education isn't there, or because they really don't care about anything beyond production.* [FG 1]

Facilitators to Implementation

To address these barriers, participants suggested strategies to enhance awareness. One participant explained:

> *I think the first step is to give informative talks about this situation and make them [outdoor companies] aware of the purpose, the importance of what we're discussing, and how to protect the worker in their own defense.* [FG 2]

Participants also recognized the importance of self-agency in sun protection, emphasizing that individuals cannot rely entirely on employers to provide guidance or resources. Cost was generally not viewed as an obstacle when sunscreen was seen as helpful and effective. Participants mentioned that clear and culturally relevant messages, identification with characters, understanding of risks, and support from family were significant factors that encouraged behavior change. They also noted that the confidence gained from watching the video would motivate them to use protective strategies.

Suggestions for Dissemination

To increase the video's impact and reach, participants offered several suggestions. They advised disseminating it via social media sites like YouTube (Google LLC) and Facebook (Meta Platforms, Inc), as these platforms are accessed by individuals of all ages, including children:

> *Now I think there is a lot of technology, and I believe that is what we are most focused on—social media. Often, I think we can use it as a way to inform people, because nowadays even the youngest children have Facebook, Instagram, and other platforms. So perhaps using social media could be an effective way to raise awareness* [FG 1]

Participants also suggested sharing it within workplaces, especially during orientations or preshift meetings (eg, toolbox talks). Use of QR codes and links sent by companies, along with follow-up question-and-answer sessions, were also encouraged to not only ensure engagement, but also comprehension.

Broader dissemination strategies included incorporating the video into community settings. Participants particularly emphasized the importance of introducing the video early, such as in daycare and schools, so that awareness of skin cancer risk and sun-protective behaviors can be established from a young age. Other sites such as clinic waiting rooms were also encouraged so viewers could further discuss any thoughts or questions with their clinicians.

Many participants also expressed strong interest in having resources like these acknowledged and supported by

XSL•FO
RenderX

occupational safety campaigns as well as organizations such as the Occupational Safety and Health Administration. According to one participant:

> *I believe that, for there to be education about skin cancer, the first focal point should be OSHA because it is the largest department here in the United States, and many Latin American countries, as well as other parts of the world, also follow its guidance.* [FG 1]

It was additionally suggested that the video be shared by, or shown in the presence of, a skin cancer survivor, as this could help viewers relate the narrative to real-life experiences and increase its impact. Participants also emphasized the importance of timing, recommending that sharing the video during the summer months, when sun exposure is highest, would be particularly important. See Tables 3 and 4 for additional quotes.

XSL•FO
RenderX

Table . Video reactions.

Category	Summary	Quotes
Satisfaction	• Appreciated that the video discusses a topic that isn't taught about despite their increased risk as outdoor workers • Highly satisfied with the video's explanation of skin cancer and the steps they can take to address it • Liked that the information is up to date	• "I think that, as [name] says, the video is very good, so that we can learn what skin cancer causes, its consequences, and how to correct it." (FG[a] 1) • "I think the video is good; it is clear with the information it has, and it is up to date with the current times." (FG 2)
Comprehension and increased knowledge	• Found the video very easy to understand • Clear message and sequence of events • Increased understanding of long-term risks of UV exposure • Didn't know skin cancer could happen up to 20 years after UV exposure • Weren't aware of the UV index • Misconceptions about UV susceptibility in the Hispanic community	• "It is very interesting how the video presented the topic, and with very, very clear language, which I think we all should know, and we should also share it with our family, and later with our work groups." (FG 1) • "I was surprised by something they mentioned there that I really had no idea about, that this cancer could appear 5, 10, even up to 20 years after having been exposed to the sun. I thought it was something earlier, right? I thought it could be something like 1 year, 2 years, or probably 3." (FG 1) • "I thought it was very important because I thought that because I am young, I didn't have the chance of getting cancer at an early age. I used to say that that only happened to old people, but I found the information they are giving us very interesting, because it makes us aware that we are exposed to radiation like ultraviolet rays at all times, and we need to protect ourselves against it." (FG 1) • "I think that what catches my attention is that people think that because of their color they are not exposed to harmful damage, and that is something that gets instilled in you from when you are little." (FG 2)
Self-efficacy	• Felt more confident applying sun-protection strategies learned from the video • Greater motivation to implement behaviors and share protection strategies with family members	• "After watching the video, how confident are you that you can follow its recommendations?" "100 %." (FG 2) • "I'm going to take it very much into account and share it with my family and all that. I didn't know about this." (FG 2)
Attraction	• Story format preferred over didactic explanations by a physician • Length and pacing were appropriate as the video covered all necessary topics without feeling rushed • Wouldn't add or take away anything	• "It's not long and it's complete. It gives the information—the process of cancer, how to treat it, and how to prevent it, which is what's important." (FG 2) • "The story resonated with me more because the doctor explains things in a way that works well." (FG 2) • "I think the video is well structured and has a sequence." (FG 1)

XSL•FO
RenderX

Category	Summary	Quotes
Cultural relevance and identification with characters	• Reflects real workplace dynamics and experiences • Appreciated that the video is in Spanish • Family support portrayed is heavily representative of Hispanic culture • Effectively explores cultural gender norms (eg, machismo) and offers an innovative approach to their associated barriers • Many participants expressed a strong desire to share the video with coworkers and family • Characters look and speak like them • Liked the trust shown between Miguel and Sofia as partners • Identified with Miguel and his father due to their strong family dynamic • Enjoyed the theme of companionship and information-sharing within the company	• "I really identify a lot with the video because in my particular case, I am someone who works almost always exposed to the sun, and I think it has been like that for a long time. I've had jobs where I spend a lot of time outside." (FG 1) • "I thought the video was very good. Especially directed toward the Hispanic community." (FG 2) • "How much do you think the video reflects the Hispanic community? Or how much do you think it can resonate with the Hispanic community? I think 100 %. Yes, 100 %." (FG 1) • "They were very well acted; they seemed like professionals." (FG 2) • "I really liked the trust between the couple because it says that between partners it is mutual to help each other." (FG 2) • "We Hispanics are like them—when someone says, 'Hey, put on sunscreen,' the response is often, 'Man, no, that's not for us.' As we say, 'hágale huevo,' meaning we just push through and act tough. The video portrayed this well, showing how people typically resist, saying, 'No, we don't use that, just do it this way,' which makes it relatable and understandable." (FG 1)
Emotions	• Evoked sadness about potential consequences of skin cancer • Ended on a positive note, leaving participants with a sense of empowerment • Emotional scenes with family were impactful and culturally resonant	• "Sometimes we don't take care of ourselves, or we don't take it seriously, and in the end, it ends up harming us." (FG 2) • "I felt good because everything ended positively." (FG 1)
Suggestions for improvement	• Include resources in a pamphlet about where they can seek treatment if needed • Have a professional available to answer questions that may arise during the video • Testimony from a real individual with melanoma could enhance motivation • Emphasize legal advocacy and integration (eg, OSHA[b] regulations) • More information on sunscreen application and type • Include images of skin cancer across a range of skin colors and ages	• "The final point, where I think it is important to reach, is that besides having awareness about skin cancer, it is also about integrating it into OSHA's safety system for work. Well, it has always been talked about only using helmets, glasses, gloves, and it would also be very good to adopt the issue of skin protection, which is never talked about." (FG 1)

[a]FG: focus group.

[b]OSHA: Occupational Safety and Health Administration.

XSL•FO
RenderX

Table . Implementation and dissemination strategies.

Category	Summary
Barriers to implementation	• Hispanic individuals are not accustomed to using sunscreen compared to non-Hispanic individuals • Most companies do not teach about sun protection or provide resources for it • Easy to forget sunscreen • Hispanic men often neglect self-care • Past allergic reactions to sunscreen • Can be inconvenient during hot days
Facilitators to implementation	• Recognize the importance of self-agency and not relying entirely on employers • Cost of sunscreen is not an issue if it is useful and effective • Clear and culturally relevant messaging • Knowledge of risks and long-term consequences • Confidence gained after watching the video • Mobile-friendly
Suggestions for dissemination	• Worksite orientations and training • QR codes • Social media (eg, YouTube, Facebook, and Instagram) • On-site informational community sessions • Schools and daycares • Clinics, contractors, and public organizations focused on occupational safety (eg, OSHA[a]) • Involve skin cancer survivors • Increased promotion during the summer months

[a]OSHA: Occupational Safety and Health Administration.

Discussion

Principal Findings

This study strongly supports the potential of a culturally tailored narrative video for skin cancer prevention in engaging Hispanic outdoor workers, supporting comprehension, and increasing confidence in adopting preventive behaviors. Building on prior work on narrative interventions in cancer education [44-47], this study addresses a critical gap in skin cancer prevention and represents one of the first reported narrative videos tailored for and developed with Hispanic outdoor workers, providing early insights into its perceived potential.

Consistent with prior research, our findings emphasize the importance of interventions that reflect the cultural values and lived experiences of the intended audience [48,49]. Participants repeatedly noted that they appreciated the video being presented in Spanish as it not only strengthened its cultural connection but also addressed a systemic gap. In addition to highlighting the portrayal of familism as central in character relatability, participants appreciated how the video addressed specific constructs that could hinder sun protective practices in and out of work. Our findings support the notion that narratives can create opportunities for participants to not only see themselves in the story but also envision themselves adopting recommended behaviors [50]. In our study, participants' connection with Miguel, his father, and other culturally resonant role models, such as soccer players, demonstrates the potency of narratives in addressing culturally specific barriers, including gender norms like machismo, while simultaneously promoting self-efficacy.

Additionally, by causing feelings of concern and introspection, the video appears to have not only enhanced message processing, but also reinforced personal relevance, especially in relation to family members. Although the video showed the potential consequences of skin cancer, participants particularly appreciated that it ended on a positive note. Beyond keeping participants engaged with the narrative, this balance of emotional intensity likely increased their immersion and subsequent self-efficacy, consistent with narrative transportation frameworks [51]. This was reflected in their responses, as many stated that the video had completely changed their perspective and motivated them to start practicing sun-protective behaviors in their everyday lives.

Participants' recognition of self-agency in sun protection despite limited employer support also highlights the importance of interventions that empower workers to take preventive action. Many participants described situations in which workplace infrastructure and/or lack of formal training made reliance on employers alone insufficient for maintaining sun safety. This focus on self-agency reflects health behavior theories that view individuals as active participants in shaping their health outcomes [52,53]. For example, Social Cognitive Theory emphasizes human agency and self-efficacy as important factors in behavior and highlights that individuals are more likely to engage in protective actions when they perceive they can successfully perform those behaviors even in constrained environments [30,31,52]. Likewise, the Theory of Planned Behavior states that perceived control over behavior is a key predictor of intentions and actions related to health, especially when structural barriers exist [53,54]. Empowerment-oriented

XSL•FO
RenderX

approaches add that improving individuals' perceived control over health decisions is a key mechanism for improving outcomes, particularly in resource-limited settings [55]. Together, these theories highlight the need for interventions that build individual capacity by increasing knowledge and confidence in practicing protective behaviors.

At the same time, our findings also stress the importance of supportive workplace policies that make these behaviors practical and sustainable. Encouraging self-efficacy with structural supports aligns with health behavior theories emphasizing that personal agency is most effective when reinforced by enabling environments [56]. This moreover supports the need for multilevel interventions for employees and employers that combine individual-level strategies with organizational policies and environmental supports to maximize adherence and long-term health outcomes. Participants identified multiple ways in which resources like these could be incorporated into their workday, particularly during orientations or preshift meetings, demonstrating both the feasibility of integrating preventive strategies into routine workflows and participants' interest in such resources. Importantly, in contexts where employer-provided resources may be limited, participants emphasized that knowledge about the importance of sun protection could best support such practices, as the cost of items like sunscreen would not be a barrier if perceived as beneficial for their health. To further address the individual-level implementation barriers identified (forgetting sunscreen, allergic reactions to sunscreen, etc), increased complementary education regarding electronic reminders and/or alternative sunscreen formulations (eg, hypoallergenic, mineral-based) for sensitive skin could be helpful.

Participants also recommended sharing the video beyond worksites such as schools, daycares, clinics, and across social media platforms. These suggestions reflect not only their positive reception of the video, but also their perception of its broader importance and relevance. For many Hispanic outdoor workers, employment is closely tied to their role as providers. Protecting their skin is not viewed as an individualistic health behavior, but as a means of safeguarding their ability to remain present and provide for their loved ones [21]. Several participants expressed that the video directly speaks to this sense of responsibility. Additionally, after learning that skin cancer most commonly develops over many years, several expressed a desire to ensure that their families and friends across all generations grow up with more accurate knowledge and preventive habits. Overall, the potential of our video is highlighted by the interest and motivation of participants in implementing and sharing the knowledge obtained from the video, particularly in a cohort in which more than half did not engage in sun-protective behaviors.

Beyond these promising findings, this study has several key strengths. First, its qualitative design enabled a thorough exploration of participants' viewpoints and provided detailed information on experiential aspects that quantitative methods would be unlikely to capture, such as cultural relevance, emotional engagement, and distribution strategies. Second, the study targeted a high-risk group with limited access to culturally-relevant skin cancer prevention resources, not only addressing a significant research gap, but also potentially improving the real-world applicability and impact of the video. Third, it produced practical recommendations to enhance the video's effectiveness and guide distribution strategies, including perceived facilitators and barriers to implementation. Fourth, the study provides insights that can guide future research, including randomized control trials and multilevel interventions with employers. Despite these strengths, several limitations should also be noted. First, this was an exploratory study with a small sample size (N=19) and a male-dominated sample (18/19, 94.7%), limiting the transferability of the study to other Hispanic outdoor workers. The small sample size also limits the range of perspectives captured and the degree to which subgroup differences are reflected in the findings. Female workers were especially underrepresented, and their views on sun protection knowledge and behaviors, as well as perceived barriers and facilitators, may differ in important ways that were not fully captured, similar to other subgroups. Second, data were self-reported, potentially introducing social desirability bias as well as recall bias. Third, with a focus group format, participants' answers may have been influenced by other participants.

Future research should more rigorously evaluate the acceptability of the video, including its impact on sustained sun-protective behaviors over time. Randomized controlled trials comparing this narrative approach to standard educational materials, including a larger sample, are needed to determine its effectiveness. Studies involving employer and occupational health engagement could also assess strategies for integrating sun protection into workplace policies. Future studies could further explore whether the impact of the video differs across outdoor work sectors (construction vs landscaping, agriculture vs construction, etc).

Conclusion

This exploratory study provides early evidence supporting the acceptability and promise of our culturally tailored narrative video intervention to promote skin cancer prevention among Spanish-speaking Hispanic outdoor workers. Feedback regarding its cultural and occupational relevance was strongly positive, with self-reported improvements in knowledge and self-efficacy. Additional research is needed to determine if this approach can meaningfully enhance sun protective behaviors to reduce skin cancer risk among Hispanic outdoor workers at scale.

Acknowledgments

This study was funded by the Georgetown-Howard Universities Center for Clinical and Translational Science (GHUCCTS). Research reported in this publication was funded by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR001409. The content is solely the responsibility of the authors and does not

XSL•FO
RenderX

necessarily represent the official views of the National Institutes of Health. Recruitment for the current study was conducted in collaboration with Georgetown Lombardi Comprehensive Cancer Center's Office of Community Outreach and Engagement, supported in part by NIH/NCI grant P30CA051008.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Hasan N, Nadaf A, Imran M, et al. Skin cancer: understanding the journey of transformation from conventional to advanced treatment approaches. Mol Cancer 2023 Oct 6;22(1):168. [doi: 10.1186/s12943-023-01854-3] [Medline: 37803407]
2. Skin cancer facts & statistics. Skin Cancer Foundation. URL: https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/ [accessed 2026-06-06]
3. Firnhaber J. Skin cancer. Prim Care 2025 Sep;52(3):499-512. [doi: 10.1016/j.pop.2025.05.003] [Medline: 40835287]
4. Our new approach to a challenging skin cancer statistic. Skin Cancer Foundation. URL: https://www.skincancer.org/blog/our-new-approach-to-a-challenging-skin-cancer-statistic/ [accessed 2026-06-06]
5. Cancer facts & figures 2026. American Cancer Society. URL: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2026-cancer-facts-figures.html [accessed 2026-06-06]
6. Cancer facts & figures 2016. American Cancer Society. URL: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2016.html [accessed 2026-06-06]
7. Watson M, Holman DM, Maguire-Eisen M. Ultraviolet radiation exposure and its impact on skin cancer risk. Semin Oncol Nurs 2016 Aug;32(3):241-254. [doi: 10.1016/j.soncn.2016.05.005] [Medline: 27539279]
8. Li Y, Wu J, Cao Z. Childhood sunburn and risk of melanoma and non-melanoma skin cancer: a Mendelian randomization study. Environ Sci Pollut Res Int 2023 Dec;30(58):122011-122023. [doi: 10.1007/s11356-023-30535-3] [Medline: 37962759]
9. Wu S, Cho E, Li WQ, Weinstock MA, Han J, Qureshi AA. History of severe sunburn and risk of skin cancer among women and men in 2 prospective cohort studies. Am J Epidemiol 2016 May 1;183(9):824-833. [doi: 10.1093/aje/kwv282] [Medline: 27045074]
10. Dennis LK, Vanbeek MJ, Beane Freeman LE, Smith BJ, Dawson DV, Coughlin JA. Sunburns and risk of cutaneous melanoma: does age matter? A comprehensive meta-analysis. Ann Epidemiol 2008 Aug;18(8):614-627. [doi: 10.1016/j.annepidem.2008.04.006] [Medline: 18652979]
11. Safe skin at work. IMPACT Melanoma. URL: https://impactmelanoma.org/safe-skin-at-work/ [accessed 2025-11-05]
12. TED: The Economics Daily. US Bureau of Labor Statistics. URL: https://www.bls.gov/opub/ted/2023/32-9-percent-of-employees-had-regular-outdoor-exposure-in-2022.htm [accessed 2026-04-04]
13. Keurentjes AJ, Kezic S, Rustemeyer T, Hulshof CTJ, van der Molen HF. Stimulating sunscreen use among outdoor construction workers: a pilot study. Front Public Health 2022;10:857553. [doi: 10.3389/fpubh.2022.857553] [Medline: 35433569]
14. de Boer FL, Vreenegoor D, Twisk JWR, et al. Evaluation of a toolbox for the prevention of skin cancer among outdoor workers: an intervention study. Front Public Health 2025;13:1579180. [doi: 10.3389/fpubh.2025.1579180] [Medline: 40552244]
15. Rocha R, Santos J, Baptista JS, Guedes J, Carvalhais C. Occupational exposure to solar ultraviolet radiation: a systematic review of protective measures. Safety 2026 Jan;12(1):10. [doi: 10.3390/safety12010010]
16. Labor force statistics from the current population survey. US Bureau of Statistics. URL: https://www.bls.gov/cps/cpsaat11.htm [accessed 2026-04-04]
17. Perez MI. Skin cancer in Hispanics in the United States. J Drugs Dermatol 2019 Mar 1;18(3):s117-s120. [Medline: 30909356]
18. Niu Z, Riley M, Stapleton JL, et al. Sunburns and sun protection behaviors among male Hispanic outdoor day laborers. Int J Environ Res Public Health 2022 Feb 22;19(5):2524. [doi: 10.3390/ijerph19052524] [Medline: 35270218]
19. Day AK, Stapleton JL, Natale-Pereira AM, Goydos JS, Coups EJ. Occupational sunscreen use among US Hispanic outdoor workers. BMC Res Notes 2015 Oct 17;8:578. [doi: 10.1186/s13104-015-1558-1] [Medline: 26477041]
20. Perez D, Sedeño-Delgado D, García CO, et al. Exposing the risk: identifying perspectives on skin cancer awareness and prevention opportunities among Hispanic outdoor workers. J Racial Ethn Health Disparities 2025 Sep 5. [doi: 10.1007/s40615-025-02623-2] [Medline: 40911153]
21. Niu Z, Rivera YM, Baskar J, et al. Barriers and facilitators to skin cancer prevention among Hispanics: a qualitative study. BMC Public Health 2024 Oct 2;24(1):2690. [doi: 10.1186/s12889-024-20000-2] [Medline: 39358786]

XSL•FO
RenderX

22. Buller DB, Walkosz BJ, Buller MK, et al. Implementation of occupational sun safety at a 2-year follow-up in a randomized trial: comparison of sun safe workplaces policy intervention to attention control. Am J Health Promot 2019 Jun;33(5):683-697. [doi: 10.1177/0890117118814398] [Medline: 30477333]
23. Houdmont J, Madgwick P, Randall R. Sun safety in construction: a UK intervention study. Occup Med (Lond) 2016 Jan;66(1):20-26. [doi: 10.1093/occmed/kqv140] [Medline: 26409050]
24. Shaffer VA, Focella ES, Hathaway A, Scherer LD, Zikmund-Fisher BJ. On the usefulness of narratives: an interdisciplinary review and theoretical model. Ann Behav Med 2018 Apr 19;52(5):429-442. [doi: 10.1093/abm/kax008] [Medline: 29684135]
25. Spencer TD, Petersen DB. Narrative intervention: principles to practice. Lang Speech Hear Serv Sch 2020 Oct 2;51(4):1081-1096. [doi: 10.1044/2020_LSHSS-20-00015] [Medline: 32776816]
26. Murphy ST, Frank LB, Chatterjee JS, et al. Comparing the relative efficacy of narrative vs nonnarrative health messages in reducing health disparities using a randomized trial. Am J Public Health 2015 Oct;105(10):2117-2123. [doi: 10.2105/AJPH.2014.302332] [Medline: 25905845]
27. Kreuter MW, Green MC, Cappella JN, et al. Narrative communication in cancer prevention and control: a framework to guide research and application. Ann Behav Med 2007 Jun;33(3):221-235. [doi: 10.1007/BF02879904] [Medline: 17600449]
28. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000 Nov;79(5):701-721. [doi: 10.1037//0022-3514.79.5.701] [Medline: 11079236]
29. Dillard AJ, Ferrer RA, Welch JD. Associations between narrative transportation, risk perception and behaviour intentions following narrative messages about skin cancer. Psychol Health 2018 May;33(5):573-593. [doi: 10.1080/08870446.2017.1380811] [Medline: 28975805]
30. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
31. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [doi: 10.1037/0003-066x.44.9.1175] [Medline: 2782727]
32. Dillard JP, Nabi RL. The persuasive influence of emotion in cancer prevention and detection messages. J Commun 2006 Aug 4;56(suppl_1):S123-S139. [doi: 10.1111/j.1460-2466.2006.00286.x]
33. Baezconde-Garbanati LA, Chatterjee JS, Frank LB, et al. Tamale Lesson: a case study of a narrative health communication intervention. J Commun Healthc 2014 Jul;7(2):82-92. [doi: 10.1179/1753807614Y.0000000055]
34. Hurtado-de-Mendoza A, Graves KD, Gómez-Trillos S, et al. Culturally targeted video improves psychosocial outcomes in Latina women at risk of hereditary breast and ovarian cancer. Int J Environ Res Public Health 2019 Nov 29;16(23):4793. [doi: 10.3390/ijerph16234793] [Medline: 31795362]
35. Garcia D, Jefferson IS, Ramirez P, et al. Video education to promote skin cancer awareness and identification in spanish-speaking patients. J Clin Aesthet Dermatol 2020 Jan;13(1):41-43. [Medline: 32082472]
36. Besch-Stokes J, Brumfiel CM, Patel MH, et al. Skin cancer knowledge, attitudes and sun protection practices in the Hispanic population: a cross-sectional survey. J Racial Ethn Health Disparities 2023 Jun;10(3):1293-1303. [doi: 10.1007/s40615-022-01314-6] [Medline: 35486349]
37. Jacobsen AA, Galvan A, Lachapelle CC, Wohl CB, Kirsner RS, Strasswimmer J. Defining the need for skin cancer prevention education in uninsured, minority, and immigrant communities. JAMA Dermatol 2016 Dec 1;152(12):1342-1347. [doi: 10.1001/jamadermatol.2016.3156] [Medline: 27626892]
38. Perez D, Gómez V, Sedeño-Delgado D, et al. Developing a culturally tailored narrative video for skin cancer prevention among Spanish-speaking Hispanic outdoor workers: co-design study. J Particip Med 2026 May 22;18:e89986. [doi: 10.2196/89986] [Medline: 42172572]
39. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [doi: 10.5993/ajhb.27.1.s3.6] [Medline: 14672383]
40. Torrez-Ruiz MS, Soto S, Lopez NV, Arredondo EM. A strength-based approach to cancer prevention in latinxs. In: Ramirez AG, Trapido EJ, editors. Advancing the Science of Cancer in Latinos: Building Collaboration for Action: Springer; 2023. [doi: 10.1007/978-3-031-14436-3_15] [Medline: 37816117]
41. Chavarria EA, Christy SM, Simmons VN, Vadaparampil ST, Gwede CK, Meade CD. Learner verification: a methodology to create suitable education materials. Health Lit Res Pract 2021 Jan 11;5(1):e49-e59. [doi: 10.3928/24748307-20210201-02] [Medline: 34251862]
42. Murphy ST, Frank LB, Chatterjee JS, Baezconde-Garbanati L. Narrative versus non-narrative: the role of identification, transportation and emotion in reducing health disparities. J Commun 2013 Feb;63(1). [doi: 10.1111/jcom.12007] [Medline: 24347679]
43. Kowalski CP, Nevedal AL, Finley EP, et al. Planning for and Assessing Rigor in Rapid Qualitative Analysis (PARRQA): a consensus-based framework for designing, conducting, and reporting. Implement Sci 2024 Oct 11;19(1):71. [doi: 10.1186/s13012-024-01397-1] [Medline: 39394597]
44. Larkey LK, Gonzalez J. Storytelling for promoting colorectal cancer prevention and early detection among Latinos. Patient Educ Couns 2007 Aug;67(3):272-278. [doi: 10.1016/j.pec.2007.04.003] [Medline: 17524595]

XSL•FO
RenderX

45. Gómez-Trillos S, Carrera P, Caballero A, et al. “How Did That Make You Feel?” Latinas’ use of genetic counseling and testing for hereditary cancer risk after watching a culturally targeted video and receiving patient navigation. Psychooncology 2025 Sep;34(9):e70261. [doi: 10.1002/pon.70261] [Medline: 40883500]
46. Borrayo EA, Rosales M, Gonzalez P. Entertainment-education narrative versus nonnarrative interventions to educate and motivate Latinas to engage in mammography screening. Health Educ Behav 2017 Jun;44(3):394-402. [doi: 10.1177/1090198116665624]
47. Lohr AM, Sauver J, O’Byrne JJ, et al. Digital storytelling interventions to promote cancer screening among Hispanic/Latino adults in primary care settings. J Canc Educ 2026 Jun;41(3):553-561. [doi: 10.1007/s13187-025-02696-x] [Medline: 40810886]
48. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
49. Noar SM, Benac CN, Harris MS. Does tailoring matter? meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
50. Nagarkar A, Martin G, Sadaniantz K, et al. Storytelling for health promotion: a scoping review. Am J Health Promot 2026 Feb;40(2):187-209. [doi: 10.1177/08901171251365366] [Medline: 40770144]
51. Isberner MB, Richter T, Schreiner C, Eisenbach Y, Sommer C, Appel M. Empowering stories: transportation into narratives with strong protagonists increases self-related control beliefs. Discourse Process 2019;56(8):575-598. [doi: 10.1080/0163853X.2018.1526032] [Medline: 32104164]
52. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
53. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
54. Ajzen I. Perceived behavioral control, self‐efficacy, locus of control, and the theory of planned behavior. J Applied Soc Pyschol 2002 Apr;32(4):665-683. [doi: 10.1111/j.1559-1816.2002.tb00236.x]
55. Wallerstein N. Empowerment and health: the theory and practice of community change. Community Dev J 1993;28(3):218-227. [doi: 10.1093/cdj/28.3.218]
56. Haughton McNeill L, Wyrwich KW, Brownson RC, Clark EM, Kreuter MW. Individual, social environmental, and physical environmental influences on physical activity among black and white adults: a structural equation analysis. Ann Behav Med 2006 Feb;31(1):36-44. [doi: 10.1207/s15324796abm3101_7] [Medline: 16472037]

Abbreviations

BCC: basal cell carcinoma
MAB: multistakeholder advisory board
SCC: squamous cell carcinoma
UVR: UV radiation

Edited by R Dellavalle; submitted 14.Apr.2026; peer-reviewed by C Strehl, J Volkman; revised version received 16.Jun.2026; accepted 17.Jun.2026; published 05.Aug.2026.

Please cite as:
Perez D, Gómez V, Sedeño-Delgado D, Orellana García C, Chacon M, Gonzalez M, Dash C, Mirabal-Beltran R, Hurtado-de-Mendoza A
Perceptions of a Culturally Tailored Narrative Video for Skin Cancer Prevention Among Spanish-Speaking Hispanic Outdoor Workers: Exploratory Study
JMIR Dermatol 2026;9:e98130
URL: https://derma.jmir.org/2026/1/e98130
doi:10.2196/98130

© Alejandra Hurtado-de-Mendoza, David Perez, Valeria Gómez, Dariana Sedeño-Delgado, Carlos Orellana García, Melida Chacon, Miguel Gonzalez, Chiranjeev Dash, Roxanne Mirabal-Beltran. Originally published in JMIR Dermatology (http://derma.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pre-Post Evaluation of a Sun Safety Social Media Campaign for Young Adults to Reduce Skin Cancer Risk

Janessa M Mendoza[1,2], BS; David Perez[2], BA; Muriel R Statman[1], BA; Mia K Price[1]; Madeleine F Brown[1]; Omar U Anwar[1], BS; Marcelo M Sleiman Jr[1], BA; Kenneth P Tercyak[1,2], PhD

[1]Lombardi Comprehensive Cancer Center, Georgetown University, 2115 Wisconsin Avenue, NW, Washington, DC, United States

[2]School of Medicine, Georgetown University, Washington, DC, United States

Corresponding Author:
Kenneth P Tercyak, PhD
Lombardi Comprehensive Cancer Center, Georgetown University, 2115 Wisconsin Avenue, NW, Washington, DC, United States

Abstract

Background: Social media is a prevalent source of health information for young adults and offers a scalable platform for skin cancer prevention messaging, including sun safety. Despite high awareness among young adults that UV radiation is a potent carcinogen, both intentional and unintentional exposure remain common. Group-targeted and tailored digital campaigns may help counter protanning norms and strengthen protective behaviors to reduce long-term risk.

Objective: This study aimed to evaluate the reach, engagement, and influence of a group-targeted and tailored Instagram-based sun safety campaign implemented as part of a university Skin Smart Campus (SSC) initiative.

Methods: A pre-post design was used during the 2024 - 2025 academic year. The social media campaign evaluation was a focused analysis within a broader sun safety implementation initiative reported elsewhere. The analysis covered the trend-aligned Instagram campaign's reach, engagement, awareness, aided thematic recall, following, and associations with sun safety behaviors. Surveys assessed demographics, sun safety behaviors, campaign awareness, and content recall. Instagram analytics quantified reach and engagement, and influence analyses included bivariate tests and logistic regression.

Results: Presurveys and postsurveys were obtained anonymously from 230 and 267 participants, respectively. A dedicated Instagram account and n=10 campus organizations shared campaign content, extending potential reach to 5711 campus followers. Weekly educational posts averaged 262 views and reached approximately 38 nonfollowers per post throughout the campaign. Postcampaign, SSC social media awareness was reported by 48.7% (130/267) of participants, compared to 24.1% (55/230) precampaign (χ^2_1=32.5; *P*<.001). There was significantly higher awareness among those with lighter skin types (93/164, 56.7%) compared to darker skin types (37/103, 35.9%; χ^2_1=10.9; *P*<.001). Followers were more likely to use on-campus sunscreen dispensers (χ^2_1=85.2; *P*<.001). Sunscreen use, protective clothing, and tanning bed risks were among the most frequently posted and recalled themes on Instagram, suggesting high salience (121/124, 97.6%; 111/124, 89.5%; and 106/124, 85.5%, respectively). Multivariable logistic regression revealed that participants with lighter skin tones (OR 2.33, 95% CI 1.19 - 4.56; *P*=.01), who followed the SSC Instagram account (OR 6.81, 95% CI 3.41 - 13.60; *P*<.001), and who used sunscreen dispensers (OR 7.78, 95% CI 3.82 - 15.84; *P*<.001) were significantly more likely to practice greater sun safety.

Conclusions: A group-targeted, tailored social media campaign embedded in a broader campus initiative demonstrated meaningful reach, engagement, message recall, and impact among young adults. Pairing digital strategies with environmental ultraviolet radiation-reducing resources may enhance young adults' sun safety behaviors. Future efforts should address differences in social media campaign awareness and strengthen messaging for groups that perceive themselves to be at lower risk for skin cancer.

(JMIR Dermatol 2026;9:e90595) doi:10.2196/90595

KEYWORDS

sun safety; social media; young adults; skin cancer prevention; health communication; digital health promotion

Introduction

Social media has rapidly become one of the most influential channels for health communication, offering opportunities to disseminate information with speed, personalization, and interactive engagement that surpass traditional media. Its scale, peer-to-peer connectivity, and algorithm-driven reach make it a particularly powerful platform for health behavior impact [1].

Young adults consistently rely on social media for health information because it is convenient, relatable, and easier to access than formal medical sources [1]. While data from 2022 indicated that over two-thirds of young adults aged 18 to 24

XSL•FO
RenderX

years used social media to learn about their health [2], this proportion has likely increased since then, as digital channels continue to outpace traditional media in speed, personalization, and accessibility. For example, Stifjell et al [3] reported that social media use among young adults outpaced search engine use when seeking health information. Video-based platforms, such as YouTube, Instagram, and TikTok, are especially appealing, as their dynamic and narrative formats facilitate understanding and enhance engagement compared with written materials [2]. Importantly, health information is encountered both actively and passively: many young adults are exposed to algorithm-driven health-related posts by chance on newsfeeds or advertisements, while others join groups or follow pages for specific health topics [2]. Beyond information-seeking, young adults also use social media for peer support and community-building, making digital platforms appear more relatable and accessible than formal medical or public health sources [4]. Thus, digital platforms act as potent tools in shaping young adults' norms, attitudes, and behaviors regarding health-related topics and public health concerns [4].

One such concern is skin cancer prevention, as skin cancer remains the most frequently diagnosed cancer worldwide, despite its largely preventable nature. Approximately 6.1 million individuals are treated annually with an estimated US $8.9 billion in associated annual health care costs in the United States alone [5]. Melanoma, the deadliest form of skin cancer, is one of the most frequently diagnosed cancers among adults aged 20 to 29 years [6,7] and is more frequently diagnosed in women before the age of 50 years [8]. These statistics reinforce the critical importance of skin cancer prevention efforts among young adults.

Because UV radiation (UVR) from sunlight or indoor tanning is a carcinogen, it remains the primary modifiable cause of skin cancer risk [9-12], and targeted, tailored behavioral interventions are critical to reducing long-term exposure and preventing disease [13]. Young women often desire tan skin to improve their appearance or self-confidence, particularly in anticipation of significant social gatherings [14,15]. The ubiquity of this beauty ideal is evidenced by the fact that over 90% of female undergraduates believe they look better with tan skin, with 1 in 5 choosing to indoor tan before special events despite acknowledging the risk of skin cancer [16]. Social media further amplifies these risks by disseminating protanning messages and misinformation about tanning benefits [17-20], consequently fostering favorable attitudes toward tanning and increased tanning behavior among young adults [21,22]. Despite widespread awareness among young adults that excess UVR is a potent skin cancer-causing agent [16,23], 51% of young adults younger than the age of 25 years report annual sunburns, accompanied by inconsistent sunscreen use and reapplication [24]. This pattern is particularly alarming because the relationship between sunburns and skin cancer is strong: even a single sunburn doubles an individual's lifetime risk of developing melanoma, resulting in substantial health care usage and high costs [5,25]. This persistent burden is exacerbated by a behavioral paradox where young adults with higher health literacy exhibit high incidental UVR exposure and low sunscreen adoption [23], highlighting the need for prevention approaches responsive to the real-world social environments where these behaviors occur.

Given these challenges, social media–based campaigns have emerged as promising strategies to counteract protanning messaging and promote sun safety behaviors among young adults. The eHealth behavior change framework emphasizes the importance of theory-driven, audience-tailored health communications [26]. Message tailoring (customizing content to individual attributes such as UVR risk awareness) has been shown to be particularly effective [27]. Combining tailored and targeted approaches, where messages are designed for specific subgroups like young adults, optimizes both resource efficiency and message impact [28]. Trend-based, visual, and memetic formats further enhance engagement and recall among young adults, leveraging perceived relevance, normative influence, and narrative persuasion [28-31].

Empirical evidence supports the effectiveness of social media health campaigns in changing behavior. For instance, Facebook-based interventions have increased condom use [32] and smoking cessation rates [33,34] among young adults. In the context of skin cancer prevention, digital interventions have led to increased sun protection, sunscreen use, and motivation to stop tanning indoors [35,36]. Social media has also served as a motivation for seeking dermatological care, with 31.7% of first-time patients undergoing dermoscopy reporting referral via social platforms [37]. In addition to influencing behavior, social media provides a feasible environment for evaluating the reach and impact of health interventions [38]. Metrics such as likes, views, shares, and impressions are commonly used to assess engagement and audience interaction [38-40]. These analytics offer valuable insights into the effectiveness of digital campaigns and may inform future strategies for public health messaging.

Taken together, evidence from prior research supports the potential of social media–based strategies in promoting sun safety and reducing skin cancer risk among young adults. The broader implementation outcomes of this university's Skin Smart Campus (SSC) initiative, including pre-post changes in sun safety knowledge, sunscreen dispenser use, and overall sun safety behavior, have been reported previously [41]. Building on that implementation evaluation, the present manuscript focuses specifically on the social media component of the SSC initiative. This study evaluates the reach, engagement, awareness, aided thematic recall, and behavioral correlates of a group-tailored, targeted, and theory-driven social media campaign in a university-based young adult population, addressing both the promise and complexity of digital health promotion in this high-risk group. Specifically, this study examines how participants interacted with campaign content, which campaign themes were recalled, how Instagram awareness and following were associated with sun safety behaviors, and how social media metrics can be used to assess the reach and influence of prevention messaging.

The broader SSC initiative included campus designation, educational web materials, small media, sunscreen dispensers, signage, and Instagram-based communication; its implementation outcomes, including knowledge change,

XSL•FO
RenderX

dispenser use, and pre-post sun safety behavior comparisons, are reported elsewhere [41]. The analysis focuses on social media–specific campaign indicators, including Instagram reach and engagement, SSC Instagram awareness, aided thematic recall, and associations between Instagram following and sun safety behaviors.

Methods

Study Design and Setting

This study used an anonymous, independent pre-post design to evaluate the impact of a multicomponent, longitudinal sun safety public health education campaign among young adults. The campaign was conducted at a private university during the 2024‐2025 academic year. The project was part of the institution's designation as an SSC, a national initiative led by the National Council on Skin Cancer Prevention that recognizes schools for their commitment to reducing UVR exposure among young adults [41]. All SSC infrastructure, including some campus sunscreen dispensers, was in place prior to the baseline assessment period. Anonymous, independent presurveys and postsurveys were administered online such that participant overlap could not be determined. The surveys collected data on respondents' sun safety knowledge, behaviors, and campaign awareness over time.

Participants and Recruitment

Eligible participants were aged 18 years or older and enrolled at the university through an undergraduate, graduate, or professional degree program. Recruitment occurred through university-managed social media, tabling events, and campus organization listserv emails. This approach was selected for its feasibility and alignment with evaluating a campus-based public health campaign in a real-world university setting. All survey data were collected using a mobile-friendly online platform (Qualtrics). To support eligibility verification while maintaining anonymity, access to the survey required a university-affiliated email login, and eligibility criteria (age and enrollment status) were self-reported. The study protocol was reviewed and deemed exempt by the university's institutional review board.

Campaign Components

Digital Environment: Group-Targeted and Tailored Social Media Campaign

Overview

The SSC social media campaign was both group-targeted and tailored [42,43]. Group-targeting is designed for a defined subgroup whose members share a characteristic relevant to the behavior or risk context, such as age, race or ethnicity, occupation, or disease risk; the same core content is typically delivered to everyone in that subgroup [43]. Relatedly, group-tailoring goes beyond selecting a subgroup audience and adapts content, delivery, framing, examples, activities, or implementation strategies to the group's assessed needs, shared identity, cultural context, barriers, resources, and mechanisms of change [42]. When used together, group targeting and group tailoring can be synergistic: targeting helps ensure that the campaign reaches a subgroup with shared needs, while tailoring increases its fit, relevance, and feasibility for that group's real-world context [44]. Campaigns that both identify an appropriate priority group and adapt strategies to that group's circumstances may have stronger effects than those that rely on targeting or tailoring alone, particularly when behavior change depends on social norms, environmental supports, and setting-specific barriers [42-44]. This distinction is important in skin cancer prevention because outdoor settings are a priority context for reducing UVR exposure, and multicomponent approaches that combine health education, behavioral supports, environmental changes, and policies have strong evidence for improving sun safety behaviors [45].

This study reports the social media–specific tailoring process at the level needed to interpret campaign content and engagement. Detailed implementation outcomes for the broader SSC initiative are reported elsewhere [41]. Briefly, baseline survey data identified communication channels used by students and highlighted sun safety knowledge and behavior gaps relevant to campaign planning, including misconceptions about UVR exposure from tanning beds, inconsistent sunscreen use and reapplication, and infrequent use of protective clothing and hats. These findings informed the selection of campaign themes and the emphasis placed on sunscreen use, sunscreen reapplication, tanning bed risks, sun-protective clothing, and UVR risk across skin tones.

Precampaign Planning and Platform Selection

Precampaign (baseline) data on participants' preferred communication channels, as well as their sun safety knowledge and behaviors, guided the group-tailored development of the SSC social media campaign to align with participants' social media usage patterns and educational needs. Instagram was selected because it is one of the most commonly used social media platforms in the United States and is particularly well-suited to short-form video content. Although YouTube is the most widely used platform among US adults, Instagram is considered especially behaviorally influential due to its highly visual interface and the ease with which users can continuously scroll through short-form content [46].

Baseline survey findings also informed the educational priorities of the campaign. For example, at baseline and as described in more detail by Perez et al [41], nearly one-fifth of respondents (18.9%) incorrectly believed that the sun emits more UVR than tanning beds. In addition, 15% reported never or rarely using sunscreen with SPF 30 or higher, 53.5% reported never or rarely reapplying sunscreen appropriately, 49.2% reported never or rarely wearing hats, and 34.2% reported never or rarely wearing clothing that covers the arms or legs. These findings informed campaign messaging emphasizing tanning bed risks, sunscreen use and reapplication, and sun-protective clothing [41]. Because these knowledge and behavior outcomes are reported previously, they are summarized here only to describe how baseline data informed the SSC Instagram campaign.

Campaign Implementation

Campaign content was thus designed to reflect participants' daily routines, lived experiences, and common misconceptions about UVR exposure. Posts were intentionally aligned with current social media consumption patterns among young adults,

XSL•FO
RenderX

incorporating concise captions and text overlays, humor-based and meme-style graphics, and short-form videos inspired by contemporary social media trends to enhance relatability, credibility, and algorithmic reach.

At the start of the fall semester and continuing through the spring semester, a total of 62 Instagram posts were shared, including 43 educational posts and 19 outreach posts. Outreach posts promoted SSC-related events, including tabling sessions designed to encourage survey participation and distribute sun safety materials (eg, educational handouts, UV bracelets, and sunscreen lip balm). Posts were distributed weekly, with increased frequency to twice weekly during relatively higher UVR months (August to October) and before and during the university's spring break week in March.

Peer Network Dissemination Strategy

To expand the reach of the social media campaign beyond the SSC Instagram account and its followers, a structured peer network dissemination strategy was implemented. A total of 85 outreach emails were sent to individuals involved in 75 (n=34 campus clubs and n=41 athletics-affiliated groups) unique academic organizations, cultural clubs, athletic teams, performance groups, and affinity groups registered with the university. Because these organizations had established digital networks, they were invited to disseminate SSC Instagram content through their own platforms. Organizations that shared SSC Instagram content were tracked using Instagram's "Story Reshares" and "Mentions" analytics, as well as manual logs of confirmed reposts.

Retrospective Content Coding

To systematically characterize the delivered campaign content, all SSC Instagram posts were retrospectively coded using a content matrix developed by the study team. Posts were individually reviewed and classified according to the dominant theme (eg, sunscreen use, sun-protective clothing, winter sun safety, skin cancer across skin tones, and indoor tanning risk) and format (post, 24-hour story, or reel). Posts could contain more than 1 theme; therefore, theme counts were not mutually exclusive. The coding approach was designed to support the present manuscript's social media evaluation by linking delivered campaign themes to Instagram engagement metrics and postcampaign aided thematic recall. Creative executions and message-framing examples, such as humor, trend-aligned audio, or appearance-related depictions of sun damage, were considered during interpretation but were not treated as separate quantitative outcome categories. Two independent coders reviewed all content, and interrater reliability was assessed using percent agreement. Discrepancies were resolved through discussion and consensus with a third reviewer. This systematic coding approach enabled subsequent evaluation of participant content recall, which served as a dependent variable in the postcampaign survey.

On-Campus Sunscreen Dispensers and Signage

Social media posts also promoted other components of the university's SSC designation, including on-campus sunscreen dispensers, signage, and online educational materials. As part of the university's SSC designation, skin cancer prevention information was included on the student health center's website, and related studies were published in the student newspaper and a campus newsletter. These materials highlighted the four publicly accessible on-campus sunscreen dispensers installed in high-traffic indoor locations (eg, dining areas) to increase access to UVR protection. These installations occurred prior to the deployment of the presurvey and before the start of the social media campaign. Given the high levels of sun exposure associated with physical activity and exercise, 3 additional sunscreen dispensers were installed at athletic facilities during the campaign period, increasing the total number of dispensers to seven. Educational signage accompanied the dispensers and was displayed throughout campus, including residence halls and academic buildings, to reinforce daily sunscreen use and promote year-round sun safety. This information was also clearly shared on the SSC Instagram account. Additional implementation details, including dispenser placement, are described by Perez et al [41].

Campaign Evaluation

Independent Variables

Demographics

In both presurveys and postsurveys, demographics were collected, including age, gender, race or ethnicity, and program affiliation (undergraduate, graduate, or professional).

Sunscreen Dispenser Use

Participants were asked whether they had ever used an on-campus sunscreen dispenser, with responses coded dichotomously (yes or no) in presurveys and postsurveys.

Sun Safety Behaviors

Sun safety behaviors were assessed in presurveys and postsurveys using a 7-item composite measure, rated on a 5-point Likert scale (0="never" to 4="always"), adapted from prior research [47]. The validation, reliability, and limitations of these measures within the SSC initiative were previously established and are described in detail elsewhere [41]. Cronbach α indicated acceptable internal consistency for the scale, with an overall value of 0.70 (0.57 at presurvey, 0.83 at postsurvey). Items included sunscreen use and its reapplication, protective clothing use, hat use, seeking shade, and checking the UV index. Higher summed scores indicated greater engagement in sun-protective behaviors.

Skin Type

In the postsurvey only, information was collected on young adults' Fitzpatrick Scale skin type ranging from I to VI, based on a person's self-reported tendency to sunburn and tan; higher values correspond to darker skin tones and a lesser tendency to sunburn [48,49].

Athletic Facility Usage

Given the locations of the new sunscreen dispensers at athletic facilities, participants reported the frequency of visiting campus athletic facilities, which was coded categorically on a 5-point Likert scale (0="never" to 4="always") in the postsurvey.

XSL•FO
RenderX

Dependent Variables

Instagram Metrics

Engagement with the SSC Instagram account was evaluated using metrics obtained from Instagram Insights, a built-in analytics tool available to users with a business or creator account. This tool provided data on how the posted content and audience were performing to optimize the SSC Instagram account's reach (exposure) and impact (interaction). Analytics included total followers and nonfollowers who engaged with content, likes per-post, views or impressions (including repeat views), and shares of content. Based on prior literature on social media health promotion campaigns [38-40], likes and views were prioritized as primary engagement indicators, with shares or comments monitored but anticipated to be less frequent in this context. These metrics reflect platform-defined measures of reach and engagement that are generated through Instagram's proprietary algorithms and reporting structures. Metrics are provided in aggregate form and summarize interactions across followers and nonfollowers. As such, these measures characterize overall content performance and audience interaction patterns at the account level rather than tracking exposure at the individual user level.

SSC and Instagram Awareness

In the precampaign survey, participants were asked about their awareness of the university's SSC initiative. In contrast, the postcampaign survey distinguished between general awareness of the initiative and awareness of the public SSC Instagram specifically; the latter was assessed using 2 items: participants were first asked whether they were aware of the university's SSC social media campaign on Instagram (yes or no) and, if so, whether they followed the public account (yes or no). SSC Instagram awareness served as the primary indicator of campaign reach. Because the SSC Instagram-specific following item was administered only postcampaign, baseline Instagram following could not be assessed.

Content Recall

In the postsurvey, participants who reported awareness of the SSC Instagram account were presented with a list of campaign content themes and were asked to select the themes they recalled seeing. This measure assessed aided thematic recall rather than unaided recall of specific posts, captions, or graphics. Aided recall was included as a social media–specific indicator of message salience and campaign penetration, particularly because Instagram analytics were available only in aggregate form and could not be linked to individual survey respondents.

Data Analysis

Cross-sectional postsurvey data were analyzed using SPSS version 29.0. Descriptive statistics were used to summarize demographic and behavioral characteristics. Chi-square tests assessed categorical variables (eg, dispenser use by Instagram followers). Independent samples *t* tests compared sun safety behaviors across other factors (eg, skin tone, Instagram following, and dispenser use). Finally, logistic regression analysis was used to identify cofactors of sun safety. The model's independent variables were selected based on their potential relevance to UVR risk perception, previous evidence that social media health campaigns impact cancer-preventive behaviors, and bivariate associations observed within the postsurvey dataset ($P<.10$) [50].

Ethical Considerations

This study was exempt from review by the Georgetown University Institutional Review Board (STUDY00008246). Informed consent was obtained from all participants involved in the study. The study protocol was reviewed and deemed exempt by the university's Institutional Review Board. All data were collected and stored in compliance with the Health Insurance Portability and Accountability Act (HIPAA).

Results

Peer Network Campaign Outreach

In fall 2024, the university enrolled approximately 7800 undergraduate and 8900 graduate or professional students, totaling an estimated 16,700 students. Of the 85 outreach emails sent to 75 unique university organizations, ultimately 10 with established Instagram accounts reposted SSC Instagram content. This extended the potential reach to 5711 Instagram followers, or roughly one-third of the student body.

Participant Characteristics

A total of 230 participants completed the precampaign survey and 267 completed the postcampaign survey (Table 1). Due to the anonymous and independent nature of the surveys, some participant overlap may have occurred. As illustrated in Table 1, presurvey and postsurvey demographic characteristics were comparable [41]; they are presented to help characterize the samples and provide context for the social media–focused analyses.

XSL•FO
RenderX

Table . Participant demographic data.

Factor	Precampaign (n=230)[a]	Postcampaign (n=267)[a]
Age (y), mean (SD)	23.7 (3.1)	23.3 (3.0)
Sex, n (%)		
Female	160 (69.6)	176 (65.9)
Male	68 (29.6)	91 (34.1)
Race, n (%)		
Asian or Pacific Islander	76 (33.0)	80 (30.0)
Black or African American	21 (9.1)	17 (6.4)
Middle Eastern	22 (9.6)	22 (8.2)
Other	32 (13.9)	27 (10.1)
White	79 (34.3)	121 (45.3)
Ethnicity, n (%)		
Non-Hispanic	212 (92.2)	242 (90.6)
Hispanic	18 (7.8)	25 (9.4)
Skin type[b], n (%)		
I (“Very fair”)	—[c]	70 (26.2)
II (“Fair”)	—	94 (35.2)
III (“Light-medium”)	—	60 (22.5)
IV (“Medium/olive”)	—	24 (9.0)
V (“Brown”)	—	15 (5.6)
VI (“Deep brown or black”)	—	4 (1.5)

[a]These samples are presumed to be independent, but cannot be entirely verified.

[b]Assessed at postcampaign only.

[c]Not applicable.

Precampaign participants had a mean age of 23.7 (SD 3.1), were predominantly female (160/230, 69.6%), and represented a diverse racial composition (Asian or Pacific Islander: 76/230, 33.0%; Black or African American: 21/230, 9.1%; Middle Eastern: 22/230, 9.6%; Hispanic: 18/230, 7.8%). More than half of the participants were enrolled in a health profession program (150/230, 65.2%). The majority reported keeping up with university news through emails (199/230, 86.5%) or university-managed social media accounts (107/230, 46.5%), highlighting the prominence of digital communication platforms in reaching this demographic.

Postcampaign demographics were comparable to those observed precampaign, with a similar average age of 23.3 (SD 3.0) and a predominantly female sample, consistent with the precampaign cohort (176/267, 65.9%; Table 1). Racial diversity was similar to that of the precampaign sample, and both samples included fewer than 10% Hispanic respondents. As in the precampaign group, more than half of participants were enrolled in a health professions program (146/267, 54.7%).

In the postcampaign sample only, athletic facility usage and skin type distribution were measured. Athletic facility usage was common, with 89.1% (238/267) reportedly visiting to some extent, as defined by selecting any response other than “never” on the 5-point Likert scale. Of these participants, more than half reported visiting very frequently, as indicated by selecting “often” (n=98) and “always” (n=32). This finding supported the strategic placement of additional sunscreen dispensers at athletic facilities across campus.

Skin type distribution in the postsurvey sample showed that 61.4% (164/267) identified as types I-II (lighter skin types) and 38.6% (103/267) as types III-VI (darker skin types). Skin type demonstrated clear variation across racial groups: among White participants (n=121), nearly all reported lighter skin tones, with 49.6% (60/121) identifying as type I; 38.0% (46/121) as type II; 12.4% (15/121) as type III; and none identifying above type III. Asian American or Pacific Islander participants (n=80) showed a broader distribution, including types II (36/80, 45.0%), III (24/80, 30.0%), and IV-V (17/80, 15.3%). The majority of Black or African American participants (n=17) reported darker skin tones, with 52.9% (9/17) identifying as type V and 23.5% (4/17) as type VI. Middle Eastern participants (n=22) predominantly reported intermediate tones (type II: 5/22, 22.7%; type III: 11/22, 50.0%). These distributions align with established patterns between race and skin phototype and support the validity of the self-reported classification within this cohort.

XSL•FO
RenderX

SSC Instagram Engagement (Instagram Insights)

Engagement metrics are summarized in Table 2. Among educational content, 24 out of 43 (55.8%) posts integrated multiple themes within a single post. Sunscreen use was the most frequently represented topic, appearing in 69.8% (30/43) of all educational posts. Educational content averaged 262.2 (493.6) views and 7 (13.3) likes per-post, compared with 214.1 (270.2) views and 5.4 (7.7) likes for outreach content. On average, 14.5% (26.5%) of viewers of educational posts were nonfollowers, indicating that each educational post reached approximately 38 new users beyond the account's existing audience. The highest performing post was a reel showing photos of an on-campus sunscreen dispenser, personified by a facial filter with isolated eyes and lips. This was subsequently followed by photos of sun-damaged skin on the face and neck overlaying popular audio saying, "Nope, I'm good, thank you," to encourage sunscreen use. This reel generated 2947 views (nonfollowers: 2720/2947, 92%) and received n=66 likes, substantially surpassing other engagement methods.

Table . Social media engagement metrics by type.

Post type or theme	Post, n	Views, mean (SD)	Likes, mean (SD)	Nonfollowers (%), mean (SD)
Educational[a]	43	262.2 (493.6)	7.0 (13.3)	14.5 (26.5)
Wearing sunscreen	30	308.1 (586.9)	8.5 (15.4)	18.5 (30.8)
Protective clothing	16	98.3 (52.6)	1.6 (1.6)	3.6 (7.5)
Sun safety in winter	16	155.1 (155.7)	4.9 (9.6)	14.2 (26.4)
Tanning beds	12	149.1 (122.6)	2.8 (3.9)	4.8 (8.2)
Skin cancer by skin type	6	172.2 (133.1)	4.7 (5.5)	6.2 (10.2)
Outreach	19	214.1 (270.2)	5.4 (7.7)	9.9 (21.0)
Total	62	279.5 (226.0)	7.9 (6.2)	15.5 (17.1)

[a]Educational posts do not sum to n=43 due to the coding of multiple themes.

SSC Instagram Awareness

At baseline, 24.1% (55/230) of precampaign participants reported awareness of the university's SSC initiative. After the campaign, postsurvey data revealed that 48.7% (130/267) of participants reported awareness of the SSC Instagram account, demonstrating a significant increase in visibility (χ^2_1=32.5; P<.001). Among those who were aware, the majority (113/130, 86.9%) indicated that they followed the account. SSC Instagram awareness varied significantly by skin type: 56.7% (93/164) of participants with skin types I-II were aware of the SSC Instagram compared with 35.9% (37/103) of participants with types III-VI (χ^2_1=10.9; P<.001). Participants who followed the SSC Instagram account were also substantially more likely to report sunscreen dispenser use (χ^2_1=85.2; P<.001); the majority of followers (102/113, 90.3%) had used a dispenser.

Content Recall

Among those who indicated aided recall of specific themes from the SSC Instagram account (n=124), 539 total recall selections were recorded across 5 educational themes. Sunscreen use was the most frequently recalled theme (121/124, 97.6%), followed by protective clothing (111/124, 89.5%), tanning bed risks (106/124, 85.5%), skin cancer by skin tone (104/124, 83.9%), and winter sun safety (97/124, 78.2%). Sunscreen was also the most frequently posted theme (30/43, 69.7% of educational posts), followed by protective clothing (16/43, 37.2% of educational posts). Notably, recall patterns did not strictly mirror posting frequency. For example, tanning bed risks appeared in only 27.9% (12/43) of education-focused posts but were recalled by 85.5% (106/124) of participants, while skin cancer by skin tone appeared in 14% (6/43) of educational posts yet achieved 83.9% (104/124) recall. These results suggest that tanning bed risks and skin cancer by skin tone may have been especially salient themes among respondents who were aware of the SSC Instagram account. However, because recall was assessed using an aided theme list and only among respondents aware of the account, these findings should be interpreted as indicators of theme salience rather than as direct evidence of knowledge gain. Winter sun safety was the third most frequently posted educational theme (16/43, 37.2% of educational posts) and was less frequently recalled (97/124, 78.2%), suggesting limited resonance.

Sun Safety Behaviors by SSC Instagram Following

There were no demographic differences between those who did and did not follow the SSC Instagram account, with the exception of skin tone: followers were more likely to report lighter (types I-II) skin tones (χ^2_1=10.9; P<.001). Respondents who reported following the social media account (113/267, 42.3%) also reported significantly higher sun safety behaviors compared to nonfollowers (Table 3). SSC Instagram followers reported greater sunscreen use (mean 4.1, SD 0.8 vs mean 3.3, SD 1.1; P<.001), more protective clothing use (mean 3.5, SD 0.8 vs mean 3.2, SD 1.0; P=.007), increased shade-seeking behaviors (mean 3.8, SD 0.8 vs mean 3.1, SD 1.1; P<.001), and more frequent UV index checking (mean 2.8, SD 0.8 vs mean 2.2, SD 1.2; P<.001). The composite sun safety behavior score was also significantly higher among followers than nonfollowers (mean 24.9, SD 4.5 vs mean 18.5, SD 4.7; P<.001).

XSL•FO
RenderX

Table . Sun safety behaviors by Skin Smart Campus (SSC) Instagram following status among postsurvey respondents (n=267).

Behavior	Instagram followers (n=113), mean (SD)	Nonfollowers (n=154), mean (SD)	*t* test (*df*)	*P* value
Sunscreen use	4.1 (0.8)	3.3 (1.1)	6.02 (265)	<.001
Protective clothing	3.5 (0.8)	3.2 (1.0)	2.73 (265)	.007
Shade seeking	3.8 (0.8)	3.2 (1.0)	5.53 (265)	<.001
UV Index checking	2.8 (0.8)	2.2 (1.2)	5.21 (265)	<.001
Total	24.9 (4.5)	18.5 (4.7)	11.12 (265)	<.001

In a multivariable logistic regression analysis adjusting for skin tone (odds ratio [OR] 2.33, 95% CI 1.19 - 4.56; *P*<.001) and sunscreen dispenser use (OR 7.78, 95% CI 3.82 - 15.84; *P*<.001), following the SSC Instagram account remained a strong independent factor associated with higher postcampaign sun safety behavior (Table 4). Participants who checked or followed the SSC Instagram account were over 6 times more likely to have a higher sun safety behavior composite score compared to nonfollowers (OR 6.81, 95% CI 3.41 - 13.60; *P*<.001). This finding should be interpreted as an association rather than evidence of causal impact, as participants who were already more interested in sun safety may have been more likely to follow the SSC Instagram account.

Table . Multivariable logistic regression of sun safety behavior at posttest (N=267)[a].

Factor	OR[b] (95% CI)	*P* value
Skin tone (lighter vs darker[c])	2.33 (1.19 - 4.56)	.01
Sunscreen dispenser use (yes vs no[c])	7.78 (3.82 - 15.84)	<.001
SSC[d] Instagram following (yes vs no[c])	6.81 (3.41 - 13.60)	<.001

[a]Sun safety behavior composite scores range from 7 - 35. High sun safety behavior is defined as a composite score >median value (22).

[b]OR: odds ratio.

[c]Model referent.

[d]SSC: Skin Smart Campus.

Discussion

This study evaluated a social media campaign implemented as part of a multicomponent sun safety initiative on a university campus, with a focus on how group-targeted and tailored digital communication shaped engagement, awareness, message recall, and early behavioral indicators among young adults. The broader implementation outcomes of the SSC initiative, including pre-post changes in knowledge, dispenser use, and overall sun safety behavior, were reported by Perez et al [41]. The analysis extends that work by focusing on the Instagram campaign's reach, engagement, aided thematic recall, and associations between Instagram following and sun safety behaviors. This anonymous pre-post analysis concentrated on how message design, theme selection, and platform-specific strategies influenced campaign reach and resonance within the young adult population.

Principal Findings

Consistent with prior work demonstrating the central role of social media as a source of health information among young adults [2,3,36], the SSC Instagram achieved substantial visibility. Trend-based posts helped embed sun protection behaviors within campus culture, increase perceived social norms around sunscreen use, and foster peer-to-peer diffusion of knowledge through humor and familiarity with current digital trends. This creative approach aimed to normalize daily sunscreen use and UVR protective practices by aligning them with the aesthetic and cultural norms of the digital environment. Rather than viewing sun safety in isolation, the campaign leveraged young adults' focus on appearance—such as by highlighting the aesthetic impact of sun damage—to reframe protective practices as culturally-relevant health behaviors.

The peer network dissemination strategy adequately leveraged trusted peer groups to boost message reach among young adults unlikely to encounter institutional channels. Campus organizations that reposted SSC Instagram content had a potential digital reach equivalent to roughly one-third of the student body. Additionally, Instagram analytics revealed that educational posts routinely reached Instagram users who did not follow the SSC account, suggesting that peer network dissemination effectively expanded exposure beyond the immediate audience. This pattern is notable given that campaigns relying solely on organic reach often struggle to penetrate beyond existing followers [38,39].

At the platform level, exposure among Instagram users who did not follow the SSC account remained a meaningful contributor to reach, with 14.5% (38/262.2) of total impressions generated by users not subscribed to the account, on average. Shares and comments were infrequent, consistent with prior young adult-focused health promotion research [38,39]. However, views and likes demonstrated consistent performance across themes. Together, these analytics indicate that the SSC

XSL•FO
RenderX

Instagram campaign achieved broad algorithmic circulation and engagement.

The high performance of sunscreen-focused reels, especially those incorporating humor and trend-based formats, highlights the importance of aligning prevention messaging with the aesthetic and cultural norms of the young adult digital environment. This environment is uniquely characterized by identity exploration in public spaces, high social visibility and constant comparison, and blended online and offline personas. Moreover, it has been described as an algorithm-shaped experience, with participatory creativity and interactivity, and one that is influenced by rapid cultural cycles and community-seeking across real and digital borders [51]. Young adult health engagement on social media is also often shaped by influencer-driven norms and intuitive assessments of credibility rather than formal expertise [2]. These dynamics can amplify both positive and negative health influences, emphasizing the need for group-tailored, trend-aligned, and inclusive messaging. The SSC Instagram attempted to leverage these characteristics to increase campaign reach and appears to have successfully done so within a university campus environment.

Content recall patterns provide insight into the salience of SSC Instagram themes among respondents who were aware of the account. Because recall was assessed using a list, these findings should be interpreted as aided thematic recall rather than unaided memory for specific posts. Although sunscreen was the most frequently posted topic, high levels of aided recall were also observed for tanning bed risks and skin cancer by skin tone, which were posted less often but may have had strong salience for this audience. These findings suggest that message novelty, perceived personal relevance, and the counter-messaging nature of content (eg, correcting misconceptions about risk across skin tones) may exert greater influence on recall than posting frequency alone. This pattern may also reflect the comparatively low baseline knowledge observed for tanning bed risks and skin cancer susceptibility across skin tones in the precampaign survey, suggesting that participants may have been more likely to notice and remember content that addressed these existing knowledge gaps. Importantly, pre-post knowledge outcomes for the broader SSC initiative are reported elsewhere [41]; the present findings focus on social media–specific recall and engagement only. While direct comparative evidence evaluating message characteristics versus posting frequency on recall is limited, prior literature demonstrates that message tailoring, narrative framing, and counter-messaging can improve salience and resonance [2,3,27,52-54].

The low recall of winter sun safety content, despite being one of the more frequently posted themes, underscores limitations in seasonal messaging for UVR prevention. Young adults may underestimate winter UVR exposure, which is consistent with previous literature showing lower perceived risk during colder months [55,56]. These findings suggest that future campaigns may need to intensify or diversify winter messaging modalities (eg, in-person prompts, location-based mobile reminders) to overcome entrenched assumptions about seasonal risk.

Patterns of campaign awareness also reveal important disparities. Young adults with lighter skin tones were significantly more likely to be aware of the SSC Instagram account, consistent with evidence that individuals with darker skin tones often perceive themselves to be at lower risk for skin cancer and are less likely to seek UVR-related health information [57,58]. This gap has substantive equity implications: without tailored efforts to engage young adults across diverse skin tones, digital prevention strategies may inadvertently reinforce disparities in UVR protection behaviors. The strong recall of the “skin cancer by skin tone” theme suggests that inclusive, identity-relevant messaging is not only feasible but impactful and should be expanded in future iterations of the campaign.

The integration of digital messaging with environmental supports also appeared meaningful, specifically with the installation of seven sunscreen dispensers. Young adults who followed the SSC Instagram account and those who visited athletic centers were more likely to use campus sunscreen dispensers, illustrating how digital and built environments can mutually reinforce one another. Similar synergistic effects have been observed in other public health interventions, such as campus-based human papillomavirus or influenza vaccination campaigns that paired social media outreach with on-site clinics, resulting in increased vaccine uptake [59,60]. Further research is warranted to elucidate the value of integrating digital and physical components to maximize healthy behavior change. Still, this multicomponent approach aligns with ecological models of behavior change, which emphasize the interaction of personal, social, and environmental factors in shaping preventive health behaviors [61].

Several implications emerge from this evaluation. First, social media campaigns targeting young adults benefit substantially from high-quality tailoring and trend-aligned design, which increase relevance and engagement. Second, theme selection should not rely solely on frequency; topics that address misconceptions, highlight disparities, or provide counter-narratives may yield disproportionately strong recall. Third, equitable campaign impact requires intentional efforts to reach populations with historically lower UVR-related risk perception, including young adults with darker skin tones. Finally, pairing digital campaigns with accessible environmental resources, such as on-campus sunscreen dispensers, may enhance both awareness and behavioral uptake.

Limitations

This study has several limitations that should be kept in mind when interpreting the findings. The anonymous, independent-samples design prevents linking participants across time and limits inferences about within-person change. Moreover, participant overlap cannot be determined. The nonexperimental design further constrains causal interpretation because exposure was not manipulated and participants were not randomly assigned to conditions; accordingly, findings should be interpreted as descriptive and associative. The observed association between Instagram account engagement and sun safety behaviors may also reflect reverse causality, such that individuals already more engaged in sun protection were more likely to seek out or attend to campaign content, and may

XSL•FO
RenderX

be influenced by common method variance given that both exposure-related and behavior-related variables were self-reported, which can inflate observed associations. More broadly, self-reported survey data are subject to recall and social desirability bias. In addition, Instagram analytics provide aggregate rather than individual-level metrics, limiting the precision of linking digital exposure to behavioral outcomes, and are influenced by platform algorithms and privacy protections that shape how content is distributed and which users are reached; as a result, engagement indicators such as likes and views may not fully capture passive exposure or reflect the true extent of audience reach. The study was also conducted at a single urban university, which may limit generalizability to other campus environments or institutional settings, and Fitzpatrick skin type was collected only in the postcampaign survey. Thus, it was not possible to evaluate pre-post sample comparability or change by skin type.

Nonetheless, this study contributes an important perspective to the literature on digital UVR prevention strategies by demonstrating how social media can function as a high-reach, low-cost, and culturally embedded medium for public health messaging among young adults. As preventive medicine increasingly adapts to digital platforms, these findings highlight the value of targeted, inclusive, and visually engaging communication strategies that resonate with contemporary media habits while supporting equitable access to cancer-preventive behaviors.

Conclusions

Taken together, this social media–focused SSC evaluation illustrates the promise and ongoing challenges of delivering sun safety campaigns within the dynamic digital landscape that shapes young adult communication and identity. By aligning message design with platform norms, cultural aesthetics, and the interaction patterns that characterize young adult social media use, the SSC Instagram campaign achieved meaningful reach, strong aided thematic recall, and tangible associations with early behavioral indicators. At the same time, our findings underscore the need for continued attention to equity, seasonal risk perception, and the integration of digital and environmental supports in order to strengthen the impact of future iterations. As universities increasingly look to social platforms as core channels for health promotion, this research demonstrates how thoughtful, audience-centered design can translate public health goals into strategies that resonate within the fast-moving and socially mediated environments of young adults.

Acknowledgments

The authors would like to express their sincere appreciation to the Office of the Vice President for Student Affairs within the Division of Student Affairs at Georgetown University, with special recognition to Kathryn Castle, PhD, and to the Dean for Medical Education at the School of Medicine, formerly Lee Jones, MD, for their support in establishing the campus as a Skin Smart Campus (SSC) and for their valuable contributions to this initiative. The authors give their heartfelt thanks to Elissa Fairbrother, Recruitment Coordinator for the National Council on Skin Cancer Prevention's SSC initiative, whose dedicated guidance and coordination were instrumental in helping each campus achieve SSC designation. The authors also wish to acknowledge Susan Sell-Haynes, Executive Director of Campus Recreation; Lee Reed, Director of Intercollegiate Athletics; and Andreya Thibault, Associate Athletic Director for Student-Athlete Health and Wellness, for their efforts in supporting the placement of sunscreen dispensers at Yates Field House and Thompson Athletic Center. Finally, the authors are grateful to Sarah W Bu, Ben W Gratz, Nesreen Shahrour, Shivali N Vora, and Talia Zamir for their valuable assistance with data collection. The authors attest that there was no use of generative artificial intelligence in the generation of text or figures contained in this manuscript.

Funding

This research was supported by a National Institutes of Health/National Cancer Institute award to Georgetown University Lombardi Comprehensive Cancer Center (CA051008).

Data Availability

Data for this study are available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: JMM, DP, MRS, KPT
Data curation: JMM, MRS, MKP, MFB, KPT
Formal analysis: JMM, MRS, KPT
Funding acquisition: KPT
Investigation: JMM, DP, MRS, MKP, MFB, OUA, MMS, KPT
Methodology: JMM, DP, MRS, KPT
Project administration: KPT
Resources: JMM, KPT
Software: JMM, MKP, MRS, KPT
Supervision: KPT
Validation: JMM, MRS, KPT

XSL•FO
RenderX

Visualization: JMM, MRS, KPT
Writing – original draft: JMM, DP, MRS, MKP, MFB, OUA, MMS, KPT
Writing – review & editing: JMM, DP, MRS, MKP, MFB, OUA, MMS, KPT

Conflicts of Interest

None declared.

References

1. Wu P, Feng R. Social media and health: emerging trends and future directions for research on young adults. Int J Environ Res Public Health 2021 Jul 31;18(15):8141. [doi: 10.3390/ijerph18158141] [Medline: 34360431]
2. Lim MSC, Molenaar A, Brennan L, Reid M, McCaffrey T. Young adults' use of different social media platforms for health information: insights from web-based conversations. J Med Internet Res 2022 Jan 18;24(1):e23656. [doi: 10.2196/23656] [Medline: 35040796]
3. Stifjell K, Sandanger TM, Wien C. Exploring online health information-seeking behavior among young adults: scoping review. J Med Internet Res 2025 Sep 9;27(1):e70379. [doi: 10.2196/70379] [Medline: 40925001]
4. Nagata JM, Huang O, Hur JO, et al. Health benefits of social media use in adolescents and young adults. Curr Pediatr Rep 2024;13(1):22. [doi: 10.1007/s40124-025-00357-7] [Medline: 40823382]
5. Health and economic benefits of skin cancer interventions. National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP). URL: https://www.cdc.gov/nccdphp/priorities/skin-cancer.html [accessed 2025-12-11]
6. Key statistics for melanoma skin cancer. American Cancer Society. URL: https://www.cancer.org/cancer/types/melanoma-skin-cancer/about/key-statistics.html [accessed 2025-12-11]
7. Adolescents and young adults with cancer. National Cancer Institute (NCI), National Institutes of Health (NIH). URL: https://www.cancer.gov/types/aya [accessed 2025-12-11]
8. Cancer facts & figures 2025. American Cancer Society. URL: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2025-cancer-facts-figures.html [accessed 2025-12-11]
9. Lauters R, Brown AD, Harrington KCA. Melanoma: diagnosis and treatment. Am Fam Physician 2024 Oct;110(4):367-377. [Medline: 39418569]
10. Tran MM, George-Washburn EA, Rhee J, Li WQ, Qureshi A, Cho E. A prospective cohort study exploring the joint influence of sunlight exposure and tanning bed use on basal cell carcinoma, squamous cell carcinoma, and melanoma risk. Arch Dermatol Res 2024 May 25;316(6):281. [doi: 10.1007/s00403-024-03132-8] [Medline: 38796657]
11. Joshi UM, Kashani-Sabet M, Kirkwood JM. Cutaneous melanoma: a review. JAMA 2025 Dec 16;334(23):2113-2125. [doi: 10.1001/jama.2025.13074] [Medline: 40853557]
12. Wehner MR. Keratinocyte carcinoma: a review. JAMA 2026 Jan 6;335(1):70-81. [doi: 10.1001/jama.2025.18749] [Medline: 41165676]
13. Skin cancer. American Academy of Dermatology (AAD). URL: https://www.aad.org/media/stats-skin-cancer [accessed 2025-12-11]
14. Seidenberg AB, Julian AK, Hartman AM, Holman DM. Sunless and indoor tanning among U.S. non-Hispanic White women ages 18-49 years. J Dermatol Nurses Assoc 2023;15(3):123-132. [doi: 10.1097/jdn.0000000000000736] [Medline: 38463180]
15. Trepanowski N, Huang L, Hartman RI. Trends in indoor tanning frequency: data from the National Health Interview Survey 2005-2015. Int J Dermatol 2023 Sep;62(9):e506-e509. [doi: 10.1111/ijd.16656] [Medline: 36999738]
16. Newton J, Ochoa L, Reinschmidt A, et al. Is beauty worth the risk? Self-confidence is the key motivating factor driving tanning bed use among undergraduate students at South Dakota universities. Int J Womens Dermatol 2024;10(2):e128. [doi: 10.1097/JW9.0000000000000128] [Medline: 38572264]
17. Waring ME, Baker K, Peluso A, May CN, Pagoto SL. Content analysis of Twitter chatter about indoor tanning. Transl Behav Med 2019 Jan 1;9(1):41-47. [doi: 10.1093/tbm/iby011] [Medline: 29474700]
18. Groshon L, Waring ME, Blashill AJ, et al. A content analysis of indoor tanning Twitter chatter during COVID-19 shutdowns: cross-sectional qualitative study. JMIR Dermatol 2024 Mar 4;7:e54052. [doi: 10.2196/54052] [Medline: 38437006]
19. Ricklefs CA, Asdigian NL, Kalra HL, et al. Indoor tanning promotions on social media in six US cities #UVTanning #tanning. Transl Behav Med 2016 Jun;6(2):260-270. [doi: 10.1007/s13142-015-0378-0] [Medline: 27356996]
20. Moreno MA, Jenkins MC, Lazovich D. Tanning misinformation posted by businesses on social media and related perceptions of adolescent and young adult White non-Hispanic women: mixed methods study. JMIR Dermatol 2021 May 19;4(1):e25661. [doi: 10.2196/25661] [Medline: 37632797]
21. Myrick JG, Noar SM, Kelley D, Zeitany AE. The relationships between female adolescents' media use, indoor tanning outcome expectations, and behavioral intentions. Health Educ Behav 2017 Jun;44(3):403-410. [doi: 10.1177/1090198116667251] [Medline: 27590838]
22. Stapleton JL, Hillhouse J, Coups EJ, Pagoto S. Social media use and indoor tanning among a national sample of young adult nonHispanic White women: a cross-sectional study. J Am Acad Dermatol 2016 Jul;75(1):218-220. [doi: 10.1016/j.jaad.2016.01.043] [Medline: 27317521]

XSL•FO
RenderX

23. Heckman CJ, Auerbach MV, Darlow S, Handorf EA, Raivitch S, Manne SL. Association of skin cancer risk and protective behaviors with health literacy among young adults in the USA. Int J Behav Med 2019 Aug;26(4):372-379. [doi: 10.1007/s12529-019-09788-1] [Medline: 31147962]
24. Watson M, Holman DM, Maguire-Eisen M. Ultraviolet radiation exposure and its impact on skin cancer risk. Semin Oncol Nurs 2016 Aug;32(3):241-254. [doi: 10.1016/j.soncn.2016.05.005] [Medline: 27539279]
25. Raimondi S, Suppa M, Gandini S. Melanoma epidemiology and sun exposure. Acta Derm Venereol 2020 Jun 3;100(11):adv00136. [doi: 10.2340/00015555-3491] [Medline: 32346751]
26. Noar SM, Harrington NG. eHealth Applications: Promising Strategies for Behavior Change: Routledge; 2012.
27. Schmid KL, Rivers SE, Latimer AE, Salovey P. Targeting or tailoring? Mark Health Serv 2008;28(1):32-37. [Medline: 18389854]
28. Agrawal A, Sharma D, Mishra MK, Shah MA, Alam M, Khandelwal V. How to cook a meme: exploring content strategies in brand and user-generated memes on Instagram. Cogent Bus Manag 2024;11(1):2366001. [doi: 10.1080/23311975.2024.2366001]
29. Occa A, Chen HY, Teffeteller KL. Using online memes to communicate about health: a systematic review. Am J Health Promot 2025 Feb;39(2):299-329. [doi: 10.1177/08901171241272075] [Medline: 39154320]
30. Engel E, Gell S, Heiss R, Karsay K. Social media influencers and adolescents' health: a scoping review of the research field. Soc Sci Med 2024 Jan;340:116387. [doi: 10.1016/j.socscimed.2023.116387] [Medline: 38039770]
31. Kostygina G, Tran H, Carter CC, Emery SL. Promotion of health-harming products on Instagram: characterizing strategies boosting audience engagement with cigar marketing messages. Int J Environ Res Public Health 2025 Aug 17;22(8):1285. [doi: 10.3390/ijerph22081285] [Medline: 40869870]
32. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media–delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474. [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
33. Baskerville NB, Azagba S, Norman C, McKeown K, Brown KS. Effect of a digital social media campaign on young adult smoking cessation. Nicotine Tob Res 2016 Mar;18(3):351-360. [doi: 10.1093/ntr/ntv119] [Medline: 26045252]
34. Lyu JC, Meacham MC, Nguyen N, Ramo D, Ling PM. Factors associated with abstinence among young adult smokers enrolled in a real-world social media smoking cessation program. Nicotine Tob Res 2024 Feb 15;26(Supplement_1):S27-S35. [doi: 10.1093/ntr/ntad170] [Medline: 38366340]
35. Heckman CJ, Mitarotondo A, Lin Y, et al. Digital interventions to modify skin cancer risk behaviors in a national sample of young adults: randomized controlled trial. J Med Internet Res 2024 Jul 2;26:e55831. [doi: 10.2196/55831] [Medline: 38954433]
36. Mays D, Alalwan MA, Long L, Atkins MB, Tercyak KP. Effects of a tailored mobile messaging intervention for indoor tanning cessation in young females: a randomized clinical trial. Cancer Epidemiol Biomarkers Prev 2025 Feb 6;34(2):332-339. [doi: 10.1158/1055-9965.EPI-24-0964] [Medline: 39535537]
37. Shkodrani E, Xhaja A, Dedej S, Shkodrani B, Hoxhallari G. The influence of social media on melanoma prevention: a retrospective analysis. Cureus 2025 Sep;17(9):e92408. [doi: 10.7759/cureus.92408] [Medline: 41111745]
38. LA Torre G, D'Egidio V, Guastamacchia S, Barbagallo A, Mannocci A. Diffusion of the Italian social media campaign against smoking on a social network and YouTube. J Prev Med Hyg 2020 Jun;61(2):E200-E204. [doi: 10.15167/2421-4248/jpmh2020.61.2.1419] [Medline: 32803006]
39. O'Kane N, McKinley MC, Gough A, Hunter RF. Investigating the feasibility and acceptability of using Instagram to engage post-graduate students in a mass communication social media-based health intervention, #WeeStepsToHealth. Pilot Feasibility Stud 2022 Dec 12;8(1):254. [doi: 10.1186/s40814-022-01207-9] [Medline: 36510310]
40. de Vere Hunt I, Cai ZR, Nava V, et al. A social media–based public health campaign to reduce indoor tanning in high-risk populations. AJPM Focus 2023 Sep;2(3). [doi: 10.1016/j.focus.2023.100123] [Medline: 37662553]
41. Perez D, Mendoza JM, Statman MR, Shahrour N, Gratz BW, Tercyak KP. Implementation outcomes of the National Skin Smart Campus initiative among young adults. Int J Environ Res Public Health 2026 Jan 28;23(2):166. [doi: 10.3390/ijerph23020166] [Medline: 41752248]
42. Fazel SS, Fenton S, Braun N, et al. Tailored sun safety messages for outdoor workers. Saf Health Work 2023 Mar;14(1):43-49. [doi: 10.1016/j.shaw.2023.01.001] [Medline: 36941943]
43. Joo JY, Liu MF. Culturally tailored interventions for ethnic minorities: a scoping review. Nurs Open 2021 Sep;8(5):2078-2090. [doi: 10.1002/nop2.733] [Medline: 34388862]
44. Kim MT, Heitkemper EM, Hébert ET, et al. Redesigning culturally tailored intervention in the precision health era: self-management science context. Nurs Outlook 2022;70(5):710-724. [doi: 10.1016/j.outlook.2022.05.015] [Medline: 35933178]
45. Preventing skin cancer: interventions in outdoor occupational settings. : Community Preventive Services Task Force (CPSTF); 2013 URL: https://stacks.cdc.gov/view/cdc/168578 [accessed 2026-06-18]
46. Americans' social media use 2025. : Pew Research Center; 2025 URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2025/11/PI_2025.11.20_Social-Media-Use_REPORT.pdf [accessed 2026-05-07]

XSL•FO
RenderX

47. Glanz K, Yaroch AL, Dancel M, et al. Measures of sun exposure and sun protection practices for behavioral and epidemiologic research. Arch Dermatol 2008 Feb;144(2):217-222. [doi: 10.1001/archdermatol.2007.46] [Medline: 18283179]
48. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 1988 Jun;124(6):869-871. [doi: 10.1001/archderm.124.6.869] [Medline: 3377516]
49. Johnson J, Johnson AR, Andersen CA, Kelso MR, Oropallo AR, Serena TE. Skin pigmentation impacts the clinical diagnosis of wound infection: Imaging of bacterial burden to overcome diagnostic limitations. J Racial Ethn Health Disparities 2024 Apr;11(2):1045-1055. [doi: 10.1007/s40615-023-01584-8] [Medline: 37039975]
50. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Predictor selection. In: Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models: Springer; 2012:395-429. [doi: 10.1007/978-1-4614-1353-0_10]
51. Matthews T, Arseneault L, Bryan BT, et al. Social media use, online experiences, and loneliness among young adults: a cohort study. Ann N Y Acad Sci 2025 Jun;1548(1):194-205. [doi: 10.1111/nyas.15370] [Medline: 40350583]
52. Kim HS, Bigman CA, Leader AE, Lerman C, Cappella JN. Narrative health communication and behavior change: the influence of exemplars in the news on intention to quit smoking. J Commun 2012 Jun 1;62(3):473-492. [doi: 10.1111/j.1460-2466.2012.01644.x] [Medline: 22736808]
53. Dixon H, Scully M, Gascoyne C, Wakefield M. Can counter-advertising diminish persuasive effects of conventional and pseudo-healthy unhealthy food product advertising on parents?: An experimental study. BMC Public Health 2020 Nov 25;20(1):1781. [doi: 10.1186/s12889-020-09881-1] [Medline: 33238936]
54. Grummon AH, Zeitlin AB, Lee CJY, et al. Countermarketing versus health education messages about sugar-sweetened beverages: an online randomized controlled trial of US adults. Am J Public Health 2024 Dec;114(12):1354-1364. [doi: 10.2105/AJPH.2024.307853] [Medline: 39361914]
55. Andersen PA, Buller DB, Walkosz BJ, et al. Environmental cues to UV radiation and personal sun protection in outdoor winter recreation. Arch Dermatol 2010 Nov;146(11):1241-1247. [doi: 10.1001/archdermatol.2010.327] [Medline: 21079060]
56. Carter OBJ, Donovan RJ. Public (mis)understanding of the UV Index. J Health Commun 2007;12(1):41-52. [doi: 10.1080/10810730601093371] [Medline: 17365348]
57. Buster KJ, You Z, Fouad M, Elmets C. Skin cancer risk perceptions: a comparison across ethnicity, age, education, gender, and income. J Am Acad Dermatol 2012 May;66(5):771-779. [doi: 10.1016/j.jaad.2011.05.021] [Medline: 21875760]
58. Gupta AK, Bharadwaj M, Mehrotra R. Skin cancer concerns in people of color: risk factors and prevention. Asian Pac J Cancer Prev 2016 Dec 1;17(12):5257-5264. [doi: 10.22034/APJCP.2016.17.12.5257] [Medline: 28125871]
59. Gerend MA, Murdock C, Grove K. An intervention for increasing HPV vaccination on a university campus. Vaccine 2020 Jan 22;38(4):725-729. [doi: 10.1016/j.vaccine.2019.11.028] [Medline: 31767468]
60. Huang JJ, Francesconi M, Cooper MH, Covello A, Guo M, Gharib SD. Community health workers on a college campus: effects on influenza vaccination. J Am Coll Health 2018;66(4):317-323. [doi: 10.1080/07448481.2018.1440582] [Medline: 29447623]
61. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
OR: odds ratio
SSC: Skin Smart Campus
UVR: UV radiation

Edited by R Dellavalle; submitted 03.Feb.2026; peer-reviewed by A Mitarotondo, H Mahler; revised version received 01.Jun.2026; accepted 10.Jun.2026; published 07.Jul.2026.

Please cite as:
Mendoza JM, Perez D, Statman MR, Price MK, Brown MF, Anwar OU, Sleiman Jr MM, Tercyak KP
Pre-Post Evaluation of a Sun Safety Social Media Campaign for Young Adults to Reduce Skin Cancer Risk
JMIR Dermatol 2026;9:e90595
URL: https://derma.jmir.org/2026/1/e90595
doi:10.2196/90595

© Janessa M Mendoza, David Perez, Muriel R Statman, Mia K Price, Madeleine F Brown, Omar U Anwar, Marcelo M Sleiman Jr, Kenneth P Tercyak. Originally published in JMIR Dermatology (http://derma.jmir.org), 7.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

ChatGPT-Generated Advice on Sun Protection and Skin Cancer Prevention Compared to American Academy of Dermatology Guidelines: Cross-Sectional Content Analysis

Hanadi Qeyam[1], MBBS; Ahmed Al-Rusan[1,2], MBBS

[1]Department of Dermatology, Faculty of Medicine, Jordan University of Science and Technology, P.O.Box 3030, Irbid, Jordan

[2]Birmingham Skin Centre, Department of Dermatology, Sandwell & West Birmingham Hospitals NHS Trust, Birmingham, United Kingdom

Corresponding Author:
Hanadi Qeyam, MBBS
Department of Dermatology, Faculty of Medicine, Jordan University of Science and Technology, P.O.Box 3030, Irbid, Jordan

Abstract

Background: Artificial intelligence tools such as ChatGPT are increasingly used by the public to seek health-related information. However, the accuracy and quality of artificial intelligence–generated dermatological guidance, particularly regarding sun protection and skin cancer prevention, have not been systematically assessed.

Objective: This study aimed to evaluate the quality of ChatGPT-generated responses to common patient questions on sun protection and skin cancer prevention by benchmarking them against guidelines from the American Academy of Dermatology.

Methods: Nine standardized questions reflecting common public inquiries were submitted to ChatGPT (GPT-4 free tier) in a single session on May 13, 2025. Responses were independently evaluated by 2 board-certified consultant dermatologists (>15 years' experience each) across 4 domains (accuracy, completeness, clarity, and relevance) using an author-developed 5-point ordinal rating scale anchored to American Academy of Dermatology guidelines. Scoring disagreements were resolved through discussion between raters until consensus was reached. Interrater reliability was assessed using the linear weighted Cohen κ and intraclass correlation coefficient.

Results: Overall mean scores were 5.0 (SD 0.0) for accuracy (ceiling effect observed), 4.1 (SD 0.6) for completeness, 5.0 (SD 0.0) for clarity (ceiling effect observed), and 4.9 (SD 0.3) for relevance, yielding an overall mean of 4.75/5.0 (SD 0.49). Interrater reliability was excellent (weighted Cohen κ=0.80; intraclass correlation coefficient=0.85; exact agreement on 33/36, 91.7% of the items). Completeness was the lowest-scoring domain (range 3.0-5.0), primarily reflecting errors of omission rather than commission.

Conclusions: ChatGPT provided largely accurate and guideline-consistent advice on sun protection and skin cancer prevention in this targeted content analysis. Its primary limitation was incomplete coverage of nuanced guideline details. While not a replacement for professional health care, ChatGPT may serve as a valuable adjunct tool for public health education on skin cancer prevention provided that its outputs are subject to ongoing, systematic evaluation.

(JMIR Dermatol 2026;9:e93839) doi:10.2196/93839

KEYWORDS

artificial intelligence; AI; skin cancer prevention; sun protection; ChatGPT; dermatology; large language models

Introduction

Epidemiology and Global Burden of Skin Cancer

Although many skin cancers are preventable, there has been a global increase in skin cancer incidence in recent years. Skin cancer accounts for one-third of all cancer diagnoses of the nonmelanoma type, with basal cell carcinomas accounting for 75%. Although melanoma represents only 2% of all skin cancer cases, it is responsible for 80% of skin cancer–related deaths. The annual incidence of melanoma has increased by 4% especially among fair-skinned populations in North America, northern Europe, Australia, and New Zealand [1-3]. In the United States, melanoma is the fifth most common cancer, with an increase in incidence of approximately 320% since 1975 [4]. Globally, melanoma cases are projected to increase from 325,000 in 2020 to approximately 510,000 by 2040—a 50% rise—and mortality is expected to increase by 68% over the same period.

This trajectory underscores the urgent need for effective sun protection strategies and enhanced public health education. UV radiation is the primary modifiable risk factor for skin cancer, causing direct DNA damage to skin cells [5]. Despite growing awareness, adherence to comprehensive sun protection

XSL•FO
RenderX

guidelines remains suboptimal [6], creating an education gap that emerging technologies may help address.

Artificial Intelligence in Patient Education

Given the documented gap between guideline recommendations and public adherence, artificial intelligence (AI) platforms such as ChatGPT represent a new opportunity to improve health communication and patient education at scale. ChatGPT enhances patient education by providing clear and accessible information about dermatological conditions, improving patient understanding and engagement [7]. However, a pilot study by Mondal et al [8] raised significant concerns regarding the accuracy of AI-generated content and potential issues related to text similarity. More broadly, evaluating conversational health AI tools requires frameworks that go beyond surface-level correctness to assess evidence grounding, clinical contextual adequacy, and usefulness to lay users [9].

Study Aims

This study aimed to evaluate the quality of responses generated by ChatGPT (OpenAI) regarding sun protection and skin cancer prevention by systematically comparing them to American Academy of Dermatology (AAD) guidelines. Using an author-developed assessment framework covering 4 domains—accuracy, completeness, clarity, and relevance—this descriptive content analysis sought to characterize ChatGPT's utility and limitations as a source of patient-facing dermatological guidance.

The findings will inform patient education strategies and contribute to evidence-based guidelines for AI integration in health care communication. This research is particularly timely given the growing reliance on AI-powered health information resources and their potential impact on public health behaviors.

Methods

Study Design

This was a descriptive content analysis comparing AI-generated responses to common patient questions on sun protection and skin cancer prevention against established clinical guidelines. Nine specific questions were developed to cover sunscreen use, protective clothing, and early skin cancer detection, reflecting common public inquiries consistent with topics addressed in previous evaluations of AI-generated dermatological content [10].

A structured summary of the study design following the METRICS (model, evaluation, timing, range/randomization, individual factors, count, and specificity of prompts and language) framework for standardized reporting of generative AI studies in health care [11] is shown in Table 1.

Table . Model, evaluation, timing, range/randomization, individual factors, count, and specificity of prompts and language (METRICS) summary of study design and reporting parameters.

METRICS component	Definition	This study
Model	AI[a] model and version evaluated	ChatGPT (GPT-4), free-tier interface; exact subversion not pinnable on the free tier
Evaluation	Approach and instrument used to assess outputs	Independent expert rating by 2 board-certified dermatologists; author-developed 5-point ordinal scale across 4 domains (accuracy, completeness, clarity, and relevance) benchmarked against AAD[b] guidelines
Timing	Dates of data collection	Single session (May 13, 2025)
Range or randomization	Query selection and randomization	Nine predefined standardized questions on sun protection and skin cancer prevention; no randomization (fixed question set)
Individual factors	Rater characteristics and interrater reliability	Two raters, each with >15 y of experience (based in Jordan and the United Kingdom); weighted Cohen κ=0.80; ICC(2,1)[c]=0.85; 91.7% exact agreement
Count	Number of outputs or data points	One output per question; 9 responses; 36 item-domain ratings
Specificity of prompts and language	Prompt design and language used	Verbatim patient-style questions submitted in English; memory enabled and not reset between questions

[a]AI: artificial intelligence.

[b]AAD: American Academy of Dermatology.

[c]ICC: intraclass correlation coefficient.

Prompt Development and Question Selection

The following nine questions were presented to ChatGPT: (1) "How to prevent skin cancer?" (2) "What sunscreen should I use?" (3) "How much sunscreen should I use, and how often should I apply it?" (4) "When should I use sunscreen?" (5) "Is a high-number SPF better than a low-number one?" (6) "What to wear to protect my skin from the sun?" (7) "What are the

XSL•FO
RenderX

common types of skin cancer?" (8) "How to perform a skin self-exam?" (9) "What are the ABCDE warning signs of melanoma?"

To maintain consistency, questions were submitted verbatim without modifications or additional context, and responses were documented exactly as provided without interpretation. The full responses are shown in Table 2.

XSL•FO
RenderX

Table . ChatGPT responses to common patient questions on sun protection and skin cancer prevention.

Question	ChatGPT response
1. How to prevent skin cancer?	Preventing skin cancer involves a combination of sun protection, lifestyle habits, and regular monitoring. 1. Protect Your Skin from UV Radiation Use Sunscreen: Apply broad-spectrum sunscreen with SPF 30 or higher every day, even on cloudy days. Reapply every 2 hours and after swimming or sweating. Avoid Peak Sun Hours: Limit direct sun exposure between 10 a.m. and 4 p.m., when UV rays are strongest. Wear Protective Clothing: Use wide-brimmed hats, sunglasses with UV protection, and tightly woven long-sleeved shirts and pants. Seek Shade: Especially when UV levels are high. 2. Avoid Tanning Beds Tanning beds emit harmful UV radiation that increases your risk of all types of skin cancer, including melanoma. 3. Check Your Skin Regularly Perform monthly self-exams to look for new or changing moles or spots. Use the ABCDE rule for melanoma. See a dermatologist annually or if you notice any suspicious changes. 4. Be Aware of Risk Factors Fair skin, a history of sunburns, excessive UV exposure, many moles, and a family history of skin cancer increase your risk. 5. Support Skin Health Stay hydrated and maintain a diet rich in antioxidants to support skin repair and immune defense.
2. What sunscreen should I use?	When choosing a sunscreen, look for the following key features: Broad-Spectrum Protection: Protects against both UVA (aging) and UVB (burning) rays. SPF 30 or Higher: SPF 30 blocks ~97% of UVB rays. Higher SPFs offer slightly more protection but still need reapplication. Water-Resistant: Especially important if you are sweating or swimming. Type of Sunscreen: Mineral (Physical): Uses zinc oxide or titanium dioxide. Good for sensitive skin. Starts working immediately. Chemical: Contains ingredients like avobenzone, octinoxate, or oxybenzone. Lightweight, better for everyday wear. Top Dermatologist-Recommended Sunscreens (as of 2024): EltaMD UV Clear Broad-Spectrum SPF 46 La Roche-Posay Anthelios Melt-in Milk SPF 100 Neutrogena Ultra Sheer Dry-Touch SPF 55 Supergoop! Unseen Sunscreen SPF 40.
3. How much sunscreen should I use, and how often should I apply it?	How Much to Use: For the whole body: About 1 ounce (30 mL), roughly a shot glass full. For the face and neck: About a nickel-sized amount or half a teaspoon. For each arm: About 1 teaspoon. For each leg, front and back of the torso: About 1 teaspoon each. How Often to Apply: Before sun exposure: Apply 15 to 30 minutes before going outside. Reapply every 2 hours, especially if outside for an extended time. Reapply immediately after swimming, sweating heavily, or towel drying. Even water-resistant sunscreen needs reapplication every 40 to 80 minutes when swimming or sweating. Bonus Tips: Use broad-spectrum SPF 30 or higher. Do not forget ears, back of the neck, tops of feet, and scalp/hairline. Check expiration dates.
4. When should I use sunscreen?	You should use sunscreen every day, even on cloudy, cool, or winter days, because up to 80% of UV rays can penetrate clouds. Apply it 15 minutes before going outdoors, and especially when exposed to sunlight during peak UV hours (10 a.m. to 4 p.m.). Sunscreen is also important if you are near water, snow, sand, or concrete, as these surfaces reflect UV rays and increase exposure. Additionally, wear sunscreen indoors if you spend a lot of time near windows, since UVA rays can penetrate glass.
5. Is a high-number SPF[a] better than a low-number one?	Yes, a high SPF (Sun Protection Factor) offers more protection from UVB rays than a low one, but the difference becomes smaller as the SPF increases: SPF 15 blocks about 93% of UVB rays SPF 30 blocks about 97% SPF 50 blocks about 98% SPF 100 blocks about 99% So while SPF 100 sounds much better than SPF 30, it is only about 1-2% more effective in blocking UVB rays. Higher SPF can be more forgiving if you do not apply enough or reapply as often as needed, but it should not replace good application habits.

XSL•FO
RenderX

Question	ChatGPT response
6. What to wear to protect my skin from the sun?	UPF Clothing: Choose clothing with a UPF (Ultraviolet Protection Factor) rating of 30 or higher. Long-Sleeve Shirts and Long Pants: Opt for loose-fitting, long-sleeve shirts and long pants made from tightly woven fabrics such as cotton, linen, or UV-protective fabrics. Wide-Brimmed Hat: A brim that is at least 3 inches wide for maximum coverage of face, neck, ears, and scalp. UV-Protective Sunglasses: Sunglasses that block 100% of UVA and UVB rays, labeled “UV400” or “100% UV protection.” Clothing with Mesh or Ventilation: For comfort in hot weather. Scarves or Shawls: For additional neck and chest coverage. Sunscreen on Exposed Skin: Apply broad-spectrum sunscreen with SPF 30 or higher to any exposed skin.
7. What are the common types of skin cancer?	Basal Cell Carcinoma (BCC): The most common type of skin cancer, typically appearing as a small, shiny, or pearly bump or a red, scaly patch. It often occurs on sun-exposed areas like the face, ears, and neck. BCC is slow-growing and rarely spreads. Squamous Cell Carcinoma (SCC): The second most common form, often appearing as a firm, red nodule or a flat lesion with a scaly, crusted surface. More likely to spread than BCC but highly treatable when caught early. Melanoma: The least common but most dangerous type, as it can spread quickly to other organs. It typically appears as an irregularly shaped mole with multiple colors. Early detection is crucial for effective treatment.
8. How to perform a skin self-exam?	Find a Well-Lit Area: Choose a brightly lit room with a full-length mirror and a handheld mirror for hard-to-see areas. Check Your Skin from Head to Toe: Start with your face and work down. Look at your scalp, neck, chest, arms, hands, underarms, back, buttocks, and legs, including the bottoms of your feet. Examine Your Moles and Birthmarks: Look for new growths or irregularities using the ABCDE criteria: Asymmetry, Border irregularity, Color variation, Diameter >6mm, Evolving changes. Look for Other Symptoms: Check for sores that do not heal, bleeding, or itching around a mole. Do Not Forget Your Scalp: Use a comb or blow-dryer to part your hair. Examine Your Nails and Soles of Feet: Look for unusual spots or discoloration. Take Notes: Keep track of any changes and consider taking photos to compare with future self-exams.
9. What are the ABCDE warning signs of melanoma?	A - Asymmetry: One half of the mole or lesion does not match the other half. B - Border: The edges are irregular, notched, or blurry. Normal moles have smooth, well-defined borders. C - Color: The color is uneven, with shades of brown, black, pink, red, white, or blue. D - Diameter: The mole is larger than 6 millimeters (about the size of a pencil eraser), although melanomas can be smaller. E - Evolving: The mole is changing in size, shape, or color over time. Any change should be evaluated by a healthcare professional.

[a]SPF: sun protection factor.

Prompt Administration and Reproducibility Conditions

All 9 questions were submitted within a single uninterrupted session on May 13, 2025, using the free-tier ChatGPT interface (GPT-4). The ChatGPT memory function and conversation history were enabled throughout the session and were not reset between questions, meaning that the model retained the full conversational context as subsequent questions were posed. This reflects the typical use pattern of a lay public user and was intentional in capturing the realistic output that a patient would receive in a real-world interaction. However, it represents a methodological limitation affecting response independence as later responses may have been influenced by earlier conversational context (this is discussed further in the Limitations section). The exact model subversion could not be pinned on the free-tier interface; GPT-4 was the version designated to free users at the time of data collection. A single response was collected per question, consistent with a realistic single-query patient use scenario [10].

Evaluation Framework

ChatGPT’s responses were evaluated against established AAD guidelines using an author-developed 5-point ordinal rating scale covering 4 domains:

- Accuracy—the factual correctness of the statements made, assessed against AAD guidelines and scientific evidence and scored on the correctness of what was stated, not on the completeness of coverage (which was captured separately under the “completeness” domain)
- Completeness—the breadth of coverage; specifically, whether all essential AAD-recommended information points were included in the response
- Clarity—the understandability and accessibility of the information for a lay audience

XSL•FO
RenderX

- Relevance—the practical applicability and directness of the response to the question asked

A 5-point scale was applied to each domain: 0="completely inadequate or incorrect"; 1="poor, significant inaccuracies or omissions"; 2="fair, partially correct with notable gaps"; 3="good, mostly correct with minor gaps"; 4="very good, only minor omissions"; and 5="excellent, fully correct and comprehensive."

Raters and Scoring Procedure

Two board-certified consultant dermatologists, each with over 15 years of clinical experience—one based in Jordan (HQ) and one based in the United Kingdom (AA-R)—independently evaluated all ChatGPT responses across the 4 domains. Scoring was completed independently without communication between raters. Score sheets were then compared; where disagreements were identified, the 2 raters discussed each item until consensus was reached without involvement of a third arbitrator. For 2 disagreements (questions 3 and 8 in the "completeness" domain), consensus was reached by averaging scores; for one disagreement (question 5 in the "completeness" domain), rater 2 (AA-R) deferred to rater 1 (HQ) after discussion. Consensus scores are reported throughout.

Interrater Reliability

Interrater reliability (IRR) was assessed prior to consensus resolution using SPSS Statistics (version 29; IBM Corp). The linear weighted Cohen κ was calculated to account for the ordinal nature of the scale, with weights proportional to the magnitude of disagreement. The intraclass correlation coefficient (ICC) was computed using a 2-way mixed-effects model, absolute agreement, single rater (ICC(2,1)). κ values were interpreted as slight (<0.20), fair (0.21 - 0.40), moderate (0.41 - 0.60), substantial (0.61 - 0.80), and almost perfect (>0.80) agreement. An ICC of 0.75 or higher was considered to indicate excellent reliability [12].

Statistical Analysis

Descriptive statistics were calculated for each evaluation domain: mean, median, SD, IQR, minimum, and maximum. Given the small sample size (n=9 questions) and ordinal scale, median and IQR values are reported alongside mean and SD values. Domains achieving a mean of 5.0 with an SD of 0.0 reflected ceiling effects and were interpreted accordingly. All analyses were performed using SPSS Statistics (version 29).

Ethical Considerations

This study did not involve human participants or patient data; therefore, institutional review board approval was not required. As this study involved only the analysis of AI-generated text against published clinical guidelines, no ethical concerns regarding human participants apply. This study was not a randomized clinical trial; accordingly, trial registration and CONSORT (Consolidated Standards of Reporting Trials) reporting requirements do not apply.

Results

IRR Results

Preconsensus rating data for all 36 item-domain combinations (9 questions × 4 domains) are shown in Table 3. Raters agreed exactly on 91.7% (33/36) of the items. All 3 disagreements occurred within the "completeness" domain and were of a magnitude of –1 point to +1 point (question 3: rater 1=4 and rater 2=5, resolved by averaging to 4.5; question 5: rater 1=4 and rater 2=3, resolved by rater 2 deferring to rater 1 [score of 4]; and question 8: rater 1=5 and rater 2=4, resolved by averaging to 4.5). No disagreements occurred in the "accuracy," "clarity," or "relevance" domains. The linear weighted Cohen κ was 0.80, indicating substantial to almost perfect interrater agreement. The ICC(2,1) was 0.85, confirming excellent reliability.

Table . Preconsensus rater scores and interrater agreement by domain[a].

Domain	Rater 1 scores (questions 1-9)	Rater 2 scores (questions 1-9)	Exact agreement, n/N (%)	Disagreements	Resolution
Accuracy	5, 5, 5, 5, 5, 5, 5, 5, 5	5, 5, 5, 5, 5, 5, 5, 5, 5	9/9 (100)	None	—[b]
Completeness	4, 4, 4, 5, 4, 3, 4, 5, 4	4, 4, 5, 5, 3, 3, 4, 4, 4	6/9 (66.7)	Questions 3, 5, and 8 (–1 point to +1 point)	Questions 3 and 8: averaged (score of 4.5); question 5: rater 2 deferred to rater 1 (score of 4)
Clarity	5, 5, 5, 5, 5, 5, 5, 5, 5	5, 5, 5, 5, 5, 5, 5, 5, 5	9/9 (100)	None	—
Relevance	5, 5, 5, 5, 5, 4, 5, 5, 5	5, 5, 5, 5, 5, 4, 5, 5, 5	9/9 (100)	None	—
Overall	—	—	33/36 (91.7)	3 items (3/36, 8.3%)	All resolved within –1 point to +1 point

[a]Weighted Cohen κ=0.80; intraclass correlation coefficient (ICC(2,1)) 2-way mixed effects, absolute agreement, single rater=0.85.

[b]Not applicable.

XSL•FO
RenderX

Comparative Analysis of ChatGPT and AAD Recommendations

Question 1: How to Prevent Skin Cancer?

ChatGPT's response was accurate (5.0/5): all preventive measures stated—including sunscreen use, sun avoidance, protective clothing, tanning bed avoidance, and skin self-examination—were factually correct and consistent with AAD guidelines [13]. Completeness obtained a score of 4.0/5; the response omitted wearing sunglasses with UV protection, extra caution near reflective surfaces (water, snow, and sand), shade as a primary prevention method, and specific guidance on protecting children's skin. These omissions, while not harmful, mean that a patient following only this advice would receive incomplete sun protection guidance. Clarity (5.0) and relevance (5.0) received perfect scores.

Question 2: What Sunscreen Should I Use?

The accuracy score was 5.0/5 as the response correctly identified broad-spectrum protection, sun protection factor (SPF) of 30 or higher, water resistance, and mineral vs chemical sunscreen distinctions [14]. Completeness obtained a score of 4.0/5, with omissions including the recommendation to combine sunscreen with other protective measures, extra precautions near reflective surfaces, and guidance on applying an adequate amount. The response included product-specific brand recommendations (EltaMD, La Roche-Posay, Neutrogena, and Supergoop!), which were considered clinically appropriate in context; however, geographic variability in product availability means that such brand-level guidance may not be universally applicable and should be interpreted as illustrative rather than prescriptive [15]. Clarity and relevance scores were both 5.0/5.

Question 3: How Much Sunscreen Should I Use, and How Often Should I Apply It?

ChatGPT demonstrated excellent accuracy (5.0/5), correctly stating the standard application amount (1 oz for the full body); the 2-hour reapplication interval; and the need to reapply after water exposure, sweating, or towel drying [14]. The completeness score was 4.5/5, with the only notable omission being lip protection as a specific AAD recommendation. Clarity and relevance scores were both 5.0/5.

Question 4: When Should I Use Sunscreen?

ChatGPT received perfect scores across all 4 domains (5.0/5 each). The response covered daily use, weather-related factors, preexposure timing (15 minutes), peak UV hours, high-reflection environments, and indoor application near windows—all consistent with AAD guidelines [14].

Question 5: Is a High-Number SPF Better Than a Low-Number One?

The response was clinically accurate (5.0/5), correctly explaining the diminishing marginal protection of higher SPFs and advising against substituting high SPF for good application habits [14]. Completeness obtained a score of 4.0/5, with omissions of the AAD's minimum SPF 30 recommendation, the fact that a higher SPF does not extend the reapplication interval, and the common problem of underapplication. Clarity and relevance scores were both 5.0/5.

Question 6: What to Wear to Protect My Skin From the Sun?

ChatGPT received the lowest scores for this question. Completeness received a score of 3.0/5, with omissions including protective footwear, warnings against ineffective options (baseball caps and straw hats with gaps), the reduced protection of wet clothing, and the substantial variation in UV protection factor (UPF) across fabric types (eg, denim with a UPF of approximately 1700 vs a white T-shirt with a UPF of approximately 7) [16]. Relevance received a score of 4.0/5 because the response did not integrate clothing into the broader multicomponent sun protection strategy. Accuracy was scored with 5.0/5 for the items that were stated; however, it should be noted that the inclusion of "clothing with mesh or ventilation" as a sun-protective option is potentially misleading as mesh and ventilated fabrics typically have lower UPF ratings due to their open weave. This item was not penalized under accuracy because the response did not explicitly claim high UV protection for such garments, but clinicians should note this as a potential point of patient misunderstanding.

Question 7: What Are the Common Types of Skin Cancer?

ChatGPT accurately described basal cell carcinoma, squamous cell carcinoma, and melanoma, achieving a score of 5.0/5 for accuracy, clarity, and relevance [17]. Completeness obtained a score of 4.0/5; the notable omission was actinic keratoses, which are clinically important precancerous lesions with progression risk to squamous cell carcinoma.

Question 8: How to Perform a Skin Self-Exam?

Accuracy received a score of 5.0/5, with a correct systematic head-to-toe approach and appropriate integration of ABCDE criteria [18]. Completeness received a score of 4.5/5; minor gaps included separate explicit mention of fingernail and toenail examination, examination between the toes, and the AAD's specific recommendation to consult a board-certified dermatologist when changes are detected. Clarity and relevance scores were both 5.0/5.0.

Question 9: What Are the ABCDE Warning Signs of Melanoma?

ChatGPT received perfect scores for accuracy, clarity, and relevance (5.0/5 each), correctly defining all 5 ABCDE criteria consistent with AAD guidelines [19]. Completeness received a score of 4.0/5; the primary clinically meaningful omission was the absence of any recommendation for the patient to document and track moles over time, which the AAD specifically recommends as a key behavior for early detection.

Quantitative Assessment

ChatGPT's responses were systematically evaluated using a 5-point ordinal rating scale across 4 domains. Descriptive statistics for each domain are summarized in Table 4, and question-level scores across all four domains are presented in Table 5.

XSL•FO
RenderX

Accuracy and clarity both demonstrated ceiling effects, with all 9 responses receiving scores of 5.0/5 and zero variance (SD 0.0; IQR 0.0). This indicates that, within the scope of the 9 questions evaluated, ChatGPT produced no factually incorrect statements. Relevance also scored highly (median 5.0, IQR 0.0; mean 4.9, SD 0.3), indicating that the responses consistently addressed the question asked with clinically applicable content. Completeness was the weakest domain (median 4.0, IQR 0.5; mean 4.1, SD 0.6; range 3.0 - 5.0), with the greatest variability between questions. The lowest-scoring response was that to question 6 (sun-protective clothing; mean 4.25 overall), which also had the lowest completeness score (3.0/5) and the only below-perfect relevance score (4.0/5). Errors across all domains were predominantly omissions rather than factual inaccuracies.

Table . Descriptive statistics of evaluation criteria scores.

Metric	Accuracy	Completeness	Clarity	Relevance
Mean (SD)	5.0 (0.0)[a]	4.1 (0.6)	5.0 (0.0)[a]	4.9 (0.3)
Median (IQR)	5.0 (5.0-5.0)	4.0 (4.0-4.5)	5.0 (5.0-5.0)	5.0 (5.0-5.0)
Range	5.0-5.0	3.0-5.0	5.0-5.0	4.0-5.0

[a]The SD of 0.0 for accuracy and clarity reflects a ceiling effect: all 9 responses received the maximum score of 5.0 in these domains.

Table . Evaluation of ChatGPT's responses on sun protection and skin cancer prevention compared to American Academy of Dermatology guidelines.

Question	Accuracy score (out of 5)[a]	Completeness score (out of 5)[b]	Clarity score (out of 5)[c]	Relevance score (out of 5)[d]	Overall score, mean (SD)[e]
1. How to prevent skin cancer?	5.0	4.0	5.0	5.0	4.75 (0.50)
2. What sunscreen should I use?	5.0	4.0	5.0	5.0	4.75 (0.50)
3. How much sunscreen should I use, and how often?	5.0	4.5	5.0	5.0	4.88 (0.25)
4. When should I use sunscreen?	5.0	5.0	5.0	5.0	5.00 (0.00)
5. Is a high-number SPF[f] better than a low-number one?	5.0	4.0	5.0	5.0	4.75 (0.50)
6. What to wear to protect my skin from the sun?	5.0	3.0	5.0	4.0	4.25 (0.96)
7. What are the common types of skin cancer?	5.0	4.0	5.0	5.0	4.75 (0.50)
8. How to perform a skin self-exam?	5.0	4.5	5.0	5.0	4.88 (0.25)
9. What are the ABCDE warning signs of melanoma?	5.0	4.0	5.0	5.0	4.75 (0.50)

[a]Mean score of 5.0 (SD 0.0) across all questions.

[b]Mean score of 4.1 (SD 0.6) across all questions.

[c]Mean score of 5.0 (SD 0.0) across all questions.

[d]Mean score of 4.9 (SD 0.3) across all questions.

[e]Mean score of 4.75 across all questions.

[f]SPF: sun protection factor.

Discussion

Summary of Main Findings

This descriptive content analysis demonstrates that ChatGPT (GPT-4) generates largely accurate and guideline-consistent responses to common patient questions on sun protection and skin cancer prevention, with an overall mean score of 4.75/5 against AAD recommendations. The primary area of weakness was completeness (mean 4.1/5, SD 0.6), reflecting a consistent pattern of errors by omission rather than commission. This distinction is clinically meaningful: ChatGPT did not produce factually incorrect information in this analysis but delivered incomplete guidance in most responses, with potential real-world

XSL•FO
RenderX

implications for patient behavior that depend on the specific omissions.

Accuracy and Guideline Alignment

ChatGPT consistently provided accurate advice across all 9 questions, with no factual errors identified. Questions concerning sunscreen use, application frequency, and melanoma warning signs received the highest combined scores, indicating reliable performance on specific, well-defined topics. Questions requiring comprehensive, multicomponent answers—particularly regarding sun-protective clothing (question 6)—received the lowest completeness score (3.0/5) and the lowest overall mean score (4.25/5).

The clinical significance of these completeness gaps should not be underestimated. A patient who receives ChatGPT's response to question 6 and acts on it exclusively would not be warned that wet clothing provides substantially reduced UV protection, would not know that mesh and ventilated garments may have lower UPF ratings than they appear, and would not receive guidance on protective footwear. For a fair-skinned patient with a high risk profile, such omissions could translate into inadequate sun protection despite perceived compliance. Similarly, the omission of guidance on protecting children's skin (question 1) is particularly important given that childhood UV exposure is a critical determinant of lifetime melanoma risk.

These findings are consistent with those of prior research demonstrating that ChatGPT provided clinically appropriate recommendations for 96% of core questions in comparable studies, with SPF guidelines and sun-protective clothing among the topics evaluated [10]. The predominance of omission over commission errors in our study aligns with the broader pattern across large language model (LLM) evaluations in clinical medicine, wherein AI-generated content is generally accurate but selectively incomplete.

Clinical Usefulness and the Accuracy-Completeness Distinction

An important methodological consideration is that surface-level guideline alignment and clinical usefulness are not equivalent. A response may receive a high accuracy score yet remain incomplete, inadequately contextualized, or potentially misleading for patients with specific risk profiles or geographic constraints. Product-specific sunscreen recommendations, for example, may be accurate in the context of widely available products but inaccessible in other health care settings. This distinction—between what is stated being correct and what needs to be stated being complete—should inform how AI-generated health content is evaluated in future research and underscores the importance of using multi-domain evaluation frameworks rather than accuracy alone [9]. Recent work evaluating LLMs for clinical classification tasks has similarly demonstrated that outputs may be technically correct yet clinically insufficient for complex or context-dependent decisions [15].

AI as an Adjunct to Clinician-Led Education

Previous research has demonstrated that AI-driven interventions such as skin aging simulations have produced long-term improvements in sun protection behaviors, particularly among younger adults [20]. This suggests that AI tools such as ChatGPT could serve not only as information sources but also as platforms that motivate adherence to sun protection recommendations. Consistent with a human-centered framework for AI in health care and education, ChatGPT is most appropriately positioned as an adjunct that supports and extends clinician-led patient education rather than a stand-alone substitute for professional guidance [12]. Patient-facing AI tools are most effective when integrated into a care model that maintains physician oversight and encourages users to seek professional confirmation of AI-generated advice.

While this study focused on preventive education, AI is also being used in clinical dermatology for diagnostic support. A recent review demonstrated that AI can accurately analyze clinical and dermoscopic images, sometimes surpassing human experts in specific settings, but emphasized that the best outcomes require combining AI tools with clinical expertise [21].

Comparison With Other AI Platforms

This study evaluated ChatGPT (GPT-4) exclusively, which was the most widely available and most extensively studied conversational AI tool for health information at the time of data collection (May 2025). Other platforms—including Google Gemini, Grok, and Anthropic's Claude—may generate meaningfully different responses to the same prompts and have not been benchmarked against AAD guidelines in this context. Multi-platform comparative evaluation represents an important direction for future research.

Limitations

Several limitations should be acknowledged. First, the 9 questions constitute a targeted content analysis rather than a comprehensive evaluation of ChatGPT's dermatological knowledge scope; follow-up questions, ambiguous prompts, and unsafe queries were not assessed. Second, all questions were submitted in a single session, with memory enabled and conversation history not reset between questions, meaning that later responses may have been influenced by prior conversational context. Third, a single output was collected per question; given the stochastic nature of LLMs, responses may vary across sessions, and findings reflect a single–time point snapshot. Fourth, ChatGPT's outputs are subject to change as underlying model versions are updated, and findings from May 2025 may not be reproduced with future versions. Fifth, 2 raters were used; although IRR was excellent (weighted κ=0.80; ICC=0.85), a larger panel would further strengthen scoring validity. Sixth, the evaluation used an author-developed rating scale without formal psychometric validation; future studies should develop and validate a standardized instrument. Seventh, the comparison was limited to AAD guidelines; international guideline variations were not assessed. Eighth, all prompts were submitted in English; because prompt language has been shown to critically shape the readability, completeness, and terminology

XSL•FO
RenderX

of AI-generated dermatological content, responses to equivalent prompts in other languages may differ, and our findings may not generalize to non–English-speaking populations [22]. Finally, product-specific brand recommendations in ChatGPT responses may reflect products not available in all geographic regions.

Conclusions

This descriptive content analysis demonstrates that ChatGPT generates largely accurate and guideline-consistent advice on sun protection and skin cancer prevention, with responses closely aligned with AAD recommendations across all 4 evaluated domains. Errors were predominantly omissions rather than factual inaccuracies, with completeness representing the primary area for improvement. The clinical significance of these omissions varies by question but warrants attention, particularly for high-risk patient groups. Although ChatGPT cannot replace professional medical advice, it can serve as a valuable adjunct tool for public health education—most effectively when integrated within a clinician-led educational framework that ensures contextual appropriateness and guideline completeness [12]. With appropriate, ongoing evaluation, AI-assisted dermatological education has significant potential to support skin cancer prevention strategies.

Funding

The authors declare that no funding was received for this study.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

Conceptualization: HQ
Data curation: HQ, AA-R
Formal analysis: HQ
Methodology: HQ, AA-R
Writing—original draft: HQ
Writing—review and editing: HQ, AA-R
Both authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Leiter U, Keim U, Garbe C. Epidemiology of skin cancer: update 2019. In: Reichrath J, editor. Sunlight, Vitamin D and Skin Cancer: Springer; 2020:123-139. [doi: 10.1007/978-3-030-46227-7_6]
2. Arnold M, Singh D, Laversanne M, et al. Global burden of cutaneous melanoma in 2020 and projections to 2040. JAMA Dermatol 2022 May 1;158(5):495-503. [doi: 10.1001/jamadermatol.2022.0160] [Medline: 35353115]
3. Urban K, Mehrmal S, Uppal P, Giesey RL, Delost GR. The global burden of skin cancer: a longitudinal analysis from the Global Burden of Disease Study, 1990-2017. JAAD Int 2021 Jan;2:98-108. [doi: 10.1016/j.jdin.2020.10.013] [Medline: 34409358]
4. Didier AJ, Nandwani SV, Watkins D, et al. Patterns and trends in melanoma mortality in the United States, 1999-2020. BMC Cancer 2024 Jul 2;24(1):790. [doi: 10.1186/s12885-024-12426-z] [Medline: 38956559]
5. Watson M, Holman DM, Maguire-Eisen M. Ultraviolet radiation exposure and its impact on skin cancer risk. Semin Oncol Nurs 2016 Aug;32(3):241-254. [doi: 10.1016/j.soncn.2016.05.005] [Medline: 27539279]
6. McKenzie C, Nahm WJ, Kearney CA, Zampella JG. Sun-protective behaviors and sunburn among US adults. Arch Dermatol Res 2023 Aug;315(6):1665-1674. [doi: 10.1007/s00403-023-02547-z] [Medline: 36790452]
7. Goktas P, Grzybowski A. Assessing the impact of ChatGPT in dermatology: a comprehensive rapid review. J Clin Med 2024 Oct 3;13(19):5909. [doi: 10.3390/jcm13195909] [Medline: 39407969]
8. Mondal H, Mondal S, Podder I. Using ChatGPT for writing articles for patients' education for dermatological diseases: a pilot study. Indian Dermatol Online J 2023;14(4):482-486. [doi: 10.4103/idoj.idoj_72_23] [Medline: 37521213]
9. Hu Y. Toward retrieval-grounded evaluation for conversational large language model-based risk assessment. JMIR AI 2026 Mar 12;5:e90759. [doi: 10.2196/90759] [Medline: 41818631]
10. Cai EM, Pathmarajah P, Daneshjou R, Ko JM, Chiou AS. Evaluating the appropriateness of skin cancer prevention recommendations obtained from an online chat-based artificial intelligence model. JAAD Int 2024;18:145-147. [doi: 10.1016/j.jdin.2024.07.028] [Medline: 39790864]
11. Sallam M, Barakat M, Sallam M. A preliminary checklist (METRICS) to standardize the design and reporting of studies on generative artificial intelligence-based models in health care education and practice: development study involving a literature review. Interact J Med Res 2024 Feb 15;13:e54704. [doi: 10.2196/54704] [Medline: 38276872]

XSL•FO
RenderX

12. Gao L, Sherwood J, Aleisa N, Damoah A, Lu Y, Qu Q. Human-centered AI agents for healthcare and education: a systematic literature review. The George Washington University. 2025. URL: https://faculty.cs.gwu.edu/xiaodongqu/papers/HCII_2025_5774_AI_Agent.pdf [accessed 2026-06-23]
13. How to prevent skin cancer. American Academy of Dermatology Association. URL: https://www.aad.org/public/diseases/skin-cancer/prevent/how [accessed 2025-05-13]
14. Sunscreen FAQs. American Academy of Dermatology Association. URL: https://www.aad.org/media/stats-sunscreen [accessed 2025-05-13]
15. Haver H, Bahl M, Chung M. Classifying the clinical significance of common breast pain symptoms using a large language model, ChatGPT (GPT-4). Clin Imaging 2025 Sep;125:110525. [doi: 10.1016/j.clinimag.2025.110525] [Medline: 40482400]
16. What to wear to protect your skin from the sun. American Academy of Dermatology Association. URL: https://www.aad.org/public/everyday-care/sun-protection/shade-clothing-sunscreen/what-to-wear-protect-skin-from-sun [accessed 2025-05-13]
17. Types of skin cancer. American Academy of Dermatology Association. URL: https://www.aad.org/public/diseases/skin-cancer/types/common [accessed 2025-05-13]
18. Find skin cancer: how to perform a skin self-exam. American Academy of Dermatology Association. URL: https://www.aad.org/public/diseases/skin-cancer/find/check-skin [accessed 2025-05-13]
19. What to look for: ABCDEs of melanoma. American Academy of Dermatology Association. URL: https://www.aad.org/public/diseases/skin-cancer/find/at-risk/abcdes [accessed 2025-05-13]
20. Gantenbein L, Cerminara SE, Maul JT, Navarini AA, Maul LV. Artificial intelligence-driven skin aging simulation as a novel skin cancer prevention. Dermatology 2025;241(1):59-71. [doi: 10.1159/000541943] [Medline: 39401496]
21. Brancaccio G, Balato A, Malvehy J, Puig S, Argenziano G, Kittler H. Artificial intelligence in skin cancer diagnosis: a reality check. J Invest Dermatol 2024 Mar;144(3):492-499. [doi: 10.1016/j.jid.2023.10.004] [Medline: 37978982]
22. Karampinis E, Zoumpourli CM, Kontogianni C, et al. Dermatology “AI Babylon”: cross-language evaluation of AI-crafted dermatology descriptions. Medicina (Kaunas) 2026 Jan 22;62(1):227. [doi: 10.3390/medicina62010227] [Medline: 41597513]

Abbreviations

AAD: American Academy of Dermatology
AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
ICC: intraclass correlation coefficient
IRR: interrater reliability
LLM: large language model
SPF: sun protection factor
UPF: UV protection factor

Edited by R Dellavalle; submitted 20.Feb.2026; peer-reviewed by E Karampinis, R Phelps, Y Hu; revised version received 13.Jun.2026; accepted 14.Jun.2026; published 30.Jun.2026.

Please cite as:
Qeyam H, Al-Rusan A
ChatGPT-Generated Advice on Sun Protection and Skin Cancer Prevention Compared to American Academy of Dermatology Guidelines: Cross-Sectional Content Analysis
JMIR Dermatol 2026;9:e93839
URL: https://derma.jmir.org/2026/1/e93839
doi:10.2196/93839

© Hanadi Qeyam, Ahmed Al-Rusan. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Remote Monitoring of Cryosurgery Response Using a Smartphone App: Prospective Study

Vanessa R Weir, MD; Emily A Cowen, MD; Trina Salvador, MD; Mary D Sun, MD; Lilly Gu, MD; Maura C Gillis, BA; Nicholas R Kurtansky, BS; Veronica Rotemberg, MD, PhD; Allan C Halpern, MD

Dermatology Service, Memorial Sloan Kettering Cancer Center, 530 E 74th St, New York, NY, United States

Corresponding Author:
Allan C Halpern, MD
Dermatology Service, Memorial Sloan Kettering Cancer Center, 530 E 74th St, New York, NY, United States

Abstract

Background: Cryosurgery is an effective treatment for benign lesions, although current unstandardized approaches may result in inadequate responses and unwanted adverse reactions. Monitoring treatment characteristics, lesion responses, and patient-reported outcomes using patient-derived mobile imaging may facilitate longitudinal treatment assessment.

Objective: This study aimed to determine the reliability of metrics for assessing the response to cryosurgery in patients with actinic and seborrheic keratoses using remote photographic monitoring.

Methods: Patients who were recommended cryosurgery by their physician for treating seborrheic and/or actinic keratoses (22 patients with 31 lesions) were enrolled. After treatment, participants took “overview” and “close up” photos of their lesion(s) and rated appearance, pain, and degree bothered on a custom smartphone app at eight posttreatment time points (days 0, 3, 7, 10, 14, 30, 60, and 90). After study completion, independent raters scored the images for local skin response (eg, erythema, scaling, crust, swelling, vesiculation, and erosion), cosmetic outcome (eg, hyperpigmentation, hypopigmentation, scarring, and atrophy), and lesion resolution.

Results: The local skin response peaked 3 days after cryosurgery, with 26% (7/27) of patients reporting pain. There was substantial agreement between raters for lesion resolution (κ=0.71, 95% CI 0.62 - 0.79), erythema (κ=0.66, 95% CI 0.57-0.74), and the local skin response index (κ=0.69, 95% CI 0.61-0.77) as measured using the quadratic-weighted Cohen κ. Overall, 77% (151/195) of submitted photos were good quality, and most image-derived metrics showed higher agreement in good-quality photos (8/14, 57% metrics had moderate-substantial κ) compared to poor-quality photos (4/14, 29% metrics had moderate-substantial κ). The peak local skin response had a moderate positive association with the lesion response at 90 days (Spearman ρ=0.556, *P*=.01).

Conclusions: This study demonstrates the utility of patient self-imaging for longitudinal assessment of the response to cryosurgery.

(JMIR Dermatol 2026;9:e63467) doi:10.2196/63467

KEYWORDS

cryosurgery; cryotherapy; keratoses; tele-dermatology; remote monitoring; local skin response; lesions; patient care; mobile imaging; erythema; skin response

Introduction

Nonsurgical therapies remain the mainstay for many symptomatic benign or precancerous lesions, including actinic or seborrheic keratoses. Cryosurgery is an effective treatment for isolated keratoses and remains the most commonly used destructive modality [1]. However, current unstandardized approaches to freezing techniques may result in inadequate responses and unwanted adverse reactions [2-4]. This highlights the need for more systematic approaches that can optimize treatment outcomes while incorporating patient preferences.

Store-and-forward mobile apps are increasingly used for therapeutic evaluation and research photo-documentation in dermatology [5-7]. This technology offers benefits by reducing geographical limitations and time constraints that make longitudinal research difficult. Beyond facilitating remote lesion monitoring, they can be utilized to document cutaneous events in response to therapeutics while decreasing barriers to follow-up. Our study aimed to evaluate the agreement of image-derived grading of lesion responses and to investigate their correlation with patient-reported adverse reactions in those with actinic and seborrheic keratoses treated with cryosurgery utilizing patient-submitted images.

XSL•FO
RenderX

Methods

Study Design

This was a prospective, single-center, observational study undertaken at the Memorial Sloan Kettering dermatology clinic in New York City from October 2021 to June 2023. Patients were included if they were at least 18 years of age with at least one seborrheic or actinic keratosis undergoing destructive treatment and who could either take a photo of their lesion themselves or have a partner do so. Exclusion criteria included not having access to an iPhone or the inability of the patient or their partner to photograph the lesion. Prior to the study, patient eligibility was assessed.

Ethical Considerations

This study was reviewed and approved by the Memorial Sloan Kettering Cancer Center's Institutional Review Board (Protocol #21‐019). All participants provided written informed consent for participation and publication of their case details, and the research was conducted in accordance with principles embodied in the Declaration of Helsinki and in accordance with local requirements. Analytic data and images submitted via the mobile app were linked to deidentified study identifiers only. Participants received no financial compensation for participation in the study.

Data Collection

Patients were trained to use a smartphone-based self-imaging app (Canfield Capture Mobile App; Canfield Scientific, Inc). Before cryosurgery, participants took baseline overview and close-up photographs of up to three lesions in clinic using the app. After receiving provider-administered cryosurgery using variable techniques, patients were asked to remotely continue photographing lesions and complete a symptom questionnaire (rating pain, degree bothered, and cosmesis) at eight posttreatment time points (days 0, 3, 7, 10, 14, 30, 60, and 90) using the app. A standardized imaging protocol emphasizing consistent lighting, positioning, and focus was suggested. Full instructions, photo quality checklists, and questionnaire items given to the patients are available in Multimedia Appendix 1.

Independently Rated Measurements

Paired pre- and posttreatment images were independently reviewed by two board-certified dermatologists using a structured scoring rubric. Reviewers assessed lesion resolution, local skin response (LSR; including erythema, crusting, swelling, vesiculation/pustulation, erosion/ulceration, and flaking/scaling), and photo quality (good vs poor) across the eight time points. Each image pair was assigned both individual scores (0-4) and a composite LSR index (range 0-24). The lesion response was classified as a binary outcome (complete vs incomplete) and using a four-point scale, which were both used to score the response. Ratings were completed using a standardized interface with anonymized, time-randomized images to reduce bias. The full scoring criteria, interface set up, and workflow can be found in Multimedia Appendix 2.

Statistical Analysis

The primary objective was to assess whether lesion resolution can be reliably evaluated through patient-captured photographs. The quadratic-weighted Cohen κ was used to measure the interrater agreement of lesion resolution (incomplete vs complete), as well as other visually determined posttreatment cutaneous gradings across all time points. The Spearman rank correlation was used to evaluate the relationships between the physician-rated skin responses and the patient-reported adverse reactions. A principal component analysis was conducted to capture significant variance and patterns in the data collected for peak response values and lesion outcomes at day 90. All statistical analysis was performed with R software (version 4.3.1; R Foundation for Statistical Computing) using the following packages: *dplyr*, *tidyverse*, *psych*, *ggplot2*, *stats*, and *table1*.

Results

Study Population

A summary of patient characteristics can be found in Table 1. We enrolled 22 patients with 31 total lesions (18 seborrheic keratoses and 13 actinic keratoses). Patients had Fitzpatrick skin types II-IV. The cryosurgery apertures used included A, B, C, 20 gauge, 22 gauge, and angiocath. Lesions were treated with liquid nitrogen using a mean distance of 1.63 (range 1.00-3.00) cm, 1 or 2 cycles, with an average spray time of 9.84 (SD 5.46) seconds. At 90 days, the participation rate was 68% (21/31 lesions).

XSL•FO
RenderX

Table . Demographic characteristics.

	Seborrheic keratosis (n=18), n (%)	Actinic keratosis (n=13), n (%)	Overall (N=31), n (%)
Age range (years)			
50 - 64	5 (27.8)	3 (23.1)	8 (25.8)
65 - 79	13 (72.2)	8 (61.5)	21 (67.7)
≥80	0 (0)	2 (15.4)	2 (6.5)
Skin type			
II	12 (66.7)	12 (92.3)	24 (77.4)
III	5 (27.8)	1 (7.7)	6 (19.4)
IV	1 (5.6)	0 (0)	1 (3.2)
Site			
Head/neck	8 (44.4)	4 (30.8)	12 (38.7)
Anterior torso	1 (5.6)	1 (7.7)	2 (6.5)
Posterior torso	2 (11.1)	0 (0)	2 (6.5)
Lateral torso	4 (22.2)	0 (0)	4 (12.9)
Upper extremity	0 (0)	5 (38.5)	5 (16.1)
Lower extremity	3 (16.7)	3 (23.1)	6 (19.4)
Lesions available for analysis[a]			
Day 0	16 (88.9)	12 (92.3)	28 (90.3)
Day 3	17 (94.4)	10 (76.9)	27 (87.1)
Day 7	17 (94.4)	9 (69.2)	26 (83.9)
Day 10	17 (94.4)	10 (76.9)	27 (87.1)
Day 14	16 (88.9)	8 (61.5)	24 (77.4)
Day 30	16 (88.9)	8 (61.5)	24 (77.4)
Day 60	17 (94.4)	5 (38.5)	22 (71.0)
Day 90	15 (83.3)	6 (46.2)	21 (67.7)

[a]Completion rate defined as the proportion of treated lesions with evaluable image submissions available at each posttreatment time point, relative to the total number of treated lesions at baseline.

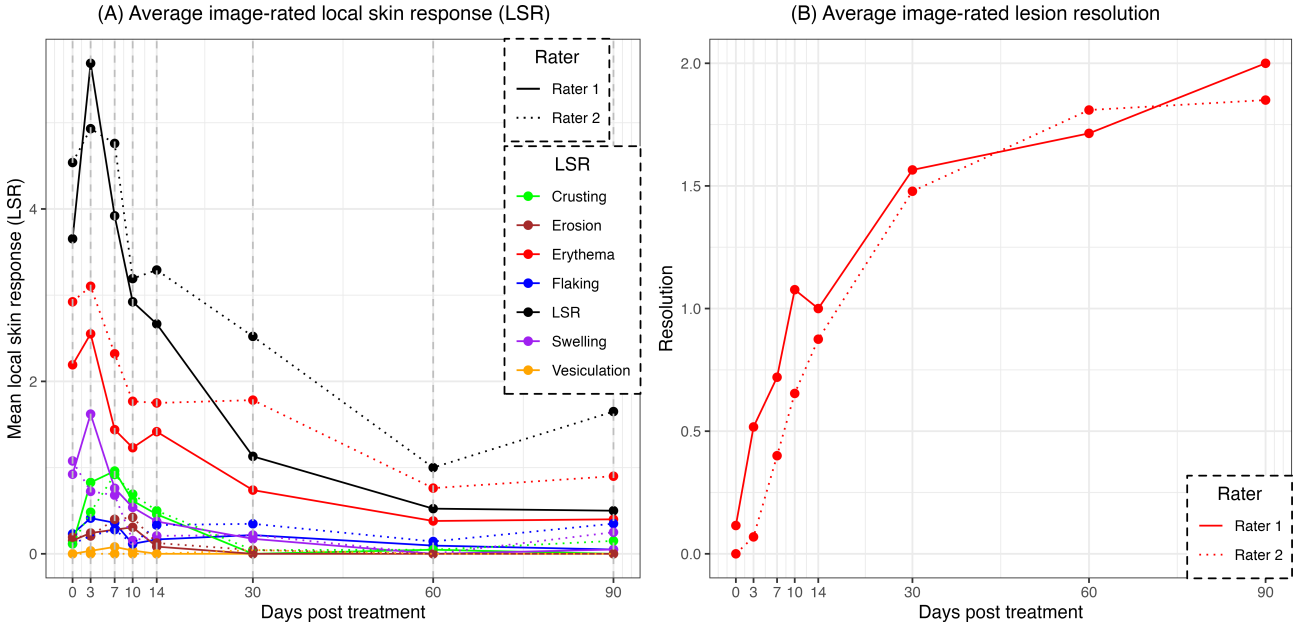
Cryosurgery Efficacy and Tolerability

The frequency of image-rated metrics across all time points is listed in Multimedia Appendix 3. Erythema, flaking, scaling, crusting, and swelling were observed in >50% of lesions, with erythema (27/29, 93%) and swelling (24/29, 83%) being the most commonly observed effects. Vesiculation, atrophy, and scarring were observed in ≤10% of lesion responses. Figure 1A shows the time course of the mean LSR for rater 1 and rater 2. The LSR peaked at day 3, with erythema having the highest mean (mean 2.58, SD 1.3 for rater 1, and mean 3.10, SD 1.5 for rater 2) and vesiculation having the lowest mean (mean 0.04, SD 0.2 for rater 1, and mean 0.00, SD 0.0 for rater 2). Figure 1B shows the average lesion response over time. At 90 days, 32% (6/19) of the lesions were considered resolved by both raters and 58% (11/19) of the lesions were considered resolved by at least one rater.

Multimedia Appendix 4 shows the patient-rated adverse reactions over time. At 3 days, 26% (7/27) of the patients reported pain, 19% (5/27) reported being bothered by adverse reactions, and 37% (10/27) reported cosmetic outcomes as “poor” or “fair.” At 90 days, 71% (15/21) of the patients reported no pain, 76% (16/21) reported very good or excellent cosmetic outcomes, and 100% (21/21) reported they were not bothered by adverse reactions.

XSL•FO
RenderX

Figure 1. Average image-rated local skin response (LSR) following cryosurgery treatment. (A) The mean image-rated LSR for rater 1 and rater 2 following cryosurgery. The individual LSR metrics were rated on a scale of 0 (none or at baseline) to 4. The LSR is a composite of the scores for crusting, erosion, erythema, scaling or flaking, swelling, and vesiculation (0 - 24). The LSR was the highest between days 0 and 7 after treatment, peaking on day 3, and largely resolved by day 60. Erythema presented as the predominant symptom, while vesiculation was the least common. (B) The mean image-rated lesion response for rater 1 and rater 2 following cryosurgery. The lesions were rated from 0 (no change) to 3 (complete resolution). At 90 days, 58% (11/19) of the lesions were considered resolved by at least one rater.



Reliability of Image-Rated Metrics

Table 2 shows the interrater agreement for image-derived metrics of the cryosurgery response. Overall, there was substantial agreement for lesion resolution using a four-point scale (κ=0.71), composite LSR (κ=0.69), and erythema (κ=0.66). Vesiculation, hyperpigmentation, hypopigmentation, and atrophy had negligible agreement.

Table 3 shows the interrater agreement by photo quality. Good-quality photos (n=151) consisted of photo sets where the quality was graded as "good" by both raters. Poor-quality photos (n=44) consisted of sets where at least one person graded the quality as "poor." Lesion resolution (scored as completed or incomplete) was more reliable for good-quality photos compared to poor-quality photos (κ=0.64 vs κ=0.14). Although several image-rated response metrics had higher agreement in good-quality photos (including the LSR index, erythema, erosion, scaling, and swelling), crusting, flaking, and hyperpigmentation had similar or slightly higher agreement in poor-quality photos.

XSL•FO
RenderX

Table . Interrater agreement for image-derived metrics of cryosurgery response (n=195).

Image derived metric	κ value[a]	95% CI
Four point scale		
Lesion resolution[b]	0.71	0.62 to 0.79
Complete versus incomplete		
Lesion Resolution	0.56	0.41 to 0.72
Local skin response metrics		
Erythema	0.66	0.57 to 0.74
Crusting	0.52	0.35 to 0.69
Local skin response composite	0.69	0.61 to 0.77
Erosion	0.47	0.20 to 0.73
Scaling	0.42	0.19 to 0.65
Swelling	0.40	0.23 to 0.58
Flaking	0.24	−0.006 to 0.48
Vesiculation	−0.006	−0.016 to 0.0042
Hyperpigmentation	−0.024	−0.043 to −0.0055
Hypopigmentation	−0.0074	−0.021 to 0.0062
Atrophy	−0.0059	−0.016 to 0.0042
Scarring	N/A[c]	N/A

[a]Interpretation of κ values: 0-0.20 (slight), 0.21-0.40 (fair), 0.41-0.60 (moderate), 0.61-0.80 (substantial), and 0.81-1.00 (almost perfect).

[b]Lesion resolution was graded at four levels (incomplete, <50%, >50%, and complete) and as a binary outcome (complete resolution vs incomplete resolution).

[c]N/A: not applicable.

XSL•FO
RenderX

Table . Interrater agreement for image-derived metrics of cryosurgery response by quality (n=195).

Image derived metric	Good-quality image (n=151)		Poor-quality image (n=44)	
	κ value[a]	95% CI	κ value	95% CI
Four point scale				
Lesion resolution[b]	0.75	0.66 to 0.84	0.50	0.278 to 0.73
Complete versus incomplete				
Lesion resolution	0.64	0.48 to 0.80	0.14	−0.23 to 0.52
Local skin response metrics				
Erythema	0.68	0.59 to 0.77	0.51	0.265 to 0.75
Crusting	0.50	0.30 to 0.70	0.60	0.29 to 0.91
Local skin response composite	0.70	0.62 to 0.79	0.62	0.402 to 0.85
Erosion	0.54	0.29 to 0.79	−0.019	−0.048 to 0.0155
Scaling	0.47	0.20 to 0.73	0.23	−0.11 to 0.57
Flaking	0.22	−0.039 to 0.48	0.31	−0.16 to 0.77
Swelling	0.42	0.211 to 0.62	0.36	0.029 to 0.68
Vesiculation	−0.0076	−0.021 to 0.0055	N/A[c]	N/A
Hyperpigmentation	−0.024	−0.047 to −0.0008	−0.031	−0.066 to 0.0040
Hypopigmentation	−0.0089	−0.025 to 0.0071	N/A	N/A
Scarring	N/A	N/A	N/A	N/A
Atrophy	−0.0067	−0.018 to 0.0045	N/A	N/A

[a]Interpretation of κ values: 0-0.20 (slight), 0.21-0.40 (fair), 0.41-0.60 (moderate), 0.61-0.80 (substantial), and 0.81-1.00 (almost perfect).

[b]Lesion resolution was graded at four levels (incomplete, <50%, >50%, and complete) and as a binary outcome (complete resolution vs incomplete resolution).

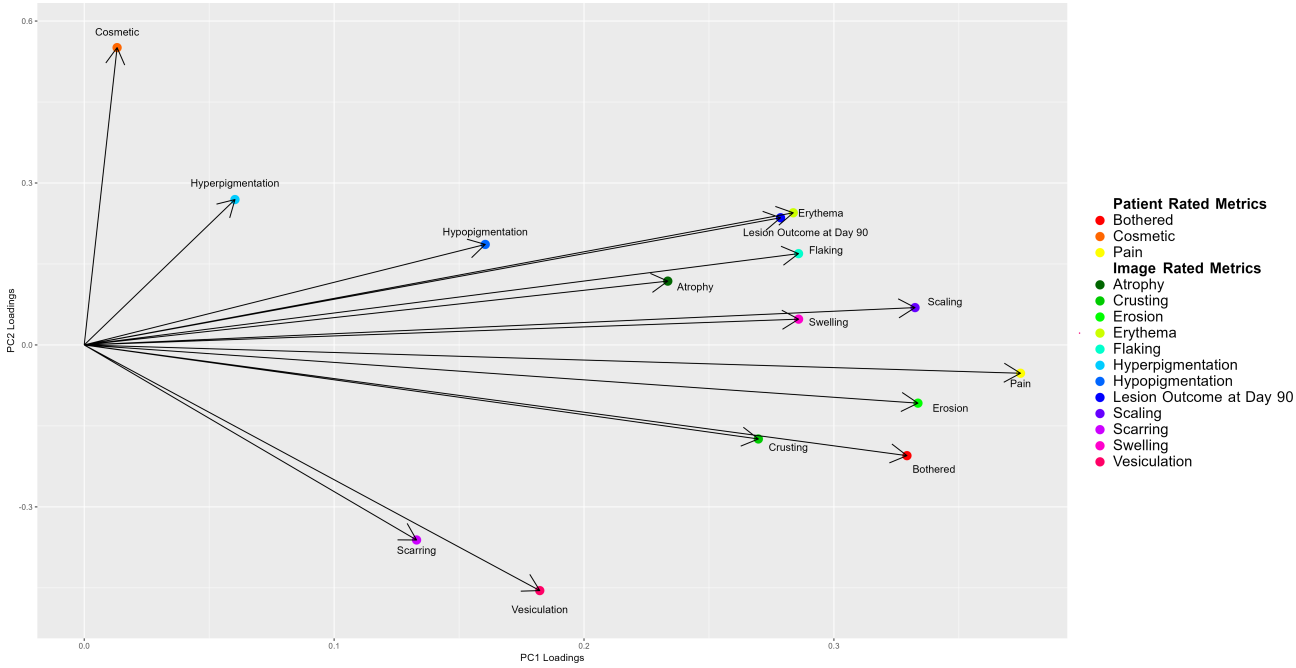
[c]N/A: not applicable.

Predictors of Lesion Response and Patient-Rated Side Effects

Peak responses for image-rated metrics were determined by averaging the highest grades assigned by rater 1 and rater 2 across all time points. Peak values for patient-rated adverse reactions were determined by identifying the maximum scores reported for cosmetic impact, level of bother, and pain across all time points. The peak LSR score had a moderate positive association with the lesion response at 90 days (Spearman ρ=0.556, *P*=.01), a moderate positive association with maximum pain (Spearman ρ=0.643, *P*<.001), and a moderate positive association with the maximum rated degree bothered (Spearman ρ=0.545, *P*=.002), all suggesting that lesions with a higher visually assessed LSR score are more likely to be painful, bothersome, and appear resolved long term. The cosmetic outcome at 90 days was also moderately associated with lesion outcome at 90 days (Spearman ρ=0.591, *P*=.008). Figure 2 shows principal component analysis of peak image-based metrics, peak patient-rated adverse reactions, and lesion resolution at day 90. A biplot of principal component 1 and principal component 2 loadings shows that the lesion outcome at day 90 is most closely clustered with peak erythema, suggesting that changes associated with increased local redness may be the most important factor for long-term lesion resolution. Pain is closely clustered with erosion, and being bothered is closely clustered with crusting.

XSL•FO
RenderX

Figure 2. Biplot of principal component analysis loadings for peak image-rated outcomes and patient-reported adverse reactions. This biplot showcases the first two principal components for principal component analysis performed on the peak local skin response values and patient-reported outcomes with the 90-day lesion resolution (26 lesions). The vector direction indicates the pattern of variance, and the length denotes the strength of contribution to the principal component. The variable "Erythema" is closely clustered with the "Lesion Outcome at Day 90," indicating its potential as a predictor of long-term resolution. Additionally, "Pain" is closely clustered with "Erosion," and "Bothered" is closely clustered with "Crusting," reflecting a shared variance among these adverse reactions. PC: principal component.



Discussion

Principal Findings and Comparison With Prior Work

The findings of this study support the feasibility of using store-and-forward photos for remotely documenting lesion response to destructive therapy. In summary, we found that the image-derived response of keratoses to cryosurgery can be reliably labeled using photos submitted by patients, and that the LSR metrics correlate with patient-reported outcomes like pain and long-term lesion resolution. Several parameters, such as lesion response, the LSR index, and erythema, showed substantial agreement between raters. However, the agreement was minimal for vesiculation, hyperpigmentation, and hypopigmentation, although these were rare events and may have been underpowered to detect any significant agreement. When comparing the agreement of image-derived metrics by photo quality, a higher concordance was noted for lesion resolution and several other response indicators for high-quality photos. However, several response metrics did not vary by photo quality. This suggests that even with a standardized grading system, both rater judgment and photo quality may influence grading of lesion response.

Of the submitted images, 77% (151/195) were deemed adequate for assessment. This surpasses a recent study where 55.1% (1985/3600) of patients submitted photos of various rashes or lesions that were of sufficient quality for medical decision-making [8]. The design of the mobile app, which provided access to baseline images and quality reminders, along with explicit instructions, likely contributed to image quality. We previously reported on the usability of the app in the first 8 patients, who reported ease of use with a mean score of 4.4 out of 5 [9].

Cryosurgery treatment is unstandardized, and our results showed a wide range of apertures, spray times, spray distances, and cycles. This may lead to inadequate treatment for patients and/or bothersome adverse reactions. In our study, 58% of lesions were considered resolved by at least one rater, slightly below the reported rates of 63% to 88% [10-15]. Still, cryosurgery was well tolerated at day 90, although it is worth noting that a subset of patients still reported long-term pain. Reported tolerability of cryosurgery varies widely, with a recent meta-analysis documenting pain or burning in 7% to 22% of cases, crusting in 6%, and discoloration or scarring in 33% [10-16]. This shows there still may be a subset of patients who experience long-term sequelae and highlights the need for personalized treatment approaches to minimize adverse effects, although this study is not powered to evaluate those specifics.

The lesion outcome at 90 days was moderately associated with the peak LSR, highlighting that more robust inflammation may lead to greater long-term resolution. As the LSR appeared to peak between day 3 and day 7, early virtual follow-up may offer insights into the probability of long-term resolution. Still, principal component analysis suggests that lesion resolution may not be closely associated with peak patient-reported outcomes of pain, cosmesis, and degree bothered, highlighting the utility of image-based approaches to evaluate treatment efficacy. Future studies may aim to utilize machine learning approaches that combine reliable image-based metrics with patient-reported outcomes to predict treatment success [17].

XSL•FO
RenderX

Limitations

The study faced limitations; most notably, there was a lack of diversity of skin type, which is important to study because both imaging and response to destructive therapy can vary based on skin type or tone. Future studies should aim to recruit patients that reflect the general population and include destructive therapies often used for patients of darker skin tones (eg, electrodessication). This study is also limited by its small sample size and the lack of in-person assessments. Specifically, a diagnosis of keratosis often relies on tactile cues that were not conducted during follow-up. Additionally, visual assessment of lesion resolution was sometimes difficult to distinguish from persistent posttreatment cosmetic adverse reactions (eg, erythema, crusting, and pigment alterations), which could have affected the grading of both lesion resolution and/or LSR. Lastly, the generalizability of the study is limited by the small number of raters and the single-center setting.

Conclusions

Our research demonstrates that certain image-derived skin response metrics can be reliably labeled from patient-submitted photos and are associated with both lesion responses and some patient-reported outcomes. These findings emphasize the potential of teledermatology for assessing the response to destructive therapies and highlight the limitations of visual grading of the LSR from non-standardized photos. These findings pave the way for future studies aimed at integrating image-based and patient-rated metrics for outcome assessment.

Acknowledgments

We would like to thank Canfield Scientific for providing use of the Canfield Capture Mobile App. Canfield was not involved in study conceptualization, implementation, or analysis.

Funding

This study was supported by the Charina Endowment Fund and NIH/NCI Cancer Center Support Grant P30-CA008748.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Authors' Contributions

Conceptualization: VW, EAC, TS, MDS, LG, MCG, NRK, VR, ACH
Data curation: VW, EAC, TS, MDS, LG
Methodology: VW, EAC, TS, MDS, LG, MCG, NRK, VR, ACH
Supervision: VR, ACH
Writing – original draft: VW

Conflicts of Interest

VW, TS, EAC, MCG, LG, MDS, and NRK have no disclosures to report. VR is an expert advisor for Inhabit Brands and Atria Institute and receives research funding from Lutris Pharma, Kaggle, and the AWS Open Data Program. ACH is a medical consultant for Canfield Scientific Inc. and Scibase and has an equity stake in SpotDoc.

Multimedia Appendix 1
Patient self-imaging protocol and survey instrument.
[DOCX File, 171 KB - derma_v9i1e63467_app1.docx]

Multimedia Appendix 2
Image-based rater grading instructions and interface.
[DOCX File, 1125 KB - derma_v9i1e63467_app2.docx]

Multimedia Appendix 3
Frequency of image-derived metrics by rater agreement.
[DOCX File, 19 KB - derma_v9i1e63467_app3.docx]

Multimedia Appendix 4
Patient-rated side effects over time.
[DOCX File, 120 KB - derma_v9i1e63467_app4.docx]

References

XSL•FO
RenderX

1. Balkrishnan R, Cayce KA, Kulkarni AS, et al. Predictors of treatment choices and associated outcomes in actinic keratoses: results from a national physician survey study. J Dermatolog Treat 2006;17(3):162-166. [doi: 10.1080/09546630600765081] [Medline: 16854758]
2. Zimmerman EE, Crawford P. Cutaneous cryosurgery. Am Fam Physician 2012 Dec 15;86(12):1118-1124. [Medline: 23316984]
3. Berman B, Shabbir AQ, MacNeil T, Knudsen KM. Variables in cryosurgery technique associated with clearance of actinic keratosis. Dermatol Surg 2017 Mar;43(3):424-430. [doi: 10.1097/DSS.0000000000000989] [Medline: 28002105]
4. Thai KE, Fergin P, Freeman M, et al. A prospective study of the use of cryosurgery for the treatment of actinic keratoses. Int J Dermatol 2004 Sep;43(9):687-692. [doi: 10.1111/j.1365-4632.2004.02056.x] [Medline: 15357755]
5. Nahm WJ, Badiavas EV, Kirsner RS, et al. Treating keratinocyte carcinomas with a combination of imiquimod, 5-fluorouracil, and tretinoin using store-and-forward telemedicine in the age of Coronavirus Disease 2019 to promote social distancing. JAAD Case Rep 2020 Sep;6(9):931-934. [doi: 10.1016/j.jdcr.2020.07.024] [Medline: 32837989]
6. Nahm WJ, Gwillim EC, Badiavas EV, et al. Treating melanoma in situ during a pandemic with telemedicine and a combination of imiquimod, 5-fluorouracil, and tretinoin. Dermatol Ther (Heidelb) 2021 Feb;11(1):307-314. [doi: 10.1007/s13555-020-00473-w] [Medline: 33458805]
7. Ali Z, Anderson K, Chiriac A, et al. High adherence and low dropout rate in a virtual clinical study of atopic dermatitis through weekly reward-based personalized genetic lifestyle reports. PLoS ONE 2020;15(7):e0235500. [doi: 10.1371/journal.pone.0235500] [Medline: 32614886]
8. Jiang SW, Flynn MS, Kwock JT, et al. Quality and perceived usefulness of patient-submitted store-and-forward teledermatology images. JAMA Dermatol 2022 Oct 1;158(10):1183-1186. [doi: 10.1001/jamadermatol.2022.2815] [Medline: 35895039]
9. Cowen EA, Sun MD, Gu L, Acevedo C, Rotemberg V, Halpern AC. Store-and-forward mobile application as an accessible method of study participant assessment. J Eur Acad Dermatol Venereol 2023 Apr;37(4):e482-e485. [doi: 10.1111/jdv.18671] [Medline: 36264231]
10. Freeman M, Vinciullo C, Francis D, et al. A comparison of photodynamic therapy using topical methyl aminolevulinate (Metvix) with single cycle cryotherapy in patients with actinic keratosis: a prospective, randomized study. J Dermatolog Treat 2003 Jun;14(2):99-106. [doi: 10.1080/09546630310012118] [Medline: 12775317]
11. Szeimies RM, Karrer S, Radakovic-Fijan S, et al. Photodynamic therapy using topical methyl 5-aminolevulinate compared with cryotherapy for actinic keratosis: a prospective, randomized study. J Am Acad Dermatol 2002 Aug;47(2):258-262. [Medline: 12140473]
12. Szeimies RM, Stockfleth E, Popp G, et al. Long-term follow-up of photodynamic therapy with a self-adhesive 5-aminolaevulinic acid patch: 12 months data. Br J Dermatol 2010 Feb 1;162(2):410-414. [doi: 10.1111/j.1365-2133.2009.09377.x] [Medline: 19804593]
13. Morton C, Campbell S, Gupta G, et al. Intraindividual, right-left comparison of topical methyl aminolaevulinate-photodynamic therapy and cryotherapy in subjects with actinic keratoses: a multicentre, randomized controlled study. Br J Dermatol 2006 Nov;155(5):1029-1036. [doi: 10.1111/j.1365-2133.2006.07470.x] [Medline: 17034536]
14. Kaufmann R, Spelman L, Weightman W, et al. Multicentre intraindividual randomized trial of topical methyl aminolaevulinate-photodynamic therapy vs. cryotherapy for multiple actinic keratoses on the extremities. Br J Dermatol 2008 May;158(5):994-999. [doi: 10.1111/j.1365-2133.2008.08488.x] [Medline: 18341663]
15. Hauschild A, Stockfleth E, Popp G, et al. Optimization of photodynamic therapy with a novel self-adhesive 5-aminolaevulinic acid patch: results of two randomized controlled phase III studies. Br J Dermatol 2009 May;160(5):1066-1074. [doi: 10.1111/j.1365-2133.2009.09040.x] [Medline: 19222455]
16. Patel G, Armstrong AW, Eisen DB. Efficacy of photodynamic therapy vs other interventions in randomized clinical trials for the treatment of actinic keratoses: a systematic review and meta-analysis. JAMA Dermatol 2014 Dec;150(12):1281-1288. [doi: 10.1001/jamadermatol.2014.1253] [Medline: 25162181]
17. Setiadi D, Islam HMM, Trisnapradika GA, Herowati W. Analyzing preprocessing impact on machine learning classifiers for cryotherapy and immunotherapy dataset. J Fut Artif Intell Tech 2024;1(1):39-50. [doi: 10.62411/faith.2024-2]

Abbreviations

LSR: local skin response

XSL•FO
RenderX

Edited by R Dellavalle; submitted 21.Jun.2024; peer-reviewed by T Schlesinger; revised version received 07.Nov.2025; accepted 09.Nov.2025; published 18.Mar.2026.

Please cite as:
Weir VR, Cowen EA, Salvador T, Sun MD, Gu L, Gillis MC, Kurtansky NR, Rotemberg V, Halpern AC
Remote Monitoring of Cryosurgery Response Using a Smartphone App: Prospective Study
JMIR Dermatol 2026;9:e63467
URL: https://derma.jmir.org/2026/1/e63467
doi:10.2196/63467

© Vanessa R Weir, Emily A Cowen, Trina Salvador, Mary D Sun, Lilly Gu, Maura C Gillis, Nicholas R Kurtansky, Veronica Rotemberg, Allan C Halpern. Originally published in JMIR Dermatology (http://derma.jmir.org), 18.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Satisfaction, Side Effects, and Other Reactions Reported by Adult Men Prescribed Compounded Topical Finasteride via a National Telehealth Platform: Retrospective Analysis of Real-World Data

Jessica Yu[1], PhD; Sachie Mochida[1], MS; Michele Emery[1], DNP; Patrick Carroll[1], MD; Justin Ko[1,2], MBA, MD; Arash Mostaghimi[1,3], MBA, MPH, MD

[1]Hims & Hers Health, Inc, 2269 Chestnut Street, Suite 523, San Francisco, CA, United States

[2]Department of Dermatology, Stanford University School of Medicine, Stanford, CA, United States

[3]Department of Dermatology, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Jessica Yu, PhD
Hims & Hers Health, Inc, 2269 Chestnut Street, Suite 523, San Francisco, CA, United States

Abstract

Background: Topical minoxidil and oral finasteride are approved by the US Food and Drug Administration (FDA) for the treatment of male androgenetic alopecia (AGA). However, concerns about adverse events related to the use of oral finasteride have led to some apprehension about the treatment. Topical finasteride, though not FDA-approved, has demonstrated efficacy and safety in a limited number of clinical trials and may be a promising alternative, such that compounding pharmacies and telehealth companies in the United States now offer access to topical finasteride for patients with AGA.

Objective: This real-world, retrospective study is, to our knowledge, the largest study to date aimed to evaluate patient satisfaction and tolerability associated with the novel combinations of topical finasteride and topical minoxidil for the treatment of male AGA.

Methods: We conducted a retrospective analysis of patient data collected during routine clinical follow-up via Hims & Hers, a direct-to-consumer health and wellness platform, between April 1, 2021 and April 30, 2025 to assess the frequency of side effects and other possible medication reactions associated with the use of compounded topical finasteride and minoxidil. Data were gathered from two sources: (1) a follow-up check-in sent to patients approximately 130 days following the initiation of treatment; (2) unprompted communications sent via in-app or web-based messaging from patients to their care team. Data about patient satisfaction with treatment, the frequency of any side effect, frequency of specific side effects, need for a higher level of care, and treatment discontinuation due to a side effect were extracted from the data sources.

Results: A total of 638,629 male patients with AGA received a prescription for a compounded topical finasteride and minoxidil product between April 1, 2021 and April 30, 2025. Of 151,352 (23.7%) patients who completed a follow-up check-in, 121,615 (80.4%) reported being satisfied with treatment and 4034 (2.7%) reported experiencing a side effect. Of all the 638,629 patients, 230 (0.04%) sent their care team a message (outside of check-ins) indicating a side effect or other possible medication reactions. No patient reported seeking a higher level of care or discontinued treatment due to such an occurrence.

Conclusions: Patients prescribed novel formulations of compounded topical finasteride and minoxidil for the treatment of AGA via a national telehealth platform reported satisfaction with the treatment and tolerated it well. The limitations of the study include the use of retrospective data and the lack of a control group, both of which preclude causal inference. Future research should include randomized controlled trials to assess the efficacy, safety, and tolerability of topical finasteride.

(JMIR Dermatol 2026;9:e84676) doi:10.2196/84676

KEYWORDS

androgenetic alopecia; topical finasteride; topical minoxidil; patient satisfaction; side effects; telehealth

Introduction

Androgenetic alopecia (AGA), commonly referred to as "male pattern baldness," is the most common form of hair loss in men. It affects approximately 50% of men worldwide [1] and an estimated 50 million in the United States alone [2]. Although AGA is considered a physically benign medical condition, it is associated with notable psychological consequences including

XSL•FO
RenderX

low self-esteem, body dissatisfaction, social anxiety, and reduced quality of life [3].

Topical minoxidil and oral finasteride are two treatments currently approved by the US Food and Drug Administration (FDA) for the treatment of AGA. Topical minoxidil is available in both 2% and 5% formulations; the 5% formulation has been shown to be significantly superior in increasing hair regrowth, with an earlier response to treatment and good tolerance [4]. Oral finasteride has been shown, in clinical trials, to be well tolerated and effective in stabilizing hair loss and promoting hair growth [5]; however, reports of certain treatment-related adverse events such as sexual side effects and depression have led to some apprehension about the treatment, which may be negatively affecting the number of individuals who could benefit from it [6]. Notably, recent studies have questioned the purported causal relationship between oral finasteride and psychiatric symptoms [7,8].

Topical finasteride may be a promising alternative to oral finasteride. Though limited in number, studies that have examined the use of topical finasteride in the treatment of AGA have found it to be an effective and safe treatment option [9]. Two randomized controlled trials (RCTs) found topical finasteride to significantly decrease the rate of hair loss and significantly improve hair count compared to the placebo, with no differences in the incidence of adverse events or treatment discontinuation between the two groups [10,11]. Plasma concentrations of finasteride were 100-fold lower with the topical application of 0.25% finasteride spray versus 1 mg oral finasteride [11]. Furthermore, a systematic review of available RCTs, prospective studies, and retrospective medical record reviews found topical finasteride, either alone or in combination with other agents including topical minoxidil, to be non-inferior to oral finasteride and well-tolerated by patients—with the authors calling for larger cohort studies to examine the potential adverse event profile of the drug [9].

Unlike oral finasteride, topical finasteride is not currently FDA-approved for the treatment of AGA. It is, however, available as a compounded medication for those who do not want to take an oral medication or might be concerned about the reported side effects associated with oral finasteride. Several compounding pharmacies and telehealth companies in the United States now offer access to topical finasteride for patients with AGA. This real-world retrospective study is, to our knowledge, the largest study to date on patient satisfaction and tolerability associated with novel combinations of topical finasteride and topical minoxidil for the treatment of male AGA. We review anonymized patient data collected during the course of routine clinical care via a direct-to-consumer telemedicine platform to understand the patient-reported satisfaction and frequency of side effects and other possible medication reactions associated with compounded topical finasteride use (compounded topical finasteride is not FDA-approved or evaluated for safety, efficacy, or quality by the FDA).

Methods

Study Overview

Hims & Hers is a direct-to-consumer health and wellness platform that aims to increase access to treatment for adults aged 18 years and older with traditionally stigmatized conditions, including hair loss. Prospective patients seeking hair loss treatment come to the platform and complete a comprehensive clinical intake. Once the intake process is complete, a licensed medical provider thoroughly reviews the information gathered during the intake process, including medical history and treatment preferences, and has the opportunity to follow-up with the patient with any questions or remaining information deemed necessary to provide care. The provider then makes an independent clinical determination as to whether treatment is appropriate, and, if appropriate, shares a diagnosis and treatment plan. All licensed medical providers furnishing care through the platform are employed or contracted by You Health, a professional corporation owned and managed by licensed health care providers, which is the provider network associated with the platform. Patients sign up for a subscription to receive their medication dispensed by a licensed pharmacy at regular intervals. With this subscription, patients have ongoing, unlimited access to their care team via messaging and are sent follow-up check-ins to assess their treatment experience.

As of June 2025, three compounded topical finasteride and minoxidil products were available via the Hims & Hers platform to treat adult men with AGA: a spray consisting of 0.3% topical finasteride and 6% minoxidil, to be sprayed four times on the individual's affected scalp area once per day; a spray consisting of 0.3% topical finasteride, 7% minoxidil, 2.2% ketoconazole, and 0.2% biotin, to be sprayed four times on the individual's affected scalp area once per day; and a serum consisting of 0.3% topical finasteride and 6% minoxidil, 1 mL of which to be massaged into the individual's affected scalp area once per day. All patients prescribed a compounded topical finasteride and minoxidil product were made aware that the product was not FDA-approved and were provided with instructions for use as well as education regarding what to expect with the treatment, common side effects, and other precautions. Patients also had access to educational treatment information via the Hims & Hers app and could contact their care team at any time with questions or concerns. In April 2025, the FDA issued an alert to health care providers, compounders, and consumers regarding potential risks associated with the use of compounded topical finasteride. This information was also shared with patients to ensure transparent communication regarding the products available through the platform.

To assess the frequency of side effects and other possible medication reactions associated with the use of compounded topical finasteride and minoxidil available via the Hims & Hers platform, we conducted a retrospective analysis of patient data collected during the course of routine clinical follow-up via the platform between April 1, 2021 and April 30, 2025. As this was an analysis of data gathered from individuals actively engaged in treatment, there was no control group.

XSL•FO
RenderX

Data Collection

The analysis included two sets of data. The first set of data consisted of responses to a follow-up check-in assessment sent to patients approximately 130 days following treatment initiation. The check-in queried patients about their treatment satisfaction and experience with side effects. To assess treatment satisfaction, patients were asked to indicate "yes" or "no" to the following prompt: "I'm happy with the way my treatment is working." To assess experience with side effects, patients were asked to respond "yes" or "no" to the following question: "Are you bothered by any side effects or other negative reactions from your treatment?" No other questions pertaining to side effects were included in the check-in.

The second set of data consisted of unprompted communications sent via in-app or web-based messaging from patients to their care team. Patients can send these unprompted messages at any time for review by the care team. These communications undergo continuous quality assurance by a clinical quality team that monitors patient messages in real-time for mention of side effects or other possible medication reactions and follows-up as appropriate. Their work includes validating the data to ensure that such events are appropriately recorded—for example, that the side effects and reactions reported are reported by patients in relation to one of the topical finasteride and minoxidil products highlighted in this analysis. Utilizing both sets of data ensured that all occurrences, both solicited and spontaneously reported by patients, were included in the analysis.

Statistical Analysis

Descriptive statistics using Google Colab (Mountain View, CA) were used to quantify the percentage of patients who reported satisfaction with treatment in their follow-up check-in, the percentage of patients who reported having been bothered by side effects or other negative reactions in their follow-up check-in, the percentage of patients who indicated experiencing a side effect or other possible medication reaction in messages to their care team, the percentage of patients who sought a higher level of care due to such a reaction, and the percentage of patients who discontinued treatment due to such a reaction. For results regarding the percentages of patients who reported treatment satisfaction and side effects in their follow-up check-in, the number of patients who completed a check-in is used as the sample size. For results regarding the percentage of patients who reported a side effect to their care team, the total number of patients prescribed a compounded topical finasteride product is used a the sample size. This is due to the fact that all patients had the ability to message their care team; thus, all patients can be included in the denominator.

Ethical Considerations

This study was approved by the WCG Institutional Review Board (Protocol 001, Review 20244102). All study procedures were conducted in accordance with the principles of the Declaration of Helsinki. The study protocol included a Waiver of Informed Consent, as all data analyzed were collected during the course of routine care and de-identified prior to analysis. Patients were not compensated for their participation in this study.

Results

Baseline Demographics

A total of 638,629 male patients with AGA received a prescription for a compounded topical finasteride and minoxidil product between April 1, 2021 and April 30, 2025. A total of 151,352 completed the follow-up check-in querying patients about their treatment satisfaction and experience with side effects.

The mean (SD) age of all patients who received a prescription for a compounded topical finasteride product (n=638,629) was 39.6 (11.9) years, while the mean (SD) age of those who completed the follow-up check-in (n=151,352) was 41.2 (11.8) years.

Treatment Satisfaction and Side Effects as Reported During Follow-Up Check-In

Overall, 121,615 (80.4%, n=151,352, 95% CI [80.2%, 80.6%]) patients who completed the follow-up check-in reported being satisfied with their treatment. A total of 4034 (2.7%, n=151,352, 95% CI [2.6%, 2.8%]) reported experiencing side effects.

Of the 151,352 patients who completed the follow-up check-in, 138,645 had been prescribed the 0.3% topical finasteride and 6% minoxidil spray; 10,774 had been prescribed the 0.3% topical finasteride, 7% minoxidil, 2.2% ketoconazole, and 0.2% biotin spray; and 1933 had been prescribed the 0.3% topical finasteride and 6% minoxidil serum. Table 1 outlines treatment satisfaction and the frequency of side effects reported by patients receiving each treatment.

Table . Treatment satisfaction and frequency of side effects reported by patients during follow-up check-ins.

	All topical finasteride treatments (n=151,352)	Topical finasteride (0.3%) and minoxidil (6%) spray (n=138,645)	Topical finasteride (0.3%), minoxidil (7%), ketoconazole (2.2%), and biotin (0.2%) spray (n=10,774)	Topical finasteride (0.3%) and minoxidil (6%) serum (n=1933)
Treatment satisfaction, n (%)	121,615 (80.4)	111,165 (80.2)	8900 (82.6)	1550 (80.2)
Experienced side effects, n (%)	4034 (2.7)	3716 (2.7)	251 (2.3)	67 (3.5)

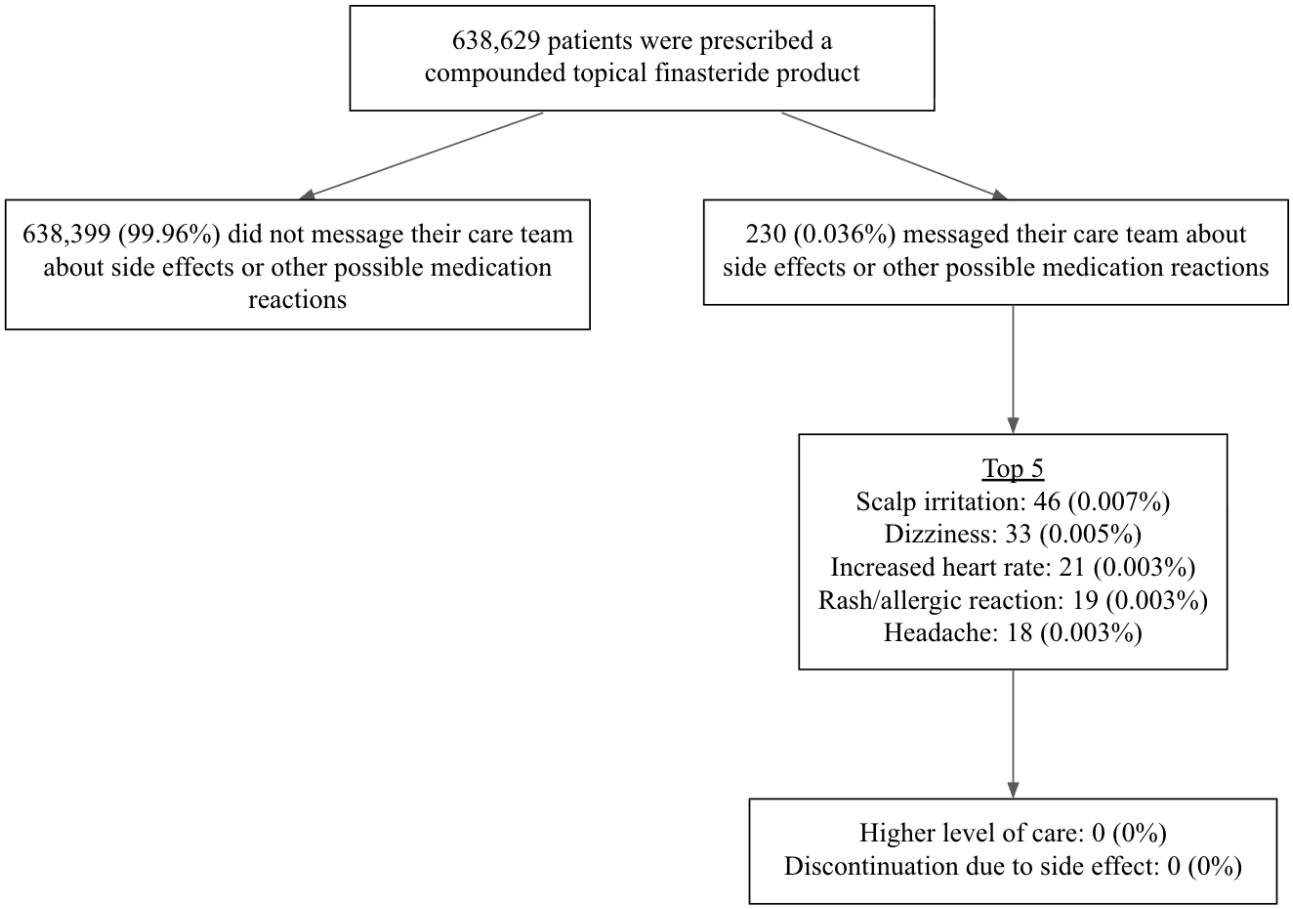
XSL•FO
RenderX

Side Effects and Other Possible Medication Reactions Reported in Patients' Communications to Their Care Team

Of the 638,629 patients prescribed a compounded topical finasteride and minoxidil product, 230 (0.04%, n=638,629, 95% CI [0.035%, 0.045%]) sent their care team messages concerning side effects or other possible medication reactions. The most commonly reported occurrences were scalp irritation (46/638,629, 0.007%, 95% CI [0.0064%, 0.0076%]), dizziness (33/638,629, 0.005%, 95% CI [0.0045%, 0.0055%]), increased heart rate (21/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]), rash or some allergic reaction (19/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]), and headache (18/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]). Sexual side effects, specifically decreased libido and erectile dysfunction, were reported by 12/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Depression was reported by 13/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Anxiety was reported by 10/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Cognitive concerns were reported by 10/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]).

No patients reported seeking a higher level of care (eg, emergency room or urgent care visit) related to a side effect or other possible medication reaction. No patients reported discontinuing treatment due to such an occurrence. During the study period, 1 spouse reported the death of a partner. Upon follow-up, no cause was identified and no causality was established. Figure 1 provides a summary of the side effects and other possible medication reactions reported by patients via messaging.

Figure 1. Schematic of side effects and other possible medication reactions reported by patients in messages to their care team.



Discussion

In this largest study of patient satisfaction and tolerability associated with the use of novel compounded formulations of topical finasteride and minoxidil, we found that 80% of those who completed a follow-up check-in reported satisfaction with treatment and less than 3% reported experiencing side effects. An additional 0.04% of patients sent their care team messages concerning side effects or other medication reactions. The most common reactions appeared to fall into one of two categories: (1) scalp irritation and rash, likely associated with the route of administration; (2) dizziness, increased heart rate, and headache, likely attributable to minoxidil acting as a vasodilator. Of note, sexual side effects, depression, anxiety, and cognitive concerns previously associated with oral finasteride were reported by just 0.002% of patients. There were no reports of "post-finasteride syndrome" [12].

Early clinical trials of 1 mg oral finasteride for the treatment of male AGA found that 3.8% of participants experienced adverse events possibly, probably, or definitely related to treatment, specifically decreased libido, erectile dysfunction, and ejaculation disorder, and 1.4% discontinued treatment due to such adverse events [13]. Trials of 2% topical minoxidil for the treatment of male AGA found that the most common adverse events were minor respiratory events such as colds and

XSL•FO
RenderX

respiratory infections (3.37% of participants), followed by dermatological reactions such as itching (1.94%) [14]. Trials of 5% topical minoxidil for the treatment of male AGA found that headache was the most frequently reported adverse drug reaction (1.7%), followed by dermatological reactions such as pruritus (1.1%) and rash (1.1%) [14].

A comparison of our findings to the findings of these historic studies reinforces the favorable tolerability profile of topical medications. Altogether, these results demonstrate that the novel compounded formulations of topical finasteride and minoxidil available to male patients with AGA via the Hims & Hers platform are associated with high satisfaction among patients and few reported side effects.

To date, few clinical trials have examined the use of topical finasteride in the treatment of male AGA [10,11,15]. A Phase III RCT by the Topical Finasteride Study Group in Europe found that 41.4% of participants reported treatment-emergent adverse events and 9.9% experienced treatment-related adverse events [10]. Another Phase III RCT in China found that 68.4% of participants reported treatment-emergent adverse events and 8.3% experienced treatment-related adverse events [11]. In both studies, the frequency of adverse events among participants using topical finasteride was similar to those using placebo. A retrospective study of 238 patients who received topical finasteride via a German direct-to-consumer teledermatology platform and completed a 6-week follow-up questionnaire found that 11.8% of patients reported adverse events after initiating the use of topical finasteride [15].

However, the aforementioned studies are methodologically limited by their relatively small sample sizes. This study, which included over 600,000 patients who were prescribed compounded topical finasteride in a real-world context, offers a much more robust and meaningful assessment of patient-reported satisfaction and tolerability associated with treatment.

There are limitations of this analysis. First, this was a retrospective analysis of data collected during the course of routine care and not an RCT, and therefore, we cannot confirm any causal relationships between patients' use of compounded topical finasteride and minoxidil and the reported outcomes. Second, we partly relied on data from an optional follow-up check-in questionnaire sent to patients approximately 130 days after treatment initiation. The rate of check-in completion was relatively low, with 23.7% of patients completing the check-in. This may indicate some selection bias, such that patients who were more engaged in or satisfied with their treatment may have been more likely to respond to the check-in and less likely to report side effects. Patients who reported side effects or other reactions to outside health care providers may not have been captured. Third, our reliance on retrospective data meant that we were unable to systematically examine other data of interest, such as the severity of and types of intervention sought for side effects and other medication reactions reported by patients.

However, our analysis also had several strengths. First, our sample size was impressive, with 638,629 patients prescribed a compounded topical finasteride product, all of whom had the ability to communicate with their care team at any time during the course of treatment, and 151,352 of whom completed the follow-up check-in that specifically queried patients about their experience with treatment and side effects. Second, our analysis utilized real-world data. The use of real-world data enables clinicians and researchers to better understand how patients experience treatment in their daily lives, thus increasing the generalizability of results. Third, in addition to relying on the optional follow-up check-in questionnaire to collect data on patient-reported side effects, we were also able to utilize unsolicited patient communications concerning side effects and other possible medication reactions. Having these additional data increased the likelihood that we were able to capture all occurrences reported by patients.

In conclusion, our analysis found that patients prescribed novel formulations of compounded topical finasteride and minoxidil for the treatment of AGA via a national telehealth platform tolerated the treatment well. The majority reported satisfaction with the treatment, and there were few reports of side effects. Future research should include RCTs to assess the efficacy, safety, and tolerability of topical finasteride. Together, this work may help provide more treatment options for those with AGA.

Acknowledgments

No generative AI or AI-assisted technologies were utilized in performing the study or preparing the manuscript.

Funding

This research was supported by Hims & Hers Health, Inc., which did not restrict data analysis, interpretation, or publication.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to commercial restrictions; however, they may be available from the corresponding author upon reasonable request.

Conflicts of Interest

JY, SM ME, and PC are full-time employees of Hims & Hers Health, Inc. JK and AM serve as advisors to Hims & Hers Health, Inc. Hims & Hers Health, Inc. had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

XSL•FO
RenderX

References

1. Lolli F, Pallotti F, Rossi A, et al. Androgenetic alopecia: a review. Endocrine 2017 Jul;57(1):9-17. [doi: 10.1007/s12020-017-1280-y] [Medline: 28349362]
2. Androgenetic alopecia. MedlinePlus. URL: https://medlineplus.gov/genetics/condition/androgenetic-alopecia/#resources [accessed 2026-01-13]
3. Aukerman EL, Jafferany M. The psychological consequences of androgenetic alopecia: a systematic review. J Cosmet Dermatol 2023 Jan;22(1):89-95. [doi: 10.1111/jocd.14983] [Medline: 35403805]
4. Olsen EA, Dunlap FE, Funicella T, et al. A randomized clinical trial of 5% topical minoxidil versus 2% topical minoxidil and placebo in the treatment of androgenetic alopecia in men. J Am Acad Dermatol 2002 Sep;47(3):377-385. [doi: 10.1067/mjd.2002.124088] [Medline: 12196747]
5. Nestor MS, Ablon G, Gade A, Han H, Fischer DL. Treatment options for androgenetic alopecia: efficacy, side effects, compliance, financial considerations, and ethics. J Cosmet Dermatol 2021 Dec;20(12):3759-3781. [doi: 10.1111/jocd.14537] [Medline: 34741573]
6. Andy G, John M, Mirna S, et al. Controversies in the treatment of androgenetic alopecia: the history of finasteride. Dermatol Ther 2019 Mar;32(2):e12647. [doi: 10.1111/dth.12647] [Medline: 30253001]
7. Gupta AK, Bamimore MA, Williams G, Talukder M. Finasteride use: evaluation of depression and suicide risk. J Cosmet Dermatol 2025 Mar;24(3):e70102. [doi: 10.1111/jocd.70102] [Medline: 40082195]
8. Neubauer Z, Ong MM, Lipner SR. No increased risk of mood disorders in male androgenetic alopecia patients treated with dutasteride and finasteride: a retrospective cohort database study. J Am Acad Dermatol 2025 Jul;93(1):197-199. [doi: 10.1016/j.jaad.2025.03.068] [Medline: 40158535]
9. Lee SW, Juhasz M, Mobasher P, Ekelem C, Mesinkovska NA. A systematic review of topical finasteride in the treatment of androgenetic alopecia in men and women. J Drugs Dermatol 2018 Apr 1;17(4):457-463. [Medline: 29601622]
10. Piraccini BM, Blume-Peytavi U, Scarci F, et al. Efficacy and safety of topical finasteride spray solution for male androgenetic alopecia: a phase III, randomized, controlled clinical trial. J Eur Acad Dermatol Venereol 2022 Feb;36(2):286-294. [doi: 10.1111/jdv.17738] [Medline: 34634163]
11. Zhou C, Yang B, Zeng H, et al. Efficacy and safety of topical finasteride spray solution in the treatment of Chinese men with androgenetic alopecia: a phase III, multicenter, randomized, double-blind, placebo-controlled study. Chin Med J (Engl) 2025 Mar 17. [doi: 10.1097/CM9.0000000000003495] [Medline: 40090937]
12. Diviccaro S, Melcangi RC, Giatti S. Post-finasteride syndrome: an emerging clinical problem. Neurobiol Stress 2020 May;12:100209. [doi: 10.1016/j.ynstr.2019.100209] [Medline: 32435662]
13. Propecia (finasteride): prescribing information. : Merck & Co, Inc; 2012 URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/020788s020s021s023lbl.pdf [accessed 2026-01-13]
14. Rogaine (minoxidil): product monograph. : Johnson & Johnson, Inc; 2017 URL: https://pdf.hres.ca/dpd_pm/00040937.PDF [accessed 2026-01-13]
15. Abeck F, Hansen I, Kött J, et al. Patient-reported outcomes of topical finasteride/minoxidil treatment for male androgenetic alopecia: a retrospective study using telemedical data. J Cosmet Dermatol 2024 Sep;23(9):2956-2963. [doi: 10.1111/jocd.16360] [Medline: 38713003]

ABBREVIATIONS

AGA: androgenetic alopecia
FDA: Food and Drug Administration
RCT: randomized controlled trial

Edited by R Dellavalle; submitted 23.Sep.2025; peer-reviewed by EM Hassan, J Jasso-Olivares; revised version received 08.Dec.2025; accepted 17.Dec.2025; published 22.Jan.2026.

Please cite as:
Yu J, Mochida S, Emery M, Carroll P, Ko J, Mostaghimi A
Patient Satisfaction, Side Effects, and Other Reactions Reported by Adult Men Prescribed Compounded Topical Finasteride via a National Telehealth Platform: Retrospective Analysis of Real-World Data
JMIR Dermatol 2026;9:e84676
URL: https://derma.jmir.org/2026/1/e84676
doi:10.2196/84676

© Jessica Yu, Sachie Mochida, Michele Emery, Patrick Carroll, Justin Ko, Arash Mostaghimi. Originally published in JMIR Dermatology (http://derma.jmir.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative

XSL•FO
RenderX

Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Community-Based Teledermatology for Urgent Suspected Skin Cancer: Health Economic Cost-Comparison and Discrete Event Simulation Study

Tim C H Hoogenboom[1], PhD; Pablo García Martínez[2], PhD; Piyush Mahapatra[1], MA, MBBS, FRCS; Nurul Ain Nizar[1], MPH

[1]Open Medical Ltd, 97-107 Uxbridge Rd, London, United Kingdom

[2]Department of Dermatology, Salford Care Organisation part of the Northern Care Alliance NHS Foundation Trust, Manchester, UK, Manchester, United Kingdom

Corresponding Author:
Nurul Ain Nizar, MPH
Open Medical Ltd, 97-107 Uxbridge Rd, London, United Kingdom

Abstract

Background: The increasing incidence and financial burden of skin cancer place immense pressure on the UK's National Health Service (NHS). Systemic challenges, including dermatologist shortages and long waiting lists, complicate timely assessment of skin lesions for patients under the urgent suspected cancer pathway. While teledermatology offers an innovative solution compared to traditional face-to-face appointments, standard teledermatology models still face limitations in addressing health care access barriers. Community-based decentralized models may reduce such barriers, but the cost and operational impact of such specific models remain largely underresearched.

Objective: This study evaluated the differences in financial cost to the NHS and patient waiting times at the Northern Care Alliance NHS Foundation Trust by comparing a community-based teledermatology model using Pathpoint eDerma against the Trust's standard-of-care for patients in the urgent suspected skin cancer pathway.

Methods: This study used an ambidirectional design involving 2 distinct analyses. The cost comparison analysis (CCA) compared costs incurred under the teledermatology model (intervention arm, n=563) against the Trust's standard care, represented by a synthetic comparator arm (n=4011). The discrete event simulation (DES) modeled the operational impact on patient waiting times over a 1-year period. Data for the intervention arm were collected prospectively from December 2022 to May 2023 for CCA and up to November 2023 for DES, while comparator data were collected retrospectively from September 2021 to December 2022. Publicly available resource costs were incorporated to ensure the robustness of the analyses.

Results: The community-based teledermatology model was associated with significant improvements in both cost to the NHS and patient waiting times. The CCA revealed a mean cost saving of £45 (£1=US $1.24) per referral (95% CI £22-£60; *P*<.001). This cost saving was associated with a 26% reduction in the proportion of patients requiring a full diagnostic biopsy, falling from 48% (1925/4011) in standard care to 22% (124/563) in the teledermatology model as well as time savings in face-to-face clinics and administration. Furthermore, the DES demonstrated that, on average, the teledermatology pathways decreased the time to reach a clinical diagnosis by 9.90 (95% CI 9.64-10.16) days; to communicate a diagnosis to patients by 54.18 (95% CI 50.76-57.61) days; and to reach a histopathological diagnosis by 62.8 (95% CI 59.76-65.83) days compared to standard care.

Conclusions: The implementation of the community-based teledermatology model appears to be a highly effective, cost-efficient strategy associated with shortened patient journeys. The intervention showed a faster initial triage phase, but the study identified the histopathology process as the next major systemic constraint that could deter further pathway efficiency. Achieving timely diagnosis for all patients, including those requiring diagnostic biopsies, will necessitate continued strategic investment in innovative technologies to accelerate this downstream process.

(JMIR Dermatol 2026;9:e86402) doi:10.2196/86402

KEYWORDS

teledermatology; urgent suspected cancer; cost-effectiveness; discrete event simulation; patient waiting times; community-based; NHS; National Health Service

XSL•FO
RenderX

Introduction

The incidence of skin cancer continues to rise steadily in the United Kingdom. The number of nonmelanoma skin cancer diagnoses surpasses the combined total of the 4 most prevalent nonskin cancers: breast, prostate, lung, and colorectal cancer [1], with malignant melanoma currently listed as the fifth most common cancer [2,3]. The increase in public awareness of skin cancers [4] and the lack of training for general practitioners (GPs) to confidently assess skin lesions [5] have also led to a surge in suspicious skin lesion referrals in the past decade [6]. Skin cancer places a significant financial strain on the National Health Service (NHS), with associated costs in England now projected to reach between £338 million (£1=US $1.24) and £465 million, a significant increase from the previous projection of over £180 million in 2020 [4]. Treatment costs for advanced skin cancer, particularly metastatic melanoma, can be substantial, exceeding £200,000 per case, while early-stage disease has a drastically lower health impact and treatment cost, highlighting the importance of early diagnosis both for the patient and the NHS [7,8].

The rising burden necessitates early detection and management, which, in turn, has placed immense pressure on dermatology services. The NHS is already grappling with significant resource constraints and a national shortage of specialist dermatologists [9]. With almost a quarter of consultant posts unfilled, over 380,000 people are waiting more than 18 weeks for a dermatology appointment [9]. The urgent suspected cancer (USC) pathway aims to provide timely access to care and minimize the risk of cancer progression following initial lesion identification [10]. The aim is achieved by prioritizing patients referred under this pathway to meet the 28-day "Faster Diagnosis Standard" (FDS) [10]; however, national statistics reveal a low conversion rate, indicating that many of these skin lesions are benign [11]. The prioritization of referrals under this pathway has strained resources and subsequently delayed care for patients with other serious, noncancerous conditions like eczema and psoriasis [12].

The systemic challenges have led to the growing adoption of teledermatology as an innovative solution to transform service delivery, although the traditional model of lesion assessment through face-to-face (F2F) clinic appointments remains common [10,13]. Teledermatology is most frequently delivered through a store-and-forward (SAF) model, which involves capturing and transmitting skin lesion images and clinical data on digital platforms for remote assessment by dermatologists [14]. A body of evidence has demonstrated the general effectiveness of this SAF model in addressing workforce constraints, reducing waiting times, and improving patient access [15,16].

However, there are operational limitations reported with the conventional SAF model, specifically within the image capture process [10,13,14,17]. Models relying on image capture performed at the secondary care provider level maintain high image quality but limit efficiency gains by requiring patients to travel to a centralized hospital facility [10,13,14,17]. Conversely, referrer (GP) image capture models often suffer from lower image quality, reduced standardization, and limited uptake due to existing constraints on GP capacity [10,13,14,17]. Consequently, the effectiveness of the traditional SAF model is still limited by existing challenges in health care access, including geographical distance and avoidant behaviors that discourage patients from visiting centralized secondary care facilities [10,13,14,17]. Single-site SAF models in particular are also limited in their ability to support regional strategic plans that seek to pool resources and support smaller sites.

To mitigate the challenges inherent in conventional teledermatology, we implemented a community-based teledermatology model, using the Pathpoint eDerma digital platform (Open Medical Ltd). This model was specifically designed to ensure high-quality image capture delivered by a dedicated health care assistant while providing timely and accessible care closer to patients' homes. This approach aligns directly with the NHS's strategic vision to shift "from hospital to community," thereby improving access, quality, standardization, and uptake of the teledermatology service [18]. The platform's data architecture and integration capabilities are also future-proofed to support delivery of more comprehensive regional care models [19].

Currently, the literature lacks a robust, real-world evaluation of a community-based model within the urgent suspected skin cancer pathway in the NHS. This study sought to address this gap by providing a comprehensive analysis of the cost implications and operational efficiency of the community-based teledermatology model from a health care system standpoint. In this study, we compare a community-based teledermatology model using Pathpoint eDerma against the trust's standard of care for patients in the urgent suspected skin cancer pathway to evaluate the associated cost implications for the Northern Care Alliance NHS Foundation Trust (NCAFT) and patients' waiting times.

Methods

Study Setting

The study was conducted within the dermatology department at NCAFT. This trust is one of the largest NHS providers in the country, which delivers health care to over 1 million people across Salford, Oldham, Rochdale, and Bury [20]. As part of a wider teledermatology project launched in December 2022, NCAFT established its first community-based teledermatology service. This model used a community diagnostic center (CDC) in Bury, allowing patients to have their skin lesions photographed closer to home rather than at the traditional location, Salford Royal Hospital. The study population included patients referred by their GPs to the urgent suspected skin cancer pathway (formerly known as the 2-week wait [2WW]), who presented with 2 or fewer suspicious lesions.

Study Design

This study used a comparative, ambidirectional design to evaluate a teledermatology model against local standard care. It consisted of 2 distinct analyses: a cost comparison analysis (CCA) to evaluate the budgetary implications for the NHS trust and a discrete event simulation (DES) to model the differences in patient waiting times. The study is considered ambidirectional

XSL•FO
RenderX

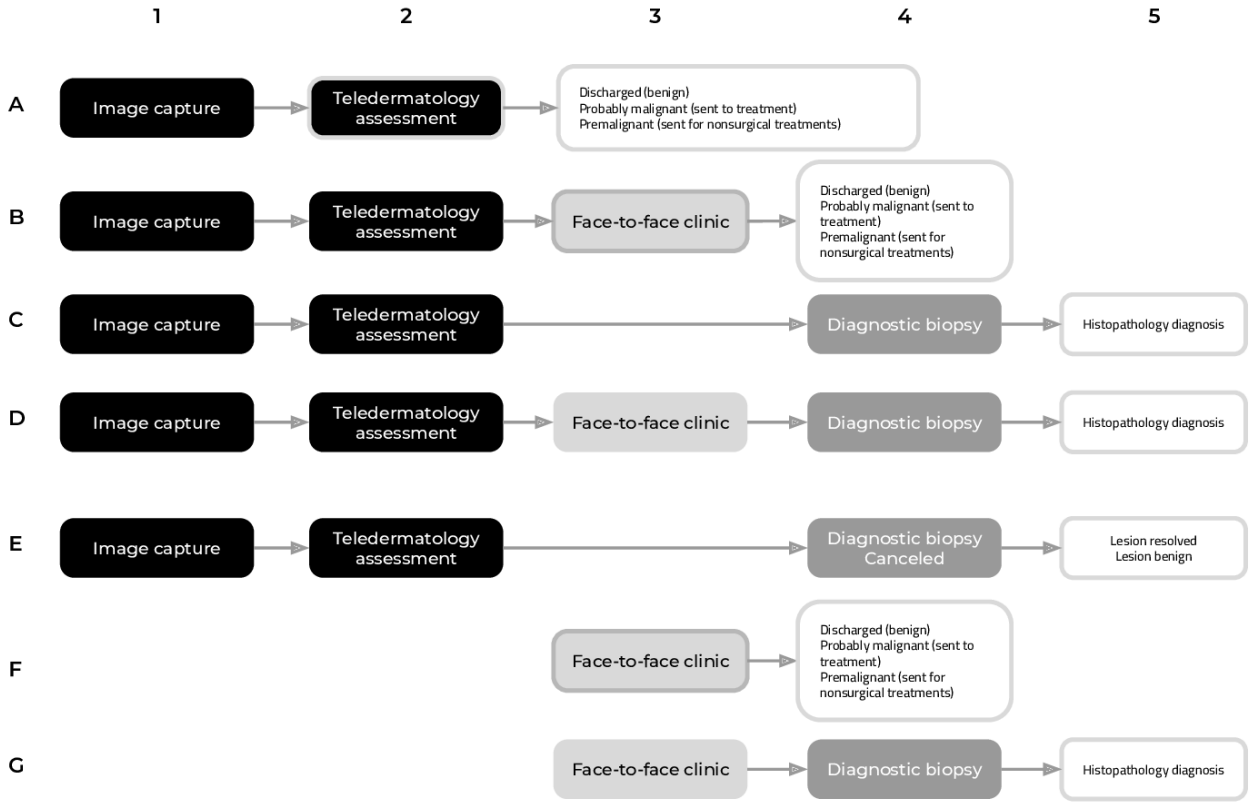
as data for the intervention arm were collected prospectively, while data for the comparator arm were collected retrospectively from historical records. Different time horizons were used in the analyses and are detailed in the data collection section.

Intervention and Comparator Arms

Each arm of the study comprised several distinct pathways representing the patient journey from referral to diagnosis. Under the intervention arm, a patient's journey began with an image capture appointment at the CDC. There, a trained health care assistant captured high-quality macroscopic and dermatoscopic images of the lesion(s) and uploaded them, along with referral notes and a digitally completed patient questionnaire, to the eDerma platform. A consultant dermatologist then remotely reviewed this information to determine the clinical outcome. This process defined 5 distinct intervention pathways (A-E).

In contrast, the comparator arm reflected the local standard-of-care pathways (F and G), where patients attended a F2F clinical appointment with a consultant dermatologist for an in-person dermoscopic examination. Following this consultation, the patient was either discharged (pathway F) or scheduled for a diagnostic biopsy (pathway G). The outcome of a canceled diagnostic biopsy (pathway E) is also a potential event for patients following the standard-of-care F2F pathway, though less frequent and was not observed in the dataset used for this study (Figure 1).

Figure 1. Potential pathways under the intervention and comparator arms.



Data Collection

For the CCA, prospective data for the teledermatology model's intervention arm were sourced from the Pathpoint eDerma platform, spanning from December 2022 to May 2023. Conversely, retrospective data for the standard care comparator arm were collected from historical NCAFT records from September 2021 to December 2022 and supplemented with public NHS protocols. This process created a synthetic control group for the comparator arm, enabling a direct cost comparison between the 2 pathways. Due to the historical nature of the records, this cohort was reconstructed using a triangulated audit approach. A manual review of 70 F2F clinical cases was conducted to determine granular pathway distributions (F and G), which were then calibrated against a larger trust-level dataset of over 5000 urgent referrals to establish representative waiting times. The comparator arm's period was also chosen to ensure it reflected a post-lockdown health care environment. National NHS data indicated that by late 2021, USC referrals for skin lesions had returned to, or exceeded, prepandemic volumes [21]. Consequently, the analysis assumes that patient reluctance to seek treatment was not a significant factor in the diagnostic timelines observed during the comparator period. Additionally, we included resource utilization data, such as staff labor and procedure costs (eg, diagnostic biopsies), which were adjusted to 2023 values for inflation using the Personal Social Services Research Unit manual [22] and the NHS National Cost Collection schedule [23].

The same retrospective data were then repurposed for the DES to inform the operational parameters of the "as-is" model, effectively defining the simulation's structure and serving as the baseline for comparison. To increase the real-world accuracy of the simulation, we integrated additional data to define specific

XSL•FO
RenderX

parameters. These included patient pathway proportions and communication methods from the eDerma platform, patient inflow rates from NHS England's Cancer Waiting Times Statistics [21], and staff numbers estimated from NHS Hospital & Community Health Service monthly workforce statistics [24]. The simulation of the intervention arm was further strengthened by incorporating an additional 6 months of prospective data, extending its data collection period up to November 2023.

Statistical Analyses

To ensure the study's independence and meet the requirements of the teledermatology project's funders, the CCA and DES were conducted by an external health technology consultancy, Health Tech Enterprise. Although the analyses were performed independently, the authors of this study provided the necessary datasets and contextual information. They also reviewed the methodology to ensure it accurately reflected the specific teledermatology project under evaluation.

The health economic CCA compared and quantified the budgetary impact of the teledermatology model against standard care from the perspective of the NHS. The analysis focused on direct medical costs incurred from the point of referral to the point of diagnosis, including health care staff labor (Bands 3, 5, 9, and Consultant levels), diagnostic biopsy procedures, and technology platform licensing. A health economic model was constructed in Microsoft Excel (Microsoft 365, v2310; Microsoft Corp) to determine the average cost per referral for both the intervention and comparator arms. This was followed by an incremental analysis to account for the SEs associated with health care staff involvement.

To test whether the difference in mean costs was statistically significant, an unpaired *t* test was performed. Further, we conducted a probability sensitivity analysis with 1000 model iterations and a deterministic sensitivity analysis to account for uncertainty in input parameters and evaluate how changes in key parameters, such as the number of biopsies, affected the results.

Moreover, we performed a DES to evaluate the impact on patient waiting times. The skin cancer screening pathways were modeled as an agent-based DES using SIMUL8 software (Professional Edition) over a 1-year time horizon. The model's capacity was based on the availability of health care staff and accounted for the seasonality of patient inflow. The simulation was run for five 1-year instances, with each patient "agent" entering the model upon referral and exiting once a diagnosis was communicated.

The simulation measured 3 key end points, representing different waiting periods: the average time from referral to clinical diagnosis, the average time from referral to histopathological diagnosis (for patients requiring a diagnostic biopsy), and the average time to final diagnosis communication.

Model Assumptions

The CCA was built upon several assumptions regarding patient characteristics and resource utilization to enable a direct cost comparison between the intervention and comparator pathways. It was assumed that patient characteristics were the same in both the intervention and comparator arms. For the standard-of-care (comparator) arm, all consultations were modeled as F2F. Specific durations were assigned to clinical activities; the initial F2F consultation with a consultant dermatologist was set at 20 minutes, while a follow-up clinic appointment after the eDerma assessment was set at 15 minutes. The resource cost model assumed that a photography appointment, conducted by an NHS Band 3 health care assistant at a local GP practice, took 20 minutes, and the remote eDerma assessment by a consultant dermatologist took approximately 4.97 minutes. Furthermore, it was assumed that any dermoscopy examination performed in the comparator arm was included within the F2F consultation time, eliminating the need for an additional photography assessment.

Similarly, the DES relied on several assumptions to model patient flow and system capacity under controlled conditions. The simulation modeled high-risk patients referred via the 2WW pathway with 2 or fewer lesions, reflecting current clinical practice and ensuring a realistic assessment of the current adoption of eDerma technology in the health care setting. This approach provides a baseline understanding of the technology's impact under current conditions and practices.

The simulation also assumed 100% patient retention (no dropouts) and no patient "no shows" for F2F appointments. The model also excluded patient mortality. Operational capacity was modeled based solely on health care staff availability, assuming fixed staffing levels throughout the simulation and excluding equipment or facility constraints. Furthermore, labor was scheduled with staff available on weekdays (8:30 AM to 6:30 PM) and 70% of staff available for reduced weekend shifts, with the exception of GP. Procedures were allocated based on a simplified first-come, first-served basis and patient risk level, disregarding individual patient availability.

Ethical Consideration

The study used anonymized standard care data. The project was reviewed by the NCAFT Research and Innovation Department (22HIP55) and was determined not to require further ethical review. Specifically, the project was registered as a Health Improvement Project (22HIP55) and was determined by the Research and Innovation office not to require NHS Health Research Authority review. This determination was made in accordance with national regulations, including the Data Protection Act [25], the General Data Protection Regulation [26], and applicable Health Research Authority [27] guidance. As this was a retrospective study using anonymized routinely collected data, the requirement for formal informed consent was not required, as confirmed by the NCA Research and Innovation department. For the Patient-Reported Experience Measure surveys, patients were informed of the evaluation's purpose, and participation was entirely voluntary. No compensation or financial incentives were provided to the patients participating in the surveys. To ensure patient privacy, all data were deidentified before being used for analysis under the formal Data Sharing and Processing Agreements. Furthermore, all researchers were issued formal Letters of Access mandating strict adherence to the trust's information governance and

XSL•FO
RenderX

confidentiality policies, and all methods in this study adhered to the ethical principles outlined in the Declaration of Helsinki.

Results

Description of Dataset

This study's dataset included 563 urgent skin lesion referrals (so-called "2WW" referrals) managed via the intervention arm. A synthetic comparator arm of 4011 referrals was generated based on referral-to-diagnosis times from 2 years of urgent skin cancer referrals (>5000 cases), supplemented with more granular information regarding the methods of diagnosis (F and G in Figure 1) that were available from 70 cases. The same data distributions were subsequently used to conduct the DES.

Within the intervention arm cohort, 324 out of 563 (57.5%) were female participants, and this cohort had a mean age of 61.5 (SD 17.6) years at the time of referral (Table 1), consistent with the higher risk of skin cancer in older populations resulting from accumulated, lifelong UV exposure [1,4,10]. While granular demographic data for the synthetic comparator arm were unavailable, clinical comparability between the 2 arms was maintained by ensuring that both cohorts consisted of adult patients referred via the USC pathway, presenting with 2 or fewer suspicious lesions at the trust. Furthermore, a parallel health inequality assessment, conducted independently of the CCA and DES but within the same trust and clinical pathway, confirmed that the local patient population remains sociodemographically consistent. The NCA cohort typically reflects a more socioeconomically advantaged profile, with a median Index of Multiple Deprivation decile of 6 (IQR 3-8) and a mode of 8.

Table . Sociodemographic and clinical characteristics of the intervention arm (N=563).

Characteristic	Value
Sex, n (%)	
Female	324 (57.5)
Male	239 (42.5)
Age (y), mean (SD)	61.5 (17.6)
Referral pathway	Urgent suspected cancer with 2 or fewer suspicious lesions

Cost Comparison Analysis

The analysis demonstrates substantial cost savings for the NCAFT through the implementation of the community-based teledermatology model compared to standard care (Table 2). The probability sensitivity analysis showed a mean savings of £45 (95% CI £22-£60) per referral with the eDerma model. This translates to a potential overall savings of £25,251 (95% CI £12,462-£34,002) for the NCAFT from December 2022 to May 2023. An unpaired *t* test confirmed that the difference in mean costs between the 2 arms was statistically significant (P<.001). Based on 1000 iterations, a cost reduction per referral was observed in 94% of cases.

A subanalysis further revealed the mechanism behind these savings by assessing the distribution of referrals and their associated unit costs (Table 3). The highest costs were incurred in pathways involving a full diagnostic biopsy, specifically pathways C, D, and G. The community-based eDerma model led to an 18% reduction in referrals needing a biopsy, as the percentage of patients in biopsy-reliant pathways fell from 49% (1965/4011) in the comparator arm (pathway G) to 31% (175/563) in the intervention arm (pathways C and D combined). A 1-way deterministic sensitivity analysis supported this finding by showing that if the eDerma system could offset the need for full biopsies, it could lead to a mean cost savings of up to £135 per referral.

Table . Potential cost savings via the eDerma intervention arm.

Economic parameter	Cost savings[a] (comparator arm–eDerma)	95% CI
Potential cost savings per referral	£45	£22-£60
Potential overall cost savings	£25,251	£12,462-£34,002

[a]£1=US $1.24.

XSL•FO
RenderX

Table . Referral percentage and unit costs in each pathway.

Arms and associated pathways	Referral percentage (%)	Unit cost (£)[a]
Intervention arm (eDerma)		
A	44	55
B	21	219
C	17	371
D	14	542
E	3	137
Comparator arm (standard-of-care)		
F	51	163
G	49	364

[a]£1=US $1.24.

Discrete Event Simulation

The DES used a slightly different pathway distribution due to the additional 6 months data for the intervention arm. The simulation demonstrated a similar and even more pronounced effect, with the biopsy rate falling from 48% (1925/4011) (pathway G) in standard care to 22% (pathway C and D) in the teledermatology model (Table 4).

Table . Pathway distribution used in the discrete event simulation.

Arms and associated pathways	Referral percentage (%)
Intervention arm	
A	53
B	24
C	5
D	17
E	1
Comparator arm	
F	52
G	48

Referral to Diagnosis Communication

To determine the overall average waiting times for the eDerma and standard-of-care arms, a weighted average was calculated by combining the average time for each pathway (A-E in the intervention arm vs F and G in the comparator arm) with the proportion of patients using each communication method. Finally, an incremental analysis was performed to compare the average waiting times of the 2 arms.

The methods of diagnosis communication included letters, emails via the eDerma platform, telephone calls, and F2F appointments. Complete results on the average time for each communication method and the weighted average of each pathway are detailed in Multimedia Appendix 1.

The analysis of weighted average times per pathway reveals that pathway A in the intervention arm has the shortest average time from referral to diagnosis communication, at 8 days. In contrast, the pathways involving diagnostic biopsies, pathways C and D from the intervention arm and pathway G from the comparator arm, show substantially longer weighted average times of 52.4 and 131.1 days, respectively.

The overall weighted average time for the intervention arm was 18.97 (SE 0.92) days. Conversely, the standard care arm had a weighted average time of 73.16 (SE 1.48) days, illustrating a more extended duration (Table 5).

Table . Comparison of the overall weighted average time from referral to diagnosis communication.

Referral to diagnosis (d)	Intervention arm	Comparator arm
Weighted average time	18.97	73.16
Weighted standard error	0.92	1.48
Maximum time	109.5	252.3

XSL•FO
RenderX

Referral to Clinical Diagnosis

The DES analysis demonstrated that the eDerma arm reached clinical diagnosis faster, with a mean of 7.38 (95% CI 7.24-7.52) days from the initial referral. In the comparator arm, the mean time to clinical diagnosis was longer, at 17.29 (95% CI 17.07-17.50) days.

Referral to Histopathological Diagnosis

For patients undergoing a biopsy, the histopathological diagnosis occurred, on average, within 66.42 (95% CI 65.33-67.50) days in the eDerma arm (Table 6). This was significantly shorter than the comparator arm, where the average waiting period for a histopathological diagnosis was 129.21 (95% CI 126.38-132.05) days (Table 6).

A subsequent incremental analysis highlighted the impact of the community-based teledermatology model in reducing patients' waiting times. On average, the teledermatology pathways were associated with a decrease in the time required to establish and communicate a skin cancer diagnosis to a patient by 54.18 (95% CI 50.76 to 57.61) days. Furthermore, the average waiting time for a clinical diagnosis and histopathological diagnosis was reduced by 9.90 (95% CI 9.64-10.16) days and 62.80 (95% CI 59.76-65.83) days, respectively. These findings are summarized in Table 7.

Table . Comparison of referral to clinical diagnosis and referral to histopathological diagnosis results.

Parameter	Intervention arm (teledermatology model)	Comparator arm (standard care)
Referral to clinical diagnosis (d)		
Mean (SE)	7.38 (0.07)	17.29 (0.11)
95% CI	7.24-7.52	17.07-17.50
Max	31.67	50.28
Referral to histopathological diagnosis (d)		
Mean (SE)	66.42 (0.55)	129.21 (1.45)
95% CI	65.33-67.50	126.38-132.05
Max	109.41	248.37

Table . Summary of discrete event simulation analysis.

Incremental analysis Δ (waiting time in current care–waiting time in eDerma care)	Mean (95% CI)
Average time to diagnosis communication (d)	54.18 (50.76-57.61)
Average time to clinical diagnosis (d)	9.90 (9.64-10.16)
Average time to histopathological diagnosis (d)	62.80 (59.76-65.83)

Discussion

Principal Findings

The study demonstrated that the community-based teledermatology model, facilitated by Pathpoint eDerma, was associated with significant cost savings of £45 per referral and a 54.18-day reduction in diagnosis communication. This efficiency is driven by the operational shift from a centralized secondary care workflow to a decentralized, community-level approach. By moving imaging to local CDCs, the model effectively diverted referrals away from hospital-based bottlenecks and addressed the "upstream" consultant-capacity shortages, allowing dermatologists to triage cases remotely.

Another driver of these efficiencies was the 26% reduction in the proportion of patients requiring biopsy-reliant pathways, falling from 48% (1925/4011) in standard care to 22% (124/563) in the teledermatology model. The analysis demonstrates that diagnostic biopsies are the most expensive and resource-intensive stage of the urgent skin cancer journey. By facilitating definitive clinical triage through high-quality imaging, the teledermatology model minimizes unnecessary surgical interventions, potentially leading to cost savings of up to £135 per referral when biopsies are offset.

The model successfully mitigates the initial bottleneck associated with consultant capacity, effectively managing high referral volumes. Our DES highlighted that this upstream optimization has, in turn, revealed a downstream constraint within the histopathology service. The massive disparity in waiting times between pathways requiring a biopsy and those that do not highlights a critical resource limitation at the histopathology level. This constraint now represents the primary rate-limiting step impacting referral-to-diagnosis time. Therefore, to realize further gains in operational efficiency and to ensure that the 28-day FDS is consistently achieved, future service improvement initiatives should be directed toward optimizing the capacity and workflow of the histopathology reporting pathway.

Operational solutions, such as the outsourcing of histopathology services, are currently performed in NCAFT following the completion of this study to increase reporting capacity. Looking forward, this constraint could also potentially be resolved through the adoption of technologies like digital and computational pathology [28,29]. Digital pathology (eg, whole

XSL•FO
RenderX

slide imaging) can enable remote review by pathologists [28]. Alternatively, computational pathology, including artificial intelligence, could accelerate diagnostic throughput by providing intelligent triage to prioritize high-risk cases and automated quantification to speed up time-intensive tasks like measuring tumor margins [29].

Policy Implications

The success of this model has significant implications for NHS health system planning. The findings support the NHS England Teledermatology Roadmap [10] by demonstrating that decentralized models can successfully achieve national targets that remain elusive under traditional F2F care. For these models to be scalable, future policies must prioritize high standards of digital interoperability where platforms can securely span the entire referral journey, enabling swift information sharing between GP, patients, and secondary care providers. Furthermore, the transition "from hospital to community" aligns with the NHS 10-year vision [18], suggesting that future capital investment should be directed toward community imaging infrastructure to sustain these efficiency gains.

Limitations and Recommendations

As an ambidirectional study, the use of nonoverlapping time horizons between the study arms introduces a potential risk of temporal confounding. Observed improvements in wait times and biopsy rates may have been partially influenced by systemic shifts in NHS dermatology protocols following the October 2022 guidelines. Nonetheless, it is essential to distinguish between the national mandate provided by these guidelines and the operational execution enabled by the intervention. Although the guidelines formalize the 28-day FDS, national dermatology waiting lists have remained at historic highs due to chronic consultant shortages [9]. Furthermore, national data indicate that by late 2021, USC referrals had already reached record-high volumes [21], suggesting that the comparator period (September 2021 to December 2022) represents a stable baseline of high systemic operational pressure. The 54.18-day reduction in diagnosis communication suggests an operational impact that far exceeds typical year-on-year service fluctuations, indicating that the digital model addressed these persistent pressures more effectively than the traditional F2F standard of care.

Another methodological limitation was the exclusion of false-negative rates and their associated long-term treatment costs from the economic model. While a longitudinal follow-up for this specific cohort was not feasible, SAF teledermatology possessed an established safety profile in clinical literature, demonstrating diagnostic accuracy and sensitivity comparable to traditional F2F triage [10,14,15,30]. To ensure a transparent and unbiased economic evaluation, these potential downstream costs were excluded consistently across both study arms. Consequently, the reported £45 mean cost saving represents a conservative estimate of the direct, front-end budgetary impact on the trust.

The DES relied on approximations for parameters such as histopathological processing times and professional review durations, which may not perfectly mirror real-world variability. Additionally, the simulation did not account for missed appointments or fluctuating staff availability, and the single-site nature of the study may limit the immediate generalizability of these findings to other regions. To address these constraints, future research could use multisite longitudinal designs, such as randomized stepped-wedge trials, to assess the long-term sustainability of these cost savings. Given that various teledermatology models are already deployed across the United Kingdom, a collaborative effort to prospectively collect standardized data for comparative health economic evaluations may represent the most pragmatic next step.

Furthermore, the use of a F2F clinic as the comparator, as opposed to a direct comparison with a traditional single-site SAF teledermatology service, was a deliberate choice to evaluate the model against current local standards. It is crucial to note that the community-based model offers inherent strategic advantages not captured by wait-time and financial metrics alone, such as improved patient access [12], reduced patient costs [31], and its unique suitability for deploying region-wide collaborative care networks that single-site models may struggle to support.

From a clinical perspective, this study focused on patients in the USC pathway with 2 or fewer suspicious lesions, a cohort primarily consisting of benign cases or early-stage malignancies. While this pathway is key for early detection, individuals with late-stage skin cancer are less likely to access care through this specific route. Therefore, further research is required to understand the journeys of patients with advanced disease to identify unique barriers and develop targeted interventions for timely access.

Ultimately, the community-based teledermatology model was developed to resolve NHS dermatology capacity constraints and to serve as a robust mitigation strategy for traditional access barriers. While the primary focus of this study was on the budgetary and operational implications, findings from a parallel Health Inequality Assessment at the same trust confirmed that diagnostic timelines remained equitable across all age and deprivation groups within this model. This suggests that decentralizing care to local CDCs mitigates the travel and mobility burdens that traditionally hinder older populations at higher risk of skin cancer [1,4,10] and those from more deprived socioeconomic backgrounds [31].

Conclusion

Our results suggest that a community-based teledermatology model, using the Pathpoint eDerma platform, can provide an efficient and cost-effective pathway for urgent suspected skin cancer referrals, benefiting both the health care system and patients. By being associated with a 26% reduction in biopsy-reliant pathways and an observed 54.18-day reduction in communication times, the model offers a viable strategy to address critical resource constraints within the NHS. Future practical interventions should focus on expanding community-based imaging infrastructure to reduce travel burdens for high-risk older populations, while exploring the potential of digital pathology and other operational solutions to resolve the remaining downstream bottlenecks identified in this analysis.

XSL•FO
RenderX

Acknowledgments

The authors would like to extend their sincere gratitude to the Open Medical Patient and Public Involvement group for their invaluable insights and feedback. Their shared experiences, particularly regarding patient waiting times, underscored the importance of this study from a patient's perspective. They also thank Atlas Naqvi and Vasilii Tereshchenko, who played an instrumental role in setting up the teledermatology service at the Northern Care Alliance National Health Service (NHS) Foundation Trust.

Funding

This work was commissioned and funded by the NHS Cancer Programme, with the support of Small Business Research Initiative for Healthcare and the NHS Accelerated Access Collaborative. The views expressed in the publication are those of the authors and not necessarily those of the NHS Cancer Programme or its stakeholders (grant SBRIC01P3030).

Data Availability

The datasets used and analyzed for the current study, along with the full discrete event simulation report that includes additional subanalyses, are available from the corresponding author upon reasonable request.

Authors' Contributions

TCHH conceptualized the study in partnership with Health Tech Enterprise and the Patient and Public Involvement and Engagement committee, collected and prepared datasets for analysis, reviewed the methodology and results, and contributed to the manuscript's write-up. NAN performed a comprehensive literature review, prepared the original manuscript, and was responsible for continuous refinement of the manuscript up to the point of publication. PGM reviewed the manuscript draft, offered practical insights into the functioning of the dermatology department at the trust, and potential external influences. PM obtained study funding. All authors have critically reviewed and accepted the final format of the manuscript.

Conflicts of Interest

Authors NAN, TCHH, and PM declare a competing interest in their employment by Open Medical Ltd.

Multimedia Appendix 1
Average time from referral to diagnosis per communication methods and weighted average time per pathway.
[PDF File, 85 KB - derma_v9i1e86402_app1.pdf]

References

1. Non-melanoma skin cancer statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-melanoma-skin-cancer [accessed 2025-10-01]
2. Melanoma skin cancer statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer#melanoma_stats0 [accessed 2025-10-01]
3. Chen K, Liu X. Global burden of skin cancer and its subtypes: a comprehensive analysis from 1990 to 2021 with projections to 2040. Front Public Health 2025;13:1610661. [doi: 10.3389/fpubh.2025.1610661] [Medline: 40977764]
4. Skin cancer campaign. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/awareness-and-prevention/be-clear-on-cancer/skin-cancer-campaign [accessed 2025-10-23]
5. Shivakumar H, Chanda UL, Onwuchekwa O. Evaluating the diagnostic accuracy and challenges of the two-week wait referral pathway for skin cancers in primary care. Cureus 2025 Jan;17(1):e77410. [doi: 10.7759/cureus.77410] [Medline: 39811725]
6. Early cancer diagnosis data hub. Early Diagnosis Hub. URL: https://crukcancerintelligence.shinyapps.io/EarlyDiagnosis/ [accessed 2025-10-23]
7. Goon PKC, Greenberg DC, Igali L, Levell NJ. Predicted cases of U.K. skin squamous cell carcinoma and basal cell carcinoma in 2020 and 2025: horizon planning for National Health Service dermatology and dermatopathology. Br J Dermatol 2017 May;176(5):1351-1353. [doi: 10.1111/bjd.15110] [Medline: 27714753]
8. Cutting VAT from sunscreen could save the NHS over £128 million per year in cancer treatment costs. Melanoma Focus. URL: https://melanomafocus.org/news-blog/cutting-vat-from-sunscreen-could-save-the-nhs-over-128-million-per-year-in-cancer-treatment-costs/ [accessed 2025-10-01]
9. Dermatology overview. : Getting It Right First Time (GIRFT) and NHS England & NHS Improvement; 2021 URL: https://www.gettingitrightfirsttime.co.uk/wp-content/uploads/2021/11/Dermatology-overview.pdf [accessed 2025-10-01]
10. A teledermatology roadmap: implementing safe and effective teledermatology triage pathways and processes. National Health Service England. URL: https://www.england.nhs.uk/long-read/

XSL•FO
RenderX

a-teledermatology-roadmap-implementing-safe-and-effective-teledermatology-triage-pathways-and-processes/ [accessed 2025-10-01]
11. Diagnostic waiting times and activity. National Health Service England. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/ [accessed 2025-09-09]
12. Referral optimisation for people with skin conditions. : National Health Service England; 2022 URL: https://www.england.nhs.uk/wp-content/uploads/2022/09/B1149-referral-optimisation-for-people-with-skin-conditions.pdf [accessed 2025-10-01]
13. Quality standards for teledermatology. : British Association of Dermatologists; 2022 URL: https://cdn.bad.org.uk/uploads/2022/02/29200021/Teledermatology-Quality-Standards.pdf [accessed 2025-10-01]
14. Jiang SW, Flynn MS, Kwock JT, Nicholas MW. Store-and-forward images in teledermatology: narrative literature review. JMIR Dermatol 2022;5(3):e37517. [doi: 10.2196/37517] [Medline: 35891983]
15. López-Liria R, Valverde-Martínez M, López-Villegas A, et al. Teledermatology versus face-to-face dermatology: an analysis of cost-effectiveness from eight studies from Europe and the United States. Int J Environ Res Public Health 2022 Feb 22;19(5):2534. [doi: 10.3390/ijerph19052534] [Medline: 35270227]
16. McKoy K, Halpern S, Mutyambizi K. International teledermatology review. Curr Dermatol Rep 2021;10(3):55-66. [doi: 10.1007/s13671-021-00333-6] [Medline: 34341713]
17. Service guidance and standards for the use of teledermatology. : British Association of Dermatologists; 2024 URL: https://cdn.bad.org.uk/uploads/2024/12/16150109/Service-Guidance-and-Standards-for-Teledermatology-2024.pdf [accessed 2025-10-01]
18. The NHS long term plan. : National Health Service England; 2019 URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20230418155402/https:/www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ [accessed 2025-10-01]
19. New digital platform set to transform urgent suspected skin cancer referrals in Lancashire and South Cumbria. Lancashire and South Cumbria Cancer Alliance. URL: https://www.lancashireandsouthcumbria.icb.nhs.uk/our-work/canceralliance/latest-news/new-digital-platform-set-transform-urgent-suspected-skin-cancer-referrals-lancashire-and-south-cumbria [accessed 2025-10-17]
20. Northern Care Alliance NHS Foundation Trust. URL: https://www.northerncarealliance.nhs.uk/about-us [accessed 2025-03-27]
21. 2022-23 monthly provider cancer waiting times statistics. National Health Service England. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/historic-cwt-data-before-september-2023/monthly-prov-cwt/2022-23-monthly-provider-cancer-waiting-times-statistics/ [accessed 2024-01-16]
22. Jones KC, Weatherly H, Birch S, et al. Unit costs of health and social care 2022 manual. : Personal Social Services Research Unit, University of Kent at Canterbury; 2022 URL: https://kar.kent.ac.uk/100519/1/Unit%20Costs%20of%20health%20and%20Social%20Care%202022%20%28amended%2013%20July%202023%29.pdf [accessed 2026-03-17] [doi: 10.22024/UNIKENT/01.02.100519]
23. National cost collection for the NHS. National Health Service England. URL: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ [accessed 2023-03-15]
24. NHS workforce statistics - July 2023. National Health Service England. 2023. URL: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/july-2023 [accessed 2023-12-15]
25. Data Protection Act 2018. legislation.gov.uk. URL: https://www.legislation.gov.uk/ukpga/2018/12/contents [accessed 2026-03-10]
26. European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). : Publications Office of the European Union URL: http://data.europa.eu/eli/reg/2016/679/oj [accessed 2026-03-19]
27. Is my study research? Health Research Authority. URL: http://www.hra-decisiontools.org.uk/research/ [accessed 2026-03-10]
28. Zarella MD, McClintock DS, Batra H, et al. Artificial intelligence and digital pathology: clinical promise and deployment considerations. J Med Imaging (Bellingham) 2023 Sep;10(5):051802. [doi: 10.1117/1.JMI.10.5.051802] [Medline: 37528811]
29. Kim I, Kang K, Song Y, Kim TJ. Application of artificial intelligence in pathology: trends and challenges. Diagnostics (Basel) 2022 Nov 15;12(11):2794. [doi: 10.3390/diagnostics12112794] [Medline: 36428854]
30. Chuchu N, Dinnes J, Takwoingi Y, et al. Teledermatology for diagnosing skin cancer in adults. Cochrane Database Syst Rev 2018 Dec 4;12(12):CD013193. [doi: 10.1002/14651858.CD013193] [Medline: 30521686]
31. Nizar NA, Farooki R, Mahapatra P, Halpern S, Hoogenboom TCH. Patient cost analysis of a community-based teledermatology service versus conventional outpatient appointments in East Kent: a retrospective study through a societal lens to reduce health inequalities. BMC Health Serv Res 2024 Dec 21;24(1):1632. [doi: 10.1186/s12913-024-12112-7] [Medline: 39709482]

Abbreviations

CCA: cost comparison analysis

XSL•FO
RenderX

CDC: community diagnostic center
DES: discrete event simulation
F2F: face-to-face
FDS: Faster Diagnosis Standard
GP: general practitioner
NCAFT: Northern Care Alliance NHS Foundation Trust
NHS: National Health Service
SAF: store-and-forward
USC: urgent suspected cancer

Edited by O Yilmaz; submitted 23.Oct.2025; peer-reviewed by H Mahler, S Heydari; revised version received 15.Feb.2026; accepted 25.Feb.2026; published 06.Apr.2026.

Please cite as:
Hoogenboom TCH, Martínez PG, Mahapatra P, Nizar NA
Community-Based Teledermatology for Urgent Suspected Skin Cancer: Health Economic Cost-Comparison and Discrete Event Simulation Study
JMIR Dermatol 2026;9:e86402
URL: https://derma.jmir.org/2026/1/e86402
doi:10.2196/86402

© Tim C H Hoogenboom, Pablo García Martínez, Piyush Mahapatra, Nurul Ain Nizar. Originally published in JMIR Dermatology (http://derma.jmir.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication at http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Interest in Stress-Induced Dermatologic Conditions and Psychodermatologic Disorders Before, During, and After the COVID-19 Pandemic: Segmented Regression Analysis

Marisa N Fat[1], BS; Hayden C Johnson[2], BA, MQE; Michael Bernas[1], MS; Cheryl L Hurd[1], MD; Robert C Smith[1], MD

[1]Department of Medicine, Anne Burnett Marion School of Medicine at Texas Christian University, 1100 W Rosedale St, Fort Worth, TX, United States
[2]Department of Graduate Education, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Marisa N Fat, BS
Department of Medicine, Anne Burnett Marion School of Medicine at Texas Christian University, 1100 W Rosedale St, Fort Worth, TX, United States

Abstract

Background: Psychological stress is known to exacerbate dermatologic conditions such as acne, eczema, and compulsive skin behaviors. The COVID-19 pandemic introduced a global stressor with widely experienced psychosocial effects and potential impacts on skin health. This study analyzes US public interest in stress-induced dermatologic conditions and psychodermatologic disorders to test for pandemic-associated shifts in 6 search terms.

Objective: This study aimed to assess whether the onset and resolution of the COVID-19 pandemic were associated with changes in Google Search interest in psychodermatologic and stress-related skin conditions in the United States.

Methods: Monthly relative search volume (RSV; 0-100 scale) for 6 terms ("skin picking," "trichotillomania," "stress rash," "eczema flare up," "dermatillomania," and "anxiety skin picking") was obtained from Google Trends for January 2018 through December 2024 (n=84 monthly observations per term); the 6 primary and 4 secondary exploratory terms were extracted in April 2025 and the 5 synonymic terms in July 2026, using the identical query window. A segmented regression model with break points at the World Health Organization's (WHO) pandemic timeline (March 2020 declaration; May 2023 end of concern) tested for level and slope changes at pandemic onset and post pandemic. Newey-West (heteroskedasticity and autocorrelation consistent) SEs corrected for residual autocorrelation, and joint *F* tests with Bonferroni and Benjamini-Hochberg (false discovery rate [FDR]) corrections assessed whether each period differed from the prior trend. A sensitivity analysis applied the identical model to 5 synonymic terms, each paired with 1 primary term, and to 4 unpaired, secondary, exploratory terms ("acne," "stress acne," "stress hives," and "hair loss from stress").

Results: A total of 2 of the 6 terms showed a statistically significant pandemic-onset effect after correction: "dermatillomania" (level shift +15.20 RSV points; Bonferroni *P*=.007; FDR *P*=.007) and "anxiety skin picking" (slope change +0.70 RSV points/month; Bonferroni *P*=.02; FDR *P*=.01). "Stress rash," "skin picking," and "eczema flare up" reached significance under FDR correction only (*P*=.04, .04, and .03, respectively); "trichotillomania" showed no effect under either correction (FDR *P*=.08). "Anxiety skin picking," "trichotillomania," and "stress rash" showed postpandemic reversals robust to both corrections. The synonymic sensitivity analysis showed weaker, inconsistent support: only 1 of 5 pairs ("dermatillomania"/"excoriation disorder") showed partial agreement under FDR correction, while the remaining 4 diverged, indicating sensitivity to search-term phrasing. The secondary exploratory analysis found no pandemic-onset effects under either correction.

Conclusions: Public search interest in a subset of psychodermatologic conditions, specifically "dermatillomania" and "anxiety skin picking," showed statistically significant shifts temporally coinciding with COVID-19 pandemic onset. A sensitivity analysis using synonym terms did not consistently reproduce these findings, indicating sensitivity to both analytic choices and search-term phrasing. These results support further research into psychodermatology and illustrate how public search behavior can reflect societal stress, though causal interpretation is unwarranted from search-volume data alone.

(JMIR Dermatol 2026;9:e93458) doi:10.2196/93458

KEYWORDS

psychodermatology; psychological stress; COVID-19 pandemic; Google Trends; skin disorders

XSL•FO
RenderX

Introduction

Psychodermatology is an emerging field that bridges psychiatry and dermatology, 2 specialties with distinct variations but notable clinical overlap [1]. It focuses on the complex interplay between the mind and skin, where psychological stress can exacerbate dermatologic conditions and skin conditions can lead to emotional distress. Psychological stress is well established as being associated with common skin-related conditions, including acne, psoriasis, urticaria, and eczema [1,2]. This is because the interaction between the skin and the immune system is affected by stress, causing dysregulation of immune cells within the skin barrier and making patients more prone to developing a skin disorder [3]. Although previous studies have detailed the likely relationship between psychological stress and dermatologic diseases, little epidemiological analysis has been performed to examine the number of those affected [3].

The COVID-19 pandemic, caused by SARS-CoV-2, was associated with psychological distress and mental health concerns among large numbers of individuals [4]. Prolonged lockdowns, reduced social interaction, and restricted access to daily routines and indoor and outdoor activities affected individuals' emotional regulation and overall mental well-being, resulting in a 25% increase in anxiety and depression diagnoses during this time [5]. Despite growing mental health concerns, the link between stress and skin conditions during this period remains underexplored [6]. The COVID-19 pandemic has been established as a possible influence on skin conditions such as acne [7], psoriasis [8], urticaria [9], and trichotillomania [10]; however, the literature is sparse.

This study aimed to evaluate changes in public interest in psychodermatologic conditions in the context of the COVID-19 pandemic in the United States, one of the most impactful public health events in recent history. Applying a segmented regression approach to formally test whether the pandemic period represented a statistically distinguishable shift from pre-existing trends provides insight into the effective use of Google Trends data in this space.

Methods

Data Source and Search Terms

A list of 6 search terms was constructed to represent commonly recognized psychodermatological conditions used in both medical and colloquial terminology [1]: "skin picking," "eczema flare up," "dermatillomania," "anxiety skin picking," "trichotillomania," and "stress rash." Using stress-qualified terms for the inflammatory and eczematous categories rather than the bare terms "rash" and "eczema" was a deliberate choice intended to capture searches more specifically attributed to a stress-related presentation. This choice avoids nonspecific causes such as allergic or infectious causes that bare dermatologic terms would capture.

Google Trends (Alphabet Inc), a publicly accessible website that analyzes the popularity of specific words or phrases entered into Google Search (Alphabet Inc), was used to obtain monthly search volumes for each term from January 2018 to December 2024 (n=84 monthly observations) [11,12]. Google Trends provides a normalized 0 to 100 relative search volume (RSV) scale rather than absolute search counts; 0 indicates a low RSV, and 100 represents the point of peak relative interest within the queried range. Google Trends does not offer absolute query counts, demographic detail, or the exact denominator population for any term [12-15]. These constraints are limitations of the platform rather than a deliberate study design, limiting our ability to report absolute counts alongside percentages that would be conventional for count-based epidemiologic data. We used specific search terms rather than Google Trends "Topics," a related feature that groups multiple related queries and entities under a single label using an internal, undisclosed classification algorithm. Because a Topic's exact query composition is not published by Google, its precise boundaries cannot be independently verified or reported. Specific search terms, by contrast, correspond to an exact, reproducible query string, which we prioritized for methodological transparency at the cost of a possibly narrower capture of related search phrasing; a parallel Topics-based analysis is a reasonable direction for future work.

Data Processing and Cleaning

The monthly data were downloaded and imported using the programming language Python (version 3.9.23; Python Software Foundation). Data were checked for missing values and outliers; no observations required exclusion across the 6 primary terms.

Segmented (Interrupted Time Series) Regression

To distinguish a pandemic-specific shift from a pre-existing linear trend, we specified a segmented regression model with 2 structural break points following the World Health Organization's (WHO) official pandemic timeline: March 2020 (pandemic onset) and May 2023 (WHO Public Health Emergency of International Concern [PHEIC]) [16,17]:

$$RSV=\beta0+\beta1(t)+\beta2(D1)+\beta3(t1)+\beta4(D2)+\beta5(t2)+\varepsilon$$

where β_1 estimates the prepandemic baseline trend. β_2 and β_3 estimate the level and slope change at pandemic onset. β_4 and β_5 estimate the level and slope change at postpandemic onset. *t* is the number of months elapsed since January 2018. *D*1 is a binary indicator for the pandemic or postpandemic period. *t*1 is the number of months elapsed since March 2020 (0 if before the pandemic). *D*2 indicates the postpandemic period as a binary variable. *t*2 is the number of months elapsed since May 2023 (0 for any month before). Newey-West (heteroskedasticity and autocorrelation consistent [HAC]) SEs were used to correct for residual autocorrelation, with a lag length of 3 months selected via the standard rule of thumb for 84 monthly observations, applied consistently across all primary and sensitivity term groups [18,19].

Each term was evaluated using a joint *F* test to test if the pandemic period and postpandemic period differed from the prior trend as a whole. This test assessed whether the pandemic period was statistically distinguishable from the pre-existing trajectory. *P* values for the pandemic effect and postpandemic effect tests were corrected across the 6 terms using both Bonferroni and Benjamini-Hochberg (false discovery rate [FDR]) procedures [20].

XSL•FO
RenderX

Sensitivity Analysis

To assess whether the 6 primary terms analyzed for a pandemic-associated pattern reflected a term selection artifact rather than a broader signal, an identical segmented regression and joint test framework was applied to 2 additional sets of terms. The first set comprised 5 colloquial synonyms of the 6 primary terms: "excoriation disorder" (synonym of "dermatillomania"), "hair pulling" (synonym of "trichotillomania"), "compulsive skin picking" (synonym of "skin picking"), "stress eczema" (synonym of "eczema flare up"), and "itching anxiety" (related to "anxiety skin picking"). Another exploratory sensitivity set comprised 4 additional stress- or skin-adjacent terms that captured a broader range than direct synonyms: "acne," "stress acne," "stress hives," and "hair loss from stress." Both sets were subjected to Bonferroni and FDR corrections to allow a better comparison with the primary 6-term group.

Results

Overview

Table 1 summarizes the segmented regression coefficients and joint significance tests for the 6 primary search terms. Table 2 summarizes the synonymic term sensitivity analysis. Table 3 summarizes the secondary exploratory sensitivity analysis.

Table . Segmented regression results for the 6 primary search terms.

Terms	Baseline β_1		Pandemic level β_2		Pandemic slope $\Delta\beta_3$		Postpandemic level β_4		Postpandemic slope $\Delta\beta_5$		Joint pandemic significance test		Joint postpandemic significance test	
	RSV points/mo	*P* value	RSV points/mo	*P* value	RSV points/mo	*P* value	RSV points/mo	*P* value	RSV points/mo	*P* value	Bonferroni *P* value	FDR[a] *P* value	Bonferroni *P* value	FDR *P* value
Skin picking	0.305	.005	−0.089	.97	0.444	.008	−3.939	.30	−0.436	.03	.19	.04	.32	.08
Eczema flare-up	0.229	.08	−3.394	.42	0.562	.003	−0.626	.82	−0.078	.77	.08	.03	>.99	.93
Dermatillomania	0.063	.48	15.198	<.001	−0.116	.45	−4.753	.35	0.332	.23	.007	.007	>.99	.43
Anxiety skin picking	−0.164	.46	13.118	.02	0.695	.01	−21.645	<.001	−0.896	.007	.02	.01	<.001	.001
Trichotillomania	0.052	.60	5.111	.03	−0.040	.80	−7.990	.06	0.206	.15	.45	.08	.02	.006
Stress rash	1.487	.003	−29.409	.009	−1.196	.01	−1.315	.70	−0.613	<.001	.21	.04	.009	.005

[a]FDR: false discovery rate.

XSL•FO
RenderX

Table . Synonymic term sensitivity analysis: segmented regression joint tests paired with primary terms.

Terms (paired primary terms)	Pandemic period			Postpandemic period			Match?[a]
	P value (raw)	Bonferroni *P* value	FDR[b] *P* value	*P* value (raw)	Bonferroni *P* value	FDR *P* value	
Excoriation disorder (dermatillomania)	.01	.06	.04	.046	.23	.08	Partial
Hair pulling (trichotillomania)	.02	.08	.04	.62	>.99	.78	No
Compulsive skin picking (skin picking)	.59	>.99	.74	<.001	.002	.002	No
Stress eczema (eczema flare-up)	.08	.41	.14	.92	>.99	.92	No
Itching anxiety (anxiety skin picking)	.86	>.99	.86	.002	.009	.004	Partial

[a]The "Match?" column indicates whether each synonym term reached the same statistical conclusion (significant vs not significant) as its paired primary term, with "Partial" meaning the two agreed on only one of the two tested effects (or agreed in direction but not in significance under both corrections), and "No" meaning they reached opposite conclusions.

[b]FDR: false discovery rate.

Table . Secondary exploratory sensitivity analysis: segmented regression joint tests for 4 additional terms.

Terms	Joint pandemic effect			Joint postpandemic effect		
	P value (raw)	Bonferroni *P* value	FDR[a] *P* value (BH[b])	*P* value (raw)	Bonferroni *P* value	FDR *P* value (BH)
Acne	.02	.09	.07	<.001	<.001	<.001
Stress acne	.16	.64	.21	.007	.03	.009
Stress hives	.04	.15	.07	.36	>.99	.36
Hair loss from stress	.72	>.99	.72	<.001	<.001	<.001

[a]FDR: false discovery rate.

[b]BH: Benjamini-Hochberg.

Skin Picking

"Skin picking" showed a significant upward prepandemic baseline trend (β_1=+0.31 RSV points/month; *P*=.005). The slope increased further during the pandemic period (β_3=+0.44 RSV points/month; *P*=.008), and the joint pandemic effect test reached significance under FDR correction (*P*=.04) but not under the stricter Bonferroni correction (*P*=.19). The postpandemic joint test was not significant under either correction (Bonferroni *P*=.32; FDR *P*=.08).

Eczema Flare-Up

"Eczema flare up" searches also trended upward across the study period, although the prepandemic baseline slope was only marginally significant (β_1=+0.23 RSV points/month; *P*=.08). The slope increased further at pandemic onset (β_3=+0.56 RSV points/month; *P*=.003), and the joint pandemic effect test reached significance under FDR correction (*P*=.03) but not under the Bonferroni correction (*P*=.08). This was the same pattern observed for "skin picking." The postpandemic joint test was not significant under either correction (*P*=.93).

Dermatillomania

"Dermatillomania" showed no significant prepandemic baseline trend (β_1=+0.06 RSV points/month; *P*=.48), but a large, statistically significant level shift at pandemic onset (β_2=+15.20 RSV points; *P*<.001). The joint pandemic effect test was significant and survived both Bonferroni (*P*=.007) and FDR (*P*=.007) corrections. This finding was significant under the stricter Bonferroni correction. No significant postpandemic reversal was detected (*P*=.36).

Anxiety Skin Picking

"Anxiety skin picking" showed a significant level shift (β_2=+13.12 RSV points; *P*=.02) and slope increase (β_3=+0.70 RSV points/month; *P*=.01) at pandemic onset, and the joint pandemic effect test survived both Bonferroni (*P*=.02) and FDR (*P*=.01) corrections. This term also showed the clearest reversal

XSL•FO
RenderX

in the dataset: a significant level drop (β_4=−21.64 RSV points; *P*<.001) and slope reversal (β_5=−0.90 RSV points/month; *P*=.007) at postpandemic onset, with the joint postpandemic effect test remaining highly significant after correction (both *P*<.001). “Anxiety skin picking” showed both a pandemic onset rise that was significant under Bonferroni correction and a significant postpandemic reversal.

Trichotillomania

“Trichotillomania” showed no significant prepandemic trend (*P*=.60) and no significant pandemic onset effect. The joint pandemic effect test did not reach significance under either correction (Bonferroni *P*=.45; FDR *P*=.08). A significant postpandemic joint effect was detected instead (joint test Bonferroni *P*=.02; FDR *P*=.006), although neither the level shift (β_4=−7.99 points; *P*=.06) nor the slope change (β_5=+0.21 RSV/points; *P*=.15) was individually significant—the reversal was detectable only when both were considered jointly.

Stress Rash

“Stress rash” exhibited a steep and statistically significant prepandemic baseline rise (β_1=+1.49 RSV points/month; *P*=.003)—largely driven by a sharp increase in 2019, prior to the pandemic. At pandemic onset, both the level (β_2=−29.41 RSV points; *P*=.009) and slope (β_3=−1.20 RSV points/month; *P*=.01) dropped significantly, and the joint pandemic effect test reached significance under FDR correction (*P*=.04) but not under the Bonferroni correction (*P*=.21). A further significant slope reversal occurred in the postpandemic period (β_5=−0.61 RSV points/month; *P*<.001; joint postpandemic effect test significant under both corrections; Bonferroni *P*=.009). Formally modeling this term showed that its prepandemic trajectory was already elevated and that the pandemic period corresponded to a decline rather than an increase. This direction differed from those observed for dermatillomania and anxiety skin picking, and the pandemic-onset finding depended on the correction threshold applied.

Sensitivity Analysis Results: Synonymic Terms

“Excoriation disorder,” the clinical synonym for “dermatillomania,” showed a pandemic onset effect significant under FDR correction (*P*=.04) but not under the Bonferroni correction (*P*=.06), suggesting only partial agreement with “dermatillomania’s” Bonferroni-robust finding. “Hair pulling,” the colloquial synonym for “trichotillomania,” showed a pandemic onset effect significant under FDR correction (*P*=.04) but not under the Bonferroni correction (*P*=.08); “trichotillomania” itself showed no significant effect under either correction, so this pair showed divergent findings. “Compulsive skin picking” showed no significant pandemic onset effect (*P*=.59), while its primary term, “skin picking,” reached FDR significance (*P*=.04). This pair also showed divergent findings. “Stress eczema” likewise showed no significant effect (FDR *P*=.14), while “eczema flare up” reached FDR significance (*P*=.03), indicating another divergence. “Itching anxiety” showed no significant pandemic onset effect (*P*=.86) and therefore did not replicate the Bonferroni-robust effect found for “anxiety skin picking”; it did, however, replicate the postpandemic reversal (Bonferroni *P*=.009; FDR *P*=.004). Overall, only 1 of 5 synonym pairs (“dermatillomania”–“excoriation disorder”) showed even partial agreement for the pandemic onset finding.

Sensitivity Analysis Results: Exploratory Terms

None of the 4 secondary exploratory terms showed a pandemic onset effect surviving either correction: “acne” (Bonferroni *P*=.09; FDR *P*=.07) and “stress hives” (Bonferroni *P*=.15; FDR *P*=.07) came closest but did not reach significance under FDR correction; “stress acne” (FDR *P*=.21) and “hair loss from stress” (FDR *P*=.72) showed no evidence of a pandemic onset effect. Postpandemic reversals, by contrast, were common and robust: “acne,” “stress acne,” and “hair loss from stress” all showed significant postpandemic shifts surviving both corrections (all Bonferroni *P*≤.03), while “stress hives” did not (*P*=.36).

Discussion

Principal Findings

This study identified statistically significant pandemic-associated changes in public search interest in dermatillomania and anxiety skin picking, whereas evidence for other search terms was weaker and more sensitive to our analytic approach. Sensitivity analyses demonstrated that several findings depended on search term selection, highlighting the importance of caution when interpreting Google Trends data.

Psychodermatology bridges the fields of psychiatry and dermatology, focusing on the complex interaction between skin health and psychological states. The American Psychiatric Association suggests that more than one-third of dermatological patients have psychological concerns [21]. The pathophysiological connection between dermatological and psychological conditions is multifactorial, involving the hypothalamic-pituitary-adrenal (HPA) axis, neuroimmune signaling, and peripheral nerve activation [22]. Psychological stress triggered by external events can alter immune function, increase proinflammatory cytokines, or disrupt the skin barrier, exacerbating conditions such as acne, eczema, psoriasis, and pruritus [3,23]. These intertwined pathways emphasize a need for treatment that addresses both dermatologic and mental health concerns in patients.

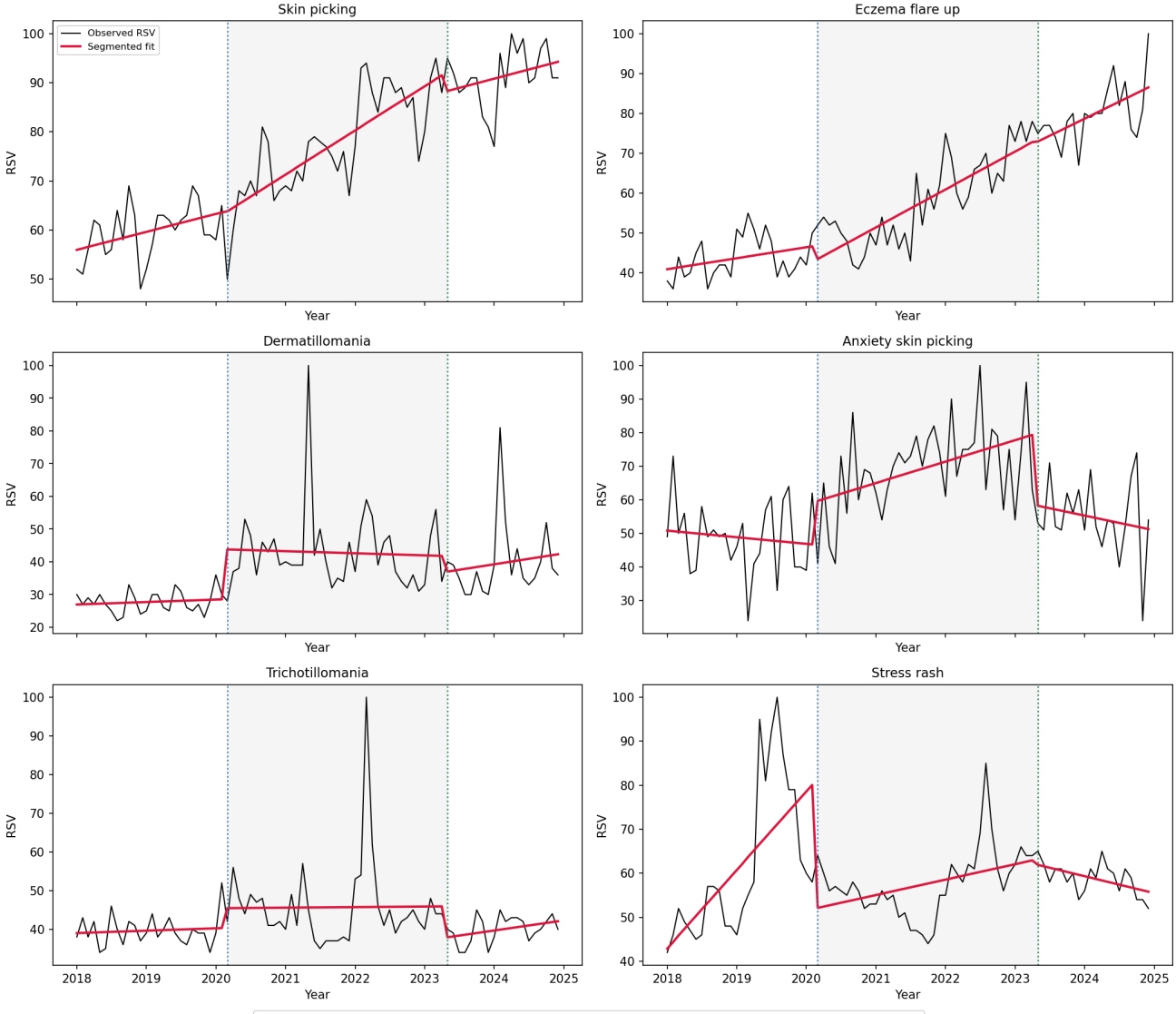
Skin Picking Behaviors

Skin picking behaviors were encompassed by the search terms “skin picking,” “dermatillomania,” and “anxiety skin picking.” “Skin picking” search interest showed a significant pandemic onset increase in RSV (+0.444 slope change) under the FDR correction but not under the Bonferroni correction. “Dermatillomania” and “anxiety skin picking” showed the strongest evidence among the 6 terms for a significant pandemic onset effect under both the FDR and Bonferroni corrections. “Dermatillomania” experienced a large jump in RSV at pandemic onset (+15.198), while “anxiety skin picking” showed an increase in the RSV slope (+0.695) and a large jump in RSV at pandemic onset (+13.118). “Anxiety skin picking” also had a significant postpandemic reversal in the RSV slope (–0.896) and a large decrease in the estimated level of RSV (–21.645). It could be that “dermatillomania” captures a more clinical

XSL•FO
RenderX

lexicon that received increased interest during the pandemic, and "anxiety skin picking" provides more specific language than "skin picking" alone. Because "skin picking" is broad, a more stable base rate of general searches is more difficult to distinguish from its prepandemic trend. Skin picking disorder has documented associations with anxiety, depression, and panic disorder [24], conditions widely reported to have increased during the pandemic [25]; this could be related to the pandemic onset of an increase in RSV, but this is purely speculative. Visual inspection of the raw "skin picking" search data (Figure 1 [16,17]) suggests a recurring decline each December; this is a descriptive observation from the unadjusted data rather than a formally tested seasonal effect (Table 1).

Figure 1. Segmented regression fits overlaid on observed monthly relative search volume (RSV) for the 6 primary search terms from January 2018 to December 2024. The shaded region indicates the pandemic period. The blue dotted line marks the World Health Organization (WHO) pandemic declaration on March 11, 2020 [16], corresponding to the model's first break point; the green dotted line marks the end of the WHO Public Health Emergency of International Concern (PHEIC) on May 5, 2023 [17], corresponding to the model's second break point.



"Skin picking" search interest showed a significant pandemic onset increase in RSV (+0.444 slope change) under the FDR correction but not under the Bonferroni correction. "Dermatillomania" and "anxiety skin picking" showed the strongest evidence among the 6 terms for a significant pandemic onset effect under both the FDR and Bonferroni corrections. "Dermatillomania" experienced a large jump in RSV at pandemic onset (+15.198), while "anxiety skin picking" showed an increase in the RSV slope (+0.695) and a large jump in RSV at pandemic onset (+13.118). "Anxiety skin picking" also had a significant postpandemic reversal in the RSV slope (–0.896) and a large decrease in the estimated level of RSV (–21.645). It could be that "dermatillomania" captures a more clinical lexicon that received increased interest during the pandemic, and "anxiety skin picking" provides more specific language than "skin picking" alone. Because "skin picking" is broad, a more stable base rate of general searches is more difficult to distinguish from its prepandemic trend. Skin picking disorder has documented associations with anxiety, depression, and panic disorder [24], conditions widely reported to have increased during the pandemic [25]; this could be related to the pandemic onset of an increase in RSV, but this is purely speculative. Visual inspection of the raw "skin picking" search data (Figure 1 [16,17]) suggests a recurring decline each December; this is

XSL•FO
RenderX

a descriptive observation from the unadjusted data rather than a formally tested seasonal effect (Table 1).

Hair-Related Stress Responses

Hair-related stress responses were encompassed by the search term "trichotillomania." "Trichotillomania" experienced no significant pandemic onset effect under either correction. The joint postpandemic test was significant under both corrections (Bonferroni P=.02; FDR P=.006), although neither the level shift (β_4=−7.99 RSV points; P=.06) nor the slope change (β_5=+0.21 RSV points/month; P=.15) was individually significant. The reversal was detectable only when both were considered jointly. Body-focused repetitive behaviors increased during the pandemic [26,27], but this dataset does not provide adequate support for the corresponding pandemic onset shift. RSV for "trichotillomania" alone is not a reliable proxy for the symptom timing (Table 1).

Stress-Influenced Inflammatory Dermatoses

Stress-influenced inflammatory dermatoses were encompassed by the search term "stress rash." "Stress rash" RSV showed a significant decrease in the RSV (–1.196) and estimated RSV level (–29.409) at the onset of the pandemic under the FDR correction, reversing its sharp increase in the preceding year, 2019 (Figure 1 [16,17]). The postpandemic era was associated with a further significant decrease under both the FDR and Bonferroni corrections (–0.613; Table 1). One possible explanation is that during the early pandemic, public attention shifted toward COVID-19–related symptoms and information seeking, which may have reduced searches for the specific term "stress rash." Alternatively, the decline may reflect changes in search terminology rather than a true reduction in stress-related dermatologic symptoms. Because Google Trends cannot determine users' motivations, these explanations remain speculative.

Stress-Influenced Chronic Eczematous Dermatoses

Stress-influenced chronic eczematous dermatoses were encompassed by the search term "eczema flare up." A significant increase in the RSV slope was found for "eczema flare up" under the FDR correction but not under the Bonferroni correction (+0.562). Frequent handwashing and psychological stress during the pandemic have been proposed as potential contributors to eczema symptom exacerbation [28,29]. Although the FDR correction found significance, the increase in RSV for "eczema flare up" could have followed the upward trend from the prepandemic era (Figure 1 [16,17]). "Eczema flare up" did not experience any significant change in the RSV slope in the postpandemic era (Table 1).

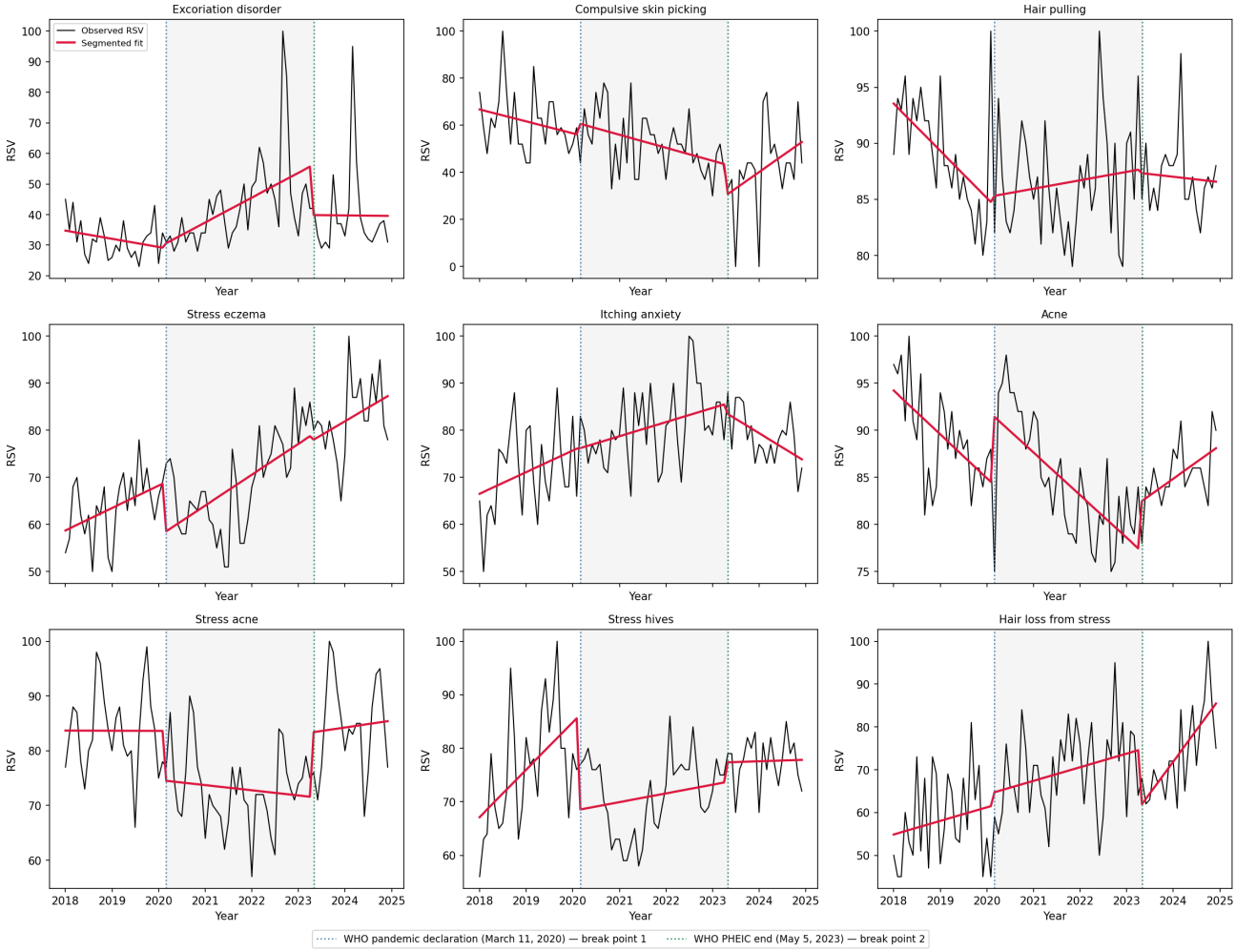
Sensitivity Analysis

A sensitivity analysis for synonyms of the original 6 terms was chosen to test if the primary findings depended on the specific phrasing chosen for each term. Only 1 of 5 synonym pairs ("dermatillomania"–"excoriation disorder") showed partial agreement with its matched primary term for the pandemic onset effect, with significance observed only under the FDR correction. The remaining 4 pairs showed divergent findings for the pandemic onset effect: "hair pulling" showed an FDR-level effect despite "trichotillomania" showing none; "compulsive skin picking" and "stress eczema" showed no effect despite their primary terms ("skin picking" and "eczema flare up") reaching FDR significance; and "itching anxiety" showed no onset effect despite "anxiety skin picking's" Bonferroni finding, although the 2 did agree on a postpandemic reversal. This sensitivity analysis argues against treating the primary pandemic onset findings as robust to search term phrasing and should be reported as a limitation rather than a validation of those findings (Table 2).

The secondary exploratory analysis used general terms to test the robustness of the primary findings. The 4 terms, "acne," "stress acne," "stress hives," and "hair loss from stress," showed no significant pandemic onset effect under either correction. "Acne," "stress acne," and "hair loss from stress" showed significant postpandemic increases (Figure 2 [16,17]). Taken together, the 2 sensitivity analyses provide less support for the robustness of the findings for the primary 6 terms. The strongest and most consistent findings remain the pandemic onset effects that survived Bonferroni correction for "dermatillomania" and "anxiety skin picking," and the postpandemic reversals for "anxiety skin picking," "trichotillomania," and "stress rash."

XSL•FO
RenderX

Figure 2. Segmented regression fits overlaid on observed monthly relative search volume (RSV) for the 9 sensitivity terms (synonymic terms and 4 secondary exploratory terms), January 2018 to December 2024. The shaded region indicates the pandemic period. The blue dotted line marks the World Health Organization (WHO) pandemic declaration on March 11, 2020 [16], corresponding to the model's first break point; the green dotted line marks the end of the WHO Public Health Emergency of International Concern (PHEIC) on May 5, 2023 [17], corresponding to the model's second break point as in Figure 1 [16,17].



Strengths and Limitations

The strengths of this study are the use of a segmented regression framework, anchored by the WHO's official pandemic timeline, that tests for pandemic-associated shifts in RSV. Newey-West SEs address residual autocorrelation, and multiple-comparison corrections were applied across the 6 primary terms. Bonferroni and FDR corrections identify the findings as sensitive to the choice of correction. The sensitivity analyses provide a more comprehensive assessment of the 6 terms to determine whether the findings were affected by specific term selection or wording. The transparency of the sensitivity analysis shows that the primary findings may be inconsistent, but this is itself a strength of reporting.

The limitations are consistent with Google Trends analyses, as the RSV metric does not provide absolute search counts and does not disclose demographic information. Furthermore, it cannot be concluded who is searching for a specific term or why. Although Google accounts for an estimated 87% of internet searches in North America [30], there are other internet search engines that are excluded from the analysis. Internet access and search behavior may vary by region, income, and other potential factors that are not measurable within this dataset. Search interest is a proxy for public attention and cannot be used to infer disease incidence, prevalence, or diagnosis. The segmented regression does not include explicit seasonal (month-of-year) controls; seasonal patterns noted in the Discussion (eg, a December decline in "skin picking" searches) are descriptive observations from the raw data rather than formally tested effects and should be interpreted accordingly. Most importantly, the pandemic onset findings for 4 of the 6 primary terms ("skin picking," "eczema flare up," "trichotillomania," and "stress rash") did not hold up consistently across correction methods or sensitivity analyses. Only "dermatillomania" and "anxiety skin picking" showed pandemic onset effects that were robust across the correction methods and sensitivity analyses, along with the postpandemic reversals for "anxiety skin picking," "trichotillomania," and "stress rash."

Conclusions

We applied a segmented regression model, which adhered to the WHO's official pandemic timeline, to Google Trends data from January 2018 through December 2024. This study finds that public search interest in "dermatillomania" and "anxiety skin picking" showed pandemic onset changes that were robust to multiple corrections. Evidence for "skin picking," "eczema

XSL•FO
RenderX

flare up," and "stress rash" was present only under the FDR correction, and "trichotillomania" showed no pandemic onset effect. Postpandemic reversals were detected for "anxiety skin picking," "trichotillomania," and "stress rash." Sensitivity analyses using synonymic and conceptually related search terms did not consistently reproduce the primary findings, indicating that most of these results are sensitive to search term phrasing and the choice of statistical correction. The findings support continued attention to the intersection of dermatology and mental health and illustrate that formal time series methods are necessary but may not always be sufficient on their own.

Acknowledgments

The authors declare that no generative AI was used in the preparation of this manuscript.

Funding

The authors declare that no external funding was received for this research.

Data Availability

The data analyzed in this study are publicly available through Google Trends.

Authors' Contributions

MNF contributed to conceptualization, methodology, investigation, interpretation of results, and drafting and revision of the manuscript. HCJ contributed to conceptualization, methodology, data acquisition and analysis, interpretation of results, and drafting and revision of the manuscript. CLH and RCS provided clinical expertise, contributed to interpretation of the findings, and critically reviewed and revised the manuscript. MB contributed to interpretation of the findings and critically reviewed and revised the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Jafferany M. Psychodermatology: a guide to understanding common psychocutaneous disorders. Prim Care Companion J Clin Psychiatry 2007;9(3):203-213. [doi: 10.4088/pcc.v09n0306] [Medline: 17632653]
2. Altalhi ER, Felimban SA, Alharbi WS, et al. Association of psychological stress with skin symptoms among the population in Saudi Arabia: a cross-sectional study. Cureus 2023 Nov;15(11):e48657. [doi: 10.7759/cureus.48657] [Medline: 38090468]
3. Zhang H, Wang M, Zhao X, Wang Y, Chen X, Su J. Role of stress in skin diseases: a neuroendocrine-immune interaction view. Brain Behav Immun 2024 Feb;116:286-302. [doi: 10.1016/j.bbi.2023.12.005] [Medline: 38128623]
4. Xiong J, Lipsitz O, Nasri F, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 1;277:55-64. [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
5. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. World Health Organization. 2022 Mar 2. URL: https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide [accessed 2026-07-30]
6. Zhang H, Gu X, Tan Y, et al. Mapping the stress-skin axis: difficulties, strategies & prospects ahead. Innov Life 2025;3(4):100167. [doi: 10.59717/j.xinn-life.2025.100167]
7. Sun M, Guo J, Li Y, et al. Post-COVID-19 syndrome-associated risk factors for acne exacerbation. Curr Med Chem 2026;33(1):210-219. [doi: 10.2174/0109298673345685241114104703] [Medline: 39727175]
8. Shah H, Busquets AC. Psoriasis flares in patients with COVID-19 infection or vaccination: a case series. Cureus 2022 Jun;14(6):e25987. [doi: 10.7759/cureus.25987] [Medline: 35855232]
9. Craffert V, Day C, Peter J. New-onset chronic spontaneous urticaria post-COVID-19 vaccination-South African case series. J Allergy Clin Immunol Glob 2023;2(4):100154. [doi: 10.1016/j.jacig.2023.100154] [Medline: 37781658]
10. Wang X, Zhang B. Trichotillomania occurs during the COVID-19 pandemic in an adolescent. World J Pediatr 2022 Jun;18(6):435-436. [doi: 10.1007/s12519-022-00520-w] [Medline: 35226334]
11. Southern MG. The top 6 search engines market share & the AI search engines to watch. Search Engine Journal. URL: https://web.archive.org/web/20260718041207/https://www.searchenginejournal.com/seo/meet-search-engines/ [accessed 2026-09-24]
12. Bishay AE, Fijany AJ, Holan C, et al. Analyzing Google Search Trends for migraine surgery and Nurtec in response to public announcements. Plast Reconstr Surg Glob Open 2024 Jul;12(7):e5996. [doi: 10.1097/GOX.0000000000005996] [Medline: 39027895]

XSL•FO
RenderX

13. Bourmand R, Olsson SE, Soleimani S, Fijany A. Lip filler versus “lip flip”: longitudinal public interest and a brief review of literature. J Cosmet Dermatol 2025 Feb;24(2):e70048. [doi: 10.1111/jocd.70048] [Medline: 39936357]
14. Olsson SE, Sreepad B, Lee T, Fasih M, Fijany A. Public interest in acetyl hexapeptide-8: longitudinal analysis. JMIR Dermatol 2024 Feb 20;7:e54217. [doi: 10.2196/54217] [Medline: 38376906]
15. Sivesind TE, Szeto MD, Kim W, Dellavalle RP. Google Trends in dermatology: scoping review of the literature. JMIR Dermatol 2021 May 25;4(1):e27712. [doi: 10.2196/27712] [Medline: 37632813]
16. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed 2020 Mar 19;91(1):157-160. [doi: 10.23750/abm.v91i1.9397] [Medline: 32191675]
17. Cheng K, Wu C, Gu S, Lu Y, Wu H, Li C. WHO declares the end of the COVID-19 global health emergency: lessons and recommendations from the perspective of ChatGPT/GPT-4. Int J Surg 2023 Sep;109(9):2859-2862. [doi: 10.1097/JS9.0000000000000521] [Medline: 37246993]
18. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017 Feb 1;46(1):348-355. [doi: 10.1093/ije/dyw098] [Medline: 27283160]
19. Newey WK, West K. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 1987 May;55:703-708. [doi: 10.2307/1913610]
20. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B Stat Methodol 1995 Jan;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
21. Psychodermatology: addressing the skin-mental health connections. American Psychiatric Association. URL: https://www.psychiatry.org/news-room/apa-blogs/psychodermatology-skin-mental-health-connections [accessed 2026-07-30]
22. Arck PC, Slominski A, Theoharides TC, Peters EM, Paus R. Neuroimmunology of stress: skin takes center stage. J Invest Dermatol 2006 Aug;126(8):1697-1704. [doi: 10.1038/sj.jid.5700104] [Medline: 16845409]
23. Chu B, Marwaha K, Sanvictores T, Awosika AO, Ayers D. Physiology, stress reaction. In: StatPearls: StatPearls Publishing; 2024. URL: https://www.ncbi.nlm.nih.gov/books/NBK541120/? [accessed 2026-09-25]
24. Grant JE, Chamberlain SR. Prevalence of skin picking (excoriation) disorder. J Psychiatr Res 2020 Nov;130:57-60. [doi: 10.1016/j.jpsychires.2020.06.033] [Medline: 32781374]
25. Delpino FM, da Silva CN, Jerônimo JS, et al. Prevalence of anxiety during the COVID-19 pandemic: a systematic review and meta-analysis of over 2 million people. J Affect Disord 2022 Dec 1;318:272-282. [doi: 10.1016/j.jad.2022.09.003] [Medline: 36096370]
26. Gu L, Pathoulas JT, Widge AS, Idnani A, Lipner SR. Exacerbation of onychophagia and onychotillomania during the COVID-19 pandemic: a survey-based study. Int J Dermatol 2022 Nov;61(11):e412-e414. [doi: 10.1111/ijd.16395] [Medline: 35948862]
27. Pathoulas JT, Olson SJ, Idnani A, Farah RS, Hordinsky MK, Widge AS. Cross-sectional survey examining skin picking and hair pulling disorders during the COVID-19 pandemic. J Am Acad Dermatol 2021 Mar;84(3):771-773. [doi: 10.1016/j.jaad.2020.11.011] [Medline: 33279648]
28. Loh ED, Yew YW. Hand hygiene and hand eczema: a systematic review and meta-analysis. Contact Dermatitis 2022 Oct;87(4):303-314. [doi: 10.1111/cod.14133] [Medline: 35460528]
29. Suárez AL, Feramisco JD, Koo J, Steinhoff M. Psychoneuroimmunology of psychological stress and atopic dermatitis: pathophysiologic and therapeutic updates. Acta Derm Venereol 2012 Jan;92(1):7-15. [doi: 10.2340/00015555-1188] [Medline: 22101513]
30. Search engine market share North America. StatCounter Global Stats. URL: https://gs.statcounter.com/search-engine-market-share/all/north-america [accessed 2026-07-30]

Abbreviations

FDR: false discovery rate
HAC: heteroskedasticity and autocorrelation consistent
HPA: hypothalamic-pituitary-adrenal
PHEIC: Public Health Emergency of International Concern
RSV: relative search volume
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Meisenheimer; submitted 13.Feb.2026; peer-reviewed by G Malafaia, S Gordon; revised version received 28.Aug.2026; accepted 08.Sep.2026; published 30.Sep.2026.

Please cite as:
Fat MN, Johnson HC, Bernas M, Hurd CL, Smith RC
Public Interest in Stress-Induced Dermatologic Conditions and Psychodermatologic Disorders Before, During, and After the COVID-19 Pandemic: Segmented Regression Analysis
JMIR Dermatol 2026;9:e93458
URL: https://derma.jmir.org/2026/1/e93458
doi:10.2196/93458

© Marisa N Fat, Hayden C Johnson, Michael Bernas, Cheryl L Hurd, Robert C Smith. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Over- and Underfunded Diseases by Comparing National Institutes of Health Funding for Skin Disease Research With US Skin Disease Burden According to 2021 Global Burden of Disease Data: Cross-Sectional Analysis

Aileen Park[1], BS; Emily Woolhiser[2,3], DO; Hannah Riva[4], MPAS; Leo Wan[5], BA; Haaris Kadri[1], BA; Elizabeth Lamberty[1*], BA; Parker Juels[1*], BA; Sandra Jaroonwanichkul[6], BA; Madison Reed[7], BS; Catherine Hegedus[8], BS; Dana Chen[9], BS; Danielle Duffle[8], BS, MS; Jessica Kirk[8], BS; Sydney Christensen[2], BA; Emma Shelby[1], BA; Robert Dellavalle[10], MD, PhD, MSPH

[1]School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[10]Department of Dermatology, School of Medicine, University of Minnesota, Phillips-Wangensteen Building, 516 Delaware St SE Suite 1-400, Minneapolis, MN, United States

[2]Kansas City University College of Osteopathic Medicine, Kansas City, MO, United States

[3]Memorial Hospital System, Pembroke Pines, FL, United States

[4]Paul L. Foster School of Medicine, Texas Tech University, El Paso, TX, United States

[5]West Virginia School of Osteopathic Medicine, Lewisburg, WV, United States

[6]School of Medicine, University of Missouri–Kansas City, Kansas City, MO, United States

[7]Kirksville College of Osteopathic Medicine, A.T. Still University, Kirksville, MO, United States

[8]Rocky Vista University College of Osteopathic Medicine, Parker, CO, United States

[9]A.T. Still University School of Osteopathic Medicine, Mesa, AZ, United States

[*]these authors contributed equally

Corresponding Author:
Robert Dellavalle, MD, PhD, MSPH
Department of Dermatology, School of Medicine, University of Minnesota, Phillips-Wangensteen Building, 516 Delaware St SE Suite 1-400, Minneapolis, MN, United States

Abstract

Background: Understanding the burden of various skin diseases can help guide funding allocation for skin disease research. A 2015 cross-sectional study found a partial correlation between US skin disease burden according to the 2010 Global Burden of Disease (GBD) study and National Institutes of Health (NIH) funding in 2012-2013.

Objective: This study aims to identify trends, correlations, and disparities in US skin disease burden and NIH research funding allocation using the latest data from the GBD 2021 and NIH funding data from the fiscal years 2021-2022.

Methods: A cross-sectional analysis was conducted to compare the disability-adjusted life years for 15 skin conditions from the GBD 2021 with NIH funding for these conditions in 2021-2022. Data were sourced from the GBD Results tool and the NIH RePORTER database.

Results: NIH funding for skin disease research and US skin disease burden according to the GBD 2021 were partially correlated, with several outliers. Malignant skin melanoma and pruritus were relatively overfunded, while psoriasis and urticaria were relatively underfunded.

Conclusions: Disease burden is just one of the many important factors that must be considered when allocating resources, including funding to encourage research efforts to improve patient outcomes and positively impact public health.

(JMIR Dermatol 2026;9:e71468) doi:10.2196/71468

KEYWORDS

epidemiology; burden of disease; disability-adjusted life years; research funding; melanoma; psoriasis; dermatitis

XSL•FO
RenderX

Introduction

The Global Burden of Disease (GBD) study aims to quantify worldwide health losses due to a wide variety of illnesses and injuries [1]. Disease burden is one of many important factors guiding decisions on policy development, disease prevention initiatives, and research funding allocation [1,2]. The GBD study quantifies disease burden using disability-adjusted life years (DALYs), a measure that accounts for both mortality due to disease (years of life lost) and years lived with decreased health and quality of life (years lived with disability; YLDs) [1]. GBD also accounts for the severity of disability (defined by any short-term or long-term loss of health) attributed to the variety of illnesses and injuries included in the study by factoring disability weights into the calculation of YLDs [1,3].

Skin conditions are ubiquitous worldwide and affect millions each year. As a result, dermatology continues to be a consistently innovative field that makes large strides in patient care thanks to a heavy research focus. Public funding is a major contributor to research and innovation in this field. In 2015, Hagstrom and colleagues [4] conducted a cross-sectional study that found a partial correlation between US skin disease burden according to the GBD 2010 and National Institutes of Health (NIH) funding in the fiscal years 2012 - 2013, identifying over- and underfunded diseases. Following this study, there have been major changes to the funding of dermatology research, with a 14.7% inflation-adjusted increase in research funding from 2015 to 2019 and fluctuations in funding after the COVID-19 pandemic [5,6]. This study reinvestigates the relationship between US skin disease burden using the latest GBD 2021 data and NIH funding data for 2021 - 2022.

Methods

Overview

A cross-sectional analysis was conducted to compare DALYs for the 15 skin conditions included in the GBD 2021 with NIH funding for these conditions in 2021 - 2022. Data were sourced from the GBD Results tool [1] and the NIH RePORTER database [7]. The search parameters used in GBD Results to obtain DALY metrics for all 15 aforementioned skin disease categories in the US were as follows: measure="DALYs," metric="number," location="United States of America," age="all ages," sex="both," and year="2021." DALY metrics were specifically gathered for the United States to facilitate a direct comparison between the US-specific burden of skin diseases measured by DALYs and funding allocated by the NIH in the United States for skin disease research.

To compile a comprehensive list of NIH-funded grants awarded for skin disease research during fiscal years 2021-2022, a total of 15 queries were entered into the NIH RePORTER database, with each query corresponding to one of the GBD skin disease categories. The following parameters were used to conduct all 15 of these search queries: fiscal year="2021 and 2022," text search logic="advanced," and limit project search="project title, project terms, and project abstracts." In the Text Search box, all *International Classification of Diseases, 10th Revision* codes categorized by the GBD 2021 under one specific skin disease category were strung with "AND," "OR," or "NOT" as determined necessary to capture all relevant NIH-funded grants.

All titles and abstracts of the grants obtained from NIH RePORTER were manually screened by two independent reviewers to determine inclusion versus exclusion (they were included if the grant studied any 1 of the 15 skin disease categories described by the GBD 2021). Following independent review, inclusion and exclusion decisions were cross-examined to identify conflicting decisions. A third reviewer served as a tie-breaker to resolve any discrepancies as needed.

Statistical analysis was performed assuming that the proportion of DALYs attributed to a disease should be the same as the proportion of NIH skin disease funding it receives (ie, if a specific disease is responsible for 25% of all US skin disease DALYs, that disease should receive 25% of all NIH skin disease funding). A one-to-one trendline was used to visualize this relationship and identify outliers representing relatively over- and underfunded skin diseases. An "observed-to-expected" ratio was calculated by dividing the true amount of funding a disease received by the amount of funding a disease could be expected to receive assuming a one-to-one relationship between DALYs and funding.

Ethical Considerations

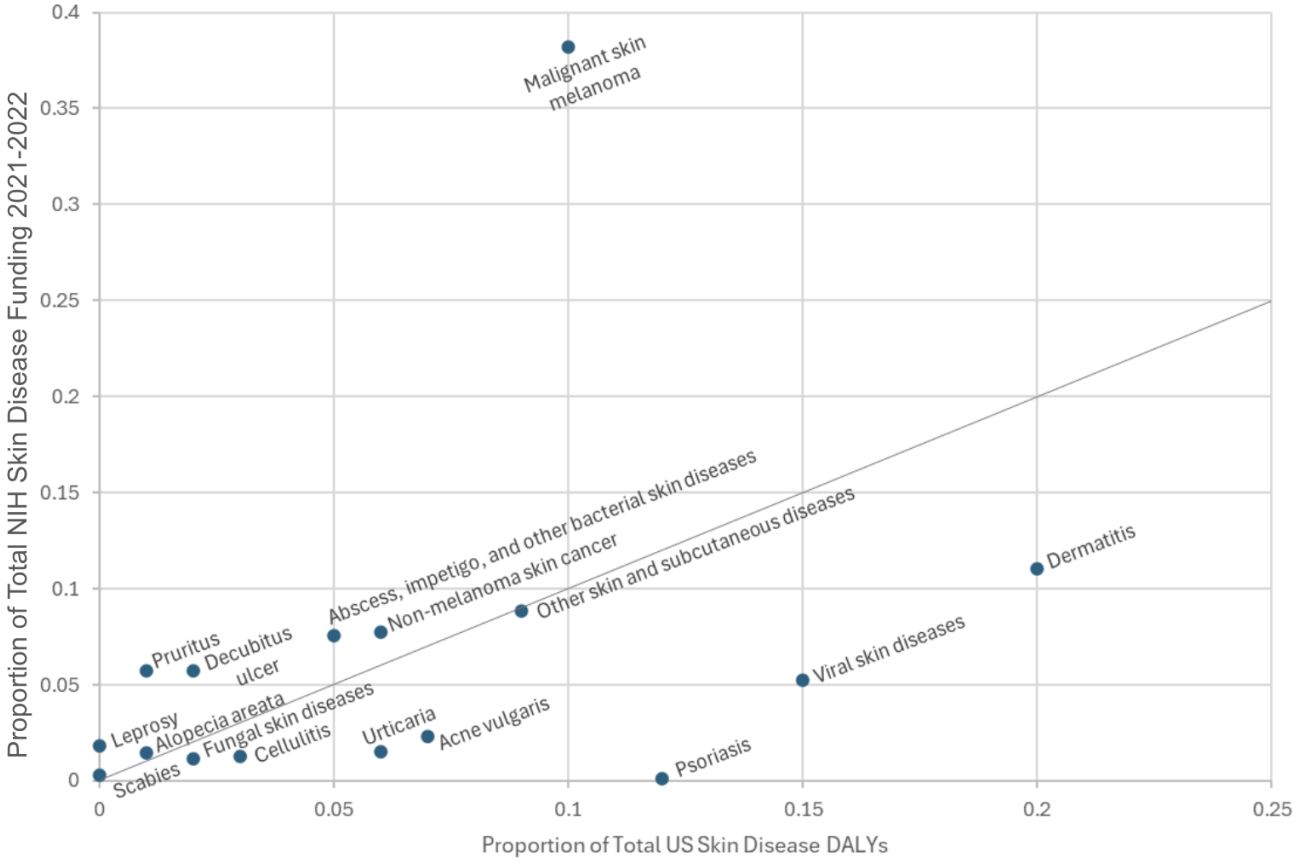
This study was exempt from review by the institutional review board, and no patient or participant consent was required or obtained, as this study did not constitute human subjects research and used publicly available data.

Results

Our analysis revealed a positive correlation between the percentage of total US skin disease DALYs in 2021 and the percentage of total NIH skin disease funding in 2021 - 2022. The correlation coefficient between these two data points was 0.3167 (95% CI 0.053626-0.579774). There were several key outliers when comparing DALYs to funding, indicating that certain skin diseases were relatively over- or underfunded in comparison to their proportion of total disease burden. Pruritus and malignant melanoma received 445% and 392% of the proportion of funding expected by their proportion of DALYs (Table 1). Other relatively overfunded diseases include leprosy, decubitus ulcers, bacterial skin diseases, and nonmelanoma skin cancer (Figure 1, Table 1).

XSL•FO
RenderX

Figure 1. Scatterplot comparing proportion of National Institutes of Health (NIH) skin disease funding received in 2021 - 2022 with the proportion of total US skin disease disability-adjusted life years (DALYs) according to the 2021 Global Burden of Disease study.



XSL•FO
RenderX

Table . Comparison of disability-adjusted life year (DALY) rank from Global Burden of Disease GBD 2010 and 2021 study data, comparison of National Institutes of Health (NIH) funding in fiscal years 2012 - 2013 (data from Hagstrom et al [4]) and 2021-2022 (data from the current analysis), and the percentage of total US skin DALYs (in 2021) and NIH skin disease funding (in 2021 - 22).

Category	US DALY rank in 2021	Proportion of total US skin disease DALYs in 2021, %	US DALY rank in 2010[a]	NIH funding rank in 2021 - 2022	Proportion of total NIH skin disease funding in 2021 - 2022, %	NIH funding rank in 2012 - 2013[a]	Observed-to-expected ratio for funding[b]
Pruritus	13	1.29	5	6	5.74	6	4.45
Malignant skin melanoma	4	9.75	3	1	38.19	1	3.92
Decubitus ulcer	11	1.87	8	7	5.74	11	3.07
Abscess, impetigo, and other bacterial skin diseases	8	4.90	13	5	7.56	9	1.54
Nonmelanoma skin cancer	6	6.10	2	4	7.73	2	1.27
Alopecia areata	12	1.36	11	12	1.48	13	1.09
Other skin and subcutaneous diseases	—[c]	9.30	—	3	8.87	3	0.95
Scabies	14	0.38	14	15	0.30	16	0.8
Dermatitis	1	19.98	1	2	11.03	5	0.55
Fungal skin diseases	10	2.17	9	14	1.14	10	0.53
Cellulitis	9	3.42	12	13	1.26	12	0.37
Viral skin diseases	2	14.61	6	8	5.22	4	0.36
Acne vulgaris	5	6.99	4	9	2.29	14	0.33
Urticaria	7	5.80	7	11	1.51	15	0.26
Psoriasis	3	12.10	10	16	0.09	7	0.0082
Leprosy[d]	15	0	15	10	1.84	8	—[d]

[a]Data obtained from Hagstrom et al [4].

[b]Percentage of funding vs percentage of DALYs.

[c]Not applicable.

[d]Ratio of funding proportion to DALY proportion could not be calculated for leprosy, as the proportion of DALYs for leprosy was 0.

Conversely, psoriasis, fungal skin diseases, cellulitis, urticaria, acne vulgaris, viral skin diseases, and dermatitis were underfunded. Notably, psoriasis received only 0.82% of the funding expected by its disease burden (Table 1). Funding for scabies, alopecia areata, and the "other skin/subcutaneous diseases" category appeared well matched to their disease burden, receiving between 80% to 110% of the funding predicted by their respective DALYs (Figure 1, Table 1).

Discussion

Principal Findings

This study reinvestigated the relationship between US skin disease burden and NIH skin disease research funding using the latest GBD 2021 data and NIH funding data from fiscal years 2021 - 2022. Compared to Hagstrom et al's [4] 2015 study, many of the same trends in relative over- and underfunding of skin diseases were observed. For example, malignant melanoma remains the most significantly overfunded skin disease relative to its disease burden (Table 1) [4]. Nonmelanoma skin cancer and leprosy also remain overfunded, while dermatitis, acne vulgaris, urticaria, fungal skin diseases, and cellulitis remain underfunded (Table 1) [4]. Interestingly, pruritus and decubitus ulcers, previously underfunded in 2015, now appear to be relatively overfunded (Table 1) [4]. Funding for psoriasis was well matched to its disease burden in 2015, but in our updated analysis, psoriasis is the most underfunded skin disease category. Similarly, viral skin diseases were well funded in 2015 and now appear underfunded (Table 1) [4].

It is important to consider disease burden when allocating research funding to ensure adequate resources are being directed toward diseases with the most significant impact. Dedicating more resources toward high-burden diseases can improve individual health and quality of life by driving the development

XSL•FO
RenderX

of innovative treatments and can also provide long-term economic benefits by reducing health care costs and increasing overall workforce productivity.

In addition to disease burden, many other factors also significantly impact resource prioritization and funding allocation. For example, more research funding is likely to be allocated to diseases with strong public awareness and advocacy campaigns, such as malignant skin melanoma. Funding is also likely influenced by disease curability and the potential for therapeutic innovation. The NIH may also prioritize funding for diseases with lower incidence or prevalence but higher mortality (ie, metastatic melanoma, metastatic nonmelanoma skin cancer) rather than diseases with lower mortality but higher incidence or prevalence (ie, dermatitis and acne vulgaris).

Limitations

It is important to keep in mind that using data strictly from the GBD study and the NIH does not fully capture all of the nuances of US skin disease burden and research funding. An important limitation of this analysis, similar to Hagstrom et al's [4] prior study, is the exclusion of industry research funding by pharmaceutical companies and other nongovernmental entities from NIH funding data [4,7]. The NIH is the largest source of public funding for biomedical research; however, a significant portion of research funding also comes from nonprofits, philanthropic organizations, and private industry [8]. Therefore, while a disease may appear underfunded relative to its disease burden using GBD and NIH data alone, additional research funding from nongovernmental agencies may be filling this perceived gap in resource allocation. For instance, although our analysis showed that psoriasis received significantly less funding from the NIH relative to its disease burden, substantial funding from pharmaceutical companies has driven the development of innovative new drugs (ie, IL-23 and IL-17 inhibitors) that have transformed the treatment of psoriasis in recent years [9]. Similarly, previous reviews have cited US $22,291,506 in nonprofit funding for dermatology research in 2019 alone and US $9.3 billion dollars of private equity investment in dermatology health care and research between 2011 and 2021 [10,11].

Conclusions

Given the wide variety of factors that must be considered in order to optimally allocate research funding, several guidelines may help ensure that funding is prioritized for research efforts that will guide clinical practice, improve patient outcomes, and positively impact public health. In addition to prioritizing high-burden diseases, prioritizing funding for translational research can help expedite the incorporation of knowledge gained from basic science research into clinical practice and patient care. Periodically evaluating the real-world impact of funded research using metrics including patient outcomes and cost-efficacy can also help ensure that funding is being distributed to research that is meaningfully impacting clinical practice. Increased funding for conditions that are impacting our patients will allow innovative solutions that improve patient quality of life. With these guidelines in mind, disease burden can easily be incorporated as one of the many important factors that should be used to inform research funding allocation, clinical practice guidelines, and health policy.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

RD is the Editor-in-Chief of *JMIR Dermatology* but was not involved in the selection of this manuscript for publication.

References

1. Institute for Health Metrics and Evaluation. Global Burden of Disease interactive data visuals. 2024. URL: https://vizhub.healthdata.org/gbd-results [accessed 2025-12-16]
2. Yakupu A, Aimaier R, Yuan B, et al. The burden of skin and subcutaneous diseases: findings from the Global Burden Of Disease study 2019. Front Public Health 2023;11:1145513. [doi: 10.3389/fpubh.2023.1145513] [Medline: 37139398]
3. Salomon JA, Vos T, Hogan DR, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 2012 Dec;380(9859):2129-2143. [doi: 10.1016/S0140-6736(12)61680-8]
4. Hagstrom EL, Patel S, Karimkhani C, et al. Comparing cutaneous research funded by the US National Institutes of Health (NIH) with the US skin disease burden. J Am Acad Dermatol 2015 Sep;73(3):383-391. [doi: 10.1016/j.jaad.2015.04.039] [Medline: 26051697]
5. Price KN, Collier EK, Atluri S, Hsiao JL, Shi VY. National Institutes of Health dermatology funding trends 2015-2019. J Invest Dermatol 2021 Jan;141(1):232-235. [doi: 10.1016/j.jid.2020.05.089] [Medline: 32464149]
6. Joshi A, Gawey L, Harrison M, et al. National Institutes of Health dermatology funding trends: 2020-2024. J Invest Dermatol 2025 Oct;145(10):2629-2632. [doi: 10.1016/j.jid.2025.03.004] [Medline: 40118168]
7. RePORTER. National Institutes of Health. URL: https://reporter.nih.gov/advanced-search [accessed 2025-12-10]
8. Grants & funding. National Institutes of Health. URL: https://www.nih.gov/grants-funding [accessed 2024-11-03]
9. Subramonian A, Walter M. Newer Biologics for the Treatment Of Plaque Psoriasis: Canadian Agency for Drugs and Technologies in Health; 2021. URL: http://www.ncbi.nlm.nih.gov/books/NBK585965/ [accessed 2024-11-03]

XSL•FO
RenderX

10. Agarwal A, Orlow SJ. Nonprofit funding in dermatology: a cross-sectional analysis in the United States, 2015 to 2019. JAAD Int 2023 Sep;12:162-164. [doi: 10.1016/j.jdin.2023.06.001] [Medline: 37404247]
11. Venkatesh K, Nambudiri VE. Characterization of venture capital investments in dermatology: a cross-sectional analysis, 2011 to 2021. J Am Acad Dermatol 2023 Mar;88(3):665-667. [doi: 10.1016/j.jaad.2022.06.1177] [Medline: 35779641]

Abbreviations

DALY: disability-adjusted life year
GBD: Global Burden of Disease
NIH: National Institutes of Health
YLD: year lived with disability

Edited by W Guo; submitted 25.Jan.2025; peer-reviewed by D Hitchcock, J Huang, M Gasmi; revised version received 17.Oct.2025; accepted 04.Dec.2025; published 06.Jan.2026.

Please cite as:
Park A, Woolhiser E, Riva H, Wan L, Kadri H, Lamberty E, Juels P, Jaroonwanichkul S, Reed M, Hegedus C, Chen D, Duffle D, Kirk J, Christensen S, Shelby E, Dellavalle R
Identifying Over- and Underfunded Diseases by Comparing National Institutes of Health Funding for Skin Disease Research With US Skin Disease Burden According to 2021 Global Burden of Disease Data: Cross-Sectional Analysis
JMIR Dermatol 2026;9:e71468
URL: https://derma.jmir.org/2026/1/e71468
doi:10.2196/71468

© Aileen Park, Emily Woolhiser, Hannah Riva, Leo Wan, Haaris Kadri, Elizabeth Lamberty, Parker Juels, Sandra Jaroonwanichkul, Madison Reed, Catherine Hegedus, Dana Chen, Danielle Duffle, Jessica Kirk, Sydney Christensen, Emma Shelby, Robert Dellavalle. Originally published in JMIR Dermatology (http://derma.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dermatologic Conditions and Incident Anxiety in Young Adults: Propensity Score–Matched Retrospective Cohort Study

Isabella Zai[1]; Adrian Zai[2], MPH, MD, PhD

[1]Westwood High School, Westwood, MA, United States

[2]Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, 368 Plantation Street, Worcester, MA, United States

Corresponding Author:
Adrian Zai, MPH, MD, PhD
Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, 368 Plantation Street, Worcester, MA, United States

Abstract

Background: Dermatologic conditions such as acne, psoriasis, and dermatitis commonly affect young adults and may contribute to psychological distress. While prior studies have suggested an association between skin disease and anxiety, longitudinal population-level evidence in young adults remains limited.

Objective: This study aimed to examine the association between common dermatologic conditions and the incidence of anxiety among young adults using a large electronic health record–based cohort.

Methods: We conducted a retrospective cohort study using the TriNetX Research Network, including young adults aged 18 to 22 years with and without a qualifying dermatologic diagnosis between 2019 and 2020. The index date was defined as the first dermatologic diagnosis for exposed individuals and a qualifying ambulatory visit for controls. Individuals with a prior diagnosis of anxiety were excluded. Propensity score matching was used to balance demographic characteristics and ambulatory visit history between cohorts. Incident anxiety diagnoses were assessed at 1, 3, and 5 years following the index date. Cumulative incidence, absolute risk differences, risk ratios, and time-to-event analyses were evaluated.

Results: After propensity score matching, 169,720 individuals were included in each cohort. Young adults with dermatologic conditions consistently exhibited a higher incidence of anxiety across all time points. At 1 year, anxiety occurred in 3.9% (5838/149,464) of individuals in the dermatologic condition group compared with 3.4% (3237/95,020) of controls. By 3 years, incidence increased to 11.8% (17,559/149,464) of individuals vs 10.4% (9906/95,020) of controls, and by 5 years to 16.8% (25,184/149,464) of individuals vs 15.5% (14,712/95,020) of controls. Absolute risk differences widened over time, from 0.50 to 1.40 percentage points. Time-to-event analyses demonstrated a modest but consistent increase in hazard, with hazard ratios ranging from 1.12 to 1.14 (all $P<.001$).

Conclusions: In this large, propensity score–matched cohort of young adults, common dermatologic conditions were associated with a small but persistent increase in the incidence of anxiety over time. Although absolute differences were modest, the consistency of findings across multiple analytic approaches highlights the importance of considering psychological well-being as part of comprehensive care for young adults with dermatologic conditions.

(JMIR Dermatol 2026;9:e90820) doi:10.2196/90820

KEYWORDS

dermatologic diseases; anxiety disorders; young adult; electronic health records; propensity score matching

Introduction

Dermatologic conditions are among the most common reasons for ambulatory health care visits in adolescents and young adults [1,2]. Disorders such as acne, psoriasis, and dermatitis frequently emerge or worsen during late adolescence and early adulthood, a developmental period characterized by significant psychological, social, and identity-related transitions [3-6]. Although the physical manifestations of skin disease are readily apparent, their potential mental health sequelae are less consistently recognized in routine clinical care [7-13].

Prior studies have suggested an association between dermatologic conditions and adverse mental health outcomes, including anxiety [7,14,15]. However, much of this literature relies on cross-sectional surveys, clinic-based samples, or patient-reported measures of psychological distress [15-19]. While these approaches provide important insights into symptom burden and quality of life, they are limited in their ability to establish temporal relationships or estimate population-level risk [20,21]. In addition, many studies focus on adults or pediatric populations, with comparatively less attention to young adults navigating the transition between adolescent and adult care [5,22-25].

XSL•FO
RenderX

Electronic health record (EHR) data offer an opportunity to examine longitudinal associations between dermatologic diagnoses and subsequent mental health outcomes at scale [26-28]. Large, multi-institutional EHR networks allow for the identification of incident diagnoses, the alignment of outcomes to clinically meaningful index events, and the evaluation of risk over extended follow-up periods [29-31]. Such data can complement survey-based and qualitative work by providing estimates of real-world diagnostic patterns and temporal trends across diverse care settings.

Despite increasing interest in psychodermatology, few studies have leveraged longitudinal EHR data to quantify the risk of incident anxiety following dermatologic diagnoses in young adults. In particular, there is limited evidence characterizing how anxiety diagnoses accumulate over time after an initial dermatologic diagnosis, or how this risk compares with peers without dermatologic conditions when baseline differences in demographics and health care use are accounted for.

The objective of this study was to evaluate the association between common dermatologic diagnoses and subsequent incident anxiety among young adults aged 18 to 22 years using a large federated EHR network. We conducted a propensity score–matched cohort study comparing cumulative incidence and time to anxiety diagnosis at 1, 3, and 5 years of follow-up. By focusing on incident anxiety and aligning outcomes to clearly defined index events, this study aimed to provide population-level estimates of anxiety risk following dermatologic diagnoses during a critical developmental period.

Methods

Study Design and Data Source

We conducted a retrospective matched cohort study using deidentified EHR data accessed through the TriNetX Research Network, a federated platform comprising multiple health care organizations across the United States [32]. The network contains longitudinal inpatient and outpatient data, including demographics, diagnoses, and encounter information. All analyses were performed within the secure analytic environment provided by TriNetX using deidentified data in accordance with institutional policies.

The study was reviewed by the institutional review board and determined to be exempt.

Study Population

The study population included young adults aged 18 to 22 years with clinical encounters recorded between January 1, 2019, and December 31, 2020. The 2019 to 2020 index period was selected to allow sufficient longitudinal follow-up of up to 5 years within the available TriNetX data while capturing a contemporary cohort of young adults receiving routine care. This period was not selected to coincide with the COVID-19 pandemic but rather reflects the most recent interval permitting extended follow-up. To ensure engagement with the health care system, patients were required to have at least one qualifying ambulatory visit during the study period.

Patients were followed longitudinally for the development of incident anxiety diagnoses through December 31, 2025, allowing for up to 5 years of potential follow-up.

Exposure Definition and Index Date

The exposed cohort consisted of patients with a diagnosis of a dermatologic condition, defined by the presence of *International Classification of Diseases, Tenth Revision* (*ICD-10*) diagnosis codes corresponding to common dermatologic disorders. The index date for exposed patients was defined as the date of the first qualifying dermatologic diagnosis during the study period.

The comparison cohort consisted of patients without any recorded dermatologic diagnosis at any time prior to or during the study period. For these patients, the index date was defined as the date of the first qualifying ambulatory visit occurring within the same calendar period.

Outcome Definition

The primary outcome was incident anxiety, defined as a new diagnosis of an anxiety disorder recorded after the index date, identified using *ICD-10* diagnosis codes for anxiety disorders.

To focus on incident anxiety, patients with any recorded anxiety diagnosis during the 1 year before the index date were excluded at the cohort definition stage. In addition, outcome analyses excluded patients with anxiety diagnoses occurring prior to the start of each analytic time window to ensure that all patients included in the risk set were at risk for incident anxiety during the specified follow-up period.

Outcome time windows began 1 day after the index date to reduce misclassification of anxiety diagnoses recorded during the same clinical encounter as the dermatologic diagnosis or index visit.

Covariates

Baseline covariates included age, sex, race, and ethnicity, as well as measures of health care use prior to the index date, including the presence of ambulatory visits. Covariates were selected a priori based on clinical relevance and their potential role as confounders of the association between dermatologic diagnoses and subsequent anxiety.

Anxiety diagnoses were not included as matching covariates because patients with prior anxiety were excluded from the study population, and postindex variables were not incorporated into propensity score estimation.

Propensity Score Matching

To reduce confounding, propensity score matching was performed to create balanced cohorts of patients with and without dermatologic diagnoses. Propensity scores were estimated using logistic regression based on baseline demographic characteristics and health care use measures.

Patients were matched one-to-one using nearest neighbor matching without replacement. Balance between cohorts was assessed using standardized mean differences, with values less than 0.1 indicating acceptable balance.

All outcome analyses were conducted using the matched cohorts.

XSL•FO
RenderX

Follow-Up and Time Horizons

Patients were followed from the index date until the earliest of the following events: first recorded anxiety diagnosis, end of the specified follow-up window, or last available encounter in the dataset.

Separate analyses were conducted using follow-up windows of 1, 3, and 5 years following the index date. As follow-up duration increased, fewer patients contributed data to later time horizons due to censoring at the end of available follow-up.

Statistical Analysis

Cumulative incidence of anxiety was calculated for each cohort at 1, 3, and 5 years following the index date. Absolute risk differences and relative risks with 95% CIs were reported.

Time to incident anxiety was additionally evaluated using Kaplan-Meier methods, with differences between cohorts assessed using the log-rank test. Hazard ratios with 95% CIs were estimated to quantify relative differences in the rate of anxiety diagnosis over time. Kaplan-Meier analyses were performed as supportive time-to-event analyses to assess consistency with cumulative incidence findings.

All statistical tests were 2-sided, and statistical significance was defined as $P<.05$.

Sensitivity Analyses

Sensitivity analyses were conducted by comparing effect estimates across multiple follow-up horizons to assess the robustness and temporal consistency of findings.

Ethical Considerations

This study was reviewed by the Institutional Review Board of UMass Chan Medical School and determined to be exempt from human subjects research (STUDY00002783). The research used deidentified EHR data accessed through the TriNetX Research Network. Because the data were deidentified prior to analysis and posed minimal risk to participants, no direct identifiers were available to the investigators, participants were not contacted, and no attempts were made to reidentify individuals, informed consent was not required in accordance with the Common Rule provision for exempt secondary research [33,34] and the HIPAA (Health Insurance Portability and Accountability Act) Privacy Rule standard for deidentification.

Data were analyzed within the secure TriNetX analytics environment, and results were reported only in aggregate form. No participant compensation was provided, as this study did not involve direct participant contact.

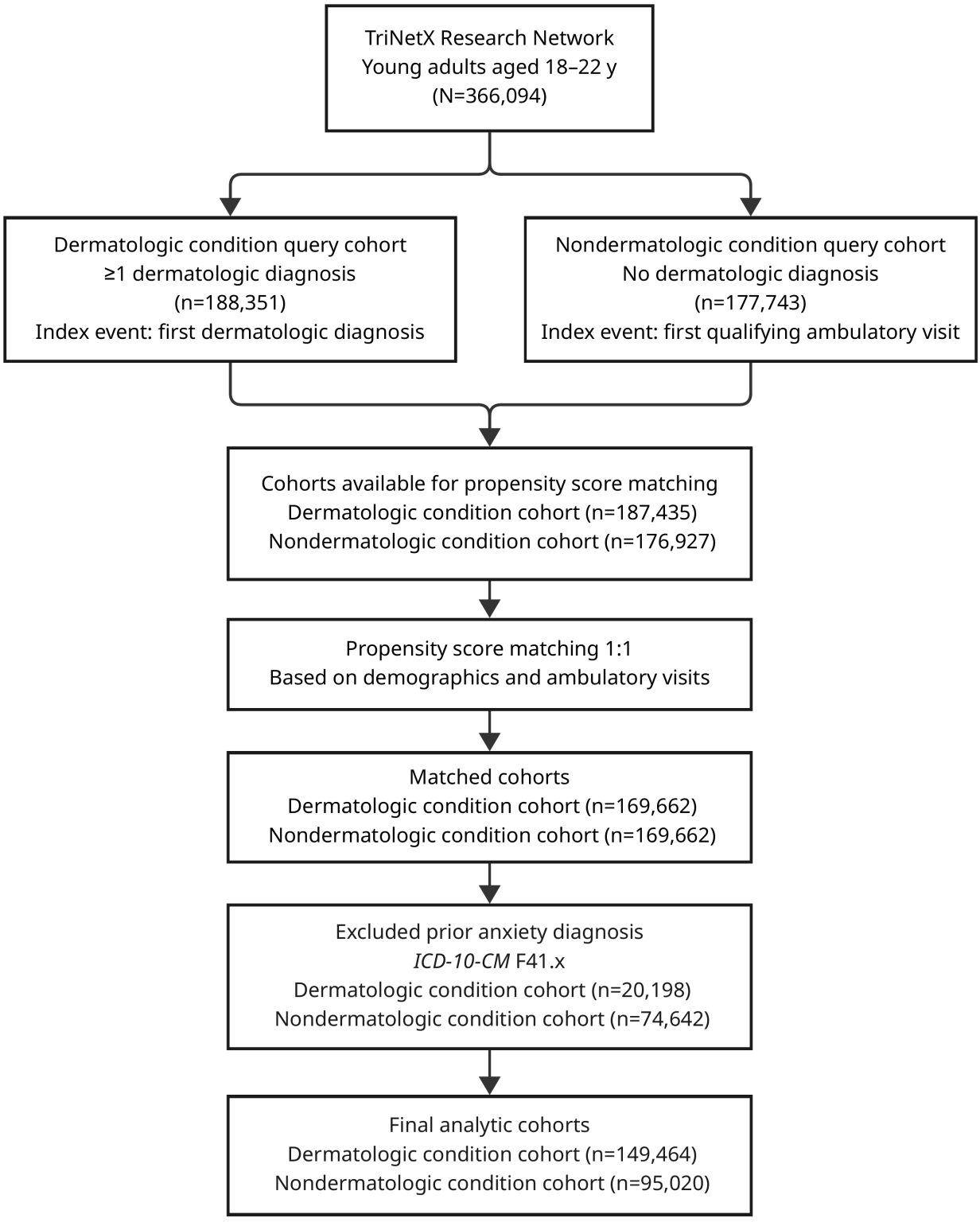
Results

Cohort Assembly and Baseline Characteristics

Figure 1 summarizes cohort assembly. Among young adults aged 18 to 22 years with a qualifying dermatologic diagnosis between January 1, 2019, and December 31, 2020, eligible individuals were matched 1:1 to controls without dermatologic diagnoses using propensity scores. After matching, 169,720 individuals were retained in each cohort.

XSL•FO
RenderX

Figure 1. Cohort selection flow diagram. *ICD-10-CM*: *International Classification of Diseases, Tenth Revision, Clinical Modification*.



Baseline demographic characteristics before and after matching are shown in Table 1. Prior to matching, modest differences were observed across race, ethnicity, sex, and ambulatory visit history. After propensity score matching, covariates were well balanced between cohorts, with standardized mean differences below 0.01 for all matched variables.

XSL•FO
RenderX

Table . Baseline characteristics of propensity score–matched young adults with and without dermatologic diagnoses[a].

Characteristic	Dermatology cohort (n=169,662), n (%)	Nondermatology cohort (n=169,662), n (%)	Standardized mean difference
Sex			
Female	96,440 (56.8)	96,257 (56.7)	0.0018
Male	73,139 (43.1)	73,325 (43.2)	0.0019
Race and ethnicity			
American Indian or Alaska Native	1018 (0.6)	10,542 (0.6)	0.0024
Asian	9289 (5.5)	9356 (5.5)	0.0017
Black or African American	27,461 (16.2)	27,451 (16.2)	0.0001
Hispanic or Latino	23,930 (14.1)	23,610 (13.9)	0.0052
Native Hawaiian or Other Pacific Islander	886 (0.5)	9198 (0.5)	0.0026
Not Hispanic or Latino	112,308 (66.2)	112,696 (66.4)	0.0047
White	107,599 (63.4)	107,415 (63.3)	0.0023
Prior ambulatory visit before index	98,656 (58.1)	98,495 (58.1)	0.0019

[a]Cohorts were matched 1:1 using propensity score matching based on sex, race, ethnicity, and prior ambulatory health care use. All standardized mean differences were less than 0.01, indicating excellent postmatching balance.

Young adults aged 18 to 22 years were identified from the TriNetX Research Network between 2019 and 2020. Patients were categorized based on the presence or absence of a qualifying dermatologic diagnosis (including acne, psoriasis, or dermatitis or eczema), and those with a prior diagnosis of anxiety were excluded. Propensity score matching was applied to create balanced cohorts with and without dermatologic conditions. The final matched cohorts were followed for incident anxiety diagnoses at 1-, 3-, and 5-year intervals after the index date.

Incident Anxiety at 1, 3, and 5 Years

Table 2 presents the cumulative incidence of newly diagnosed anxiety disorders at 1, 3, and 5 years following the index date. Across all time horizons, young adults with dermatologic conditions experienced a consistently higher incidence of anxiety compared with matched controls. The absolute difference between cohorts increased over time, while the relative increase in risk remained modest across follow-up periods.

Figure 2 illustrates these cumulative incidence estimates over time, demonstrating a parallel increase in anxiety diagnoses in both cohorts with a persistently higher absolute risk among individuals with dermatologic conditions.

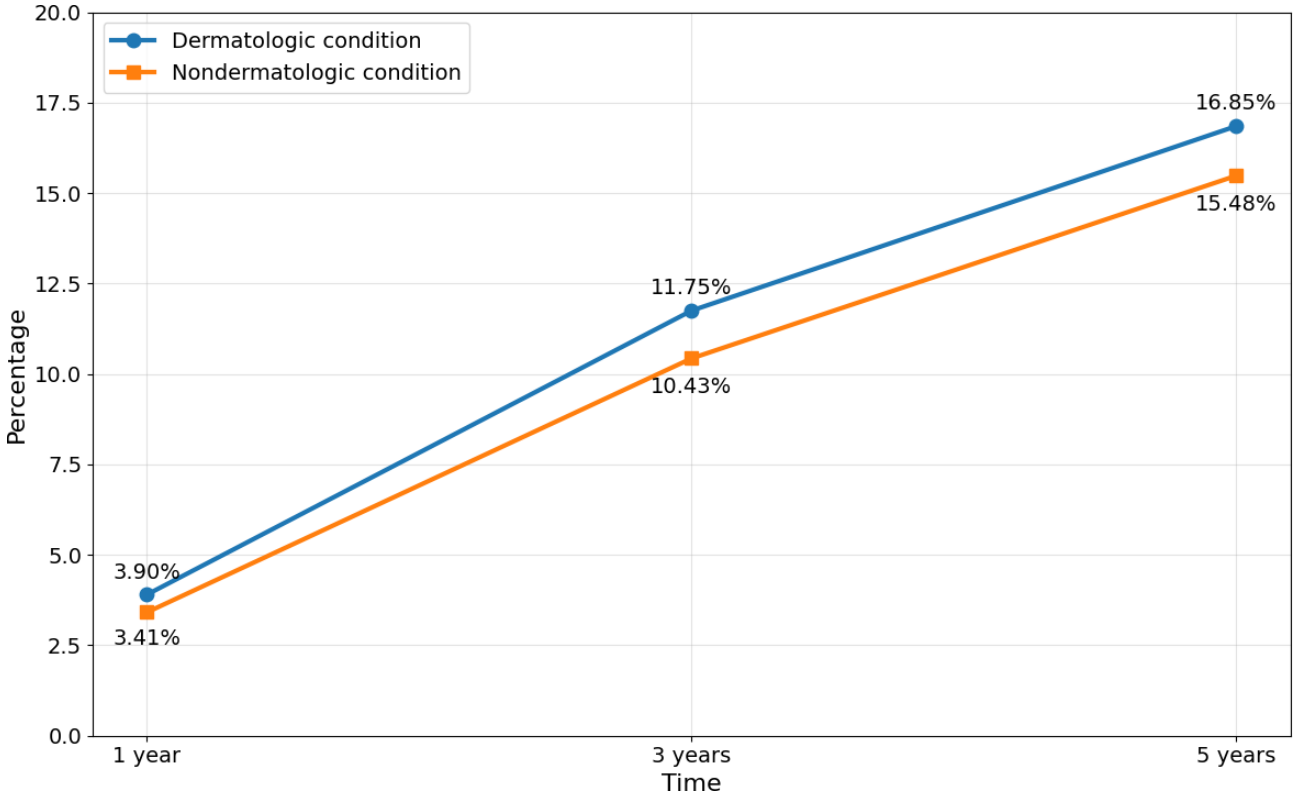
Table . Risk of incident anxiety among young adults with vs without dermatologic conditions[a].

Follow-up period and cohort	Incident anxiety, n (%)	Risk difference (%; 95% CI)	Risk ratio (95% CI)	Odds ratio (95% CI)
1 year				
With a dermatologic condition (n=149,464)	5838 (3.9)	0.50 (0.30 - 0.70)	1.147 (1.099 - 1.196)	1.153 (1.103 - 1.204)
Without a dermatologic condition (n=95,020)	3237 (3.4)	Reference	Reference	Reference
3 years				
With a dermatologic condition (n=149,464)	17,559 (11.8)	1.32 (1.07 - 1.58)	1.127 (1.101 - 1.153)	1.144 (1.114 - 1.174)
Without a dermatologic condition (n=95,020)	9906 (10.4)	Reference	Reference	Reference
5 years				
With a dermatologic condition (n=149,464)	25,184 (16.8)	1.40 (1.10 - 1.70)	1.088 (1.068 - 1.109)	1.106 (1.082 - 1.131)
Without a dermatologic condition (n=95,020)	14,712 (15.5)	Reference	Reference	Reference

[a]Patients with a prior diagnosis of anxiety before the index date were excluded. Analyses were conducted after propensity score matching.

XSL•FO
RenderX

Figure 2. Cumulative incidence of anxiety among young adults with and without dermatologic conditions. Cumulative incidence of new-onset anxiety at 1, 3, and 5 years after the index date among propensity score–matched young adults aged 18 to 22 years with dermatologic conditions diagnosed between 2019 and 2020 and matched peers without dermatologic conditions. Percentages represent the proportion of individuals with a new anxiety diagnosis within each follow-up interval, excluding those with anxiety prior to the index date.



Time-to-Event Analysis

Time-to-event analyses using Kaplan-Meier methods further supported these findings. Individuals with dermatologic conditions exhibited a modestly increased hazard of incident anxiety across follow-up periods.

The hazard ratio for incident anxiety was 1.135 (95% CI 1.088 - 1.185) at 1 year, 1.121 (95% CI 1.094 - 1.149) at 3 years, and 1.096 (95% CI 1.074-1.119) at 5 years, consistent with the cumulative incidence findings in Table 2; log-rank tests were significant across all analyses (all $P<.001$).

Despite statistical significance, Kaplan-Meier curves showed substantial overlap between cohorts, reflecting the modest magnitude of the association and reinforcing the importance of interpreting relative measures alongside absolute risk differences.

Discussion

In this large, propensity score–matched cohort study of young adults, we found that individuals with common dermatologic conditions experienced a consistently higher incidence of anxiety compared with matched peers without dermatologic conditions. Although the magnitude of the association was modest, the excess risk was observed across several follow-up intervals and persisted over time. These findings suggest that dermatologic conditions in early adulthood are associated with a small but sustained increase in the likelihood of subsequent anxiety diagnoses. The consistency of findings across cumulative incidence, relative risk, and time-to-event analyses supports the robustness of the observed association.

The absolute differences in anxiety incidence were relatively small, ranging from approximately one-half of a percentage point at 1 year to just over 1 percentage point at 5 years. However, given the high prevalence of dermatologic conditions among young adults, even modest increases in risk may have meaningful implications at the population level. Importantly, the consistency of the association across 1-, 3-, and 5-year follow-up intervals strengthens confidence that the observed relationship is not driven by short-term diagnostic clustering or surveillance bias alone.

The relative effect size attenuated with longer follow-up, while the absolute difference in cumulative incidence increased over time. This pattern is expected in large observational cohorts and reflects the increasing background incidence of anxiety in young adulthood. From a clinical perspective, the absolute risk difference may be more informative than relative measures, as it better contextualizes the magnitude of excess risk attributable to dermatologic conditions. Our findings underscore the importance of interpreting statistically significant associations in the context of absolute effect sizes.

Several mechanisms may plausibly explain the observed association. Dermatologic conditions, such as acne, psoriasis, and dermatitis, often emerge or worsen during adolescence and early adulthood, a period characterized by heightened psychosocial vulnerability. Visible skin disease may contribute to anxiety through effects on self-image, social interactions, and perceived stigma. In addition, chronic or recurrent

XSL•FO
RenderX

symptoms, such as pruritus, pain, or flares, may increase psychological distress over time. While our study was not designed to identify causal pathways, the persistence of the association across multiple years is consistent with a cumulative psychosocial burden rather than a transient response to diagnosis alone.

Our findings are consistent with prior literature reporting associations between dermatologic disease and mental health outcomes, while extending this work in several important ways. First, we focused specifically on young adults, a population in which both dermatologic conditions and anxiety commonly emerge but remain underrepresented in large-scale longitudinal studies. Second, by using propensity score matching and excluding individuals with prior anxiety diagnoses, we sought to minimize confounding and focus on incident outcomes. Third, the use of multiple follow-up intervals allowed us to characterize the temporal evolution of risk rather than relying on a single time point estimate.

This study has several limitations. As an observational analysis based on EHR data, it is subject to residual confounding despite propensity score matching. The study period included the onset of the COVID-19 pandemic, which may have influenced background rates of anxiety diagnoses as well as patterns of health care use. Although both cohorts were drawn from the same calendar period and matched on baseline characteristics and ambulatory visit history, residual confounding related to pandemic-era behavioral and health system changes remains possible. However, because both cohorts were derived from the same period, such effects are likely to be nondifferential and unlikely to materially affect the internal validity of the comparison. Anxiety diagnoses may be underascertained or misclassified, and diagnostic practices may vary across health care organizations. Additionally, patients with dermatologic conditions may have more frequent health care encounters, potentially increasing opportunities for anxiety diagnosis. Although matching on ambulatory visit history partially addresses this concern, differential health care use cannot be fully excluded. Finally, the TriNetX platform does not allow for detailed assessment of disease severity, symptom burden, or patient-reported outcomes, which may further modify the relationship between dermatologic conditions and anxiety.

Despite these limitations, the strengths of this study include its large, geographically diverse population, rigorous matching strategy, and consistent findings across multiple analytic approaches. The concordance between cumulative incidence estimates and time-to-event analyses supports the robustness of the observed association.

In conclusion, young adults with common dermatologic conditions experience a small but persistent increase in the incidence of anxiety compared with matched peers without dermatologic conditions. These findings highlight the importance of holistic care approaches that consider both physical and psychological well-being in patients with dermatologic disease. Future studies incorporating patient-reported outcomes and disease severity measures may help clarify which subgroups are at greatest risk and inform targeted screening or intervention strategies.

Acknowledgments

The authors acknowledge the TriNetX Research Network for providing access to deidentified electronic health record data used in this study.

Funding

This study did not receive any external funding.

Data Availability

The data used in this study were obtained from the TriNetX Research Network. These data are available only under license for this study and are not publicly available. Qualified researchers may access data through the TriNetX platform, subject to institutional agreements, data use approvals, and compliance with applicable privacy and security requirements. Aggregate results supporting the findings of this study are included within the paper.

Authors' Contributions

Conceptualization: IZ
Data curation: IZ
Formal analysis: IZ
Methodology: AZ
Project administration: AZ
Supervision: AZ
Validation: AZ
Writing—original draft: IZ
Writing—review and editing: AZ
Both authors approved the final version of the manuscript for submission and agree to be accountable for all aspects of the work.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Ho T, Taylor MT, Marathe KS, Lucky AW, Barbieri JS. Most common pediatric skin conditions managed in outpatient dermatology clinics in the United States stratified by race and ethnicity. Pediatr Dermatol 2021 Nov;38 Suppl 2:129-131. [doi: 10.1111/pde.14693] [Medline: 34339074]
2. Chang J, Nock MR, Cohen JM, Bunick CG. Acne accounts for an almost 2.5-fold higher proportion of dermatology visits among adult females compared to adult males in the United States: a study of the national ambulatory medical care survey from 2002-2016. PLoS One 2023;18(9):e0290763. [doi: 10.1371/journal.pone.0290763] [Medline: 37733776]
3. Potterton R, Austin A, Robinson L, Webb H, Allen KL, Schmidt U. Identity development and social-emotional disorders during adolescence and emerging adulthood: a systematic review and meta-analysis. J Youth Adolesc 2022 Jan;51(1):16-29. [doi: 10.1007/s10964-021-01536-7] [Medline: 34783954]
4. Branje S, de Moor EL, Spitzer J, Becht AI. Dynamics of identity development in adolescence: a decade in review. J Res Adolesc 2021 Dec;31(4):908-927. [doi: 10.1111/jora.12678]
5. De Vere Hunt IJ, Howard E, McPherson T. The impact of chronic skin disease in adolescence and the need for specialist adolescent services. Clin Exp Dermatol 2020 Jan;45(1):5-9. [doi: 10.1111/ced.14021] [Medline: 31236992]
6. Fournier H, Calcagni N, Morice-Picard F, Quintard B. Psychosocial implications of rare genetic skin diseases affecting appearance on daily life experiences, emotional state, self-perception and quality of life in adults: a systematic review. Orphanet J Rare Dis 2023 Feb 23;18(1):39. [doi: 10.1186/s13023-023-02629-1] [Medline: 36823650]
7. Salari N, Heidarian P, Hosseinian-Far A, Babajani F, Mohammadi M. Global prevalence of anxiety, depression, and stress among patients with skin diseases: a systematic review and meta-analysis. J Prev (2022) 2024 Aug;45(4):611-649. [doi: 10.1007/s10935-024-00784-0]
8. Balieva F, Abebe DS, Dalgard FJ, Lien L. Risk of developing psychiatric disease among adult patients with skin disease: a 9-year national register follow-up study in Norway. Skin Health Dis 2023 Dec;3(6):e294. [doi: 10.1002/ski2.294] [Medline: 38047256]
9. Cortés H, Rojas-Márquez M, Del Prado-Audelo ML, Reyes-Hernández OD, González-Del Carmen M, Leyva-Gómez G. Alterations in mental health and quality of life in patients with skin disorders: a narrative review. Int J Dermatol 2022 Jul;61(7):783-791. [doi: 10.1111/ijd.15852] [Medline: 34403497]
10. Hughes O, Bewley A. Is it really ever “just acne”? Considering the psychodermatology of acne. Br J Dermatol 2023 Oct 30;189(Suppl 1):i11-i16. [doi: 10.1093/bjd/ljad251] [Medline: 37903071]
11. Sood S, Jafferany M, Vinaya Kumar S. Depression, psychiatric comorbidities, and psychosocial implications associated with acne vulgaris. J Cosmet Dermatol 2020 Dec;19(12):3177-3182. [doi: 10.1111/jocd.13753] [Medline: 33006820]
12. Baker N, Billick SB. Psychiatric consequences of skin conditions: multiple case study analysis with literature review. Psychiatr Q 2022 Sep;93(3):841-847. [doi: 10.1007/s11126-022-09991-6] [Medline: 35771406]
13. Wu JJ, Feldman SR, Koo J, Marangell LB. Epidemiology of mental health comorbidity in psoriasis. J Dermatolog Treat 2018 Aug;29(5):487-495. [doi: 10.1080/09546634.2017.1395800] [Medline: 29057684]
14. Storer B, Kershaw KA, Braund TA, et al. The prevalence of anxiety disorders in dermatology outpatients: a systematic review and meta-analysis. J Eur Acad Dermatol Venereol 2023 Apr 28. [doi: 10.1111/jdv.19151] [Medline: 37118899]
15. Dalgard FJ, Gieler U, Tomas-Aragones L, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015 Apr;135(4):984-991. [doi: 10.1038/jid.2014.530] [Medline: 25521458]
16. Dalgard FJ, Svensson Å, Halvorsen JA, et al. Itch and mental health in dermatological patients across Europe: a cross-sectional study in 13 countries. J Invest Dermatol 2020 Mar;140(3):568-573. [doi: 10.1016/j.jid.2019.05.034] [Medline: 31491369]
17. Schuster B, Peifer C, Ziehfreund S, et al. Happiness and depression in psoriasis: a cross-sectional study in Germany. Qual Life Res 2022 Jun;31(6):1761-1773. [doi: 10.1007/s11136-021-02991-2] [Medline: 34532835]
18. Pärna E, Aluoja A, Kingo K. Quality of life and emotional state in chronic skin disease. Acta Derm Venereol 2015 Mar;95(3):312-316. [doi: 10.2340/00015555-1920] [Medline: 24978135]
19. Picardi A, Abeni D, Pasquini P. Assessing psychological distress in patients with skin diseases: reliability, validity and factor structure of the GHQ-12. J Eur Acad Dermatol Venereol 2001 Sep;15(5):410-417. [doi: 10.1046/j.1468-3083.2001.00336.x] [Medline: 11763380]
20. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 2020 Jul;158(1S):S65-S71. [doi: 10.1016/j.chest.2020.03.012]
21. Grujičić S, Nikolić A. Cross-section studies: advantages and disadvantages. Zdravstvena Zaštita 2021;50(4):43-54. [doi: 10.5937/zdravzast50-35574]
22. Zhou H, Roberts P, Dhaliwal S, Della P. Transitioning adolescent and young adults with chronic disease and/or disabilities from paediatric to adult care services - an integrative review. J Clin Nurs 2016 Nov;25(21-22):3113-3130. [doi: 10.1111/jocn.13326] [Medline: 27145890]

XSL•FO
RenderX

23. Hendrickx G, De Roeck V, Maras A, et al. Challenges during the transition from child and adolescent mental health services to adult mental health services. BJPsych Bull 2020 Aug;44(4):163-168. [doi: 10.1192/bjb.2019.85] [Medline: 31931898]
24. Marchini S, Laroche MA, Nemorin H, et al. From adolescence to adulthood: understanding care trajectories in an early detection and intervention centre in France. Early Interv Psychiatry 2025 Jan;19(1):e13605. [doi: 10.1111/eip.13605] [Medline: 39223792]
25. De Vere Hunt I, Chapman K, Wali G, et al. Establishing and developing a Teenage and Young Adult dermatology clinic with embedded specialist psychological support. Clin Exp Dermatol 2019 Dec;44(8):893-896. [doi: 10.1111/ced.13950] [Medline: 30784103]
26. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health 2016;37:61-81. [doi: 10.1146/annurev-publhealth-032315-021353] [Medline: 26667605]
27. Huguet N, Kaufmann J, O'Malley J, et al. Using electronic health records in longitudinal studies: estimating patient attrition. Med Care 2020 Jun;58:S46-S52. [doi: 10.1097/MLR.0000000000001298] [Medline: 32412953]
28. Li Q, Patrick MT, Sreeskandarajan S, et al. Large-scale epidemiological analysis of common skin diseases to identify shared and unique comorbidities and demographic factors. Front Immunol 2024;14:1309549. [doi: 10.3389/fimmu.2023.1309549]
29. Manemann SM, St Sauver JL, Liu H, et al. Longitudinal cohorts for harnessing the electronic health record for disease prediction in a US population. BMJ Open 2021 Jun 8;11(6):e044353. [doi: 10.1136/bmjopen-2020-044353] [Medline: 34103314]
30. Madhavan S, Bastarache L, Brown JS, et al. Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. J Am Med Inform Assoc 2021 Feb 15;28(2):393-401. [doi: 10.1093/jamia/ocaa287] [Medline: 33260207]
31. Simon GJ, Peterson KA, Castro MR, Steinbach MS, Kumar V, Caraballo PJ. Predicting diabetes clinical outcomes using longitudinal risk factor trajectories. BMC Med Inform Decis Mak 2020 Jan 8;20(1):6. [doi: 10.1186/s12911-019-1009-3] [Medline: 31914992]
32. TriNetX. URL: https://trinetx.com [accessed 2026-05-27]
33. eCFR: 45 CFR 46.104 -- exempt research. Office for Human Research Protections, Department of Health and Human Services. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.104 [accessed 2026-06-01]
34. eCFR: 45 CFR 164.514 other requirements relating to uses and disclosures of protected health information. Office for Human Research Protections, Department of Health and Human Services. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-C/part-164/subpart-E/section-164.514 [accessed 2026-06-01]

Abbreviations

EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
***ICD-10*:** *International Classification of Diseases, Tenth Revision*

Edited by S Subhadarshani; submitted 04.Jan.2026; peer-reviewed by S Gulliver; revised version received 19.Apr.2026; accepted 20.May.2026; published 03.Jun.2026.

Please cite as:
Zai I, Zai A
Dermatologic Conditions and Incident Anxiety in Young Adults: Propensity Score–Matched Retrospective Cohort Study
JMIR Dermatol 2026;9:e90820
URL: https://derma.jmir.org/2026/1/e90820
doi:10.2196/90820

© Isabella Zai, Adrian Zai. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging AI Large Language Models for Writing Clinical Trial Proposals in Dermatology: Instrument Validation Study

Megan Hauptman[1], MD; Daniel Copley[2], BSc; Kelly Young[1], MD, PhD; Tran Do[1], MD, PhD; Joseph S Durgin[1], MD; Albert Yang[1], MD; Jungsoo Chang[1], MD; Allison Billi[1], MD, PhD; Mio Nakamura[1], MD; Trilokraj Tejasvi[1], MD

[1]Department of Dermatology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI, United States

[2]Anduril, Costa Mesa, CA, United States

Corresponding Author:
Trilokraj Tejasvi, MD
Department of Dermatology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI, United States

Abstract

Background: Large language models (LLMs) are becoming increasingly popular in clinical trial design but have been underused in research proposal development.

Objective: This study compared the performance of commonly used open access LLMs versus human proposal composition and review.

Methods: A total of 10 LLMs were prompted to write a research proposal. Six physicians and each of the LLMs assessed 11 blinded proposals for capabilities and limitations in accuracy and comprehensiveness.

Results: ChatGPT-o1 and Llama 3.1 were rated the most and least accurate, respectively, by human scorers. LLM scorers rated ChatGPT-o1 and DeepSeek R1 as the most accurate. ChatGPT-o1 and Llama 3.1 were rated as the most and least comprehensive, respectively, by human and LLM scorers. LLMs performed poorly on scoring proposals and, on average, rated proposals 1.9 points higher than humans for both accuracy and comprehensiveness.

Conclusions: Paid versions of ChatGPT remain the highest-quality and most versatile option of the available LLMs. These tools cannot replace expert input but serve as powerful assistants, streamlining the development process and enhancing productivity.

(JMIR Dermatol 2026;9:e76674) doi:10.2196/76674

KEYWORDS

artificial intelligence; AI; large language model; research proposal; clinical research; clinical trials; deep learning; machine learning; research design

Introduction

Advancements in artificial intelligence (AI) have led to the development of large language models (LLMs) using algorithms that learn from data and recognize patterns to make decisions based on all available data within a training set [1]. However, AI is limited by the data it is trained on and an inability to account for the nuanced contexts of individual research studies [2]. Researchers are increasingly using LLMs in clinical trial design to improve patient selection, cohort composition, and recruitment [3]. In contrast, the use of LLMs in research proposal development is largely unexplored, and thus, they are perhaps underused. This study aimed to address this gap by comparing the performance of LLMs versus the current gold standard of human proposal composition and review. Our goals were 3-fold: to rate LLMs in composing clinical trial proposals, assess LLMs in scoring clinical trial proposals, and evaluate the ease of using LLMs (including usability and efficiency).

Methods

Overview

Commonly used open access AI platforms (DeepSeek R1, ChatGPT-o3-mini [OpenAI], ChatGPT-o1 [OpenAI], ChatGPT-4o [OpenAI], Claude Sonnet [Anthropic], Claude Opus [Anthropic], OpenEvidence, Grok 2 [xAI], Gemini Advanced [Google], and Llama 3.1 [Meta AI]) were evaluated for use in research proposal drafting. We requested each of the models to do the following:

> *Write a research proposal for a study looking at the use of narrowband-ultraviolet B phototherapy for psoriasis treatment for psoriasis patients of varying skin pigmentation with 3 aims: 1. To understand the factors that affect the response of NB-UVB in psoriasis patients of varying skin pigmentation. 2. Evaluate adverse effects of NB-UVB and their impact on psoriasis patients of varying skin pigmentation. 3. Compare the acute immunologic response to NB-UVB*

XSL•FO
RenderX

in psoriasis patients of varying skin pigmentation using bulk and single-cell RNA sequencing. Include the following sections: 1 page 'Specific Aims' with details on each of the 3 aims, 1/2 page background and significance of the topic, 1 page of 'preliminary data/studies' relevant to the study, 1 page 'experimental design' (include summary of study, inclusion and exclusion criteria, study visits and procedures with an associated table describing specifics of study visits), 1/2 page of 'statistical methods, power calculations and bioinformatic analyses' specific for each aim, 1/4 page of 'potential problems and alternative strategies.' Please have approximately 30 references from reputable sources. Make the proposal a total of 7 pages long in paragraph form, in formal scientific language and at a graduate level.

To assess the outputs, each of the 11 blinded proposals (n=10, 90.9% LLM generated and n=1, 9.1% human written) was systematically reviewed and scored by 6 independent physician evaluators, all with strong research backgrounds. Each evaluator used a standardized Likert scale ranging from 1 to 5 (1="strongly disagree"; 5="strongly agree") to rate each proposal for capabilities and limitations in the LLMs' accuracy and comprehensiveness (Table 1).

Table . Criteria for assessing the accuracy, usability, comprehensiveness, and efficiency of large language models (LLMs).

Domain	Assessment criteria	Scoring methodology
Accuracy	Raters systematically fact-checked all proposal content. Only proposals with fully correct and verified factual information (including cited data, statistics, and conclusions) were rated highly. All references were checked for verifiability, relevance, and reputable source quality.	Rated independently by each evaluator on a Likert scale from 1 to 5 (1="strongly disagree: not accurate"; 5="strongly agree: fully accurate"). Scores were aggregated by calculating the mean of all raters' scores for each proposal.
Comprehensiveness	Assessed by evaluating inclusion and completeness of required proposal sections: specific aims, background and significance, preliminary data and studies, experimental design with inclusion and exclusion criteria and study visits and procedures, statistical methods, power calculations and bioinformatic analyses, and potential problems and alternative strategies. Proposals were further checked to meet format requirements: approximately 7 pages in length and 30 reputable references.	Rated independently on a Likert scale from 1 to 5. The mean score was calculated for all evaluators per proposal.
Usability	Assessed qualitatively based on researchers' (MH and DC) experience using each LLM. Criteria included intuitiveness of the interface, clarity of documentation, and ease of generating proposals without technical guidance.	Rated by 2 nontechnical investigators on a Likert scale from 1 to 5; scores were descriptively summarized.
Efficiency	The time from user input to final output was measured in minutes. Minimal delays and rapid response were rated favorably.	The time (minutes and seconds) for the LLM to complete the query was recorded.

For each domain assessed by human reviewers, individual scores were first tabulated. Scores from the 6 evaluators for each proposal were then aggregated by calculating the mean domain score, yielding an overall mean score per domain for each proposal. These aggregated scores provided a quantitative measure of each proposal's performance relative to evaluator consensus. No additional weighting was applied; each evaluator's score carried equal weight in the final aggregation.

In addition to scientific content review, LLM usability and efficiency, including description of pros and cons, were evaluated by 2 investigators. These qualitative evaluations were collected separately and did not contribute to the aggregated proposal scores.

Ethical Considerations

The authors have adhered to local, national, regional, and international law and regulations regarding protection of personal information, privacy, and human rights. This study did not involve human participants, identifiable private information, or interactions requiring human subjects protections. Accordingly, formal human ethics review approval was not required, and informed consent was not necessary. All data used in this study were deidentified prior to analysis to ensure participant confidentiality. No compensation was provided for participation in this study. These determinations are in accordance with University of Michigan policies and federal regulations (45 CFR 46) governing human research [4]. The research was conducted in compliance with the University of Michigan's guidelines on research ethics.

Results

LLMs Composing Proposals

The human-written proposal obtained a score of 5 for accuracy and comprehensiveness across all human scorers and remained the gold standard (Table 2). Human scorers rated ChatGPT-o1 as the most accurate and Llama 3.1 as the least accurate. When

XSL•FO
RenderX

assessed in scoring LLM-derived clinical trial proposals, LLM scorers rated ChatGPT-o1 and DeepSeek R1 as the most accurate (Multimedia Appendix 1). ChatGPT-o1 and Llama 3.1 were found to be the most and least comprehensive, respectively, by both human and LLM scorers.

Table . Full scores by evaluation criterion for each proposal and model.

Proposal and model	Accuracy (1-5), mean (SD)	Comprehensiveness (1-5), mean (SD)	Usability (1-5), mean (SD)	Efficiency
ChatGPT-4o	2.2 (1.2)	1.8 (1.4)	5.0 (0.0)	1 min, 37 s
Claude Opus	3.3 (1.4)	2.7 (0.6)	5.0 (0.0)	1 min, 30 s
ChatGPT-o1	3.5 (1.6)	4.3 (0.5)	3.5 (0.7)	1 min
ChatGPT-o3-mini	2.8 (1.7)	4.0 (0.6)	4.0 (0.0)	30 s
Claude Sonnet	2.0 (1.3)	1.8 (0.8)	4.0 (0.0)	28 s
DeepSeek R1	3.2 (1.5)	3.3 (1.4)	4.0 (0.0)	1 min, 23 s
OpenEvidence	2.3 (1.5)	1.3 (0.5)	3.5 (0.7)	45 s
Grok 2	3.2 (1.5)	3.0 (0.6)	4.0 (0.0)	1 min, 15 s
Gemini Advanced	2.5 (1.0)	1.5 (0.5)	4.5 (0.7)	37 s
Llama 3.1	1.7 (1.0)	1.5 (0.8)	4.5 (0.7)	20 s
Human proposal	5.0 (0.0)	5.0 (0.0)	N/A[a]	N/A (>10 working d)

[a]N/A: not applicable.

Mean and SD scores per criterion are reported for each proposal and model as assessed by 6 independent physician raters (except for usability, which was rated by 2 nontechnical investigators). Efficiency is reported as actual proposal generation time.

All raw scores are available in Multimedia Appendix 1.

LLMs Scoring Proposals

Overall, LLMs performed poorly on scoring proposals and, on average, rated proposals 1.9 points higher than humans for both accuracy (range 1.3-2.8) and comprehensiveness (range 0.7-3). The Claude Sonnet proposal showed the largest discrepancy between human and LLM scoring, with an average difference of 2.8 (SD 3.4) points for accuracy and 3 (SD 4.2) points for comprehensiveness. Interestingly, the ChatGPT-o1 and DeepSeek proposals both received top scores of 5 for both accuracy and comprehensiveness from all LLMs versus human averages of 4.3 (SD 2.2) and 3.3 (SD 1.9), respectively. The absence of variance at the top of the range (and wide variance in the middle of the range) suggests that the discriminatory power of the LLMs plateaued at the top LLM quality.

Ease of Using LLMs

All open access LLMs were highly efficient and ran in a matter of seconds to minutes (minimum of 20 seconds for Llama 3.1 and maximum of 1 minute and 37 seconds for ChatGPT-4o). When assessed for ease of use, ChatGPT-4o and Claude Opus offered the most intuitive interfaces and were highly usable for researchers (DC and MH) without computer science backgrounds.

Discussion

Principal Findings

LLMs offer powerful tools to assist humans in clinical trial proposal creation. LLMs take only minutes to generate proposals, whereas prior investigations into time commitment for generation of proposals by humans have reported estimates of 116 principal investigator hours, 55 coinvestigator hours, and 38 working days [5,6]. Therefore, judicious use of LLMs in proposal development allows researchers to save significant time in organizing sections, formatting, and ensuring coherence.

To provide guidance for readers, we performed a direct comparison of the tested LLMs, highlighting meaningful differences in performance, usability, and application. Table 3 summarizes these findings, with clear delineation of unique strengths and limitations for each model.

XSL•FO
RenderX

Table . Pros and cons of open access large language models (LLMs).

LLM (AI platform)	Pros	Cons
Overall	• Generally reliable, very user-friendly, and highly comprehensive and efficient	• Occasional factual inaccuracies and hallucinations (eg, fabricated references) • Lack of access to the most recent studies due to their training data cutoffs[a]
ChatGPT	• Most advanced and versatile option of the available LLMs • GPT-4o is the lowest-latency[b] and cheapest model	• Offers more advanced, paid "reasoning" models (GPT-o1 and GPT-o3), but they are computationally expensive and slower
Claude	• Designed with emphasis on alignment with human values • Tends to be more cautious about controversial or sensitive topics	• Models less tailored to clinical contexts compared to ChatGPT
DeepSeek	• Fully open source, promoting transparency and community contributions • Does not have associated license fees	• Struggles with fine-tuning on dialogue • Large models (eg, DeepSeek-Coder-33B) require large amounts of GPU[c] memory
Gemini	• Gemini 1.5 Pro boasts the largest context window[d] as a part of Google's ecosystem • Gemini 1.5 Flash is one of the fastest models	• Struggles to produce quality responses without significant prompt engineering • Concerns about data privacy and use with integration into various Google services
Grok 2	• Integration into X's (formerly known as Twitter) ecosystem allows Grok to stay up-to-date with current events and trends • Offers conversational capabilities tailored for social interaction	• Remains suboptimal compared to Claude 3.5 or GPT-4o • As a result of being directly linked to X, a platform with frequent user-generated content, Grok struggles to moderate sensitive or controversial interactions
Llama 3.1	• Llama 3.2 is one of the fastest models (along with Gemini 1.5) • Optimized for efficiency with lower computational requirements compared to other models	• Technical expertise required for it to run properly • Less user-friendly for researchers without technical support
OpenEvidence	• Offers access to the most recently curated medical research • Most robust and relevant citations	• Weaker reasoning capabilities than those of leading frontier models

[a]LLM training data cutoffs: October 2023 for ChatGPT, April 2024 for Claude Sonnet and July 2024 for Claude Haiku, December 2023 for Llama 3.1, May 2024 for Gemini, and unknown for OpenEvidence and Grok.

[b]Time to first token of tokens received, in seconds, after the application programming interface request is sent.

[c]GPU: graphics processing unit.

[d]Maximum number of combined input and output tokens.

ChatGPT-o1 and ChatGPT-o3-mini demonstrated the highest overall accuracy and comprehensiveness, delivering well-structured proposals with robust citations and high scientific rigor. Llama 3.1 and Gemini Advanced were notably efficient, reliably delivering full proposals with rapid turnaround times, but occasionally produced less nuanced sections in preliminary data or limited discussion. Regarding ease of use, ChatGPT-4o and Claude Opus feature intuitive interfaces and require minimal learning curves, making them ideal for researchers new to AI-powered tools. In contrast, Llama 3.1 and OpenEvidence ranked the lowest in usability as their technical requirements and specialized interfaces can be challenging for new users.

All open access LLMs can aid in initial outlining and creation of research proposals. They can assist in initial brainstorming of a clear researchable question and generating hypotheses based on existing literature. LLMs are useful in literature review and can summarize existing studies related to the proposal topic and identify gaps in current knowledge. Furthermore, all open access LLMs can propose data collection methods, define eligibility criteria based on study objectives, recommend appropriate statistical tests based on study design, and help draft proposal sections. They also allow for iterative refinements, enabling tailored outputs to meet specific requirements or needs. While human verification is always required, LLMs can greatly improve time spent on initial proposal drafting and aid in

XSL•FO
RenderX

mundane tasks associated with proposal writing, including proofreading and revisions, writing administrative sections, and optimizing citations.

Limitations to Consider

All LLMs operate similarly to traditional autocomplete and work by using available contextual clues and a statistical model to predict the most likely next "token" or word. Due to the training data cutoffs of AI models, researchers must manually incorporate the latest literature findings. AI researchers are working on incorporating more access to real-time data, for example, generative pretrained transformer actions [6], but these solutions come with their own trade-offs. Another limitation is that users must verify citations as the model may "hallucinate" or fabricate realistic-sounding but false information. Finally, although AI models such as DALL-E (or others) can create images, they are less effective at producing accurate, clinical-grade figures.

Additionally, current LLMs were largely unable to score proposals and should not replace human review for quality control. The high scores from the LLM raters indicate that the LLMs were unable to detect entire missed protocol sections. Other than Gemini Advanced (who self-scored its written proposal with 3 for accuracy and comprehensiveness), Claude Sonnet, and Llama 3.1, all the LLMs self-scored their own proposals with 5 for both accuracy and comprehensiveness, suggesting overlapping "blind spots" in LLM proposal generation and evaluation.

One limitation of this study is that the order in which the proposals were sent for respondents to review was not randomized. Additionally, the "gold standard" (human proposal) was last, and question order likely played a role, with kinder grading of the LLM-derived proposals before reviewing the human-written proposal. Had the human proposal been first, this would have highlighted missing components of LLM-derived proposals and likely led to harsher human grading of the latter.

Another important limitation is the rapid and frequent versioning of LLM platforms, which poses challenges for scientific reproducibility. As models are updated, their performance and outputs can meaningfully change over time, making it difficult to reproduce results or maintain consistency in studies that rely on AI-generated content. Researchers should document model versions and use dates to mitigate this issue and ensure transparency.

Conclusions

The future of AI in clinical research is expected to be transformative and far-reaching. As AI algorithms continue to evolve, they are likely to become more accurate, comprehensive, efficient, and interpretable, enabling researchers to leverage AI-driven insights for personalized medicine, disease prevention, and improved patient outcomes. In the coming years, AI is anticipated to play a crucial role in optimizing clinical trial design and accelerating drug discovery [7]. The integration of AI with other emerging technologies, such as blockchain and the Internet of Medical Things, could further revolutionize clinical research by improving data security, privacy, and real-time patient monitoring [8]. As these advancements continue to unfold, AI has the potential to democratize access to novel therapies, reduce health care costs, and, ultimately, usher in an era of precision medicine [9].

LLMs offer a transformative approach to drafting research proposals [10]. Paid versions of ChatGPT (ChatGPT-o3-mini and ChatGPT-o1) currently remain the highest-quality (as determined by the Artificial Analysis Quality Index) and most versatile option of the available LLMs, balancing usability, speed, accuracy, and customization [11]. While these tools cannot entirely replace expert input, they serve as powerful assistants, streamlining the development process and enhancing productivity. For optimal results, researchers should combine AI-generated content with their expertise, ensuring precision and adherence to the latest research standards.

Funding

The authors received no specific funding for this work.

Data Availability

The data used in this study can be made available upon request to the corresponding author.

Authors' Contributions

Study conception and design were completed by MH and TT. Material preparation and data collection were performed by MH, DC, KY, TD, JSD, AY, JC, AB, and MN. Data analysis was performed by MH. The first draft of the manuscript was written by MH and DC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Human and large language model (LLM) scoring of LLM performance on accuracy and comprehensiveness.
[DOCX File, 21 KB - derma_v9i1e76674_app1.docx]

XSL•FO
RenderX

References

1. Fliorent R, Fardman B, Podwojniak A, et al. Artificial intelligence in dermatology: advancements and challenges in skin of color. Int J Dermatol 2024 Apr;63(4):455-461. [doi: 10.1111/ijd.17076] [Medline: 38444331]
2. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial intelligence applied to clinical trials: opportunities and challenges. Health Technol (Berl) 2023;13(2):203-213. [doi: 10.1007/s12553-023-00738-2] [Medline: 36923325]
3. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. Trends Pharmacol Sci 2019 Aug;40(8):577-591. [doi: 10.1016/j.tips.2019.05.005] [Medline: 31326235]
4. eCFR :: 45 CFR Part 46 -- Protection of Human Subjects. Electronic Code of Federal Regulations (eCFR). URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46 [accessed 2026-01-05]
5. von Hippel T, von Hippel C. To apply or not to apply: a survey analysis of grant writing costs and benefits. PLoS ONE 2015;10(3):e0118494. [doi: 10.1371/journal.pone.0118494] [Medline: 25738742]
6. Herbert DL, Barnett AG, Clarke P, Graves N. On the time spent preparing grant proposals: an observational study of Australian researchers. BMJ Open 2013 May 28;3(5):e002800. [doi: 10.1136/bmjopen-2013-002800] [Medline: 23793700]
7. GPT actions: customize ChatGPT with GPT Actions and API integration. OpenAI Platform. URL: https://platform.openai.com/docs/actions/introduction [accessed 2025-01-07]
8. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
9. AI: the future of clinical research. ICON. URL: https://careers.iconplc.com/blogs/2024-2/ai-the-future-of-clinical-research [accessed 2025-03-19]
10. Markey N, El-Mansouri I, Rensonnet G, van Langen C, Meier C. From RAGs to riches: utilizing large language models to write documents for clinical trials. Clin Trials 2025 Oct;22(5):626-631. [doi: 10.1177/17407745251320806] [Medline: 40013826]
11. Independent analysis of AI. Artificial Analysis. URL: https://artificialanalysis.ai [accessed 2025-01-15]

Abbreviations

AI: artificial intelligence
LLM: large language model

Edited by R Alhusayen; submitted 28.Aug.2025; peer-reviewed by RQ Wang, T Karen; revised version received 07.Nov.2025; accepted 06.Dec.2025; published 12.Jan.2026.

Please cite as:
Hauptman M, Copley D, Young K, Do T, Durgin JS, Yang A, Chang J, Billi A, Nakamura M, Tejasvi T
Leveraging AI Large Language Models for Writing Clinical Trial Proposals in Dermatology: Instrument Validation Study
JMIR Dermatol 2026;9:e76674
URL: https://derma.jmir.org/2026/1/e76674
doi:10.2196/76674

© Megan Hauptman, Daniel Copley, Kelly Young, Tran Do, Joseph S Durgin, Albert Yang, Jungsoo Chang, Allison Billi, Mio Nakamura, Trilokraj Tejasvi. Originally published in JMIR Dermatology (http://derma.jmir.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Corrigendum: Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study

JMIR Publications Editorial Office[1]

JMIR Publications, Toronto, ON, Canada

Corresponding Author:
JMIR Publications Editorial Office
JMIR Publications
130 Queens Quay E, Suite 1100
Toronto, Ontario M5A 0P6
Canada
Phone: 1 416 583 2040
Email: ed-support@jmir.org

Related Article:

Correction of: https://derma.jmir.org/2020/1/e23724/

(JMIR Dermatol 2026;9:e106381) doi:10.2196/106381

Following the publication of "Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study" [1], concerns were raised regarding potential conflicts of interest between reviewers Alex Gu and Vikas Kotha and the authors. These concerns included prior co-publication with members of the author list and active professional relationships between the reviewers and an author. These conflicts were not disclosed to the JMIR Publications Editorial Office as per journal policy.

This article has been subsequently reassessed by the JMIR Publications Editorial Office, who determined that the content of these reviews did not unduly influence the decision to accept the article. The decision to accept the submission was based primarily on the remaining reviewer's comments and the editor in chief's contemporary appraisal of the manuscript.

As such, the JMIR Publications Editorial Office issues this corrigendum to update the Conflicts of Interest statement:

> *Reviewers Alex Gu and Vikas Kotha were found to have undisclosed conflicts of interest with members of the author list under the terms of JMIR Publications policy. The JMIR Publications Editorial Office determined that the content of the reviews did not unduly influence the decision to accept the article. The decision to accept the submission was based primarily on the remaining reviewer's comments and the editor in chief's contemporary appraisal of the manuscript.*

We regret that these issues were not identified before publication.

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Wei C, Fong A, Quan T, Gupta P, Friedman A. Assessment of altmetrics and plumX metrics scoring as mechanisms to evaluate the top 100 trending hidradenitis suppurativa articles on social media: cross-sectional study. JMIR Dermatol 2020 Nov 19;3(1):e23724. [doi: 10.2196/23724]

XSL•FO
RenderX

Submitted 06.Jul.2026; this is a non–peer-reviewed article;accepted 07.Jul.2026; published 18.Aug.2026.

Please cite as:
Editorial Office JMIRP
Corrigendum: Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study
JMIR Dermatol 2026;9:e106381
URL: https://derma.jmir.org/2026/1/e106381
doi:10.2196/106381
PMID:

©JMIR Publications Editorial Office. Originally published in JMIR Dermatology (http://derma.jmir.org), 18.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of a Self-Test Technology for Chlamydia and Gonorrhea Testing: Observational Before-and-After Study

Bettina Trettin[1,2,3], PhD; Ida Wagner Josefsson[3], MSc; Nadja Trier Munk[1], MSN; Tine Vestergaard[1,2], PhD; Tue Kjølhede[2,3], PhD; Mette Maria Skjøth[1,3], PhD

[1]Department of Dermatology and Allergy Centre, Odense University Hospital, J.B. Winsløws Vej 4, Odense, South Denmark, Denmark

[2]Clinical Institute, Health Sciences, University of Southern Denmark, Odense, South Denmark, Denmark

[3]Centre for Innovative Medical Technology, Odense University Hospital, Odense, South Denmark, Denmark

Corresponding Author:
Bettina Trettin, PhD
Department of Dermatology and Allergy Centre, Odense University Hospital, J.B. Winsløws Vej 4, Odense, South Denmark, Denmark

Abstract

Background: Chlamydia remains highly prevalent, and the prevalence of gonorrhea is increasing. Declining testing rates alongside rising positivity in Denmark highlights the need for accessible testing solutions. Self-test technologies (STTs) may reduce stigma and improve testing uptake.

Objective: This study aimed to examine the effects of implementing an STT on activity-related outcomes, including test volume, testing frequency, and positivity rates for sexually transmitted infections (chlamydia and gonorrhea) at a sexual health clinic.

Methods: A retrospective before-and-after study evaluated the implementation of an STT located in a sexual health clinic at a university hospital over the period of 2022 to 2025. The evaluation consisted of 3 phases: preimplementation phase (phase 0), self-test with booking (phase 1), and drop-in self-test (phase 2). Outcomes included testing volume, frequency, positivity rates, staff resource use, and costs, assessed using descriptive statistics and chi-square tests.

Results: The monthly tests increased from a mean of 163 (SD 51) before implementation to 246 (SD 42) and 283 (SD 78) in phases 1 and 2, respectively. The proportion of patients tested multiple times increased significantly (P=.001). Positivity rates declined from 15.7% (180/1144) to approximately between 12% (652/5417) and 12.6% (464/3689; P=.003). Weekly staff time decreased from 102 to 63.75 hours, and annual personnel costs declined by 37%. Despite increased testing volume, the cost per test decreased from DKK 1179 (US $185) to DKK 501 (US $79).

Conclusions: Implementation of an STT increased testing activity and efficiency while reducing costs and staff resource use, suggesting that self-testing is a scalable and resource-efficient strategy for sexually transmitted infection diagnostics in clinical settings.

(JMIR Dermatol 2026;9:e99694) doi:10.2196/99694

KEYWORDS

sexually transmitted infections; chlamydia; self-testing; evaluation; health technology

Introduction

Chlamydia remains the most prevalent bacterial sexually transmitted infection (STI) worldwide, whereas gonorrhea is the second most common, and its incidence is increasing. Both infections may lead to pelvic inflammatory disease and serious reproductive complications, including ectopic pregnancy and infertility, if left untreated [1-3]. In Denmark, testing rates are declining, whereas the positivity rate is on the rise [4], which emphasizes the need to increase testing uptake and make an effort to reduce transmission. In Denmark, most patients, who are mainly young sexually active people, are tested at their general practitioner. Additionally, there are 6 sexual health clinics offering STI testing. Both options are free of charge. Furthermore, internet-ordered postal home-sampling kits exist, but patients have to wait for days, sometimes weeks, for a result and, if it is positive, have to attend a face-to-face consultation at their general practitioner [5]. This may be a problem as STI testing is associated with embarrassment, stigma, and shame [6-8], and thus, patients request discreet and anonymous testing, preferably without any face-to-face contact with a health care professional.

There is only limited evidence on how acceptable home testing is. Internet-based self-sampling may not be the preferred testing by users, and there seems to be a need for alternative testing strategies [5]. However, it prevents cases of pelvic inflammatory disease and is cost-effective compared to traditional clinic-based screening [9]. To address feelings of embarrassment and stigma

XSL•FO
RenderX

and increase testing, a self-test technology (STT) that allows patients to be tested at a sexual health clinic through self-collected sampling without a face-to-face consultation was developed and tested in a previous study [10]. Thus, patients could attend the clinic as a drop-in at any time during its opening hours. On arrival, they had to scan their social security card at the STT, which verified their identity and linked the test to their electronic medical record. Instructions for self-collection were provided through written guides and video. The STT was co-designed and developed with end users, and an evaluation of its use consisting of qualitative research methods showed that it minimized feelings of shame and awkwardness [10]. The STT was pilot-tested and implemented into daily operation. Using the STT at a sexual health clinic located at a hospital empowered patients as it gave them control to make choices about when to take a test [10] and the opportunity to be tested easily and quickly and obtain a result within 1 or 2 days. In addition to minimized feelings of shame and awkwardness, patients experienced it as an easy solution that both saved time and allowed for flexibility to plan their visits. This resulted in greater confidence and feelings of self-determination and, therefore, has the potential to increase testing uptake [10]. Thus, self-testing may facilitate earlier detection and treatment, thereby reducing onward transmissions and long-term complications such as pelvic inflammatory disease. While improving access to testing is a key objective of STTs, easier access may also have unintended consequences. Reduced barriers to testing could encourage individuals to undergo testing more frequently than clinically indicated, resulting in a higher proportion of negative tests and increased repeated testing among the same individuals. Such changes could have implications for health care resource use and cost-effectiveness. Therefore, in addition to evaluating whether implementation of the STT increased testing uptake, it was important to examine whether improved accessibility was associated with changes in testing frequency and positivity rates that might indicate potential overuse.

The STT was located in a sexual health clinic in a discrete setting at a university hospital and was fully implemented in clinical practice, with over 3400 users per year. Although the qualitative evaluation indicated that the STT has the potential to increase testing rates, it remains important to examine the clinical, organizational, and economic effects of its implementation. Thus, the aim of this study was to examine the effects of implementing an STT on activity-related outcomes, including test volume, testing frequency, and positivity rates for STIs (chlamydia and gonorrhea) at a sexual health clinic. In addition, the consequences for staff resource use and overall economic implications were assessed.

Methods

Study Design and Setting

This retrospective observational before-and-after study was conducted at the Department of Dermatology and Allergy Centre, Odense University Hospital, Denmark. The aim was to evaluate the implementation of the STT for chlamydia and gonorrhea, focusing on its impact on test activity (volume, frequency, and positivity rates), staff resource use, and associated costs.

Implementation Phases

Overview

Data were analyzed across 3 distinct periods:

- Phase 0 (preimplementation phase): February 19, 2022, to October 23, 2022
- Phase 1 (self-test with booking): October 24, 2022, to September 1, 2024
- Phase 2 (drop-in self-test): September 2, 2024, to October 20, 2025

These phases were defined according to the introduction and operational workflow of the STT, which allowed patients to get tested either after a scheduled booking or as drop-in without direct health care professional involvement

Phase 0: Preimplementation Phase

Before the implementation of the STT, all tests were performed at face-to-face consultations held by either 2 nurses or a nurse and a physician. All STI tests are performed using nucleic acid amplification tests (NAATs), which are considered a “gold standard” and used in laboratory diagnosis. Patient-collected samples for chlamydia and gonorrhea using NAATs are feasible and acceptable, with very high agreement between health care professional-collected and self-collected tests [11].

Phase 1: Self-Test With Booking

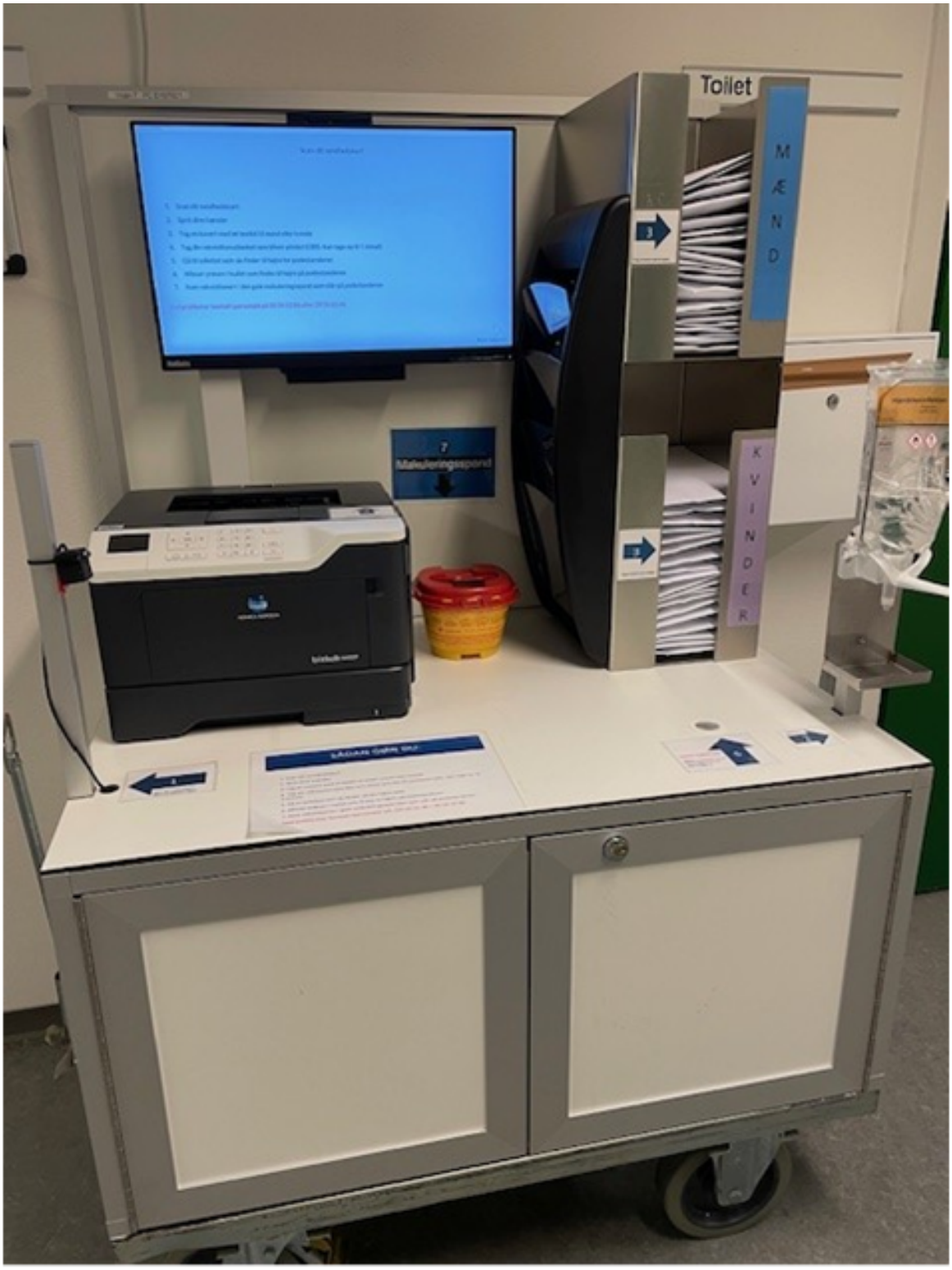
To use the STT, patients had to call a nurse trained in venerology, who conducted a brief interview and entered patient information into the electronic medical record. Patients were then given access to the STT using their personal identification number. Men were instructed to self-collect a urine sample, and women were instructed to self-collect a vaginal swab. Patients had a window of 14 days to use the STT during clinic drop-in hours and were informed that positive results would be sent electronically to their secure digital mailbox, whereas negative results could be checked via the Danish national portal for patient communication[12].

Phase 2: Drop-In Self-Test

The STT was further developed so that patients could access it without any contact with a health care professional and, consequently, without the need for prior booking (Figure 1). Textbox 1 provides further details on the process.

XSL•FO
RenderX

Figure 1. The self-test technology.



XSL•FO
RenderX

Textbox 1. Details on the process of self-testing with drop-ins.

- Upon arrival, patients scanned their national health insurance card at the self-test technology (STT), which verified their identity and linked the test to their electronic health record.
- Patients then collected a self-sampling kit from the STT. Instructions for self-sampling were provided through both written guidance and instructional videos.
- The sample was collected by the patient in a nearby restroom and subsequently deposited into a locked compartment in the self-testing station.
- Samples were collected daily and transported to the laboratory for analysis.
- Patients were informed that positive test results would be delivered electronically to their secure digital mailbox, whereas negative results could be accessed through the Danish national health portal [12].

Data Collection

Data were extracted from electronic health records through a collaborative process involving health technology assessment experts, clinical staff, and data specialists. Two meetings were held to define the data extraction specifications, including variables and inclusion criteria, and review a preliminary dataset for quality and completeness. Extracted variables included test date (year and month), test type (clinical vs self-test), and test result (positive or negative). Months with incomplete data or overlapping phases were excluded to ensure valid comparisons. The data extract contained only the overall test result across the entire dataset. Consequently, the analysis did not distinguish between chlamydia and gonorrhea.

Staff resource use data were collected from the Department of Dermatology and Allergy Centre detailing the number of hours that physicians, nurses, and administrative staff dedicated in a week to working in the STI testing pathway. Staff time included all activities directly related to testing, including patient guidance where required; specimen handling; communication of test results; and follow-up of patients with positive test results, including treatment when indicated. Patients with more complex health care needs requiring additional clinical assessment or management were not included as the study population comprised only individuals who completed testing using the STT. Patients with more complex health care needs were managed through the standard clinical pathway and were therefore outside the scope of the resource analysis. Weekly staff hours were converted into annual full-time equivalents (FTEs) and associated personnel costs based on average gross salaries. Material costs per test (including the self-test kit and other consumables) and laboratory analysis costs were also included. Annual personnel costs were calculated using the annual salaries for each professional group (DKK 986,243 for a senior physician, DKK 499,610 for a nurse, and DKK 489,036 for a medical secretary). All costs are reported in 2025 DKK. For international reference, equivalent values in US dollars can be obtained using the historical exchange rate of July 1, 2025 (DKK 6.3651=US $1).

Data Analysis

Descriptive statistics were used to summarize testing volume, test frequency, test results, and resource use across the study periods. Because the implementation phases differed in duration, testing activity was summarized both as total number of tests and as the mean monthly number of tests, reported with SDs and minimum and maximum values.

Test frequency was assessed at the patient level within a 1-year observation period. To ensure comparability across implementation phases of differing duration, test frequency was evaluated within calendar years rather than across the duration of each phase. Patients were categorized according to the number of STI tests performed (1 test, 2 - 5 tests, and >5 tests), and the distribution of these categories was compared between the pre- and postimplementation periods using the Pearson chi-square test.

Test positivity was calculated as the proportion of positive tests among all tests for chlamydia and gonorrhea. Overall positivity rates were compared across the 3 implementation phases using the Pearson chi-square test. To explore differences by testing modality, positivity rates were also described separately for clinic-based testing and self-testing.

Resource use was assessed as weekly staff time (hours per week) for physicians, nurses, and administrative staff. On the basis of these estimates, FTEs, annual personnel costs, monthly personnel costs, total monthly costs, and costs per test were calculated using the assumptions described above.

Statistical significance was defined as $P<.05$. Analyses were conducted using Stata (version 17; StataCorp).

Ethical Considerations

In Denmark, register studies do not require formal approval from ethics committees according to legislation [13]. This study was approved by the management of the Department of Dermatology and Allergy Centre and by the Danish Data Protection Agency (journal number 22/30101). Furthermore, all data were available in an anonymized format so that specific individuals could not be identified.

Results

Testing Volume

During the defined period (February 19, 2022 - October 20, 2025), a total of 10,250 STI tests were conducted. These tests comprised both clinic-based testing and self-testing and were distributed across the 3 implementation phases shown in Table 1. As the implementation phases differed in duration, direct comparison of the total number of tests across phases should be interpreted with caution. However, Table 1 provides an overview of the distribution of clinic-based testing and self-testing across the study periods and illustrates that the proportion of self-tests relative to clinic-based tests was higher in phase 2 following the introduction of drop-in self-testing.

XSL•FO
RenderX

Table . Number of tests conducted by test modality and implementation phase.

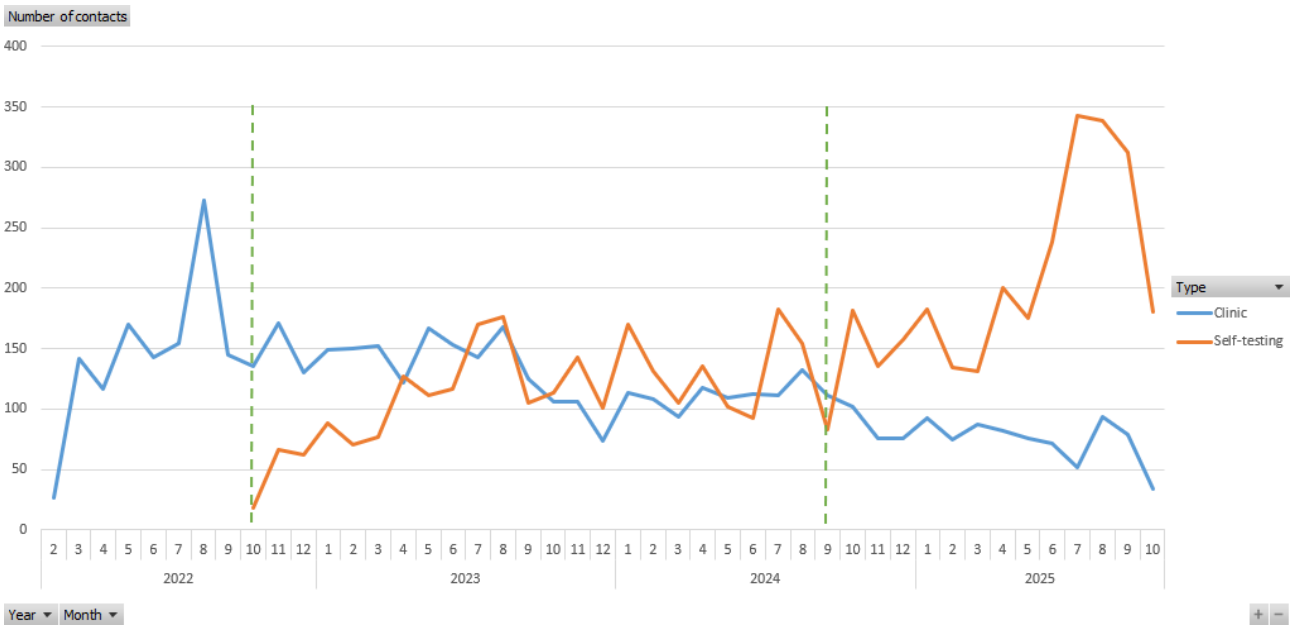
Periods	Clinic-based testing (n=5034), n (%)	Self-test (n=5216), n (%)	Total (n=10,250), n (%)	Tests per month, mean (SD; range)	Number of months
Phase 0 (preimplementation phase)	1144 (22.7)	—[a]	1144 (11.2)	163 (51; 117 - 273)	7
Phase 1 (self-test with booking)	2815 (55.9)	2602 (49.9)	5417 (52.8)	246 (42; 175 - 344)	22
Phase 2 (drop-in self-test)	1075 (21.4)	2614 (50.1)	3689 (36)	283 (78; 197 - 433)	13

[a]Not applicable.

The number of tests performed throughout the study is illustrated in Figure 2. The figure shows the monthly number of tests conducted in the clinic and using the STT. It demonstrates a declining trend in clinic-based testing over the study period (blue line), whereas the number of tests conducted using the STT increased. The green dashed vertical lines mark the beginning of phase 1 and phase 2.

Prior to the implementation of the STT, a mean of 163 (SD 51) tests were conducted per month. Following the introduction of the STT (phase 1), testing activity increased to a mean of 246 (SD 42) tests per month and further to 283 (SD 78) tests per month in phase 2 (drop-in; Table 1).

Figure 2. Number of sexually transmitted infection tests conducted in the clinic and via the self-test technology. The vertical markers identify changes from phase 0 to phase 2.



Test Frequency

Test frequency was compared between the pre- and postimplementation periods using calendar year–based data. Patients were categorized according to the number of STI tests performed (1 test, 2 - 5 tests, and >5 tests). The distribution of patients across these test frequency categories changed after the introduction of the STT (Table 2). A chi-square test demonstrated a statistically significant difference between the pre- and postimplementation periods (P=.001), indicating that the distribution of patients across these categories was no longer the same. Prior to implementation, 82.8% (813/982) of patients were tested only once, 16.8% (165/982) were tested 2 to 5 times, and 0.4% (4/982) were tested more than 5 times. After implementation, the corresponding proportions were 77.3% (5466/7068), 22.2% (1569/7068), and 0.5% (33/7068), respectively. The observed statistical difference is primarily driven by a decrease in the proportion of patients tested only once and an increase in those tested 2 to 5 times. The group with more than 5 tests remained very small and changed only marginally, thus contributing minimally to the chi-square result.

XSL•FO
RenderX

Table . Distribution of patients according to the number of sexually transmitted infection tests performed within a 1-year observation period before and after the introduction of the self-test technology (STT).[a]

Frequency	Before[b] (n=982), n (%)	After (n=7068), n (%)	Total (n=8050), n (%)
1 test	813 (82.8)	5466 (77.3)	6279 (78)
2 - 5 tests	165 (16.8)	1569 (22.2)	1734 (21.5)
>5 tests	4 (0.4)	33 (0.5)	37 (0.5)

[a]Pearson χ^2_2=15.0; *P*=.001.

[b]"Before" includes only patients tested in the clinic, whereas "after" includes patients tested both in the clinic and via the STT.

Test Results

Table 3 shows the overall distribution of positive and negative results across the 3 implementation phases irrespective of test modality. The purpose of this analysis was to assess whether the overall positivity rate changed following implementation of the STT. The overall positivity rate was 12.6% (1296/10,250). The proportion of positive tests decreased from 15.7% (180/1144) prior to implementation to 12% (652/5417) in phase 1 and 12.6% (464/3689) in phase 2. A Pearson chi-square test demonstrated a statistically significant difference in the overall distribution of positive and negative test results across the 3 implementation phases (χ^2_2=11.7; *P*=.003).

Table . Distribution of test results and proportion of positive tests by implementation phase.

Periods	Negative, n (%)	Positive, n (%)
Preimplementation phase (n=1144)	964 (84.3)	180 (15.7)
Phase 1 (self-test with booking implemented; n=5417)	4765 (88)	652 (12)
Phase 2 (drop-in self-test implemented; n=3689)	3225 (87.4)	464 (12.6)
Total (n=10,250)	8954 (87.4)	1296 (12.6)

To further explore whether this overall pattern differed by testing modality, positivity rates were examined separately for clinic-based testing and self-testing (Table 4). Prior to implementation, the positivity rate for clinic-based testing was 15.7% (180/1144). In phase 1, the overall positivity rate decreased to 12% (652/5417), with a lower positivity rate observed for self-testing (243/2602, 9.3%) than for clinic-based testing (409/2815, 14.5%). In phase 2, the overall positivity rate increased slightly to 12.6% (464/3689), with similar positivity rates for clinic-based testing (135/1075, 12.6%) and self-testing (329/2614, 12.6%). Notably, positivity rates for clinic-based testing decreased throughout the study period, whereas positivity rates for self-testing increased from phase 1 to phase 2.

Table . Positivity rates by implementation phase and testing modality.

Periods and testing modality	Positive tests, n/N (%; 95% CI)
Preimplementation phase	
Clinic-based testing	180/1144 (15.7; 13.7-18.0)
Phase 1	
Clinic-based testing	409/2815 (14.5; 13.3-15.9)
Self-testing	243/2602 (9.3; 8.9-10.5)
Phase 2	
Clinic-based testing	135/1075 (12.6; 10.7-14.7)
Self-testing	329/2614 (12.6; 11.4-13.9)

Personnel Costs and Resource Use

Table 5 shows an overview of testing activity, staff resource use, and associated costs across the 3 implementation periods. Total weekly staff time decreased from 102 hours prior to implementation to 63.75 hours in phase 2. Reductions were observed across all staff groups, most prominently among physicians (36%) and administrative staff (88%). This reflects that administrative tasks and physician-led clinical activities were increasingly handled automatically or performed by patients themselves through the STT.

Resource use was converted into annual costs based on the assumptions described in the Methods section. Prior to the implementation of self-testing, testing activities corresponded to a total of 3.45 FTEs distributed across professional groups as follows: approximately 0.74 FTEs for physicians, 2.03 FTEs for nurses, and 0.68 FTEs for administrative staff. This resulted in a total annual personnel expenditure of DKK 2,077,177.

XSL•FO
RenderX

Following the introduction of self-testing with appointment booking (phase 1), personnel usee declined to 2.57 FTEs, comprising approximately 0.61 FTEs for physicians, 1.88 FTEs for nurses, and 0.09 FTEs for administrative staff. Annual personnel costs were reduced to approximately DKK 1,576,966, corresponding to a cost reduction of approximately 24% compared with the period prior to the implementation of the STT.

In the current operational model with drop-in self-testing (phase 2), total personnel use amounted to 2.16 FTEs, including approximately 0.47 FTEs for physicians, 1.60 FTEs for nurses, and 0.09 FTEs for administrative staff. This resulted in a total annual personnel cost of approximately DKK 1,305,088, corresponding to a reduction of approximately 37% compared with the period prior to the implementation of the STT.

Using data on analytical costs and expenditures for test kits together with the average number of tests per month reported in Table 1, monthly costs and costs per test were estimated (Table 5). Prior to the implementation of self-testing, an average of 163 (SD 51) tests were performed per month. Total monthly personnel costs amounted to approximately DKK 173,098, whereas consumables and laboratory analyses accounted for approximately DKK 19,027. Overall, this resulted in a total monthly cost of approximately DKK 192,125, corresponding to an estimated cost of DKK 1179 per test.

With the introduction of self-testing with appointment booking (phase 1), test volume increased to an average of 246 (SD 42) tests per month.

Table . Activity, resource use, and costs across the 3 implementation periods. Values are presented as monthly or annual estimates as indicated. Full-time equivalents (FTEs) include physicians, nurses, and administrative staff. Annual personnel costs were calculated using profession-specific salary levels.

Variables	Before self-testing	Phase 1 (self-testing with booking)	Phase 2 (drop-in self-testing)
Activity			
Tests per month, mean (SD)	163 (51)	246 (42)	283 (78)
Resource use (h per wk)			
Physicians	22.0	18.0	14.0
Nurses	60.0	55.5	47.25
Administrative staff	20.0	2.5	2.5
Total resource use (h per wk)	102.0	76.0	63.75
FTEs	3.45	2.57	2.16
Personnel costs (DKK; DKK 6.3651=US $1 as of July 1, 2025)			
Annual personnel costs	2,077,177	1,576,966	1,305,088
Monthly personnel costs	173,098	131,414	108,757
Other costs (DKK per month)			
Materials and laboratory costs	19,027	28,716	33,035
Total costs (DKK per month)	192,125	160,130	141,792
Cost per test (DKK)	1179	651	501

Discussion

Principal Findings

This evaluation examined the operational, clinical, and economic consequences of implementing an STT for the diagnosis of STIs at the Department of Dermatology and Allergy Centre, Odense University Hospital, and demonstrated that the implementation of the STT increased testing activity while reducing personnel costs and lowering the cost per test. The intervention was introduced in two phases: (1) as appointment-based self-testing (phase 1) and (2) as a drop-in solution (phase 2), allowing for assessment of both gradual and fully decentralized implementation.

Across the study period, a substantial increase in overall testing activity was observed, with mean monthly test volume rising from 163 (SD 51) prior to implementation to 246 (SD 42) in phase 1 and 283 (SD 78) in phase 2. Overall, this indicates a clear increase in testing activity after the introduction of the STT and again following the transition to a drop-in model, which allowed patients to attend without prior appointment or contact with a health care professional. This pattern is consistent with the World Health Organization's self-care intervention framework, which conceptualizes self-testing as a means of lowering structural, temporal, and psychological barriers to care by transferring parts of the diagnostic pathway from the health care system to the individual [14]. By reducing reliance on appointment scheduling, clinical encounters, and direct staff involvement, self-testing enhances accessibility and convenience, which is reflected in the increased uptake observed in this study.

The shift from clinic-based testing to self-testing after the introduction of the drop-in model suggests that self-testing became integrated as a primary testing modality rather than a supplementary service. From an implementation perspective, this development aligns with normalization process theory,

XSL•FO
RenderX

which emphasizes how new practices become embedded in routine care once they are coherent, manageable, and perceived as beneficial by both users and staff. The transition from appointment-based to drop-in self-testing appears to have facilitated such normalization [15].

The analysis demonstrated a shift in patient testing patterns, with a decrease in individuals tested only once and an increase in those tested 2 to 5 times. This change can be interpreted using the behavioral model of health service use by Andersen [16], where improved enabling factors such as ease of access, reduced waiting time, and lower perceived stigma lead to increased use independent of changes in clinical need. Thus, self-testing may have reduced practical and psychological barriers to repeat testing, thereby facilitating more regular engagement with STI screening among users. However, the proportion of individuals being tested more than 5 times remained small and changed only marginally across phases. This suggests that expanded access through self-testing did not result in widespread overuse; instead, the findings indicate a targeted behavior change primarily among individuals who benefited from improved access to repeated testing. Thus, self-testing is a way for people to take responsibility for their own health and places a responsibility to integrate people into existing health care systems, allowing people to make a well-informed decision [17]. This is consistent with the qualitative evaluation of the STT that showed that patients were more satisfied with using the STT rather than attending face-to-face consultations and that using the STT facilitated self-care [10].

The overall positivity rate declined from 15.7% (180/1144) prior to implementation to between 12% (652/5417) and 12.6% (464/3689) during the self-testing phases. This was higher than the national positivity rate for chlamydia in Denmark in 2024 (8.5%), which was published in 2025 [18]. This indicates that the technological solution is an attractive option for individuals who might otherwise not undergo testing and, therefore, has the potential to increase testing uptake. From a screening perspective, this reduction is a well-recognized phenomenon often referred to as risk dilution: when barriers to testing are lowered, individuals at lower average risk are more likely to participate, which reduces the overall positivity rate even if absolute case detection increases [19]. Thus, the observed decline should not be interpreted as reduced clinical relevance or effectiveness of testing. Furthermore, in phase 2, the positivity rate among self-tests increased to a level comparable to that of clinic-based testing. This suggests that, over time, self-testing attracted a patient population with a risk profile similar to that of those traditionally tested in outpatient clinics. From an implementation perspective, this may reflect a stabilization phase in which early exploratory use is replaced by sustained, need-driven use, consistent with normalization process theory [15].

The implementation of the STT had substantial organizational implications. Total weekly personnel time decreased from 102 hours prior to self-testing to 63.75 hours in phase 2. The largest reduction was observed among administrative staff, whose workload was previously driven by appointment booking, telephone contact, and coordination. This pattern is consistent with task shifting and task automation frameworks in which administrative and procedural tasks are either automated or transferred to patients themselves [20]. Reductions were also observed among physicians and nurses, reflecting decreased need for clinician involvement in routine, asymptomatic STI testing. Hence, from a task shifting perspective, self-testing can be understood as a form of patient-mediated task redistribution in which parts of the diagnostic process are safely delegated to users with minimal loss of quality and substantial gains in efficiency. It should be noted that examining potential variation in the weekly number of hours allocated to testing activities over a longer period would be relevant as such variation may influence annual personnel costs.

Despite increased test volume, total monthly costs decreased markedly following implementation, driven primarily by reductions in personnel expenses. Consequently, the cost per test declined from approximately DKK 1179 before self-testing to DKK 501 in phase 2. This finding reflects basic health economic principles, particularly substitution of semifixed personnel costs with variable material and laboratory costs. Moreover, the results are consistent with the lowest effective cost level principle as testing tasks were redistributed to align with the lowest appropriate level of competence, enabling more efficient use of health care personnel resources [21]. As testing volumes increase, average costs per test fall, resulting in improved cost efficiency.

These findings are consistent with those of prior economic evaluations of STI self-testing, which have demonstrated that reductions in staff time and clinic resource use can outweigh increased laboratory expenditures, leading to net cost savings [22]. However, the findings of Wilson et al [22] are based on the principles of pooling NAATs during laboratory analysis, whereas our study used NAATs with laboratory diagnosis of each test. There is little knowledge on using NAATs and STTs in an STI clinical setting and, therefore, little to compare our findings with regarding cost-effectiveness. Thus, the present study adds real-world evidence from a hospital-based setting, reinforcing the economic viability of self-testing as a scalable diagnostic strategy.

Limitations and Implications for Future Research

Several limitations must be considered. This study used an observational before-and-after design based on aggregated electronic patient record data, which precludes causal attribution. Changes in testing behaviors, positivity rates, and volume may also have been influenced by external factors such as seasonal variation, shifts in population behaviors, or broader public health trends.

Personnel resource use was estimated rather than measured through direct time tracking, introducing potential uncertainty. Although estimates were developed by staff with detailed knowledge of workflows, some degree of over- or underestimation is possible. Additionally, reliance on routinely collected data carries an inherent risk of registration errors that cannot be fully identified in aggregated analyses.

Future research should include longer follow-up periods to assess whether testing volume, testing frequency, and positivity rates stabilize over time, as well as more granular data on patient

XSL•FO
RenderX

characteristics and risk profiles. Systematic time registration would further strengthen economic evaluations and support more precise estimates of resource redistribution.

Conclusions

This evaluation and analysis of activity and resource data demonstrates that the implementation of the STT increased testing activity while reducing personnel costs and lowering the cost per test. Despite methodological limitations, the findings indicate that self-testing represents an effective and resource-efficient solution capable of accommodating higher testing volumes without requiring additional staff resources.

Acknowledgments

The authors used the AI tool Microsoft Copilot provided by the University of Southern Denmark to assist in some of the translation of our original work into a second language for inclusion in the manuscript.

Funding

This study was funded by a grant received from the Innovations Fund of Southern Denmark.

Authors' Contributions

Conceptualization: BT
Data curation: BT, IWJ, TK, NTM
Formal analysis: IWJ, TK
Funding: BT, NTM
Investigation: BT, NTM
Project administration: BT, NTM, MMS, TV
Supervision: BT, MMS, TV
Writing—original draft: BT
Writing—review and editing: BT, NTM, IWJ, TK, TV, MMS

Conflicts of Interest

None declared.

References

1. Hoenderboom BM, van Oeffelen AA, van Benthem BH, et al. The Netherlands Chlamydia Cohort Study (NECCST) protocol to assess the risk of late complications following Chlamydia trachomatis infection in women. BMC Infect Dis 2017 Apr 11;17(1):264. [doi: 10.1186/s12879-017-2376-y] [Medline: 28399813]
2. Davies B, Turner KM, Frølund M, et al. Risk of reproductive complications following chlamydia testing: a population-based retrospective cohort study in Denmark. Lancet Infect Dis 2016 Sep;16(9):1057-1064. [doi: 10.1016/S1473-3099(16)30092-5] [Medline: 27289389]
3. Whelan J, Abbing-Karahagopian V, Serino L, Unemo M. Gonorrhoea: a systematic review of prevalence reporting globally. BMC Infect Dis 2021 Nov 11;21(1):1152. [doi: 10.1186/s12879-021-06381-4] [Medline: 34763670]
4. Uge 38 - 2019 [Website in Danish]. Statens Serum Institut. 2018. URL: https://www.ssi.dk/aktuelt/nyhedsbreve/epi-nyt/2019/uge-38---2019 [accessed 2026-08-27]
5. Sacks-Davis R, Gold J, Aitken CK, Hellard ME. Home-based chlamydia testing of young people attending a music festival--who will pee and post? BMC Public Health 2010 Jun 28;10:376. [doi: 10.1186/1471-2458-10-376] [Medline: 20584287]
6. Trettin B, Vestergaard T, Stensgaard A. Understanding young people's barriers to sexually transmitted disease screening and meeting their needs: a focus group study. J Nurs Educ Pract 2015;5(6):81-86. [doi: 10.5430/jnep.v5n6p81]
7. Theunissen K, Bos AE, Hoebe C, et al. Chlamydia trachomatis testing among young people: what is the role of stigma? BMC Public Health 2015 Jul 14;15:651. [doi: 10.1186/s12889-015-2020-y] [Medline: 26169173]
8. Lee AS, Cody SL. The stigma of sexually transmitted infections. Nurs Clin North Am 2020 Sep;55(3):295-305. [doi: 10.1016/j.cnur.2020.05.002] [Medline: 32762851]
9. Huang W, Gaydos CA, Barnes MR, Jett-Goheen M, Blake DR. Cost-effectiveness analysis of Chlamydia trachomatis screening via internet-based self-collected swabs compared with clinic-based sample collection. Sex Transm Dis 2011 Sep;38(9):815-820. [doi: 10.1097/OLQ.0b013e31821b0f50] [Medline: 21844736]
10. Trettin B, Skjøth MM, Munk NT, Vestergaard T, Nielsen C. Shifting grounds-facilitating self-care in testing for sexually transmitted infections through the use of self-test technology: qualitative study. J Particip Med 2024 Aug 14;16:e55705. [doi: 10.2196/55705] [Medline: 39141903]

XSL•FO
RenderX

11. Navas ME, Musoke L, Van Epps P, et al. Extragenital self-collection testing for gonorrhea and chlamydia: a feasibility study for expanding STI screening in the Veterans Health Administration. Am J Clin Pathol 2025 Mar 8;163(3):406-410. [doi: 10.1093/ajcp/aqae135] [Medline: 39394953]
12. sundhed.dk [Website in Danish]. URL: https://www.sundhed.dk/ [accessed 2026-09-01]
13. Hvad skal IKKE anmeldes? [Website in Danish]. VMK og NVK. 2024. URL: https://www.videnskabsetik.dk/ansoegning-til-etisk-komite/overblik-over-anmeldelsespligten/hvad-skal-ikke-anmeldes [accessed 2026-08-27]
14. Self-testing implementation toolkit for HIV, hepatitis and sexually transmitted infections. : World Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240102323 [accessed 2026-08-27]
15. May CR, Mair F, Finch T, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29. [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
16. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
17. den Oudendammer WM, Broerse JE. Towards a decision aid for self-tests: users' experiences in the Netherlands. Health Expect 2019 Oct;22(5):983-992. [doi: 10.1111/hex.12903] [Medline: 31195432]
18. Klamydia - opgørelse over sygdomsforekomst 2024 [Website in Danish]. Statens Serum Institut. 2025. URL: https://www.ssi.dk/sygdomme-beredskab-og-forskning/sygdomsovervaagning/k/klamydia---opgoerelse-over-sygdomsforekomst-2024 [accessed 2026-08-27]
19. Wilson JM, Jungner G. Principles and Practice of Screening for Disease: World Health Organization; 1968.
20. Das S, Grant L, Weller D. Conceptualisation of the SHIFT–SHARE: a new strategic healthcare implementation framework for task shifting, sharing and resource enhancement. Glob Implement Res Appl 2024;4:394-403. [doi: 10.1007/s43477-024-00132-8]
21. Bukh PN, Skovvang Christensen K. Effektiv opgaveløsning med de rette kompetencer til de rette opgaver [Article in Danish]. Økonomi Politik 2023;96(1). [doi: 10.7146/okonomi-og-politik.v96i1.136935]
22. Wilson JD, Wallace HE, Loftus-Keeling M, et al. Swab-yourself trial with economic monitoring and testing for infections collectively (SYSTEMATIC): part 2. A diagnostic accuracy and cost-effectiveness study comparing rectal, pharyngeal, and urogenital samples analyzed individually, versus as a pooled specimen, for the diagnosis of gonorrhea and chlamydia. Clin Infect Dis 2021 Nov 2;73(9):e3183-e3193. [doi: 10.1093/cid/ciaa1546] [Medline: 33044490]

Abbreviations

FTE: full-time equivalent
NAAT: nucleic acid amplification test
STI: sexually transmitted infection
STT: self-test technology

Edited by M Kirchberger; submitted 29.Apr.2026; peer-reviewed by O Stirrup, R Hart-Malloy; revised version received 06.Jul.2026; accepted 18.Aug.2026; published 15.Sep.2026.

Please cite as:
Trettin B, Josefsson IW, Munk NT, Vestergaard T, Kjølhede T, Skjøth MM
Evaluation of a Self-Test Technology for Chlamydia and Gonorrhea Testing: Observational Before-and-After Study
JMIR Dermatol 2026;9:e99694
URL: https://derma.jmir.org/2026/1/e99694
doi:10.2196/99694

© Bettina Trettin, Ida Wagner Josefsson, Nadja Trier Munk, Tine Vestergaard, Tue Kjølhede, Mette Maria Skjøth. Originally published in JMIR Dermatology (http://derma.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effects of Outreach Initiatives on Equitable Access to Dermatology Services: Single-Center Retrospective Analysis

J Mark Albrecht[1], MD; Justin David Lyon[1], BMedSci; Michael Birdsall[1], MD; Eliza Broadbent[1], MPH, MD; Nicole Ufkes[2], MD; Stephanie Zone[2], MD; Anneli Bowen[2], MD; Christopher Hull[2], MD; Lauren Madigan[2], MD; Ashley M Snyder[2,3,4], MPH, PhD; Chaorong Wu[5], PhD; Bethany K Lewis[2], MPH, MD

[1]School of Medicine, University of Utah, Salt Lake City, UT, United States

[2]Department of Dermatology, University of Utah Hospital, 30 Mario Capecchi Drive, HELIX Building, Salt Lake City, UT, United States

[3]Department of Population Health, University of North Dakota, Grand Forks, ND, United States

[4]Department of Population Health Sciences, University of Utah, Salt Lake City, UT, United States

[5]Division of Epidemiology, Department of Internal Medicine, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Bethany K Lewis, MPH, MD
Department of Dermatology, University of Utah Hospital, 30 Mario Capecchi Drive, HELIX Building, Salt Lake City, UT, United States

Abstract

Background: Maldistribution of providers, differences in insurance coverage and acceptance, and underrepresentation of underserved patient populations in dermatology all affect patient access to dermatology services. Few studies have evaluated institutional models to address these pervasive gaps.

Objective: The aim of this study was to evaluate one academic institution's efforts to comprehensively expand access to dermatology services.

Methods: A retrospective analysis was performed using data from the University of Utah Health (UUH) Department of Dermatology's outreach initiatives compared with general dermatology ambulatory clinics on UUH's main campus in Salt Lake City, Utah, over 3 years. Descriptive statistics were reported for patient demographics, insurance coverage, and diagnoses at the encounter level. Statistical methods assessed differences between clinic types. Outreach initiatives were categorized by delivery method: rural in-person clinics, urban in-person charity care clinics, live video teledermatology in rural and incarcerated settings, and asynchronous eConsults (store-and-forward teledermatology).

Results: Outreach initiatives served 2283 distinct patients in 3471 encounters. The general dermatology clinics at UUH served 207,493 patients in 1,141,280 encounters. Overall, UUH outreach initiatives served a more diverse, younger, and Medicaid-insured patient population with a higher prevalence of inflammatory/autoimmune conditions than its general ambulatory dermatology clinics. eConsult and charity care clinics more commonly served patients younger than 65 years of age, and eConsults in particular served a more racially and ethnically diverse patient population compared with the general dermatology clinics.

Conclusions: This study confirms that UUH Dermatology's efforts to enhance health care access effectively reach targeted underserved patient populations that are underrepresented at UUH's general dermatology clinics.

(JMIR Dermatol 2026;9:e94872) doi:10.2196/94872

KEYWORDS

disparities; dermatology; outreach; access; rural; underserved

Introduction

Between 1974 and 2000, office visits to dermatologists in the United States approximately doubled, increasing from about 18 million to about 36 million [1]. Density of dermatology providers also increased overall. However, pervasive gaps in dermatology access continue to exist for underserved populations [2]. Underserved populations are defined as those lacking adequate access to health care, including racial and ethnic minority groups, rural populations, persons of lower socioeconomic status, incarcerated populations, children, LGBTQ+ (lesbian, gay, bisexual, transgender, queer) populations, and older adults, among others [3]. While the factors surrounding these gaps are complex, dermatology workforce distribution and insurance acceptance policies are likely substantial contributors.

Challenges in accessing care are likely shaped by the uneven geographic distribution of providers. In this article, providers

XSL•FO
RenderX

are defined as board-certified dermatologists and nonphysician midlevel providers. In 2013, urban counties had approximately 50 times the density of dermatology providers as rural counties [4]. Geographic disparities are particularly pronounced within pediatric dermatology, with 96.4% of pediatric dermatologists practicing in urban areas and 3.6% in rural areas [5]. A survey of over 2000 members of the American Academy of Dermatology revealed that nearly half of respondents in rural communities endorsed a need for more dermatologists, while over one-third of urban and suburban dermatologists reported an oversupply in their communities [6]. A higher density of dermatologists has been associated with improved survival of more severe conditions, such as Merkel cell carcinoma [7]. Specialist shortages, by contrast, have been linked to longer visit wait times and greater travel distances, with concomitant negative impacts on treatment adherence and follow-up [8].

Even if provider density is adequate, insurance status affects access to care. Patients with Medicaid are less than half as likely as those with private insurance to receive a skin-related diagnosis from a dermatologist [9]. Creadore et al [10] conducted a secret shopper study of 611 dermatology clinics and found that patients with Medicaid experienced less success in securing an appointment and longer wait times than those with private insurance or Medicare. An analysis of the National Medical Expenditure Panel Survey found that patients with Medicaid or Medicare and patients with no insurance had reduced odds (0.75 and 0.39, respectively) of receiving outpatient dermatology care compared with patients with private coverage [11]. Publicly insured and uninsured populations accordingly make up a disproportionately small proportion of the general dermatologic patient base [12].

Dermatology workforce idiosyncrasies and insurance acceptance policies likely influence who accesses dermatologic services, possibly affecting the outcomes of underserved patients. Individuals from racial and ethnic minority groups are less likely to visit an outpatient dermatologist and bear a disproportionate burden of skin cancer morbidity and mortality [9,13-15]. Rural patients experience greater travel distances, lower rates of follow-up, and higher mortality rates [7]. Public insurance has been associated with reduced visit frequency, lower acceptance rates, and longer wait times [9,16]. Incarcerated populations have been reported to experience prolonged time to treatment for melanoma [17]. In summary, these findings suggest that significant portions of the US population lack access to dermatologic specialist care.

Academic medical institutions can play a key role in addressing these disparities. These institutions may have the administrative, workforce, and intellectual resources to create programs targeting underserved populations while teaching the next generation of dermatology providers the skills to treat these populations. However, the literature on possible models for academic institutions to expand access to underserved populations is limited and lacks a comprehensive approach. A cross-sectional study assessing community outreach initiatives across dermatology residency programs found that 43% of outreach offerings were focused on skin cancer prevention and screening, while charity care clinics comprised the second most common outreach initiative at 32.4% [18]. Clearly, large pockets of underserved populations—for example, racial and ethnic minority patients in whom skin cancer rates are low overall, rural patients who live hours away from a free clinic or skin cancer screening offering, or incarcerated patients who would not qualify for care via a free clinic—would not be reached by these highly circumscribed initiatives. Over the past 2 decades, University of Utah Health (UUH) has engaged in efforts to provide comprehensive and equitable access to dermatology services. This review aims to summarize and analyze our institution's outreach efforts in the hope of contributing a potential model for academic institutions to address this pressing issue.

The UUH Department of Dermatology has analyzed the demographics, insurance status, and diagnostic data from most of its outreach efforts to assess whether it is reaching the intended populations and to identify the most common dermatologic diagnoses across different venues.

Methods

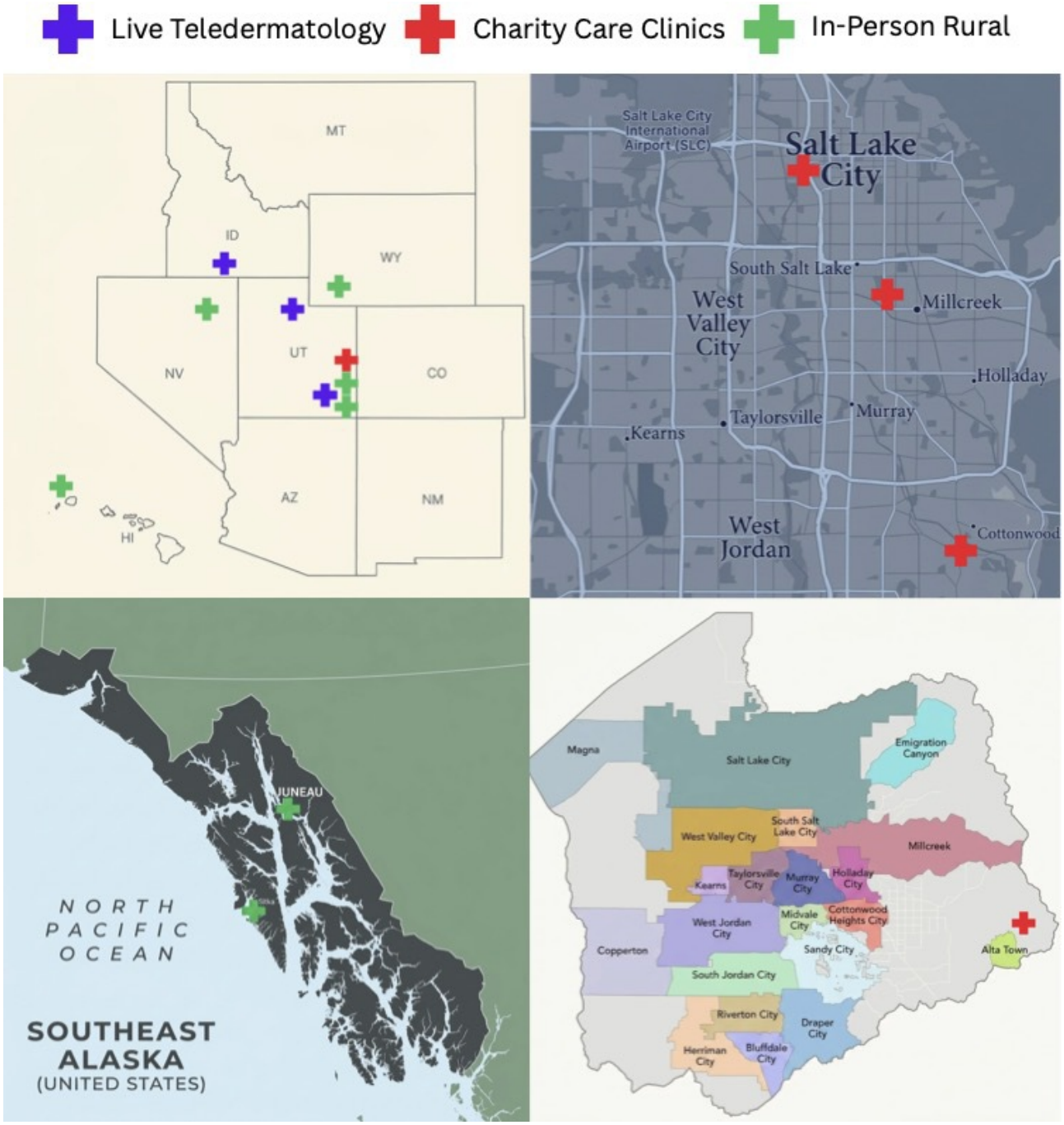
Descriptive Review of UUH Dermatology Outreach Efforts

Overview

Outreach requires community collaboration and works to bridge the health care access gaps observed in disadvantaged populations [19]. UUH seeks to increase equitable access to health care through a mixed model of in-person and technology-assisted outreach. UUH is located in Salt Lake County, Utah, which had an immediate surrounding population of 1.19 million people as of 2022 [20,21]. As a uniquely positioned tertiary care center, UUH's reach extends to a large rural geographic area known as the Intermountain West (Figure 1). Salt Lake County is predominantly White (70.7%), and 8% of its population is enrolled in Medicaid [20,21]. Demographic data indicate that the patient population of UUH's general dermatology clinics skews White, female, older, and Medicare-insured (unpublished internal data; Table 1) [20,21].

XSL•FO
RenderX

Figure 1. Maps of University of Utah Health (UUH) outreach initiative locations by category. (A) and (C) show locations of outreach initiatives in 6 states. (B) and (D) show outreach initiative locations in the Salt Lake City metropolitan area and adjacent Summit County areas. Map generated using Gemini Flash 3.6.



XSL•FO
RenderX

Table . Comparison of demographic characteristics and insurance status between the populations of Salt Lake County and the University of Utah Health (UUH) general dermatology clinics.

Demographic characteristics	Salt Lake County population (N=1,186,257), n (%)[a]	UUH general dermatology clinic population, n/N (%)[b]
Age (years)		
<18	296,884 (25.0)	61,912/997,869 (6.2)
18 - 64	746,293 (62.9)	522,970/997,869 (52.4)
≥65	143,080 (12.1)	412,987/997,869 (41.4)
Sex		
Female	584,135 (49.2)	534,328/997,864 (53.5)
Race[c]		
White	839,228 (70.7)	847,091/870,944 (97.3)
Asian	52,103 (4.4)	13,699/870,944 (1.6)
Pacific Islander	21,429 (1.8)	2,421/870,944 (0.3)
Black	21,707 (1.8)	5,075/870,944 (0.6)
Ethnicity[d]		
Hispanic	233,780 (19.7)	43,399/900,971 (4.8)
Insurance status[e]		
Medicare	109,254 (9.2)	354,326/917,240 (38.6)
Private	705,822 (59.5)	534,774/917,240 (58.3)
Medicaid	96,561 (8.1)	28,140/917,240 (3.1)

[a]US census data, 2022 [21].

[b]Denominators differ due to exclusion of missing data for some variables; percentages may not add up to 100 because the "other," "unknown," and "self-pay/uninsured" categories are excluded.

[c]The "unknown" category is excluded from UUH general dermatology clinics data. The "other" category is included in the denominator but not reported here given lack of specificity or direct comparison to county census data.

[d]The "unknown" category is excluded in UUH general dermatology clinic data.

[e]The "self-pay/uninsured" category is not included.

In-Person Initiatives: Urban Charity Care Clinics and Rural In-Person Clinics

In-person clinics include charity care clinics, which are primarily located in urban centers in Utah, and rural clinics located throughout the Intermountain West and beyond. Textbox 1 provides a concise overview of UUH Dermatology outreach initiatives by type, location, and targeted populations. Charity care clinics are staffed by faculty dermatologists, residents, and/or medical students. Fourth Street Clinic specifically serves persons experiencing homelessness (Salt Lake City, Utah), and Maliheh Clinic (Millcreek, Utah) targets underinsured, uninsured, and Hispanic and Latino patient populations. Midvale Clinic (Midvale, Utah) and People's Health Clinic (Summit County, Utah) similarly target underinsured, uninsured, and Hispanic patients, and Moab Free Clinic (Moab, Utah) targets underinsured, uninsured, and rural patients; however, data are not yet available from the latter 3 sites. Six rural in-person clinics provide dermatology care to rural populations in the Intermountain West and beyond. Providers fly or drive to these clinics, generally on a monthly basis. Providers spend from 1 day to up to 2 weeks on-site providing care and then manage follow-up care by working either with on-site staff or with UUH Dermatology staff and resources. Currently, clinics take place in Moab, Utah (population of 5317); Rock Springs, Wyoming (population of 23,196); Elko, Nevada (population of 20,613); Montezuma Creek, Utah (population of 324, located in San Juan County, with a population of 18,557); Princeville, Kauai, Hawaii (population of 1918); and Sitka and Juneau, Alaska (populations 8968 and 32,255, respectively). Data from the Montezuma Creek, Kauai, Sitka, and Juneau clinics are not yet available due to restrictions on data access or the new status of these clinics. All site partnerships with both rural in-person and urban charity care clinics arose from the members and staff of the aforementioned clinics, who sought regular dermatologic care from UUH Dermatology; each site has its own unique clinic flow, logistical arrangements, and contractual agreements. Of note, no interstate licensure exists, so multiple state licensures for providers must be applied for and maintained.

XSL•FO
RenderX

Textbox 1. University of Utah Health (UUH) Dermatology outreach initiatives by type, location, and targeted populations.

Urban charity care clinics in Utah:

- Fourth Street Clinic, Salt Lake City: persons experiencing homelessness
- Maliheh Clinic, Millcreek: urban under- and uninsured persons and persons of Hispanic ethnicity
- Midvale Clinic, Midvale: urban under- and uninsured persons and persons of Hispanic ethnicity
- People's Health Clinic, Park City: under- and uninsured persons and persons of Hispanic ethnicity
- Moab Free Clinic, Moab: rural population, under- and uninsured persons

Rural in-person clinics:

- Elko, Nevada: rural population
- Rock Springs, Wyoming: rural population
- Moab, Utah: rural population
- Kauai, Hawaii: rural population
- Montezuma Creek, Utah: rural population, Native American persons
- Sitka and Juneau, Alaska: rural population, Native Alaskan persons

Live video teledermatology:

- Salt Lake County Jail, Utah: incarcerated population
- Utah Prison, Utah: incarcerated population
- Blanding, Utah: rural population
- Gooding, Idaho: rural population

Store-and-forward eConsults:

- Medicaid users, children, racial and ethnic minority groups

Telehealth Initiatives: Live Video Teledermatology and Store-and-Forward eConsults

Teledermatology provides several unique methods of outreach, with studies indicating similar clinical outcomes as in-person dermatology services [22,23]. The use of teledermatology (synchronous live video and asynchronous eConsults) in the UUH Dermatology outreach program has expanded the breadth of coverage to both rural and incarcerated settings. For rural populations, live video teledermatology appointments are offered in Blanding, Utah (population of 3319), and Gooding, Idaho (population of 3713). Similar programs have been established at the Utah Prison (occurring monthly since 2003) and Salt Lake County Jail (occurring weekly since 2020). It should be emphasized that incarcerated persons would have severely limited access to dermatologic specialty services without live video teledermatology. Additionally, store-and-forward teledermatology has been offered under the UUH umbrella by means of our eConsult service since 2019; notably, eConsults are covered by Medicaid in Utah. UUH providers submit a history and photographs of the dermatologic issue of interest, which is reviewed by a faculty dermatologist who provides recommendations. This method has been shown by other institutions to decrease patient appointment wait times and provide cost savings to patients and health care systems without sacrificing diagnostic accuracy or patient satisfaction [24].

Study Design

Data from UUH outreach initiatives and general dermatology clinics at the encounter level over a 3-year period were extracted and compared. Differences between outreach initiatives and general dermatology clinics were examined in terms of demographics, insurance status, and diagnostic categories. Diagnostic codes from patient encounters were categorized and classified as benign growth, malignant lesion, premalignant lesion, infectious process, inflammatory/autoimmune process, wound/burn, or other. Total numbers and percentages were reported at the encounter level for insurance, gender, ethnicity, race, age group, and diagnostic category. *t* tests were used to assess continuous variables. χ^2 tests or Fisher exact tests were used, as appropriate, to investigate categorical variables. Outreach initiatives were further classified and analyzed based on delivery method, distinguishing between (1) rural in-person clinics (Elko, Moab, and Rock Springs), (2) live video teledermatology appointments (subcategorized as rural [Gooding and Blanding] and incarcerated [Salt Lake County Jail and Utah Prison]), (3) charity care clinics (Maliheh Clinic and Fourth Street Clinic), and (4) eConsults. All analyses were performed with R software (version 4.3.1; R Foundation for Statistical Computing).

Ethical Considerations

This retrospective chart review was reviewed by the University of Utah Institutional Review Board and determined to be exempt from institutional review board approval (IRB #76927). The

XSL•FO
RenderX

study was conducted in accordance with applicable institutional and ethical standards for research involving human subjects.

As this study involved a retrospective review of existing medical records, informed consent was not obtained from participants. Patient privacy and confidentiality were maintained throughout the study. Data were deidentified prior to analysis, and no names, medical record numbers, or other directly identifying information were included in the study dataset or manuscript. Study data were securely stored and accessible only to authorized members of the research team. No participant compensation was provided.

Figure 1 was generated using a generative AI image model, Gemini (Flash 3.6; Google), accessed through the Health Insurance Portability and Accountability Act (HIPAA)–secure University of Utah Workspace environment. The image was produced through an iterative prompting process: an initial prompt requested a generic map of (map/topic), drawing on user-supplied reference examples and publicly available online sources. Subsequent prompts were used to refine the output (eg, narrowing the geographic scope, removing highway labels, and adjusting background color) until the final version met the intended presentation requirements. No patient data, protected health information, or identifiable images were used in generating this figure. The AI-generated image was reviewed by the authors for accuracy prior to inclusion in the manuscript.

Results

In total, 3471 encounters (2283 distinct patients) occurred across all UUH outreach initiatives, while the university's general dermatology clinics reported 1,141,280 encounters (207,493 patients). A comparison of demographics and insurance coverage between outreach settings and general dermatology clinics at the encounter level is reported in Table 2. Results are significant (P<.001) unless otherwise noted. Overall, encounters in outreach settings were more commonly associated with patients covered by Medicaid than encounters at UUH general dermatology clinics (n=205, 8.3% vs n=28,140, 2.5%). Gender differences were also observed, with male patients seen in significantly more encounters at outreach settings than in general dermatology clinics (n=1886, 54.3% vs n=463,536, 46.5%). Furthermore, compared with general dermatology clinics, significantly more encounters in outreach settings were with patients of Hispanic/Latino (n=309, 11.1% vs n=43,399, 3.8%), American Indian (n=156, 4.7% vs n=2658, 0.3%), Black/African American (n=74, 2.2% vs n=5075, 0.5%), or Native Hawaiian or Pacific Islander race or ethnicity (n=31, 0.9% vs n=2421, 0.2%). Consequently, an overall lower proportion of encounters in outreach settings were with White/Caucasian patients (n=2570, 77.5%) compared with general dermatology clinics (n=847,091, 85.1%). The mean age of patients in encounters across all outreach initiatives was younger than in general dermatology clinics (49.4, SD 18.8 years vs 55.5, SD 20.9 years), with a greater percentage falling in the age range of 18 to 64 years (n=2041, 71.3% vs n=522,970, 52.4%). Diagnostically, a higher proportion of inflammatory/autoimmune diseases (n=1169, 28.0% vs n=60,951, 5.6%) and a lower proportion of malignant (n=318, 7.6% vs n=168,136, 15.5%), premalignant (n=551, 13.0% vs n=204,657, 18.9%), and benign growths (n=871, 20.5% vs n=305,619, 28.2%) were seen across all outreach settings versus general dermatology clinics (Figure 2).

XSL•FO
RenderX

Table . Demographics and insurance comparison between patient encounters in University of Utah Health (UUH) outreach and general clinic settings.

	Outreach encounters (n=3471)	UUH general clinic encounters (n=1,141,280)	Total encounters (N=1,144,751)	*P* value
Age (years), mean (SD)[a]	49.4 (18.8)	55.5 (20.9)	55.4 (20.8)	<.001
Age (years), n (%)[a]				<.001
<18	191 (6.7)	61,912 (6.2)	62,103 (6.2)	
18 - 64	2041 (71.3)	522,970 (52.4)	525,011 (52.5)	
≥65	631 (22.0)	412,987 (41.4)	413,618 (41.3)	
Insurance, n (%)[b]				<.001
Medicaid	205 (8.3)	28,140 (2.5)	28,345 (2.5)	
Medicare	771 (31.3)	354,326 (31.0)	355,097 (31.0)	
Private	1168 (47.5)	534,774 (46.9)	535,942 (46.9)	
Self-pay/uninsured	316 (12.8)	224,040 (19.6)	224,356 (19.6)	
Gender, n (%)[c]				<.001
Female	1585 (45.7)	534,328 (53.5)	535,913 (53.5)	
Male	1886 (54.3)	463,536 (46.5)	465,422 (46.5)	
Ethnicity, n (%)[d]				<.001
Hispanic/Latino	309 (11.1)	43,399 (3.8)	43,708 (3.8)	
Not Hispanic/Latino	2051 (73.8)	857,572 (75.1)	859,623 (75.1)	
Unknown/declined to answer	419 (15.1)	240,309 (21.1)	240,728 (21.0)	
Race, n (%)[e]				<.001
American Indian or Alaskan Native	156 (4.7)	2658 (0.3)	2814 (0.3)	
Asian	25 (0.8)	13,699 (1.4)	13,724 (1.4)	
Black/African American	74 (2.2)	5075 (0.5)	5149 (0.5)	
Native Hawaiian or other Pacific Islander	31 (0.9)	2421 (0.2)	2452 (0.2)	
Other	128 (3.9)	38,073 (3.8)	38,201 (3.8)	
Unknown	330 (10.0)	86,128 (8.7)	86,458 (8.7)	
White/Caucasian	2570 (77.5)	847,091 (85.1)	849,661 (85.1)	

[a]Missing: n=1011.

[b]Missing: n=143,416.

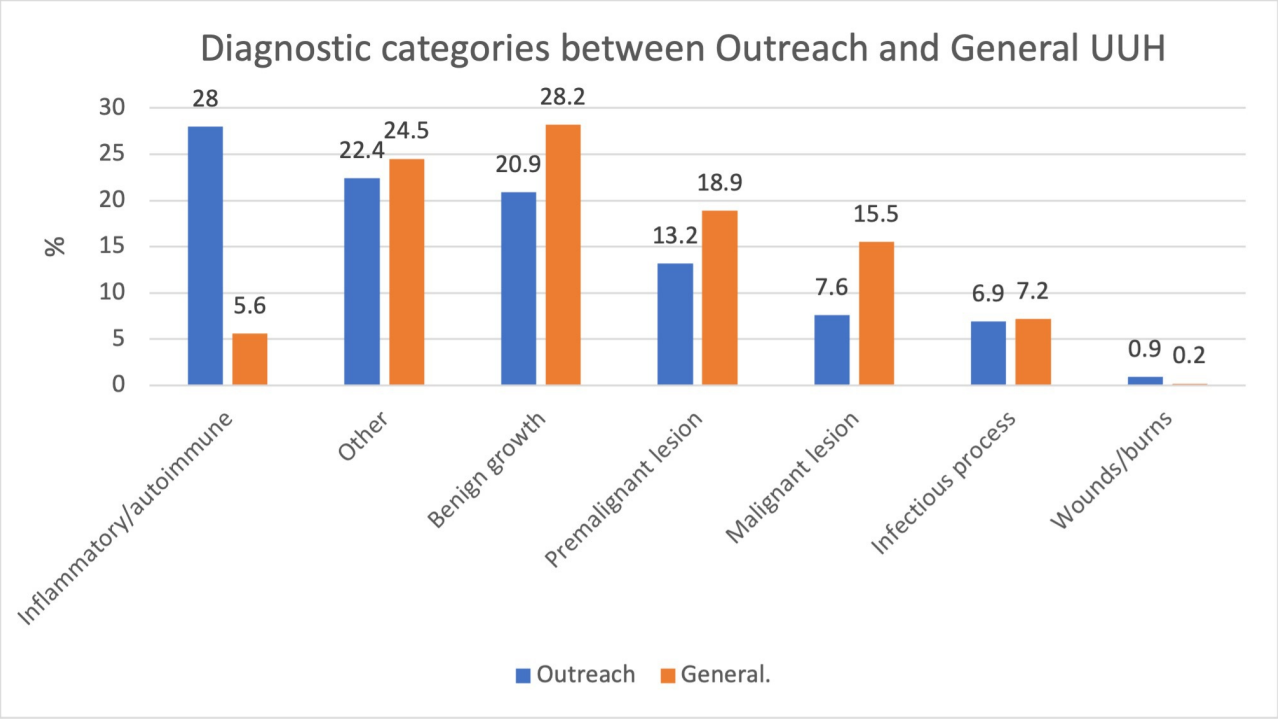
[c]Missing: n=692.

[d]Missing: n=146,292.

[e]Missing: n=144,019.

XSL•FO
RenderX

Figure 2. Percentage of diagnoses by diagnostic category among University of Utah Health (UUH) outreach initiative methods and general dermatology clinics.



Distinctions in demographics and insurance coverage between methods of outreach (teledermatology vs in person) to similar rural locations and general dermatology clinics were also observed (Table 3). Findings are significant (P<.001) unless otherwise noted. Medicaid use was proportionally higher in live video teledermatology encounters (Blanding: 34/150, 25.0%; Gooding: 24/103, 27.0%) compared with general dermatology clinics (28,140/1,141,280, 2.5%). Furthermore, live video teledermatology encounters were more commonly with American Indian and Alaska Native patients (Blanding: 69/158, 43.7%; Gooding: 45/103, 43.7%) compared with general dermatology clinics (2658/1,141,280, 0.3%). Conversely, significantly higher proportions of rural in-person clinic encounters involved non-Hispanic patients (Moab: 641/762, 84.1%; Rock Springs: 460/528, 87.1%) compared with general dermatology clinics (857,572/1,141,280, 75.1%). Notably, patients younger than 18 years of age were more commonly seen in encounters via live video teledermatology (Blanding: 23/158, 14.6%; Gooding: 20/103, 19.4%) than rural in-person clinics (Elko: 44/427, 10.3%; Moab: 36/762, 9.1%; Rock Springs: 21/528, 7.3%) or general dermatology clinics (61,912/1,141,280, 6.2%).

Regarding diagnostic categories at the encounter level, rural in-person clinics more commonly saw a higher proportion of benign growths (Elko: 193/926, 20.8%; Moab: 172/263, 22.6%; Rock Springs: 140/528, 26.5%) compared with live video teledermatology services to rural locations (Blanding: 8/158, 5.1%; Gooding: 5/103, 4.9%), while live video teledermatology services to rural locations diagnosed more inflammatory/autoimmune processes (Blanding: 55/158, 34.8%; Gooding: 35/103, 34.0%) compared with general dermatology clinics (60,951/1,408,744, 5.6%). Findings are summarized in Figure 3 and expanded upon in Table 4.

XSL•FO
RenderX

Table . Demographics and insurance coverage of patient encounters at each outreach initiative location and the University of Utah Health (UUH) general clinics.

	Elko (n=427)	Moab (n=762)	Rock Springs (n=528)	Blanding (n=158)	Gooding (n=103)	Utah Prison (n=336)	SLC[a] Jail (n=64)	eConsult (n=233)	Maliheh (n=234)	Fourth Street (n=626)	General UUH (n=1,141,280)	*P* value
Insurance, n (%)[b]												<.001
Medicaid	63 (15.2)	19 (2.5)	12 (2.3)	34 (25.0)	24 (27.0)	—[c]	1 (1.6)	52 (22.3)	—	—	28,140 (2.5)	
Medicare	126 (30.4)	347 (45.5)	184 (34.8)	49 (36.0)	29 (32.6)	—	4 (6.2)	32 (13.7)	—	—	354,326 (31.0)	
Private	225 (54.3)	352 (46.2)	305 (57.8)	53 (39.0)	36 (40.4)	—	59 (92.2)	138 (59.2)	—	—	534,774 (46.9)	
Self-pay/uninsured	0 (0)	44 (5.8)	27 (5.1)	0 (0)	0 (0)	—	—	11 (4.7)	234 (100.0)	—	224,040 (19.6)	
Gender, n (%)[d]												<.001
Female	264 (61.8)	415 (54.5)	304 (57.6)	68 (43.0)	45 (43.7)	32 (9.5)	2 (3.1)	114 (48.9)	148 (63.2)	193 (30.8)	534,328 (53.5)	
Male	163 (38.2)	347 (45.5)	224 (42.4)	90 (57.0)	58 (56.3)	304 (90.5)	62 (96.9)	119 (51.1)	86 (36.8)	433 (69.2)	463,536 (46.5)	
Ethnicity, n (%)[e]												<.001
Hispanic/Latino	40 (9.4)	10 (1.3)	27 (5.1)	3 (1.9)	2 (1.9)	43 (12.8)	9 (14.1)	40 (17.2)	135 (57.7)	—	43,399 (3.8)	
Not Hispanic/Latino	301 (70.5)	641 (84.1)	460 (87.1)	127 (80.4)	82 (79.6)	186 (55.4)	37 (57.8)	185 (79.4)	32 (19.0)	—	857,572 (75.1)	
Unknown/declined	86 (20.1)	111 (14.6)	41 (7.8)	28 (17.7)	19 (18.4)	107 (31.8)	18 (28.1)	1 (0.6)	0 (0)	—	240,309 (21.1)	
Race, n (%)[f]												<.001
American Indian or Alaskan Native	17 (4.0)	2 (0.3)	2 (0.4)	69 (43.7)	45 (43.7)	1 (0.4)	—	2 (0.9)	1 (0.7)	17 (2.7)	2658 (0.3)	
Black/African American	0 (0)	3 (0.4)	0 (0)	1 (0.6)	1 (1.0)	6 (2.3)	5 (7.8)	7 (3.0)	9 (6.0)	42 (6.7)	5075 (0.5)	
Native Hawaiian or other Pacific Islander	2 (0.5)	0 (0)	0 (0)	4 (2.5)	4 (3.9)	1 (0.4)	10 (15.6)	6 (2.6)	2 (1.3)	2 (0.3)	2421 (0.2)	
Other	5 (1.2)	8 (1.0)	19 (3.6)	6 (3.8)	4 (3.9)	42 (16.0)	9 (14.1)	35 (15.0)	—	—	38,073 (3.8)	
Unknown	76 (17.8)	109 (14.3)	37 (7.0)	27 (17.1)	18 (17.5)	30 (11.5)	18 (28.1)	6 (2.6)	1 (0.7)	8 (1.3)	86,128 (8.7)	
White/Caucasian	327 (76.6)	639 (83.9)	467 (88.4)	51 (32.3)	31 (30.1)	178 (67.9)	22 (34.4)	165 (70.8)	138 (91.4)	552 (88.2)	847,091 (85.1)	
Age (years), mean (SD)[g]	51.1 (21.4)	54.5 (20.5)	56.5 (20.1)	54.7 (23.6)	51.7 (24.6)	41.0 (13.1)	44.3 (11.6)	38.4 (20.9)	48.7 (16.7)	49.5 (10.4)	55.5 (20.9)	<.001
Age (years), n (%)[g]												<.001

XSL•FO
RenderX

	Elko (n=427)	Moab (n=762)	Rock Springs (n=528)	Blanding (n=158)	Gooding (n=103)	Utah Prison (n=336)	SLC[a] Jail (n=64)	eConsult (n=233)	Maliheh (n=234)	Fourth Street (n=626)	General UUH (n=141,280)	*P* value
<18	44 (10.3)	36 (9.1)	21 (7.3)	23 (14.6)	20 (19.4)	0 (0.0)	0 (0.0)	37 (15.9)	10 (4.3)	0 (0.0)	61,912 (6.2)	
18 - 64	244 (57.1)	196 (49.4)	137 (47.9)	79 (50.0)	50 (48.5)	318 (94.9)	62 (96.9)	170 (73.0)	183 (78.2)	602 (96.2)	522,970 (52.4)	
≥65	139 (32.6)	165 (41.6)	128 (44.8)	56 (35.4)	33 (32.0)	17 (5.1)	2 (3.1)	26 (11.2)	41 (17.5)	24 (3.8)	412,987 (41.4)	

[a]SLC: Salt Lake County.

[b]Missing: n=1011.

[c]Not applicable.

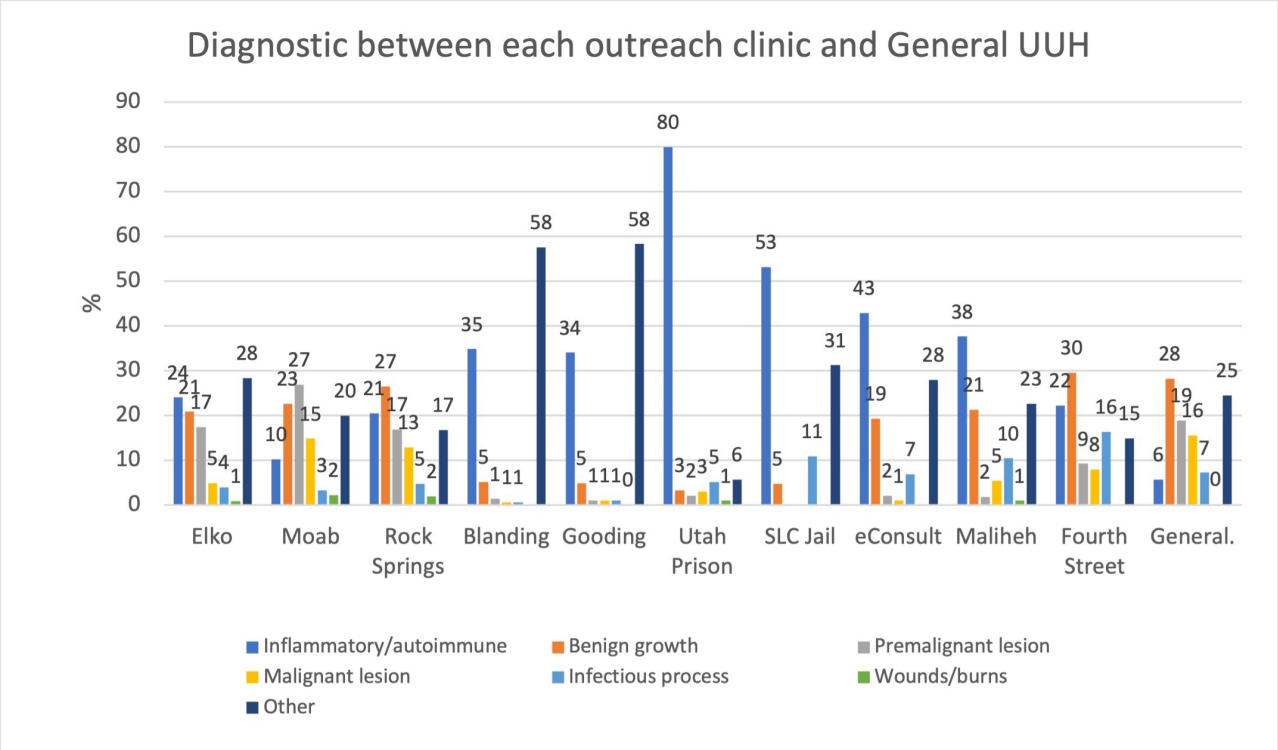
[d]Missing: n=143,416.

[e]Missing: n=692.

[f]Missing: n=146,292.

[g]Missing: n=144,019.

Figure 3. Prevalence of diagnostic categories at individual outreach initiative locations and general dermatology clinics. UUH: University of Utah Health; SLC: Salt Lake County.



XSL•FO
RenderX

Table . Diagnoses by category at individual outreach initiatives and University of Utah Health (UUH) general clinics (N=1,089,488; missing, n=323,495).

	Elko	Moab	Rock Springs	Blanding	Gooding	Utah Prison	SLC[a] Jail	eConsult	Maliheh	Fourth Street	UUH general clinics	*P* value
Total diagnoses, n (%)	926 (0.07)	762 (0.05)	528 (0.04)	158 (0.01)	103 (0.01)	336 (0.02)	64 (0.00)	233 (0.02)	291 (0.02)	838 (0.06)	1,408,744 (99.70)	—[b]
Diagnostic category, n (%)												<.001
Benign growth	193 (20.8)	172 (22.6)	140 (26.5)	8 (5.1)	5 (4.9)	11 (3.3)	3 (4.7)	45 (19.3)	47 (21.3)	247 (29.5)	305,619 (28.2)	
Infectious process	36 (3.9)	25 (3.3)	25 (4.7)	1 (0.6)	1 (1.0)	17 (5.1)	7 (10.9)	16 (6.9)	23 (10.4)	137 (16.3)	78,607 (7.2)	
Inflammatory/autoimmune process	222 (24.0)	78 (10.2)	108 (20.5)	55 (34.8)	35 (34.0)	268 (80.0)	34 (53.1)	100 (42.9)	83 (37.6)	186 (22.2)	60,951 (5.6)	
Malignant lesion	45 (4.9)	113 (14.8)	68 (12.9)	1 (0.6)	1 (1.0)	10 (3.0)	0 (0)	2 (0.9)	12 (5.4)	66 (7.9)	168,136 (15.5)	
Premalignant lesion	161 (17.4)	205 (26.9)	89 (16.9)	2 (1.3)	1 (1.0)	7 (2.1)	0 (0)	5 (2.1)	4 (1.8)	77 (9.2)	204,657 (18.9)	
Wounds/burns	7 (0.8)	17 (2.2)	10 (1.9)	0 (0)	0 (0)	3 (0.9)	0 (0)	0 (0)	2 (0.9)	0 (0)	1773 (0.2)	
Other	262 (28.3)	152 (19.9)	88 (16.7)	91 (57.6)	60 (58.3)	19 (5.7)	20 (31.2)	65 (27.9)	50 (22.6)	125 (14.9)	265,577 (24.5)	

[a]SLC: Salt Lake County.

[b]Not applicable.

Encounters via live video teledermatology for incarcerated populations (Utah Prison and Salt Lake County Jail) and eConsults were also compared with UUH general dermatology in terms of demographics and insurance status (Table 3) and diagnostic data (Figure 3, Table 4). Results are significant (*P*<.001) unless otherwise noted. A significant predominance of encounters in incarcerated settings were with male patients (Utah Prison: 304/336, 90.5%; Salt Lake County Jail: 62/64, 96.9%) and patients between the ages of 18 and 64 years (Utah Prison: 318/336, 94.9%; Salt Lake County Jail: 62/64, 96.9%) compared with general dermatology clinics (463,536/1,141,280, 46.5%, and 522,970/997,869, 52.4%, respectively). In incarcerated settings, diagnostic categorization revealed a significantly higher proportion of inflammatory/autoimmune processes (Utah Prison: 268/336, 80%; Salt Lake County Jail: 34/64, 53.1%) compared with general dermatology clinics (60,951/1,408,744, 5.6%). Encounters via asynchronous eConsult services were more commonly with patients covered by Medicaid (52/233, 22.3%) compared with general dermatology clinics (28,140/1,141,280, 2.5%), and a larger proportion of eConsult encounters were with Hispanic/Latino patients (40/233, 17.2% vs 43,399/1,141,280, 3.8%). Of note, 100 of 233 (42.9%) diagnoses made via eConsult were of inflammatory/autoimmune processes, a significantly higher proportion than at general dermatology clinics.

Finally, an analysis of demographics and insurance coverage (Table 3) and diagnostic data (Figure 3, Table 4) from charity care clinics revealed that encounters at Maliheh Clinic were more commonly with Hispanic/Latino patients (135/234, 57.7%) compared with general dermatology clinics (43,399/1,141,280, 3.8%). A higher percentage of encounters at Fourth Street Clinic compared with general dermatology clinics were with American Indian or Alaskan Native patients (17/626, 2.7% vs 2658/1,141,280, 0.3%) and Black/African American patients (42/626, 6.7% vs 5075/1,141,280, 0.5%); findings are statistically significant (*P*<.001) unless otherwise noted. Encounters at both charity care clinics were disproportionately with patients 18 to 64 years of age (Maliheh: 183/234, 78.2%; Fourth Street Clinic: 602/626, 96.2%) compared with general dermatology clinics (522,970/1,141,280, 52.4%). Further, encounters at the charity care clinics were with a greater proportion of patients with inflammatory/autoimmune conditions (Maliheh: 83/291, 37.6%; Fourth Street Clinic: 186/838, 22.2%) and infectious processes (Maliheh: 23/291, 10.4%; Fourth Street Clinic: 137/838, 16.3%) compared with general dermatology clinics (60,951/1,408,744, 5.6%, and 78,607/1,408,744, 7.2%, respectively).

XSL•FO
RenderX

Discussion

Principal Findings

This retrospective analysis demonstrates that the UUH Department of Dermatology's outreach initiatives serve a meaningfully distinct patient population from that of the institution's general dermatology clinics. Outreach patients were younger (mean age 49.4, SD 18.8 years vs 55.5, SD 20.9 years), more racially and ethnically diverse, and more likely to be Medicaid-insured, differences that reflect the deliberate design of our outreach locations and partnerships. However, when interpreting these findings, we must acknowledge that the data are skewed by the predominantly White, non-Hispanic, female, older than 65 years, and Medicare-insured population served by our main academic medical campus. This problem is not unique to Utah. Access inequities in dermatology are well documented throughout the United States and are likely amplified by provider distribution and insurance acceptance policies within the field of dermatology. Urban areas commonly have a higher concentration of dermatologists, creating care deserts in rural communities. While the literature has thoroughly characterized these disparities, practical models for academic medical institutions to systematically address them remain scarce.

UUH Dermatology has pursued a multifaceted strategy to fill this gap, deliberately extending care to rural populations, racial and ethnic minority groups, and populations who are uninsured, Medicaid-insured, or incarcerated. These groups are of particular interest given that they are rarely captured in national dermatology use data and often experience the steepest barriers to access [9]. The higher proportion of Hispanic/Latino patients across outreach encounters compared with general dermatology clinics (309/3471, 11.1% vs 43,708/1,144,751, 3.8%; $P<.001$) reflects the intentional design of specific outreach initiatives rather than aggregate coincidence. Maliheh Clinic serves under- and uninsured persons (234/234, 100%) and Hispanic/Latino populations (135/234, 57.7%). Aggregate outreach statistics, therefore, should not be interpreted as uniform across sites. Rather, the targeted nature of individual clinics must be considered when drawing conclusions from pooled data.

This interpretive nuance extends in the other direction as well. Given that the rural Intermountain West skews White, non-Hispanic, older, and Medicare-insured and that rural outreach encounters outnumber charity care clinic encounters in absolute terms, the degree of demographic diversification achieved across outreach sites is even more notable than the raw data may initially suggest. The regional population is predominantly rural and includes a high proportion of Medicare users; the outreach program's success in reaching people outside of these groups makes the findings more noteworthy, not less.

The diagnostic distribution across outreach settings likewise reflects structural realities. The markedly higher proportion of inflammatory and autoimmune diagnoses (28.0% vs 5.6%) made through outreach initiatives is unsurprising for several reasons. Teledermatology, which comprises a substantial portion of outreach activity, is less suited to lesion-of-concern evaluations and opportunistic skin checks, which often correlate with neoplastic diagnoses in traditional clinic settings. Additionally, more racially and ethnically diverse patient populations carry a lower baseline burden of skin cancer and are more likely to present with inflammatory dermatoses. This pattern is prevalent at the site level: the highest rates of inflammatory and autoimmune diagnoses were found at the Utah State Prison (268/336, 80.0%) and Salt Lake County Jail (34/64, 53.1%), both of which serve predominantly male patients aged 18 to 64 years. Similarly, inflammatory and autoimmune conditions accounted for 100 of 233 (42.9%) diagnoses made through the eConsult service, consistent with the younger, Medicaid-insured, and racially and ethnically diverse patient population it serves. These patterns have direct implications for clinical planning, informing provider expertise, formulary access, and downstream referral resources tailored to each outreach site.

With a model that uses a variety of outreach methods, including rural in-person clinics, urban charity care clinics, live video teledermatology for rural and incarcerated populations, and store-and-forward eConsults, the UUH Department of Dermatology delivers care to pockets of the population that lack access to more traditional clinics. Future goals include expanding this analysis to encompass all current outreach services, including clinics in Montezuma Creek, Utah; Kauai, Hawaii; and Sitka and Juneau, Alaska, after data transfer agreements have been established. Beyond descriptive analyses, future studies may address clinical outcomes and patient-reported experiences, evaluate the cost-effectiveness of each delivery modality, and determine if sustained outreach contributes to reducing delayed diagnosis rates for dermatologic conditions in underserved communities. Longitudinal data would also shed light on continuity of care and adherence to follow-up recommendations, metrics that are essential for understanding whether intentional outreach translates to improved health equity in a concrete and measurable manner.

Limitations

This study carries several limitations. As a single-center, retrospective, descriptive analysis, these findings reflect the particular geography, infrastructure, and patient population of the Intermountain West. This region largely comprises a White, non-Hispanic, older, Medicare-insured population, which shapes outreach demographics in ways that may not generalize to institutions in more urban or racially diverse settings.

Furthermore, data from newer outreach partnerships and sites without active data transfer agreements were not included in this study, limiting the comprehensiveness of the current analysis. Additionally, eConsult services are currently reimbursed by Medicaid in Utah, which facilitates access for publicly insured patients in this state. However, Medicaid coverage of eConsults varies widely by state, which may limit the direct applicability of this model to institutions in states without such reimbursement privileges.

Race and ethnicity data were self-reported and collected using varied instruments across outreach locations, which may limit direct comparison between sites. Race and ethnicity were recorded as separate variables; patients identifying as Hispanic under ethnicity may also have identified as White/Caucasian

XSL•FO
RenderX

under race, which could have contributed to the apparent overrepresentation of White patients in aggregate data.

This analysis captures demographic, insurance coverage, and diagnostic data at the encounter level without long-term clinical outcomes, patient-reported satisfaction, or measures of downstream health care use. These factors would be necessary to complete a comprehensive overview of the contributions these outreach initiatives make to health equity. Finally, given the retrospective, descriptive nature of this study, the analysis is designed to characterize outreach populations rather than to establish causal relationships or evaluate the efficacy of specific outreach modalities. The aim of future work is to provide a comprehensive evaluation to allow for the fullest and most accurate picture of the populations served, diagnoses seen, and impact on reducing access barriers in underserved populations.

Conclusions

This analysis of the UUH Department of Dermatology's clinical outreach services demonstrates that these initiatives reach distinct underserved patient populations compared with those seen in the institution's general dermatology ambulatory clinics. The breadth and deliberate design of these initiatives offer a working model for how academic medical institutions can move beyond describing health disparities and toward actively addressing them.

Academic dermatology departments are particularly well positioned to lead this work because they can forge community partnerships and have the administrative infrastructure and workforce depth necessary to operate across multiple delivery modalities simultaneously. We present this model and its evaluation as a starting point for cross-institutional dialogue about how dermatology providers can more equitably reach the patients who need these services the most.

Acknowledgments

The authors would like to acknowledge the University of Utah Health (UUH) Department of Dermatology's Community and Global Engagement group. The authors disclose the use of generative AI (Gemini, Flash 3.6; Google) within the Health Insurance Portability and Accountability Act (HIPAA)–secure University of Utah Workspace environment to generate Figure 1, a map created through an iterative series of author-directed prompts. All prompts and generated outputs used to create this figure have been retained on file and are available upon request. The authors take full responsibility for the content and accuracy of the final image.

Funding

No grant funding was obtained specifically for this study. This investigation was supported by the University of Utah Study Design and Biostatistics Center, with funding in part from the National Center for Research Resources and the National Center for Advancing Translational Sciences of the National Institutes of Health under award 8UL1TR000105 (formerly UL1RR025764). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

All data generated or analyzed during this study are included within the published article.

Conflicts of Interest

None declared.

References

1. Buster KJ, Stevens EI, Elmets CA. Dermatologic health disparities. Dermatol Clin 2012 Jan;30(1):53-59. [doi: 10.1016/j.det.2011.08.002] [Medline: 22117867]
2. Glazer AM, Rigel DS. Analysis of trends in geographic distribution of US dermatology workforce density. JAMA Dermatol 2017 May 1;153(5):472-473. [doi: 10.1001/jamadermatol.2016.6032] [Medline: 28296988]
3. Serving vulnerable and underserved populations - archive. US Department of Health and Human Resources. 2018 Jun 18. URL: https://web.archive.org/web/20210729144344/https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/006_Serving_Vulnerable_and_Underserved_Populations.pdf [accessed 2026-04-01]
4. Feng H, Berk-Krauss J, Feng PW, Stein JA. Comparison of dermatologist density between urban and rural counties in the United States. JAMA Dermatol 2018 Nov 1;154(11):1265-1271. [doi: 10.1001/jamadermatol.2018.3022] [Medline: 30193349]
5. Sinha S, Lin G, Zubkov M, Wu R, Feng H. Geographic distribution and characteristics of the pediatric dermatology workforce in the United States. Pediatr Dermatol 2021 Nov;38(6):1523-1528. [doi: 10.1111/pde.14824] [Medline: 34647352]
6. Ehrlich A, Kostecki J, Olkaba H. Trends in dermatology practices and the implications for the workforce. J Am Acad Dermatol 2017 Oct;77(4):746-752. [doi: 10.1016/j.jaad.2017.06.030] [Medline: 28784330]

XSL•FO
RenderX

7. Criscito MC, Martires KJ, Stein JA. A population-based cohort study on the association of dermatologist density and Merkel cell carcinoma survival. J Am Acad Dermatol 2017 Mar;76(3):570-572. [doi: 10.1016/j.jaad.2016.10.043] [Medline: 28212766]
8. Ashrafzadeh S, Peters GA, Shi CR, Nambudiri VE. Demographic and medical school characteristics associated with urban versus rural dermatology practice: a national cross-sectional study. Dermatol Online J 2021 Jan 15;27(1):13030/qt4bn5t4b3. [Medline: 33560785]
9. Mulcahy A, Mehrotra A, Edison K, Uscher-Pines L. Variation in dermatologist visits by sociodemographic characteristics. J Am Acad Dermatol 2017 May;76(5):918-924. [doi: 10.1016/j.jaad.2016.10.045] [Medline: 28069298]
10. Creadore A, Desai S, Li SJ, et al. Insurance acceptance, appointment wait time, and dermatologist access across practice types in the US. JAMA Dermatol 2021 Feb 1;157(2):181-188. [doi: 10.1001/jamadermatol.2020.5173] [Medline: 33439219]
11. Tripathi R, Knusel KD, Ezaldein HH, Scott JF, Bordeaux JS. Association of demographic and socioeconomic characteristics with differences in use of outpatient dermatology services in the United States. JAMA Dermatol 2018 Nov 1;154(11):1286-1291. [doi: 10.1001/jamadermatol.2018.3114] [Medline: 30267073]
12. Resneck JS, Isenstein A, Kimball AB. Few Medicaid and uninsured patients are accessing dermatologists. J Am Acad Dermatol 2006 Dec;55(6):1084-1088. [doi: 10.1016/j.jaad.2006.07.012] [Medline: 17097404]
13. Harvey VM, Paul J, Boulware LE. Racial and ethnic disparities in dermatology office visits among insured patients, 2005-2010. J Health Dispar Res Pract 2016;9(2) [FREE Full text]
14. Gloster HM, Neal K. Skin cancer in skin of color. J Am Acad Dermatol 2006 Nov;55(5):741-760. [doi: 10.1016/j.jaad.2005.08.063] [Medline: 17052479]
15. Sanchez DP, Maymone MBC, McLean EO, et al. Racial and ethnic disparities in melanoma awareness: A cross-sectional survey. J Am Acad Dermatol 2020 Oct;83(4):1098-1103. [doi: 10.1016/j.jaad.2020.04.137] [Medline: 32380221]
16. Alghothani L, Jacks SK, Vander Horst A, Zirwas MJ. Disparities in access to dermatologic care according to insurance type. Arch Dermatol 2012 Aug;148(8):956-957. [doi: 10.1001/archdermatol.2012.804] [Medline: 22911201]
17. Swigert A, Majidian M, Chen L, Vick G, Murina A. Skin cancer in the incarcerated population—a single-center study. Dermatol Surg 2022 Jan 1;48(1):17-20. [doi: 10.1097/DSS.0000000000003247] [Medline: 34608091]
18. Fritsche M, Maglakelidze N, Zaenglein A, Lam C. Community outreach initiatives in dermatology: cross-sectional study. Arch Dermatol Res 2023 Nov;315(9):2693-2695. [doi: 10.1007/s00403-023-02629-y] [Medline: 37243854]
19. Shin HY, Kim KY, Kang P. Concept analysis of community health outreach. BMC Health Serv Res 2020 May 13;20(1):417. [doi: 10.1186/s12913-020-05266-7] [Medline: 32404160]
20. Salt Lake County, UT. Data USA. URL: https://datausa.io/profile/geo/salt-lake-county-ut#health [accessed 2026-08-05]
21. Salt Lake County population 2022 - Census Bureau tables. Census Bureau Data. URL: https://data.census.gov/table?q=salt+lake+county+population+2022 [accessed 2026-08-05]
22. Zakaria A, Maurer T, Su G, Amerson E. Impact of teledermatology on the accessibility and efficiency of dermatology care in an urban safety-net hospital: a pre-post analysis. J Am Acad Dermatol 2019 Dec;81(6):1446-1452. [doi: 10.1016/j.jaad.2019.08.016] [Medline: 31415834]
23. Wootton R, Bloomer SE, Corbett R, et al. Multicentre randomised control trial comparing real time teledermatology with conventional outpatient dermatological care: societal cost-benefit analysis. BMJ 2000 May 6;320(7244):1252-1256. [doi: 10.1136/bmj.320.7244.1252] [Medline: 10797038]
24. Zakaria A, Miclau TA, Maurer T, Leslie KS, Amerson E. Cost minimization analysis of a teledermatology triage system in a managed care setting. JAMA Dermatol 2021 Jan 1;157(1):52-58. [doi: 10.1001/jamadermatol.2020.4066] [Medline: 33206146]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
LGBTQ+: lesbian, gay, bisexual, transgender, queer
SLC: Salt Lake County
UUH: University of Utah Health

Edited by R Dellavalle; submitted 07.Mar.2026; peer-reviewed by LE Gracey, V Long; revised version received 08.Jul.2026; accepted 09.Jul.2026; published 11.Sep.2026.

Please cite as:
Albrecht JM, Lyon JD, Birdsall M, Broadbent E, Ufkes N, Zone S, Bowen A, Hull C, Madigan L, Snyder AM, Wu C, Lewis BK
The Effects of Outreach Initiatives on Equitable Access to Dermatology Services: Single-Center Retrospective Analysis
JMIR Dermatol 2026;9:e94872
URL: https://derma.jmir.org/2026/1/e94872
doi:10.2196/94872

XSL•FO
RenderX

© J Mark Albrecht, Justin David Lyon, Michael Birdsall, Eliza Broadbent, Nicole Ufkes, Stephanie Zone, Anneli Bowen, Christopher Hull, Lauren Madigan, Ashley M Snyder, Chaorong Wu, Bethany K Lewis. Originally published in JMIR Dermatology (http://derma.jmir.org), 11.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficiency and Sustainability in the Skin Cancer Excision Pathway: Life Cycle Assessment and Process Map

Zahra Ahmed[1], MBBS; Alexander Zargaran[1,2], PhD; Sara Moeini[1], BSc; Sara Sousi[1], MRes; David Zargaran[1,2], FRCS (Plast); Ann-Marie Quigley[2], FRCR; Norbert Kang[2], FRCS (Plast); Afshin Mosahebi[1,2], FRCS (Plast)

[1]Division of Surgery and Interventional Science, University College London, Gower Street, London, England, United Kingdom

[2]Department of Plastic and Reconstructive Surgery, Royal Free London NHS Foundation Trust, London, England, United Kingdom

Corresponding Author:
Zahra Ahmed, MBBS
Division of Surgery and Interventional Science, University College London, Gower Street, London, England, United Kingdom

Abstract

Background: Skin cancer is the most commonly diagnosed cancer. In secondary care, patients enter a complex pathway, often involving excision. This multistage process likely has a significant carbon footprint, which has not been previously evaluated.

Objective: The aims of this study were to quantify the carbon footprint of the skin cancer excision pathway and to identify opportunities for pathway optimization to improve both sustainability and efficiency.

Methods: A retrospective process mapping and life cycle analysis of 64 patients undergoing skin cancer excision were performed for patient-facing and organizational steps from diagnosis to follow-up. Estimates for carbon dioxide emissions and areas for pathway remodeling were generated.

Results: The carbon footprint of a patient undergoing skin cancer excision was 118.5 kg carbon dioxide equivalent (CO2eq). The perioperative phase had the highest emissions (53.1 kgCO2eq, 44.8%), but the combined diagnostic and follow-up periods had greater emissions (65.4 kgCO2eq, 55.2%). Imaging and nuclear medicine were responsible for over a third of emissions (42.5 kgCO2eq, 35.9%).

Conclusions: This is the first study to evaluate the carbon footprint of the secondary care skin cancer excision pathway. Multiple inefficiencies were identified, and opportunities for improvement, pathway stratification, diagnostic justification, and adoption of circular economy principles were suggested to increase sustainability and improve standards of care.

(JMIR Dermatol 2026;9:e90087) doi:10.2196/90087

KEYWORDS

skin cancer; carbon footprint; life cycle analysis; sustainability; health care pathways; environmental impact

Introduction

Climate change has been described as the defining issue of our time by the United Nations [1]. The Paris Agreement states that limiting the rise in global temperature to no more than 1.5 °C can preserve a climate suitable for human habitation [2,3]. However, current targets are not being met, with May 2024 being declared as the hottest May on record by the European Commission's Copernicus Climate Change Service [4]. To limit climate change, every industrial sector must be scrutinized, especially those with the highest emissions.

If the health care sector were a country, it would be the fifth largest polluter on the planet [5]. In the United Kingdom, the National Health Service (NHS) accounts for 40% of public sector emissions [6], and despite fewer than 5% of hospital inpatients undergoing surgery, operating theaters contribute up to 25% of hospitals' carbon emissions. Therefore, if the NHS is to reach its net zero by 2045 target, surgical practices must be evaluated and adapted [7,8].

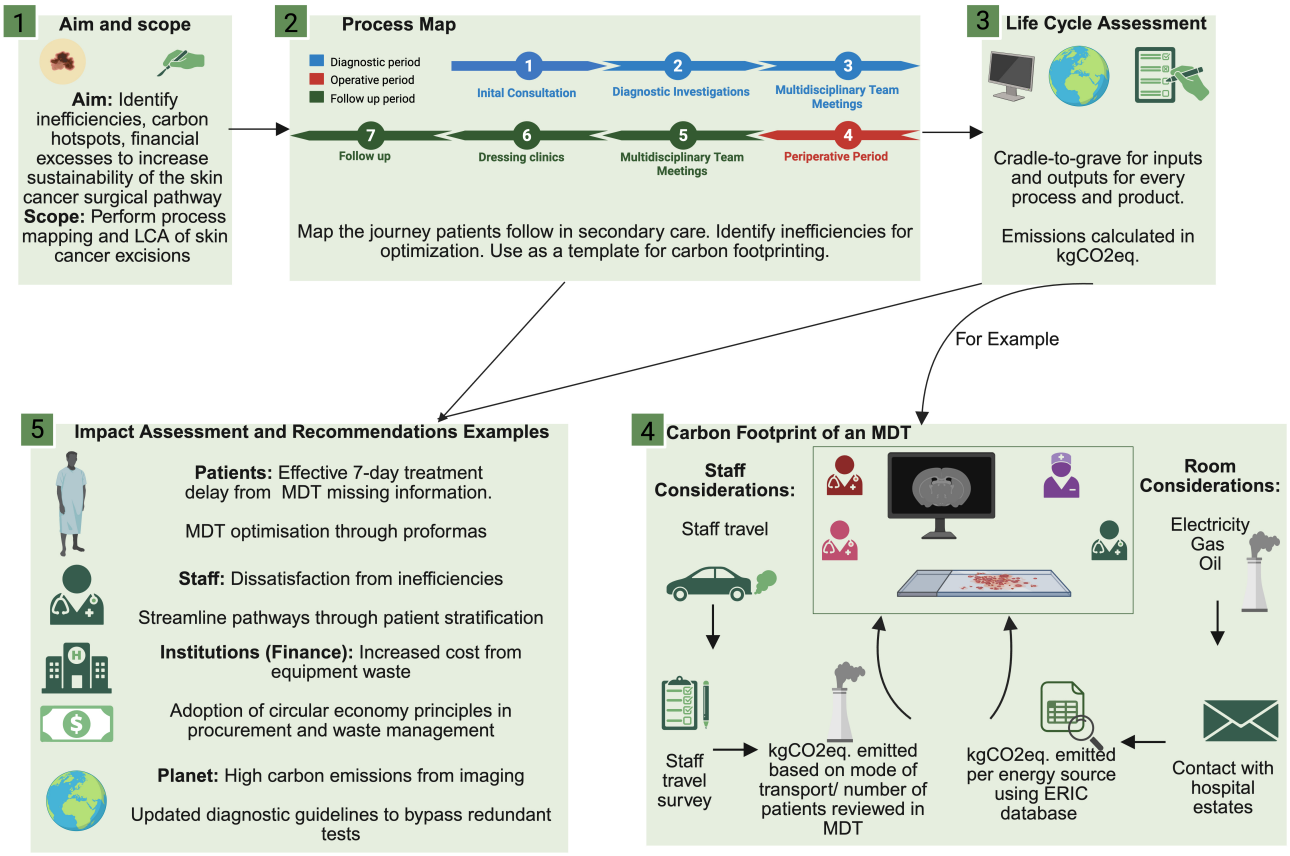
To identify how improvements can be made to reduce emissions, sustainability audits should be performed. Sustainability audits assess the performance of an organization in terms of its sustainable business practices through analyzing internal and external factors [9]. Furthermore, life cycle assessments enable us to evaluate a product or service's environmental consequences throughout its course [10]. Process mapping is a powerful tool that can help guide life cycle assessments, allowing visual representation of patient journeys in a series of consecutive events or steps [11]. Combining process mapping and life cycle assessment techniques within surgical pathways has been described by Glynou et al [12] as a way to understand and recognize opportunities for improvement through identification of carbon hotspots and has been used in reconstructive breast surgery [13].

Skin cancer is the most commonly diagnosed cancer worldwide, with an estimated 1.5 million new diagnoses in 2022 [14]. Treatment options for skin cancer include surgery, radiotherapy, photodynamic therapy, and chemotherapy [15]; however,

XSL•FO
RenderX

between 2013 and 2021, surgery was performed in more than 85% of cases [16]. Due to the high rate of excisional procedures performed annually, the cumulative carbon footprint attributable to the skin cancer surgery service is likely to be high. This study aims to quantify the carbon footprint of skin cancer excisions through process mapping and performing a life cycle assessment. This will identify carbon hotspots and suggest targets to increase the sustainability of the skin cancer surgical pathway (Figure 1).

Figure 1. Material flow analysis of the skin cancer pathway. ERIC: Environmental Reporting Information Centre; LCA: life cycle assessment; MDT: multidisciplinary team.



Methods

Study Design and Population

Following audit approval at our trust, a retrospective service evaluation was performed using data collected from all patients undergoing skin cancer excision from November 7, 2023, to March 1, 2024, at a large teaching hospital in London. Process mapping was performed, with inclusion criteria comprising all patients undergoing elective skin cancer excision with or without lymph node biopsy, using primary wound closure without skin grafts or flaps, and not requiring a postoperative inpatient stay.

Data were categorized into 3 phases (diagnostic, perioperative, and follow-up) encompassing 8 time points, including patient-facing and organizational steps, all of which comprised the skin cancer pathway.

Ethical Considerations

Institutional review board approval was not required for this study, but audit approval was obtained from the Royal Free London NHS Foundation Trust (no reference number available).

Data Collection

A hybrid study was performed. Bottom-up calculations were performed where possible through data collection from electronic patient records, raw data collection intraoperatively, surveys, and communication with hospital departments including procurement, the central sterilization unit, and estates, and contact with external suppliers. Where raw data were not available, reasonable estimates were made using existing literature as benchmarks. Table S1 in Multimedia Appendix 1 details data variables and sources.

Carbon Footprint Calculations

The carbon footprint for each phase of the patient journey on the process map was calculated and summed in units of carbon dioxide equivalent (CO2eq) using multiplication factors from sources including government reports, previous studies, and online resources (Table S2 in Multimedia Appendix 1).

Boundary Setting

A study of this kind requires clear limits to be set and assumptions to be made. Our study followed patients from their first encounter with plastic surgery or dermatology for skin cancer excision following referral to their final dressing clinic. Follow-up beyond dressing clinics was not included as most patients did not require further treatment (44/64, 68.8%). A top-down approach was used for calculations regarding imaging, histopathology, and nuclear medicine due to a lack of required internal and external data for a bottom-up approach.

XSL•FO
RenderX

Results

Patient Pathway Characteristics

There were 64 patients who underwent skin cancer excision in the skin cancer pathway between November 7, 2023, and March 1, 2024, at our center (Table 1 and Figure 2). Most patients were referred into the pathway by dermatologists (n=39, 60.9%). Most other referrals into the pathway came from general practitioners (n=10, 15.6%) and plastic surgeons (n=11, 17.2%) with a minority group (n=4, 6.3%) being referred from other specialties including oncology. The most common method of anesthesia was general anesthesia with total intravenous anesthesia, and no maintenance gases were used.

Table . Variables used to calculate carbon dioxide equivalent emissions for patients undergoing skin cancer excision with or without sentinel lymph node biopsy (N=64).

Variables	Values
Referral type, n (%)	
General practitioner	10 (15.6)
Dermatology	39 (60.9)
Plastic surgery	11 (17.2)
Other[a]	4 (6.3)
Preoperative imaging, n (%)	
Computed tomography	18 (28.1)
Magnetic resonance imaging	28 (43.8)
Other[b]	6 (25)
Number of in-person consultations before surgery, mean	1
Number of preoperative MDT[c] meetings, mean	2
Distance from patient postcode to hospital (km), mean	26.90
Timings for operative phase (min), mean	
Total	247
Anesthetic room	17
Theater	85
Recovery area to discharge	145
Distance from equipment supplier postcode to hospital (km), mean	211.63
Weight of waste, kg (%)	
Total	4.48 (100)
Sharps	0.13 (2.9)
Noninfectious offensive waste	3.85 (85.9)
Dry mixed recyclable waste	0.50 (11.2)

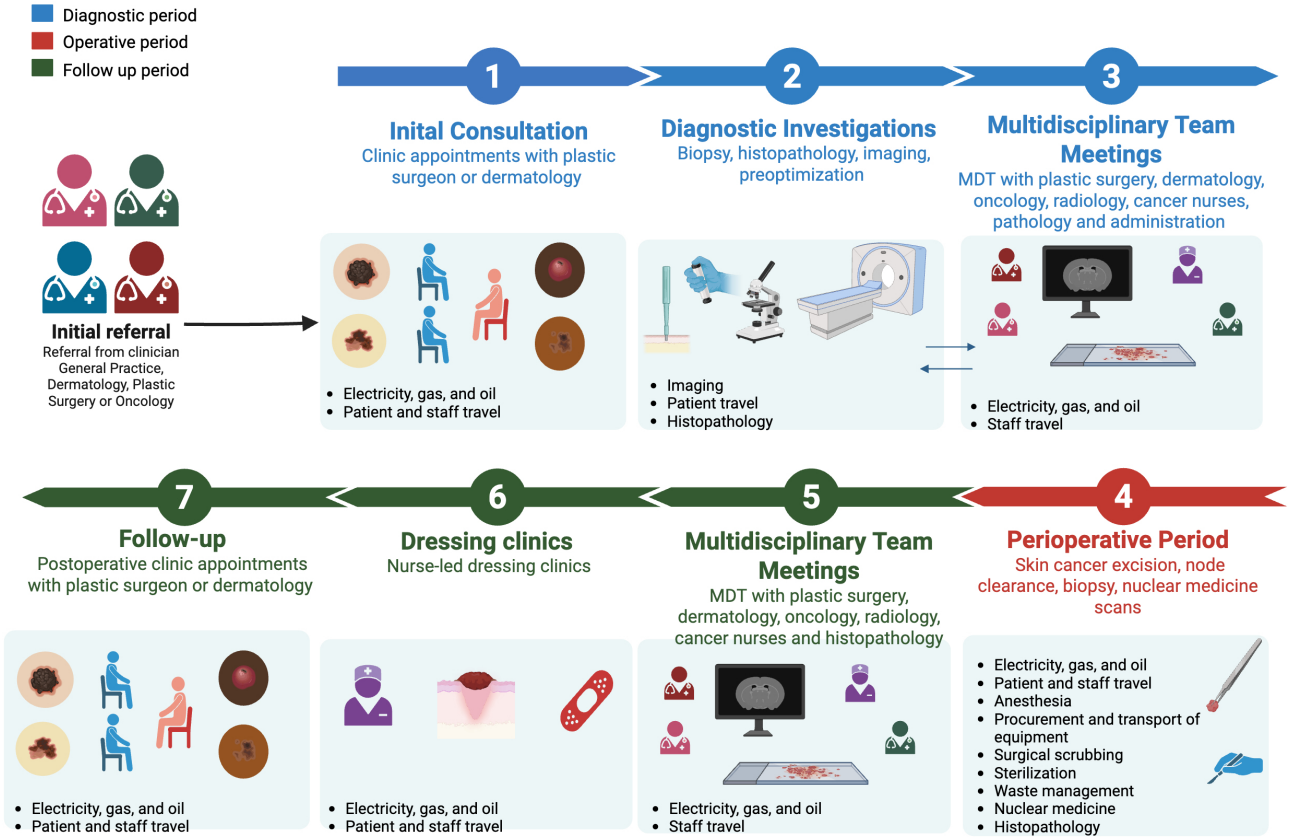
[a]Other referrals were from oncology (n=1) or an unknown source.

[b]Other imaging included ultrasound (n=7) and positron emission tomography (n=2).

[c]MDT: multidisciplinary team.

XSL•FO
RenderX

Figure 2. An illustration of the secondary care skin cancer patient pathway. MDT: multidisciplinary team.



During the preoperative diagnostic phase of the pathway, 67.2% (43/64) of patients had at least one form of imaging, the most common modalities of which were magnetic resonance imaging (MRI; n=28, 43.8%) and computed tomography (CT; n=18, 28.1%). Patients had an average of 1 in-person consultation with either a plastic surgeon or dermatologist before skin cancer excision, requiring an average travel distance of 26.9 km each way to our center. Five (7.8%) patients had virtual preoperative consultations. Patients were discussed an average of 2 times in multidisciplinary team (MDT) meetings. MDTs were hybrid, with 8 staff present in person and 2 histopathologists joining virtually via video conferencing software (Table S1 in Multimedia Appendix 1). Staff traveled an average of 12.8 km to work each day using a variety of transport methods.

During the perioperative phase, patients spent the longest time in the recovery area (145/247 min, 57.8% perioperative phase), and the average time from induction of anesthesia to discharge was 247 minutes. The most common method of anesthesia was a general anesthetic using total intravenous anesthesia and no maintenance gases. There was an average of 6 staff involved in the perioperative phase. In addition to skin cancer excision, 57.8% (37/64) of patients had simultaneous sentinel lymph node biopsy (SLNB) with an average of 2 lymph nodes being sampled. The locations of skin cancer excisions and SLNBs varied and are shown in Multimedia Appendix 2. Nuclear medicine most commonly involved 4 injections of Tc-99m nanocolloid with dynamic and static imaging and intraoperative sentinel node location. using a gamma probe.

Different suppliers were used for equipment procurement, and equipment was transported an average of 211.6 km from the manufacturer to our center. There was a total of 4.48 kg of waste produced from the skin cancer excision procedure. Waste was divided among 3 streams: sharps, noninfectious offensive waste, and dry mixed recyclable waste, the vast majority of which was attributable to noninfectious offensive waste (3.85/4.48, 85.9% total weight), with only 11.2% being recycled.

Postoperatively, patients were discussed in an average of 1 MDT; however, some patients were discussed in multiple MDTs. Patients attended an average of 2 nurse-led dressing clinics and one follow-up clinic with a plastic surgeon; 39.1% (25/64) of patients had virtual follow-up consultations. Most patients did not require further treatment (44/64, 68.8%) but those who did had adjuvant therapy, oncology referrals, or further surveillance.

Overall Carbon Footprint

The total carbon footprint for a patient undergoing skin cancer excision from referral into the skin cancer pathway to their final dressing clinic postoperatively was calculated to be 118.5 kgCO2eq and involved a multitude of contributing factors (Table 2, Figure 3, and Multimedia Appendix 3). The perioperative period had the highest individual contribution to the carbon footprint out of the 3 pathway phases (53.1 kgCO2eq, 44.8%); however, the combination of the diagnostic and follow-up period was greater than this (65.4 kgCO2eq, 55.2%).

XSL•FO
RenderX

Table . Carbon dioxide equivalent (kgCO2eq) for each sector in each pathway period (total emissions=118.46 kgCO2eq).

	Carbon dioxide equivalent emissions (kgCO2eq)	Overall emissions, %
Diagnostic period	27.82	23.4
Average per in-person patient consultations[a]	9.32	7.9
Diagnostic investigations[b,c]	18.09	15.3
MDTs[d,e]	0.28	0.2
Operative period	53.10	44.8
Electricity, gas, oil, and water	2.18	1.8
Patient travel	4.57	3.9
Staff travel	1.67	1.4
Equipment transport	1.16	1
Production of equipment	8.74	7.4
Anesthesia[f]	2.00	1.7
Surgical scrubbing	0.05	0.1
Sterilization of instruments	0.27	0.2
Waste management	1.05	0.9
Laundry of gowns	0.51	0.4
Nuclear medicine[b]	25.00	21.1
Histopathology	5.89	5
Follow-up period	37.67	31.8
Patient consultations	18.64	15.7
MDTs	0.14	0.1
Dressing clinics	18.90	16

[a]Patient consultations included electricity, gas, oil, and staff and patient travel.

[b]Calculated from top-down approach.

[c]Preoperative diagnostic investigations included imaging and histopathology.

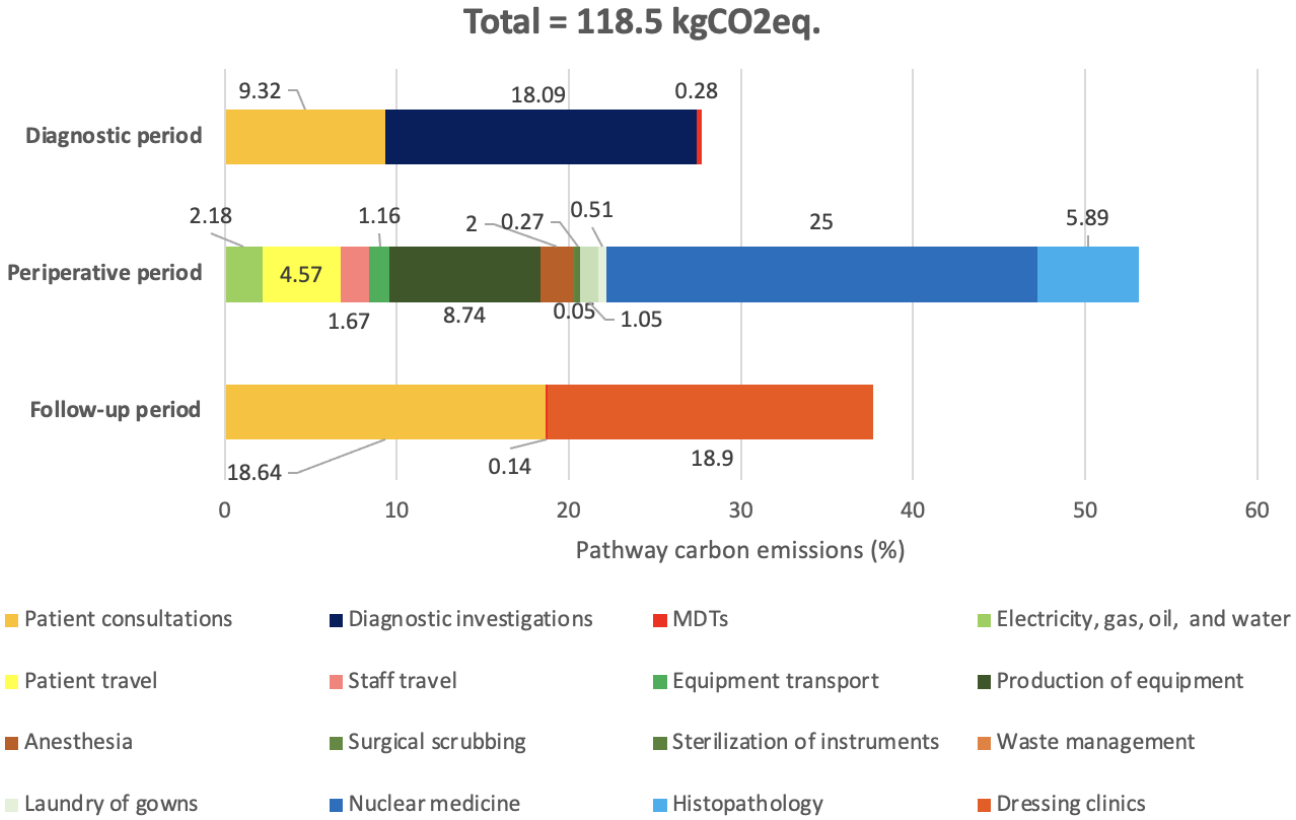
[d]MDT: multidisciplinary team.

[e]MDTs included electricity, gas, oil, and staff travel.

[f]Anesthesia included induction, maintenance, procurement, and transport for anesthetics.

XSL•FO
RenderX

Figure 3. A stacked bar chart showing the contribution of each pathway phase and factors to the overall carbon footprint of the skin cancer patient pathway. MDT: multidisciplinary team.



Diagnostic Period Contributors

The diagnostic period consisted of both patient-facing and organizational components and contributed 27.8 kgCO2eq (23.4%) to the carbon footprint (the equivalent of driving 163.2 km in a petrol-powered car). Diagnostic investigations, including imaging and histopathology, contributed 65.3% of diagnostic period emissions and 15.3% (18.09 kgCO2eq) of overall pathway emissions. In-person patient consultations contributed an average of 9.32 kgCO2eq (7.9%) per patient, and preoperative MDTs contributed 1% (0.28 kgCO2eq) of diagnostic period emissions.

Perioperative Contributors

The perioperative period was the greatest contributor to emissions across the skin cancer pathway (53.1 kgCO2eq, 44.8%), which was the equivalent of driving 311.6 km in a petrol-powered car. Nuclear medicine was responsible for almost half of total perioperative emissions (25.0 kgCO2eq, 47.1%) and 21.1% of overall pathway emissions. The second greatest contributor to perioperative emissions was the production of instruments (8.7 kgCO2eq, 16.5% perioperative emissions), followed by histopathology and patient travel (11.1% and 8.6% of perioperative emissions, respectively). Other contributors to perioperative period emissions included electricity, gas, oil, and water use; staff travel, equipment transport, anesthesia, surgical scrubbing, waste management, and laundry.

Follow-Up Period Contributors

The follow-up period contributed 37.6 kgCO2eq (31.8%) to overall pathway emissions, almost all of which were attributable to in-person nurse-led dressing clinics and follow-up clinics (37.5 kgCO2eq, 99.7% of the follow-up period emissions). Remaining emissions came from postoperative MDTs (0.14 kgCO2eq per patient).

Discussion

This study used process mapping and performed a life cycle assessment for patients on the skin cancer patient pathway in secondary care. Through consideration of both patient-facing and organizational steps, this study estimated the approximate carbon footprint from referral to the final dressing clinic to be 118.5 kgCO2eq. This is equivalent to driving a petrol-powered car a distance of 472.2 km or flying at 1317 km, approximately the distance from London to Bilbao, Spain [17]. All 3 phases of the pathway had significant contributions to emissions, despite the perioperative phase having the highest emissions (53.1 kgCO2eq, 44.8%).

Process mapping was first used in the manufacturing industry and has since been adopted by other industries, such as business and health care. Within health care, process mapping enables direct visualization and measurement of the delivery of complex patient pathways at each point in time [18]. Process mapping, therefore, allows inefficiencies in pathways to be identified and improved. This study found multiple inefficiencies within the skin cancer pathway, many of which, if addressed, have the potential to increase patient throughput, improve patient outcomes, and increase staff satisfaction. A recent report by Cancer Research UK found that in 7% of MDTs, decisions were deferred due to missing information or team members [19]. Our study found substantially higher rates of missing information, with an average of 2 MDTs before surgery per patient due to

XSL•FO
RenderX

missing imaging, pathology, and antigen testing results and a lack of coordination between departments. This led to patients being rescheduled for discussion at the next MDT, causing an effective 7-day delay in the treatment plan and a wasted slot for discussion of other patients. Suggestions for improvement to the MDT include a streamlined protocol and standardized proforma to enable more time for discussion of more complex patients (an issue ranked as the second highest priority among the changes necessary for the skin cancer MDT) [20]. Clinical guidance from the National Institute for Health and Care Excellence (NICE) dictates that core MDT members should attend more than 50% of the total meetings annually [21]. The skin cancer MDT in this study was a hybrid meeting with histopathologists joining virtually. This enabled improved meeting attendance by all clinicians and simultaneously reduced carbon emissions from staff travel.

The skin cancer patient pathway is currently multistaged and the same for all patients regardless of varying diagnostic uncertainty on patient diagnosis and/or treatment. We propose a pathway restructure through stratification after the initial consultation. Restructuring surgical pathways and process maps via risk stratification has been used within perioperative medicine by anesthetists but has yet to be used for surgical decision-making [22]. Within the skin cancer pathway, low-risk patients with limited uncertainty could follow a simpler pathway, bypassing potentially unnecessary logistical hurdles, and only those with higher risk or uncertainty would be referred using the traditional pathway with multiple MDTs to allow collaboration between specialties. Furthermore, lesions that mimic skin cancer account for a significant proportion of all excisions. The NICE reported in 2025 that 94% of excisions were classified as noncancer or nonurgent cases [23]. This burden therefore carries a significant environmental footprint due to the cumulative volume of these removals.

Medical imaging departments are the greatest contributors to greenhouse gas emissions, responsible for approximately 1% of global emissions [24]. Hanneman et al [25] reported that 1 imaging department covering 5 hospital sites had comparable energy use to that of 422 single-family homes or the average greenhouse gas emissions generated by 770 gasoline-powered passenger vehicles driven for 1 year. A life cycle assessment by Thiel et al [26] found that energy consumption from clinical use of imaging equipment accounted for more than half of departmental greenhouse gas emissions. In this study, imaging and nuclear medicine contributed more than one-third of total carbon emissions across the skin cancer pathway (42.5 kgCO2eq, 35.9%). The most common form of preoperative imaging in the diagnostic period was MRI, the modality reported to have the highest carbon emission output [27]. Clinicians should be advised to use toolkits such as that proposed by Brown et al [28], whose recommendations include using and regularly updating clinical decision-making support tools to justify imaging requirements and to limit the number of unnecessary images, which increase energy use even after scanner use due to the need for data storage servers and cooling systems. Furthermore, most patients in this study had sentinel node scintigraphy as part of their procedure (37/64, 57.8%); similar pressures to reduce nuclear medicine use where possible exist. Potential strategies include further involvement of patients in the decision-making process by including the carbon cost of imaging as part of informed consent and carbon offset payments by governments or insurance companies [29].

This study is the first to calculate an approximate carbon footprint for the skin cancer patient pathway. Ang et al [30] found the mean carbon emissions for nonmelanoma skin cancer surgical procedures ranged from 29.82 to 34.31 kgCO2eq. Our study found a similar estimate of 28.1 kgCO2eq (23.7% of overall pathway emissions) for the perioperative period (after excluding nuclear medicine). This therefore shows that calculating the emissions solely during the perioperative phase, while useful, underestimates the overall contribution of these procedures. We therefore recommend that future studies in other pathways do not overlook the stages surrounding the procedure itself.

The proportion of theater waste that is potentially recyclable is estimated to be 55% by weight [31]. In this study, only 11.2% of the total weight of waste was recycled. While waste minimization, recovery, and reuse should be priorities in line with circular economy principles, recycling remains important for increasing sustainability where waste is unavoidable. Suggestions to increase recycling proposed in the Royal College of Surgeons Green Surgery Report include decontamination of infectious waste beyond that already used for surgical instruments, increased use of high-quality recyclable materials in manufacturing that are easy to disassemble, and adoption of a closed-loop recycling process enabling recycled material to be reused for the same market purpose [32]. Adoption of circular economy principles and increased recycling also bring financial benefits, which can be used to incentivize hospitals to promote sustainability [33].

In the United Kingdom, the NHS' long-term plan and the Royal College of Physicians (RCP) have urged a radical change to the way in which outpatient consultations are conducted [34,35]. An RCP survey found that 28% of doctors said that patients could have been seen virtually, and despite this, only 35% of outpatient clinics facilitate alternatives to face-to-face appointments. In this study, there were 5 times as many virtual appointments postoperatively as preoperatively. This likely represents a decreased need to examine patients postoperatively following histopathology results. While a good starting point, we propose a change in approach to outpatient clinics to make virtual consultations postoperatively the default, with clinical justification required for in-person consultations, to reduce emissions from patient travel and decrease pressure on waiting times through increased appointment availability.

Health care has the potential to become a model for sustainability through the implementation of practices to mitigate the dangers of the climate crisis while improving public health. To do so, studies must be performed to identify ways in which health care can become more sustainable. This research often requires contact with external departments not routinely encountered within clinical practice, such as estates, procurement, and even external suppliers. This can therefore lead to data limitations due to the inherent reliance on their carbon emission data. This hybrid study used a top-down

XSL•FO
RenderX

approach for estimates of carbon emissions from diagnostic investigations (imaging, nuclear medicine, and histopathology) due to a lack of the required data for bottom-up calculations. This was a key limitation of the study; hence, the main takeaways from this study should be centered on process mapping rather than a specific numerical value for the overall carbon footprint of the pathway. Furthermore, the heterogeneity of patient care meant that further breakdown of carbon contributions per pathway stage could not be provided. Future work should include bottom-up calculations for diagnostics using frameworks set out by a recent life cycle assessment by Thiel et al [26] to further dissect the carbon hot spots of the skin cancer patient pathway and other surgical pathways, alongside the design of carbon calculators whereby patients and clinicians can calculate the carbon contribution of their care. In particular, data were sparse on life cycle assessments for single-photon emission CT–CT and this should be an area of focus.

This study process mapped and performed a life cycle assessment for the skin cancer patient pathway and estimated the carbon footprint to be 118.5 kgCO2eq. Process mapping highlighted multiple inefficiencies within the pathway and key carbon hot spots. Emissions came not just from the operative period but also from the diagnostic and follow-up periods. There were many contributing factors to the carbon footprint, the greatest of which were imaging and diagnostic investigations. Proposed strategies to increase sustainability focused on pathway modification, increased efficiency of MDT meetings, justification of imaging, adoption of circular economy principles in waste management, and use of virtual consultations. Future work should focus on performing eco audits on other cancer pathways and closed-loop audits of these interventions to assess the success of these strategies in decreasing carbon emissions, improving efficiency, and increasing health care sustainability.

Funding

The authors declared no financial support was received for this work.

Data Availability

Data will be made available upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average staff variables used to calculate carbon dioxide equivalent for 64 patients undergoing skin cancer excision with or without sentinel lymph node biopsy.
[DOCX File, 26 KB - derma_v9i1e90087_app1.docx]

Multimedia Appendix 2
An illustration of (A) the location of skin cancer excision and (B) locations of sentinel lymph node biopsy sites.
[PNG File, 521 KB - derma_v9i1e90087_app2.png]

Multimedia Appendix 3
A flowchart showing the complex interplay of factors contributing to the overall carbon footprint of the skin cancer patient pathway.
[PNG File, 242 KB - derma_v9i1e90087_app3.png]

References

1. Climate change. United Nations. URL: https://www.un.org/en/global-issues/climate-change [accessed 2026-08-09]
2. What is climate change? United Nations. URL: https://www.un.org/en/climatechange/what-is-climate-change [accessed 2026-08-09]
3. The Paris Agreement. United Nations Climate Change. URL: https://unfccc.int/process-and-meetings/the-paris-agreement [accessed 2026-08-09]
4. Secretary-General's special address on climate action "a moment of truth" [as delivered]. United Nations. URL: https://www.un.org/sg/en/content/sg/statements/2024-06-05/secretary-generals-special-address-climate-action-moment-of-truth-delivered [accessed 2026-08-10]
5. Smith L. The nexus between climate change and healthcare. The Health Policy Partnership. 2022. URL: https://www.healthpolicypartnership.com/the-nexus-between-climate-change-and-healthcare/#:~:text=The%20healthcare%20sector%20is%20responsible,fifth%20largest%20polluter%20on%20Earth [accessed 2026-08-10]
6. More support needed to help the NHS reach net zero. British Medical Association. 2024. URL: https://www.bma.org.uk/what-we-do/population-health/protecting-people-from-threats-to-health/

XSL•FO
RenderX

more-support-needed-to-help-the-nhs-reach-net-zero#:~:text=The%20health%20service%20has%20an,of%20the%20public%20sector's%20emissions [accessed 2026-08-10]
7. University Hospitals Birmingham: a world first in carbon net zero surgery. NHS England. URL: https://www.england.nhs.uk/greenernhs/whats-already-happening/university-hospitals-birmingham-a-world-first-in-carbon-net-zero-surgery/#:~:text=Operating%20theatres%20are%20an%20important,of%20hospital%20inpatients%20undergoing%20surgery [accessed 2026-08-10]
8. New report sets the groundwork for reducing the carbon footprint of surgical care. UK Health Alliance on Climate Change. 2023. URL: https://ukhealthalliance.org/news-item/new-report-sets-the-groundwork-for-reducing-the-carbon-footprint-of-surgical-care/ [accessed 2026-08-10]
9. How to conduct a sustainability audit and why it's important. EcoCart. 2023. URL: https://ecocart.io/sustainability-audit/ [accessed 2026-08-10]
10. Life cycle assessment. European Environment Agency. URL: https://www.eea.europa.eu/en/analysis/publications/gh-07-97-595-en-c [accessed 2026-08-10]
11. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. BMJ 2010 Aug 13;341:c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]
12. Glynou SP, Ahmed Z, Zargaran A, Zargaran D, Mosahebi A. Innovating in sustainability research: grassroot approaches to carbon footprinting. Bulletin 2024 Sep;106(S1):34-36. [doi: 10.1308/rcsbull.TB2024.8]
13. Ahmed Z, Zargaran A, Zargaran D, et al. Sustainability in reconstructive breast surgery: an eco-audit of the deep inferior epigastric perforator flap pathway. Plast Reconstr Surg Glob Open 2024 Dec;12(12):e6374. [doi: 10.1097/GOX.0000000000006374] [Medline: 39726822]
14. Skin cancer. International Agency for Research on Cancer. URL: https://www.iarc.who.int/cancer-type/skin-cancer/ [accessed 2026-08-10]
15. Treatment for non melanoma skin cancer. Cancer Research UK. URL: https://www.cancerresearchuk.org/about-cancer/skin-cancer/treatment [accessed 2026-08-10]
16. Treatment breakdown by demographics, 2013-2021. National Disease Registration Service. URL: https://nhsd-ndrs.shinyapps.io/incidence_and_mortality/ [accessed 2026-08-10]
17. Distance London, England, GBR → Bilbao, Vizcaya, País-Vasco, ESP. Distance.to. URL: https://www.distance.to/London,England,GBR/Bilbao,Vizcaya,Pa%C3%ADs-Vasco,ESP [accessed 2026-08-10]
18. Marriott RD. Process mapping - the foundation for effective quality improvement. Curr Probl Pediatr Adolesc Health Care 2018 Jul;48(7):177-181. [doi: 10.1016/j.cppeds.2018.08.010] [Medline: 30205944]
19. Meeting patients' needs: improving the effectiveness of multidisciplinary team meetings in cancer services. : Cancer Research UK; 2020 URL: https://assets.ctfassets.net/u7vsjnoopqo5/2DzGDhcOMXDDrK3sOw7s0X/f8d6f390cb9980aead1cef14be9fa0f9/executive_summary_meeting_patients_needs_improving_the_effectiveness_of_multidisciplinary_team_meetings.pdf [accessed 2026-08-10]
20. Ali SR, Dobbs TD, Jovic M, Hutchings HA, Whitaker IS. Improving the effectiveness of multidisciplinary team meetings on skin cancer: analysis of the National Cancer Research UK survey responses. J Plast Reconstr Aesthet Surg 2023 Jul;82:141-151. [doi: 10.1016/j.bjps.2023.01.002] [Medline: 37167715]
21. Improving outcomes for people with skin tumours including melanoma. National Institute for Health and Care Excellence. 2006. URL: https://www.nice.org.uk/guidance/csg8 [accessed 2026-08-10]
22. Grocott MP, Plumb JO, Edwards M, Fecher-Jones I, Levett DZ. Re-designing the pathway to surgery: better care and added value. Perioper Med (Lond) 2017;6:9. [doi: 10.1186/s13741-017-0065-4] [Medline: 28649376]
23. Artificial intelligence (AI) technologies for assessing and triaging skin lesions referred to the urgent suspected skin cancer pathway: early value assessment. National Institute for Health and Care Excellence. 2025. URL: https://www.nice.org.uk/guidance/hte24 [accessed 2026-08-10]
24. Picano E, Mangia C, D'Andrea A. Climate change, carbon dioxide emissions, and medical imaging contribution. J Clin Med 2022 Dec 27;12(1):215. [doi: 10.3390/jcm12010215] [Medline: 36615016]
25. Hanneman K, McKee H, Nguyen ET, Panet H, Kielar A. Greenhouse gas emissions by diagnostic imaging modality in a hospital-based radiology department. Can Assoc Radiol J 2024 Nov;75(4):950-953. [doi: 10.1177/08465371241253314] [Medline: 38742437]
26. Thiel CL, Vigil-Garcia M, Nande S, et al. Environmental life cycle assessment of a U.S. hospital-based radiology practice. Radiology 2024 Nov;313(2):e240398. [doi: 10.1148/radiol.240398] [Medline: 39589247]
27. Martin M, Mohnke A, Lewis GM, Dunnick NR, Keoleian G, Maturen KE. Environmental impacts of abdominal imaging: a pilot investigation. J Am Coll Radiol 2018 Oct;15(10):1385-1393. [doi: 10.1016/j.jacr.2018.07.015] [Medline: 30158086]
28. Brown M, Schoen JH, Gross J, Omary RA, Hanneman K. Climate change and radiology: impetus for change and a toolkit for action. Radiology 2023 May;307(4):e230229. [doi: 10.1148/radiol.230229] [Medline: 37070994]
29. Currie GM, Hawk KE, Rohren EM. Challenges confronting sustainability in nuclear medicine practice. Radiography (Lond) 2024 Jun;30 Suppl 1:1-8. [doi: 10.1016/j.radi.2024.04.026] [Medline: 38797115]

XSL•FO
RenderX

30. Ang KL, Jovic M, Malin I, Ali SR, Whitaker S, Whitaker IS. Carbon footprint of non-melanoma skin cancer surgery. BJS Open 2024 Sep 3;8(5):zrae084. [doi: 10.1093/bjsopen/zrae084] [Medline: 39417641]
31. McGain F, Jarosz KM, Nguyen MN, Bates S, O'Shea CJ. Auditing operating room recycling: a management case report. A A Case Rep 2015 Aug 1;5(3):47-50. [doi: 10.1213/XAA.0000000000000097] [Medline: 26230308]
32. Green surgery: reducing the environmental impact of surgical care. : UK Health Alliance on Climate Change; 2023 URL: https://www.atachcommunity.com/resources/resource-repository/green-surgery-reducing-the-environmental-impact-of-surgical-care/ [accessed 2026-08-10]
33. Babu MA, Dalenberg AK, Goodsell G, Holloway AB, Belau MM, Link MJ. Greening the operating room: results of a scalable initiative to reduce waste and recover supply costs. Neurosurgery 2019 Sep 1;85(3):432-437. [doi: 10.1093/neuros/nyy275] [Medline: 30060055]
34. Online version of the NHS Long Term Plan. National Health Service. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20250506042228/https://www.longtermplan.nhs.uk/online-version/ [accessed 2026-08-10]
35. Outpatients: the future – adding value through sustainability. British Dermatological Nursing Group. URL: https://bdng.org.uk/outpatients/ [accessed 2026-08-10]

Abbreviations

CO2eq: carbon dioxide equivalent
CT: computed tomography
MDT: multidisciplinary team
MRI: magnetic resonance imaging
NHS: National Health Service
NICE: National Institute for Health and Care Excellence
RCP: Royal College of Physicians
SLNB: sentinel lymph node biopsy

Edited by R Dellavalle; submitted 22.Dec.2025; peer-reviewed by E Karampinis, NY Vidal; revised version received 15.Jul.2026; accepted 03.Aug.2026; published 19.Aug.2026.

Please cite as:
Ahmed Z, Zargaran A, Moeini S, Sousi S, Zargaran D, Quigley AM, Kang N, Mosahebi A
Efficiency and Sustainability in the Skin Cancer Excision Pathway: Life Cycle Assessment and Process Map
JMIR Dermatol 2026;9:e90087
URL: https://derma.jmir.org/2026/1/e90087
doi:10.2196/90087

© Zahra Ahmed, Alexander Zargaran, Sara Moeini, Sara Sousi, David Zargaran, Ann-Marie Quigley, Norbert Kang, Afshin Mosahebi. Originally published in JMIR Dermatology (http://derma.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Quality and Readability of ChatGPT and Gemini Responses to Patient Questions About Systemic Isotretinoin: Comparative Evaluation Study

Elif Bal Avcı[1], MD; Buğra Burç Dağtaş[2], MD; Dilara İlhan Erdil[3], MD; Ozan Erdem[4], MD; Andaç Salman[5], Prof Dr; Selçuk Korkmaz[6], Prof Dr; Sezgi Sarıkaya Solak[1], Prof Dr

[1]Department of Dermatology, Faculty of Medicine, Trakya University, Trakya University Balkan Campus, Iskender, Edirne, Türkiye

[2]Department of Dermatology, Istanbul Training and Research Hospital, Istanbul, Türkiye

[3]Department of Dermatology, Etlik City Hospital, Ankara, Türkiye

[4]Department of Dermatology, Medeniyet University Faculty of Medicine, Göztepe Prof. Dr. Suleyman Yalcin City Hospital, Istanbul, Türkiye

[5]Department of Dermatology, School of Medicine, Acıbadem Mehmet Ali Aydınlar University, Istanbul, Türkiye

[6]Department of Biostatistics and Medical Informatics, Trakya University Faculty of Medicine, Edirne, Türkiye

Corresponding Author:
Elif Bal Avcı, MD
Department of Dermatology, Faculty of Medicine, Trakya University, Trakya University Balkan Campus, Iskender, Edirne, Türkiye

Abstract

Background: The use of AI in health care is rapidly expanding due to its accessibility, with growing reliance among both physicians and patients.

Objective: This study evaluated ChatGPT (OpenAI) and Gemini (Google) responses to patient questions on systemic isotretinoin.

Methods: A total of 40 commonly anticipated patient questions were developed based on online sources and patient education materials and categorized into 5 groups. Responses were generated using the January 2025 free versions of ChatGPT (GPT-4o) and Gemini 2.0 Flash. A total of 5 experienced dermatologists evaluated the responses in a blinded manner using a 5-point Likert scale assessing accuracy, clarity, completeness, usefulness, and overall quality score (Global Quality Score [GQS]). Readability was measured using Flesch-Kincaid Grade Level (FKGL), Flesch Reading Ease, and Simple Measure of Gobbledygook (SMOG) index scores. Mixed-effects modeling was performed to adjust for reviewer and question variability, and intraclass correlation coefficients (ICCs) were calculated to assess interrater reliability.

Results: ChatGPT achieved significantly higher ratings in accuracy, completeness, usefulness, and GQS, while clarity showed a nonsignificant trend. When questions were analyzed by subtype, no statistically significant differences were observed between models across evaluation parameters. Mixed-effects modeling confirmed higher odds of superior ratings for ChatGPT across domains (odds ratios [ORs] 1.82 - 5.88; all $P<.001$). However, interrater reliability was lower for ChatGPT (global ICC 0.401) compared with Gemini (global ICC 0.677). Readability analyses demonstrated that ChatGPT required a higher reading level (FKGL 12.6 vs 11.2; SMOG 11.3 vs 10.0).

Conclusions: ChatGPT outperformed Gemini in terms of data accuracy and quality, while Gemini offered superior readability and consistency. Both models exceeded recommended readability thresholds, highlighting a gap between informational quality and linguistic accessibility in AI-generated health communication.

(JMIR Dermatol 2026;9:e99780) doi:10.2196/99780

KEYWORDS

artificial intelligence; AI; large language models; patient education; readability; systemic isotretinoin

Introduction

Systemic isotretinoin is among the most effective treatments for moderate-to-severe acne [1]. However, its use involves significant adverse effects, strict monitoring, and complex counseling. As a result, patients frequently seek detailed information on efficacy, safety, laboratory monitoring, teratogenic risks, and long-term outcomes [2]. As health information becomes increasingly digitalized, patients are increasingly using online sources to supplement information provided by physicians [3]. AI fulfills diverse roles in dermatology, such as facilitating diagnosis, monitoring disease progression, guiding treatment selection, providing patient information, and supporting medical education [4,5]. Numerous

XSL•FO
RenderX

studies across various medical specialties have used AI language models to evaluate responses to patient questions. However, standardized guidelines for performance assessment are lacking, and the methodological details of these studies often remain unclear [6]. Several studies have assessed the performance of large language models (LLMs) in responding to patient questions about acne and systemic isotretinoin, with mixed results for accuracy and quality [7-11]. Due to isotretinoin's teratogenicity, monitoring needs, and medicolegal considerations, it is essential that the information provided is both accurate and easily understandable.

The study aimed to comparatively evaluate ChatGPT and Gemini in their responses to patient questions regarding systemic isotretinoin, assessing performance across multiple quality domains, interrater reliability, and readability metrics.

Methods

Study Design

This cross-sectional comparative study assessed the quality of responses generated by 2 LLMs to patient-oriented questions concerning systemic isotretinoin. The evaluation criteria included accuracy, clarity, completeness, usefulness, Global Quality Score (GQS), and readability.

Question Development

Forty patient-oriented questions were developed based on information from patient information brochures and online patient forum queries related to isotretinoin. To prevent confusion between topical and systemic formulations, the standardized term "oral isotretinoin" was included in each question. The questions were categorized into 5 domains: general information, efficacy and treatment process, side effects, follow-up and precautions, and cosmetic concerns. Six experienced dermatologists (EBA, BBD, OE, DİE, AS, and SSS) reviewed and validated all questions to ensure content relevance and representativeness prior to inclusion. The complete list of the 40 isotretinoin-related questions evaluated in this study is provided in Multimedia Appendix 1.

Evaluation of AI Models and Query Procedures

Free access versions of Gemini 2.0 Flash (Google Inc) and GPT-4o (OpenAI) were evaluated through their standard web-based interfaces to reflect real-world patient use. All queries were conducted in January 2025. The models were not accessed via APIs; therefore, API-level parameters, including temperature, were not accessible or configurable by the investigators. Each question was entered verbatim, without additional prompts or role instructions. Queries were submitted without logging into a user account, using incognito browsing mode, and each question was asked in a separate browser session to minimize the influence of prior conversational context. Each question was submitted only once to each model, and the first generated response was used for evaluation, with no regeneration or repeated querying, to simulate a typical real-world patient interaction. Responses were recorded verbatim without modification.

Blinded Expert Assessment (Accuracy, Clarity, Completeness, Usefulness, and the GQS)

The questions were entered into the 2 chatbots by 1 dermatologist (EBA) to generate the responses, and the resulting outputs were independently evaluated by 5 experienced dermatologists (≥5 years of experience; BBD, OE, DİE, AS, and SSS) in a blinded fashion. Interrater agreement was assessed using the intraclass correlation coefficient (ICC). Responses were scored on a 5-point Likert scale (1=very poor, 5=excellent) across 5 domains: accuracy, clarity, completeness, usefulness, and the GQS. A standardized scoring rubric was provided to evaluators prior to assessment. Evaluations were performed by 5 experienced dermatologists, informed by contemporary acne treatment guidelines and clinical practice standards. Readability was assessed using Readability Formulas [12]. The following indices were calculated for each response: Flesch Reading Ease (FRE), Flesch-Kincaid Grade Level (FKGL), and Simple Measure of Gobbledygook (SMOG) index. Source citation was not included as a comparative parameter because the reference provision differed structurally between models: Gemini frequently generated citations, while ChatGPT did not consistently provide explicit references.

Statistical Analysis

Data were analyzed using R (version 4.3.2; R Foundation for Statistical Computing). Descriptive statistics are presented as mean, median, and range (minimum-maximum), as appropriate. Ratings assigned to ChatGPT and Gemini responses to the same questions by the same reviewer panel were compared using the Wilcoxon signed-rank test for paired samples. These comparisons were performed for the pooled dataset and separately within each of 5 predefined clinical domains (general information, efficacy and treatment process, side effects, follow-up and precautions, and cosmetic concerns), resulting in 30 pairwise comparisons (5 metrics across 6 analyses). Within each analysis, Bonferroni correction was applied across the 5 metric-level comparisons (adjusted P=min[raw $P \times 5$, 1.0]) to control the family-wise error rate. A Bonferroni-adjusted $P<.05$ was considered statistically significant. To account for the hierarchical structure of the ratings and repeated assessments by evaluators and across questions, model performance for accuracy, clarity, completeness, usefulness, and GQS was additionally evaluated using ordinal mixed-effects regression models, with evaluator and question included as random intercepts. Effect estimates are reported as odds ratios (ORs) with 95% CIs. Interrater reliability was assessed using 2-way mixed-effects ICCs for absolute agreement. All tests were 2-sided, with a significance level set at $P<.05$.

Ethical Considerations

This study did not involve human participants, animal subjects, or patient data. The analyzed data consisted solely of responses generated by publicly available AI chatbots to questions prepared by the authors. Therefore, ethics committee approval was not required.

XSL•FO
RenderX

Results

Expert Assessment of AI Responses

ChatGPT demonstrated consistent numerical outperformance across all assessed domains (Table 1).

In the domain-specific analysis, no statistically significant differences between ChatGPT and Gemini were detected across the 5 predefined domains (general information, efficacy and treatment process, side effects, follow-up and precautions, and cosmetic concerns) for any of the assessed metrics after Bonferroni adjustment (all adjusted P>.05). Although ChatGPT showed numerically higher mean scores across all metrics in the cosmetic concerns domain, including accuracy (4.60 vs 3.80) and completeness (4.60 vs 3.60), these differences were not statistically significant after adjustment. The complete results of the domain-specific comparisons are presented in Table 2.

Table . Performance comparison of ChatGPT and Gemini across key quality metrics.

Metric	ChatGPT, mean (SD; range)	ChatGPT, median (IQR)	Gemini, mean (SD; range)	Gemini, median (IQR)	*P* value	Adjusted *P* value
Accuracy	4.37 (0.71; 3-5)	5 (4-5)	4.20 (0.65; 3-5)	4 (4-5)	.004	.02
Clarity	4.42 (0.72; 3-5)	5 (4-5)	4.27 (0.62; 3-5)	4 (4-5)	.01	.07
Completeness	4.39 (0.75; 3-5)	5 (4-5)	3.77 (0.75; 1-5)	4 (3-4)	<.001	<.001
Usefulness	4.42 (0.70; 3-5)	5 (4-5)	4.04 (0.69; 2-5)	4 (4-4)	<.001	<.001
GQS[a]	4.42 (0.71; 3-5)	5 (4-5)	4.05 (0.70; 2-5)	4 (4-5)	<.001	<.001

[a]GQS: Global Quality Score.

XSL•FO
RenderX

Table . Domain-specific comparison of performance metrics between ChatGPT and Gemini across 5 predefined domains.

Domain and metric	ChatGPT, mean (SD; range)	ChatGPT, median (IQR)	Gemini, mean (SD; range)	Gemini, median (IQR)	*P* value	Adjusted *P* value
General information						
Accuracy	4.00 (0.00; 4-4)	4 (4-4)	4.00 (0.00; 4-4)	4 (4-4)	—[a]	—
Clarity	4.00 (0.00; 4-4)	4 (4-4)	4.33 (0.58; 4-5)	4 (4-4.5)	>.99	>.99
Completeness	3.67 (0.58; 3-4)	4 (3.5-4)	3.33 (0.58; 3-4)	3 (3-3.5)	>.99	>.99
Usefulness	4.00 (0.00; 4-4)	4 (4-4)	4.00 (0.00; 4-4)	4 (4-4)	—	—
GQS[b]	4.00 (0.00; 4-4)	4 (4-4)	4.00 (0.00; 4-4)	4 (4-4)	—	—
Efficacy and treatment process						
Accuracy	3.89 (0.33; 3-4)	4 (4-4)	3.56 (0.53; 3-4)	4 (3-4)	.15	.75
Clarity	4.00 (0.71; 3-5)	4 (4-4)	4.00 (0.50; 3-5)	4 (4-4)	>.99	>.99
Completeness	3.78 (0.44; 3-4)	4 (4-4)	3.44 (1.01; 1-4)	4 (3-4)	.37	>.99
Usefulness	3.89 (0.33; 3-4)	4 (4-4)	3.67 (0.71; 2-4)	4 (4-4)	.35	>.99
GQS	3.78 (0.44; 3-4)	4 (4-4)	3.56 (0.73; 2-4)	4 (3-4)	.35	>.99
Side effects						
Accuracy	3.82 (0.40; 3-4)	4 (4-4)	3.91 (0.30; 3-4)	4 (4-4)	.77	>.99
Clarity	3.91 (0.30; 3-4)	4 (4-4)	4.09 (0.70; 3-5)	4 (4-4.5)	.48	>.99
Completeness	3.91 (0.54; 3-5)	4 (4-4)	3.55 (0.52; 3-4)	4 (3-4)	.20	>.99
Usefulness	3.91 (0.30; 3-4)	4 (4-4)	3.82 (0.40; 3-4)	4 (4-4)	.77	>.99
GQS	4.00 (0.00; 4-4)	4 (4-4)	3.91 (0.54; 3-5)	4 (3-4)	.77	>.99
Follow-up and precautions						
Accuracy	4.00 (0.43; 3-5)	4 (4-4)	4.00 (0.00; 4-4)	4 (4-4)	>.99	>.99
Clarity	4.50 (0.52; 4-5)	4.5 (4-5)	4.50 (0.52; 4-5)	4.5 (4-5)	>.99	>.99
Completeness	4.25 (0.45; 4-5)	4 (4-4.25)	3.92 (0.29; 3-4)	4 (4-4)	.07	.36
Usefulness	4.25 (0.45; 4-5)	4 (4-4.25)	4.08 (0.29; 4-5)	4 (4-4)	.42	>.99
GQS	4.50 (0.52; 4-5)	4.5 (4-5)	4.33 (0.49; 4-5)	4 (4-5)	.53	>.99
Cosmetic concerns						
Accuracy	4.60 (0.55; 4-5)	5 (4-5)	3.80 (0.45; 3-4)	4 (4-4)	.07	.36
Clarity	4.40 (0.55; 4-5)	4 (4-5)	3.80 (0.45; 3-4)	4 (4-4)	.15	.75
Completeness	4.60 (0.55; 4-5)	5 (4-5)	3.60 (0.89; 2-4)	4 (4-4)	.09	.45
Usefulness	4.60 (0.55; 4-5)	5 (4-5)	3.80 (0.45; 3-4)	4 (4-4)	.07	.36
GQS	4.60 (0.55; 4-5)	5 (4-5)	3.80 (0.45; 3-4)	4 (4-4)	.07	.36

[a]*P* value was not calculated because there was no variation in the corresponding scores.

[b]GQS: Global Quality Score.

Ordinal mixed-effects models revealed that ChatGPT was significantly more likely than Gemini to receive higher ratings across all metrics (all $P<.001$). ChatGPT outperformed Gemini in accuracy (OR 1.94, 95% CI 1.28 - 2.94), clarity (OR 1.82, 95% CI 1.19 - 2.76), and GQS (OR 3.55, 95% CI 2.32 - 5.43). The most substantial advantages were observed in completeness (OR 5.88, 95% CI 3.84 - 9.02) and usefulness (OR 3.74, 95% CI 2.44 - 5.74). Representative examples of responses generated by ChatGPT and Gemini are provided in Multimedia Appendix 2.

Interrater Reliability

Global ICC estimates indicated “moderate” consistency for Gemini (ICC 0.677) and “poor” consistency for ChatGPT (ICC 0.401; Table 3). Although both models exceeded chance agreement ($P<.05$), Gemini demonstrated significantly more stable evaluator consensus.

XSL•FO
RenderX

Table . Comparison of interrater reliability (intraclass correlation coefficient [ICC]) between ChatGPT and Gemini across key evaluation domains.

Metric	ChatGPT, ICC (95% CI)	Gemini, ICC (95% CI)	Interpretation (ChatGPT)	Interpretation (Gemini)
Accuracy	0.435 (0.101 - 0.671)	0.554 (0.291 - 0.740)	Poor	Moderate
Clarity	0.389 (0.028 - 0.644)	0.629 (0.410 - 0.784)	Poor	Moderate
Completeness	0.379 (0.012 - 0.638)	0.527 (0.248 - 0.725)	Poor	Moderate
Usefulness	0.393 (0.035 - 0.647)	0.593 (0.353 - 0.763)	Poor	Moderate
GQS[a]	0.462 (0.145 - 0.687)	0.636 (0.421 - 0.788)	Poor	Moderate
Global ICC	0.401 (0.048 - 0.651)	0.677 (0.486 - 0.812)	Poor	Moderate

[a]GQS: Global Quality Score

Readability

Gemini demonstrated superior accessibility across multiple readability metrics (Table 4). Significant reductions in grade-level requirements (mean 11.2 vs 12.6; adjusted *P*=.002) and SMOG index scores (adjusted *P*=.001) were observed for Gemini compared with ChatGPT. Although Gemini also achieved higher raw FRE scores, this difference did not reach statistical significance after Bonferroni adjustment (*P*=.09).

Table . Comparison of readability indices between ChatGPT and Gemini.

Index	ChatGPT, mean (SD; range)	ChatGPT, median (IQR)	Gemini, mean (SD; range)	Gemini, median (IQR)	*P* value	Adjusted *P* value
FRE[a]	37.9 (10.26; 17-72)	38.00 (31.00-42.25)	42.03 (12.62; 20-65)	42.5 (32.00-51.75)	.03	.09
FKGL[b]	12.62 (1.95; 7.7-16.59)	12.44 (11.48-13.51)	11.16 (2.13; 7.15-15.14)	10.805 (32.00-51.75)	<.001	.002
SMOG[c]	11.31 (1.65; 7.78-15.37)	11.31 (10.40-12.10)	10.03 (1.63; 7.24-13.62)	9.865 (8.79-11.12)	<.001	.001

[a]FRE: Flesch Reading Ease.

[b]FKGL: Flesch-Kincaid Grade Level.

[c]SMOG: Simple Measure of Gobbledygook.

Discussion

In this study, ChatGPT demonstrated significantly higher performance across evaluation metrics, including accuracy, completeness, clarity, and overall quality, while Gemini showed greater interrater reliability and superior readability. Accordingly, the choice of LLM may depend on whether the priority is high-quality medical content or improved patient comprehension.

Recent evaluations indicate that LLMs generate more contextually appropriate medical responses than traditional internet-based resources [13]. ChatGPT has demonstrated potential as a public dermatology resource, particularly for common dermatological diseases, with 88% of its responses rated as appropriate [14]. However, the clinical utility of LLMs remains constrained by their propensity to generate fictitious citations [15,16]. Their inability to account for individual patient factors and age-related nuances may also lead to misleading guidance, potentially resulting in adverse outcomes based on unverified data [11].

The literature further demonstrates that LLM accuracy and comprehensiveness ratings vary considerably across dermatologic diseases and topics [11]. For example, ChatGPT's accuracy has ranged from 36% to 90% across various topics [6,14,17,18]. ChatGPT showed higher accuracy in dermatological topics with well-established evidence, such as melanoma and nonmelanoma skin cancers, rosacea, and alopecia. In contrast, lower accuracy was observed in areas requiring individualized approaches or where scientific consensus remains limited, including the effects of diet on skin diseases, nonpharmacological treatments for eczema and psoriasis, antiaging interventions, skincare routines, and wound healing [14]. In a study examining isotretinoin-related queries, ChatGPT's accuracy varied by domain: daily life achieved 72% reliability, while the treatment process yielded 64% with outdated or incorrect data [7]. In contrast, our findings showed no significant differences across query categories.

Our findings contradict previous research indicating that Gemini provided more accurate answers than ChatGPT on isotretinoin questions [8], as well as studies asserting Gemini's superiority in comprehensiveness in melanoma-related queries [18]. Recent evidence regarding Mohs surgery indicates that GPT-4 outperforms Gemini (Google Bard) in clinical relevance and accuracy, providing answers that closely align with official American College of Mohs Surgery recommendations [19], suggesting that differences in topic may significantly influence the performance of LLMs. Furthermore, such discrepancies may result from the absence of standardized evaluation criteria and differences among LLM models and versions.

Consistent with prior research on isotretinoin, which reported a low Fleiss κ (κ=0.045) [7], our study also demonstrated

XSL•FO
RenderX

similarly weak interrater agreement. Consistent with our findings, Gemini demonstrated higher interrater agreement than ChatGPT in a study of common dermatological conditions [20]. This low interrater agreement may be partially attributed to heterogeneity in physicians' clinical communication styles and expectations in patient education. To address inconsistent evaluations, it has been proposed that LLM outputs be assessed using standardized, guideline-based frameworks [21]. Nevertheless, the complex dynamics of physician-patient interactions and the subjective nature of clinical judgment continue to influence evaluation outcomes.

Previous studies suggest that AI-generated health information often exceeds recommended readability levels, potentially limiting accessibility for individuals with low health literacy [21-23]. Recent studies have shown that successive iterations of LLMs exhibit progressive improvement in generating more readable and accessible content [24,25]. In our study, both ChatGPT and Gemini produced patient information at readability levels exceeding recommended thresholds for public education materials. Consistent with our results, some studies report higher readability for Gemini [18], whereas others find no significant readability differences between Gemini and ChatGPT [8]. High readability demands, incomplete incorporation of updated approaches, and limited actionable guidance preclude the use of LLM outputs as standalone patient education tools [7,8]. For high-risk medications such as isotretinoin, patients should be explicitly informed that LLM-generated outputs constitute preliminary information and do not substitute for professional medical guidance. Given the substantial patient concerns about isotretinoin [26-28], clear and accessible educational materials may mitigate misconceptions, reduce treatment resistance, and improve adherence.

Additionally, ChatGPT can be directed to create patient brochures at predefined reading levels, thereby improving accessibility for diverse patient populations [18,29]. Effective prompt engineering enables LLMs to deliver clearer, more targeted responses [8,30-32]. Incorporating an audience-adaptive response model in which users identify as health care professionals or laypersons prior to output generation could facilitate flexible adjustment of terminology and grade-level complexity.

Gemini consistently advised consultation with a health care professional, whereas this was not observed in our study with ChatGPT, despite prior reports of such disclaimers [33,34]. This suggests that Gemini may offer a comparatively safer profile for patient safety.

The strengths of this study include methodological transparency, a comprehensive multidimensional evaluation approach, a comparative design, and the provision of actionable recommendations. This study has several limitations. The selected questions may not accurately represent real-world patient inquiries. The study was conducted exclusively in English, and language-dependent variability in AI model training data and source retrieval may limit generalizability across different languages. Moreover, despite blinded assessment, clinicians familiar with AI language models may have inferred the model's identity, potentially introducing observer bias. Another limitation is that our analysis included only ChatGPT and Gemini. Therefore, the findings may not be generalizable to other commercially available or open-source LLMs. Given the rapid evolution of LLMs, the findings of this study reflect the performance of GPT-4o and Gemini 2.0 Flash at the time of evaluation. As newer model iterations are released with enhanced reasoning and language capabilities, ongoing comparative studies are required to determine whether these advancements yield measurable improvements in clinical quality, reliability, and readability. Furthermore, because LLMs are frequently updated, their responses and performance may change over time. Future studies should include a broader range of models and evaluate changes across model versions longitudinally.

In conclusion, there is a clear distinction between ChatGPT's technical accuracy and Gemini's linguistic accessibility. For these models to function as reliable clinical tools, transparent source attribution and patient-centered, intelligible outputs are imperative. Adopting standardized, objective evaluation frameworks is also crucial for enabling robust comparisons of outputs and tracking model performance over time, thereby mitigating interrater variability. However, the complexity and heterogeneity of patient characteristics make it unlikely that clinicians will be fully substituted for by AI, particularly when it comes to contextualizing clinical evidence for individual patients.

Acknowledgments

During the preparation of this manuscript, ChatGPT (OpenAI, GPT-5 series) was used solely for English language editing, including grammar, spelling, and clarity improvements. Different versions within the GPT-5 series may have been used at different stages of manuscript preparation, as automatic model selection was enabled. The authors reviewed and edited all AI-assisted revisions and take full responsibility for the final content of the manuscript.

Funding

This study received no external funding.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request and with appropriate justification.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of questions.
[DOCX File, 17 KB - derma_v9i1e99780_app1.docx]

Multimedia Appendix 2
Example questions and answers.
[DOCX File, 34 KB - derma_v9i1e99780_app2.docx]

References

1. Reynolds RV, Yeung H, Cheng CE, et al. Guidelines of care for the management of acne vulgaris. J Am Acad Dermatol 2024 May;90(5):1006.e1-1006.e30. [doi: 10.1016/j.jaad.2023.12.017] [Medline: 38300170]
2. Sadeghzadeh-Bazargan A, Ghassemi M, Goodarzi A, Roohaninasab M, Najar Nobari N, Behrangi E. Systematic review of low-dose isotretinoin for treatment of acne vulgaris: focus on indication, dosage, regimen, efficacy, safety, satisfaction, and follow up, based on clinical studies. Dermatol Ther 2021 Jan;34(1):e14438. [doi: 10.1111/dth.14438] [Medline: 33085149]
3. De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health 2023 Apr 25;11:1166120. [doi: 10.3389/fpubh.2023.1166120] [Medline: 37181697]
4. Sohail N, Puyana C, Zimmerman L, Tsoukas MM. Artificial intelligence in dermatology: bridging the gap in patient care and education. Clin Dermatol 2024 Sep;42(5):434-436. [doi: 10.1016/j.clindermatol.2024.06.009] [Medline: 38936639]
5. Gabashvili IS. ChatGPT in dermatology: a comprehensive systematic review. medRxiv. Preprint posted online on Jun 12, 2023. [doi: 10.2196/preprints.48979]
6. Wei Q, Yao Z, Cui Y, Wei B, Jin Z, Xu X. Evaluation of ChatGPT-generated medical responses: a systematic review and meta-analysis. J Biomed Inform 2024 Mar;151:104620. [doi: 10.1016/j.jbi.2024.104620] [Medline: 38462064]
7. Küçük R, Güngör CB, Topal İ, Küçük RB, Karadağ AS. Accuracy, readability, understandability, and actionability of ChatGPT responses to patient questions on systemic isotretinoin treatment. Indian Dermatol Online J 2026 Jan 1;17(1):27-32. [doi: 10.4103/idoj.idoj_1243_24] [Medline: 41437721]
8. Soysal M. Performance of artificial intelligence large language models (LLMs) in answering frequently asked questions about isotretinoin. Cutan Ocul Toxicol 2026;45(1):58-63. [doi: 10.1080/15569527.2025.2601639] [Medline: 41410102]
9. Frederickson KL, Gui H, Barbieri JS, Daneshjou R. Artificial intelligence use in acne diagnosis and management-a scoping review. Int J Dermatol 2026 Mar;65(3):437-443. [doi: 10.1111/ijd.70110] [Medline: 41199593]
10. Deoghare S. An interesting conversation with ChatGPT about acne vulgaris. Indian Dermatol Online J 2023 Nov 7;15(1):137-140. [doi: 10.4103/idoj.idoj_77_23] [Medline: 38283020]
11. Lakdawala N, Channa L, Gronbeck C, et al. Assessing the accuracy and comprehensiveness of ChatGPT in offering clinical guidance for atopic dermatitis and acne vulgaris. JMIR Dermatol 2023 Nov 14;6:e50409. [doi: 10.2196/50409] [Medline: 37962920]
12. Readability Formulas. URL: https://readabilityformulas.com/#google_vignette [accessed 2026-10-11]
13. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 1;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
14. Ferreira AL, Chu B, Grant-Kels JM, Ogunleye T, Lipoff JB. Evaluation of ChatGPT dermatology responses to common patient queries. JMIR Dermatol 2023 Nov 17;6:e49280. [doi: 10.2196/49280] [Medline: 37976093]
15. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
16. Thirunavukarasu AJ, Ting DS, Elangovan K, Gutierrez L, Tan TF, Ting DS. Large language models in medicine. Nat Med 2023 Aug;29(8):1930-1940. [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]
17. D'Agostino M, Feo F, Martora F, et al. ChatGPT and dermatology. Ital J Dermatol Venerol 2024 Oct;159(5):566-571. [doi: 10.23736/S2784-8671.24.07928-3] [Medline: 39039954]
18. Kamminga NC, Kievits JE, Plaisier PW, et al. Do large language model chatbots perform better than established patient information resources in answering patient questions? A comparative study on melanoma. Br J Dermatol 2025 Jan 24;192(2):306-315. [doi: 10.1093/bjd/ljae377] [Medline: 39365602]
19. Robinson MA, Belzberg M, Thakker S, et al. Assessing the accuracy, usefulness, and readability of artificial-intelligence-generated responses to common dermatologic surgery questions for patient education: a double-blinded comparative study of ChatGPT and Google Bard. J Am Acad Dermatol 2024 May;90(5):1078-1080. [doi: 10.1016/j.jaad.2024.01.037] [Medline: 38296195]

XSL•FO
RenderX

20. Podder I, Pipil N, Dhabal A, Mondal S, Pienyii V, Mondal H. Evaluation of artificial intelligence-based chatbot responses to common dermatological queries. Jordan Med J 2024;58(3):271-278. [doi: 10.35516/jmj.v58i2.2960]
21. Huo B, Boyle A, Marfo N, et al. Large language models for chatbot health advice studies: a systematic review. JAMA Netw Open 2025 Feb 3;8(2):e2457879. [doi: 10.1001/jamanetworkopen.2024.57879] [Medline: 39903463]
22. Chen AI, Ferris LK, Nambudiri VE, Piette EW. ChatRx: ChatGPT's potential to educate patients on medication adverse effects. J Am Acad Dermatol 2024 Mar;90(3):669-670. [doi: 10.1016/j.jaad.2023.11.008] [Medline: 37952568]
23. Andrikyan W, Sametinger SM, Kosfeld F, et al. Artificial intelligence-powered chatbots in search engines: a cross-sectional study on the quality and risks of drug information for patients. BMJ Qual Saf 2025 Jan 28;34(2):100-109. [doi: 10.1136/bmjqs-2024-017476] [Medline: 39353736]
24. Lambert R, Choo ZY, Gradwohl K, Schroedl L, Ruiz De Luzuriaga A. Assessing the application of large language models in generating dermatologic patient education materials according to reading level: qualitative study. JMIR Dermatol 2024 May 16;7:e55898. [doi: 10.2196/55898] [Medline: 38754096]
25. Rames MM, O'Hern K, Rames JD, et al. ChatGPT as a resource for patient education in cosmetic dermatological procedures: a boon or a bane? J Cosmet Dermatol 2024 Mar;23(3):1085-1086. [doi: 10.1111/jocd.16166] [Medline: 38168902]
26. Al-Harbi M. Concerns and awareness of acne patients about isotretinoin in Qassim region of Saudi Arabia. Int J Health Sci (Qassim) 2010 Jan;4(1):47-51. [Medline: 21475525]
27. Jakobi AU, Bircher AJ, Pagnamenta A, Terrani I. Isotretinoin concerns in Switzerland: a student-based transversal study. J Clin Med 2025 Mar 7;14(6):1801. [doi: 10.3390/jcm14061801] [Medline: 40142609]
28. Kara Polat A, Akin Belli A, Ergun EZ, et al. Knowledge levels and concerns about oral isotretinoin treatment in the parents of adolescent acne patients. Dermatol Ther 2020 Jul;33(4):e13669. [doi: 10.1111/dth.13669] [Medline: 32459383]
29. Breneman A, Trager MH, Gordon ER, Samie FH. Readability rescue: large language models may improve readability of patient education materials. Arch Dermatol Res 2024 Oct 10;316(9):669. [doi: 10.1007/s00403-024-03441-y] [Medline: 39387945]
30. Tian S, Jin Q, Yeganova L, et al. Opportunities and challenges for ChatGPT and large language models in biomedicine and health. Brief Bioinform 2023 Nov 22;25(1):bbad493. [doi: 10.1093/bib/bbad493]
31. Russe MF, Reisert M, Bamberg F, Rau A. Improving the use of LLMs in radiology through prompt engineering: from precision prompts to zero-shot learning. Rofo 2024 Nov;196(11):1166-1170. [doi: 10.1055/a-2264-5631] [Medline: 38408477]
32. Meskó B. Prompt engineering as an important emerging skill for medical professionals: tutorial. J Med Internet Res 2023 Oct 4;25:e50638. [doi: 10.2196/50638] [Medline: 37792434]
33. Dulai A, Tarmaster A, Sivamani R. Assessment of an artificial intelligence model's capability of responding to queries in the realm of integrative dermatology. JOID 2025 Jan;1(1). [doi: 10.64550/joid.0hy8wz51]
34. Reynolds K, Tejasvi T. Potential use of ChatGPT in responding to patient questions and creating patient resources. JMIR Dermatol 2024 Mar 6;7:e48451. [doi: 10.2196/48451] [Medline: 38446541]

Abbreviations

FKGL: Flesch-Kincaid Grade Level
FRE: Flesch Reading Ease
GQS: Global Quality Score
ICC: intraclass correlation coefficient
LLM: large language model
OR: odds ratio

Edited by J Meisenheimer; submitted 04.May.2026; peer-reviewed by A Snyder, M Chen; revised version received 09.Sep.2026; accepted 28.Sep.2026; published 30.Oct.2026.

Please cite as:
Bal Avcı E, Dağtaş BB, İlhan Erdil D, Erdem O, Salman A, Korkmaz S, Sarıkaya Solak S
Quality and Readability of ChatGPT and Gemini Responses to Patient Questions About Systemic Isotretinoin: Comparative Evaluation Study
JMIR Dermatol 2026;9:e99780
URL: https://derma.jmir.org/2026/1/e99780
doi:10.2196/99780

© Elif Bal Avcı, Buğra Burç Dağtaş, Dilara İlhan Erdil, Ozan Erdem, Andaç Salman, Selçuk Korkmaz, Sezgi Sarıkaya Solak. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Oct.2026. This is an open-access article distributed under

XSL•FO
RenderX

the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Lived Experience of Acne in the United States and the United Kingdom: Social Media Analysis

John S Barbieri[1], MBA, MD, FAAD; Valentine Delore[2], PharmD; Isabelle Vieillard[2], MS; Sandrine Samson[2], PhD; Esteban Puentes[2], MSc, MPH; Joelle Malaab[3], MS, MPH; Melanie Sambrès[3], MS; Manissa Talmatkadi[3], MS; Stéphane Schück[3], MPH, MD

[1]Department of Dermatology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[2]Global Medical, Vaccines, Sanofi, Lyon, France

[3]Kap Code, 4 rue de Cléry, Paris, France

Corresponding Author:
Joelle Malaab, MS, MPH
Kap Code, 4 rue de Cléry, Paris, France

Abstract

Background: Acne is a chronic skin condition that primarily affects adolescents and young adults but can persist into adulthood. It can have repercussions on physical and mental health, self-esteem, and body image. The increasing use of social media for health information and peer support offers an opportunity to explore real-life experiences with acne.

Objective: This study aims to analyze social media messages from users in the United States and the United Kingdom using artificial intelligence to assess the impact of acne on quality of life (QoL), identify discussion topics, and explore unmet needs.

Methods: The data were extracted from public platforms using a query containing the word “acne” between January 1 and December 31, 2024. Data cleaning and filtering were performed using natural language processing, machine learning methods, and algorithms. Biterm topic modeling was used to identify the main discussion topics, and QoL impact was assessed using a deep learning algorithm adapted from the EuroQol 5-Dimension Questionnaire or the 36-Item Short Form Health Survey. Unmet needs were identified through manual annotation using the saturation method.

Results: A total of 646,809 messages posted by 432,234 users were identified. The main topics included skincare routines and product recommendations (n=154,907, 23.9%), acne scars (n=135,643, 21%), and general treatment information (n=97,177, 15%). Engagement varied across topics and platforms. On Instagram, dietary and nutritional strategies (0.16%, SD 6.36%) showed the highest mean engagement, followed by skincare routines and product recommendations (0.11%, SD 4.81%). In general, engagement scores were higher in the United Kingdom compared to the United States across all topics. On TikTok, content about makeup and acne had the highest mean engagement score (3.03%, SD 92.65%). Overall, 52.9% (228,613/432,234) of the users expressed at least 1 QoL impact, most frequently related to signs and symptoms (175,604/228,613, 76.8%), social functioning (n=149,234, 65.3%), mental health (n=107,155, 46.9%), and cost (n=62,008, 27.1%). Of 3200 annotated messages, 582 contained unmet needs, including effective solutions for hormonal acne (111/582, 19.1%), clarity in identifying acne triggers (n=84, 14.4%), treatment guidance (n=68, 11.7%), and psychological support (n=68, 11.7%).

Conclusions: This study revealed the significant physical, psychological, social, and financial impact of acne on QoL and identified several unmet needs. Given the growing role of social media, these findings highlight opportunities for dermatologists and health professionals to educate and engage with the acne community through digital platforms.

(JMIR Dermatol 2026;9:e91126) doi:10.2196/91126

KEYWORDS

acne; dermatology; social media listening; natural language processing; machine learning; infodemiology

Introduction

Acne is a chronic skin condition that is estimated to affect 230 million people worldwide [1]. Although most prevalent among young adults and adolescents [2], up to 50% of women and 30% of men experience acne that begins or persists into adulthood [3-6].

Acne can have meaningful effects on individuals’ mental health, self-esteem, and body image [7,8], partly due to the stigmatization they may face at school, in the workplace, and in their personal lives [9]. The global burden of acne in terms of disability-adjusted life years surpasses that of psoriasis and melanoma [10].

XSL•FO
RenderX

Acne is primarily caused by four interrelated factors: excess sebum production, follicular hyperkeratinization, overgrowth of *Cutibacterium acnes* bacteria, and inflammation [11-14]. Treatment options include topical and systemic medications, procedures, diet [15], and lifestyle interventions [16].

In recent years, social media has been increasingly used for various purposes. As of 2025, there are 5.24 billion users, accounting for around 64% of the global population [17]. There is growing interest in social media for health-related purposes, including sharing experiences, seeking advice, accessing medical information, supporting health campaigns, and connecting with others through online communities and support groups [18-22]. Especially given challenges in accessing care, individuals with acne may turn to social media as a source of information and treatment advice [23,24]. At the same time, the growing emphasis on image and beauty trends can be concerning, particularly for adolescents and young adults, who may be more vulnerable to these influences [25].

Given the substantial amount of data that these platforms offer, social media provides a unique opportunity to explore users' experiences and gain insights into how acne affects their daily lives. While prior studies have evaluated the educational quality of content on social media and focused on limited platforms [26,27], little is known about how individuals with acne engage with social media more broadly, including the effects of acne on quality of life (QoL), topics of discussion, engagement with content, and unmet needs. This study aims to fill this gap by using artificial intelligence techniques, including natural language processing (NLP) and machine learning (ML) to analyze social media messages from users in the United States and the United Kingdom and better understand the effects of acne on users' QoL, identify key discussion topics, and uncover their unmet needs.

Methods

Study Design and Data Sources

This observational, retrospective, real-world study analyzed English-language social media posts in the United States and the United Kingdom between January 1 and December 31, 2024. The data were extracted using the data broker Brandwatch (Cision Ltd.) [28] for all sources except X (formerly Twitter), Instagram, and TikTok, which were extracted via the data broker Visibrain [29]. To perform the extraction, Brandwatch scanned publicly available sources across the internet and retrieved messages that contained the word "acne." In this approach, platforms and websites are not preselected; rather, the data broker automatically identifies relevant content from publicly accessible sources. These sources include public websites such as online forums, blogs, and health-related discussion platforms.

Data Processing

Messages were collected using a query consisting of the word "acne" along with their metadata (eg, author and publication date) and compiled into Microsoft Excel. This keyword-based approach was selected to maximize specificity and retrieve messages explicitly discussing acne-related experiences. However, it may have excluded posts referring to acne indirectly, through slang, misspellings, abbreviations, or subtype-specific terminology without explicit mention of the term "acne."

Geolocation was performed using Brandwatch and Visibrain and verified via a lexical analysis based on country-specific spelling variations (eg, "color" vs "colour"). Thus, 2 distinct datasets were created: 1 for the United States and 1 for the United Kingdom.

The data were cleaned by removing posts with ≤5 words or more than 10,000 characters to optimize text processing. Duplicates, non-English content, and irrelevant or unreliable sources (eg, advertising websites, forums about cars, or animals) were also removed. A supervised ML algorithm using Extreme Gradient Boosting (XGBoost) [30] classifiers (F_1-score: 87%) was applied to identify posts from individuals directly affected by acne.

Following these processing steps, the final "clean" dataset was obtained, consisting exclusively of posts from individuals with acne.

Ethical Considerations

This study included data from publicly available sources; private groups or web pages were therefore excluded from our data extraction process. We did not seek approval, as users automatically grant their consent for the reuse of their data when they post on public platforms. Furthermore, the results of this study do not contain any identifiable information and are presented in aggregate. Information such as names, usernames or handles, geographic locations, or any other sensitive data was not included. The names of the brands were also removed and replaced with a generic (BRAND) placeholder in the messages.

Data Analysis

Demographics

We identified users' gender through explicit mentions or inferred using regular expressions (ie, searching usernames for first names, as well as gender-related suffixes and prefixes, and scanning posts for gender-specific vocabulary). Based on these indicators, a scoring system then determined the most likely gender [31].

Age was determined through explicit mentions or predicted using a probabilistic ML model based on features such as syntax, emotional tone, and platform of expression [31].

The remaining cases were classified as "undetermined."

Topics of Discussion

Biterm topic modeling (BTM; via the BTM R package [32]) was applied to all posts. BTM is an NLP, text mining approach that analyzes large volumes of text and clusters similar text based on common topics [32]. The model automatically groups messages into different categories—each representing a specific topic—in descending order of frequency.

In this study, we applied BTM without prior knowledge about the topics that might emerge to ensure an unbiased analysis. As a result, BTM automatically organized the posts into categories,

XSL•FO
RenderX

ranked by the frequency with which they appeared in the dataset. Posts could belong to multiple topics. Through human interpretation, these lists of words were used to label the topics, and the associated posts were thoroughly reviewed to ensure correct interpretation.

Engagement Rate Analysis by Content Topic and Social Media Platform

For each topic, the engagement rate of every message was calculated using the available engagement parameters, according to the following formula:

> Engagement rate per message=Likes+Comments on that messageFollowers at time of posting×100

We then calculated the average engagement rate per topic by computing the mean of all engagement rates from the messages assigned to that topic. Engagement rate data were only available for Instagram and TikTok, so this analysis was limited to these 2 platforms.

Impact on QoL

A hybrid NLP approach, adapted from the method developed by Renner et al [33], was used to assess the impact on QoL. The original framework was developed using a manually annotated corpus of 1399 health-related social media messages extracted from online forums discussing multiple diseases and treatments. Messages were manually labeled by trained annotators according to the presence or absence of QoL impact and categorized into predefined QoL dimensions.

The methodology was based on an adaptation of validated QoL questionnaires, the EuroQol 5-Dimension Questionnaire [34] and 36-Item Short Form Health Survey [35], and included 5 dimensions of impact: physical or signs and symptoms, mental health, social functioning, activity-related functioning, and financial burden. In this study, activity-related impacts were integrated within the social functioning dimension.

Consistent with the original methodology, the algorithm incorporates multiple types of features, including lexical fields specific to QoL dimensions, sentiment-related features (eg, expressions of sadness, fear, or negativity), grammatical patterns (eg, pronoun usage, verb tense), and contextual markers of user experiences. The framework also incorporates a rule-based approach using regular expressions (Regex) to identify explicit mentions of QoL-related concepts through predefined linguistic patterns and keywords. During the development of the original framework, several ML algorithms were evaluated, including K-nearest neighbors, support vector machines, random forest, multilayer perceptron, and XGBoost. XGBoost demonstrated the best overall performance and was selected for the final supervised classification models [33]. Separate models were developed for overall QoL impact detection and for each QoL dimension using 5-fold cross-validation and hyperparameter optimization, with the synthetic minority oversampling technique applied to address class imbalance [33].

In this study, the previously developed and validated algorithm was applied to the acne-related dataset. Messages were first classified according to the presence or absence of QoL impact, then assigned to one or more QoL dimensions: signs and symptoms, mental health, social functioning, and cost. The final categorization combined outputs from both the supervised ML models and the Regex-based approach, which enabled the identification of both explicit and implicit QoL impacts.

Classifications were additionally reviewed manually by 2 evaluators (MS and JM) to ensure consistency and relevance in the context of acne-related discussions.

The workflow can be illustrated using the following example:

> *My acne makes me feel really anxious about my appearance, and it's starting to affect my confidence at work and my relationships.*

This message would first be included in the dataset as originating from an individual directly affected by acne. BTM would then assign it to a topic related to self-image and psychological well-being. The QoL algorithm would subsequently classify the message as having an impact on QoL and assign it to the "mental health" and "social functioning" dimensions.

Unmet Needs

Two evaluators (MS and JM) annotated messages using an extensive grid consisting of 54 distinct unmet needs. While there are no established guidelines for determining saturation in the context of social media content, this approach is widely accepted in qualitative methodology as a means to determine when further data collection becomes unnecessary [36-38]. To apply this method, successive random samples of messages were annotated until data saturation was reached, defined as ≤2 new unmet needs identified across 2 consecutive samples. Saturation was then confirmed with 2 additional samples.

Results

Population and Posts

A total of 1,220,014 messages written by 745,116 internet users discussing acne were retrieved. After data cleaning, 646,809 messages (n=294,786 from the United States; n=352,023 from the United Kingdom) posted by 432,234 users (n=196,911 from the United States; n=235,323 from the United Kingdom) were retained (Multimedia Appendix 1).

A total of 711 sources were included in this study (Multimedia Appendix 2). Over 90% (596,242/646,809) of the messages originated from X (n=304,347, 47.1%), Instagram (n=183,655, 28.4%), and TikTok (n=108,240, 16.7%). Sources contributing individually to less than 0.1% (n≤312) of the dataset, including smaller online forums and community websites, were grouped into the category "Others" (n=28,614, 4.4%). Subsequent analyses were not segmented by source.

Most users self-identified as female (246,373/432,234, 57.0%) compared to male (n=88,176, 20.4%); the rest were undetermined (n=97,685, 22.6%).

Regarding age, 16.0% (69,363/432,234) stated they were under 20 years and 26.5% (n=114,365) were aged 20 to 29 years. Overall, users under 30 represented 42.5% (183,728/432,234) of the dataset, with TikTok users showing the highest proportion in this age group (36,186/79,356, 45.6%). The 30 - to 39-year

XSL•FO
RenderX

category was also prominent (152,728/432,234, 35.3%), particularly on Instagram (45,876/121,688, 37.7%). The mean age was 30.8 (SD 11.24) years (Table 1).

Table . Characteristics of users posting about acne on the top 3 data sources analyzed (United States and United Kingdom, 2024).

Characteristics	All sources	X/Twitter	Instagram	TikTok	Other
Users, n (%)	432,234 (100)	205,008 (47.4)	121,688 (28.2)	79,356 (18.4)	26,182 (6.0)
Posts, n (%)	646,809 (100)	304,347 (47.1)	183,655 (28.4)	108,240 (16.7)	50,567 (7.8)
User gender, n (%)					
Female	246,373 (57.0)	115,214 (56.2)	71,431 (58.7)	45,154 (56.9)	14,574 (55.7)
Male	88,176 (20.4)	42,642 (20.8)	23,364 (19.2)	16,347 (20.6)	5823 (22.2)
Undetermined	97,685 (22.6)	47,152 (23)	26,893 (22.1)	17,855 (22.5)	5785 (22.1)
User age (y), n (%)					
0 - 19	69,363 (16.0)	33,006 (16.1)	18,253 (15.0)	14,284 (18.0)	3819 (14.6)
20-29	114,365 (26.5)	56,582 (27.6)	31,761 (26.1)	21,902 (27.6)	4120 (15.7)
30-39	152,728 (35.3)	70,113 (34.2)	45,876 (37.7)	25,473 (32.1)	11,266 (43.0)
40-49	73,882 (17.1)	34,851 (17.0)	19,713 (16.2)	13,649 (17.2)	5668 (21.6)
>50	13,451 (3.1)	6355 (3.1)	3772 (3.1)	2539 (3.2)	784 (3.0)
Undetermined	8445 (2.0)	4100 (2.0)	2312 (1.9)	1508 (1.9)	525 (2.0)
Age (y), mean (SD)	30.8 (11.24)	30.6 (11.26)	31.0 (11.03)	30.3 (11.45)	32.6 (11.24)

Topics of Discussion

BTM identified various discussion topics (Table 2).

Table . List of the main identified topics of discussion about acne among US and UK internet users and their proportion of messages.

Title	US users, n (%)	UK users, n (%)	Total of all messages, n (%)
Discussion about skincare routine and product recommendation	73,492 (24.9)	81,415 (23.1)	154,907 (24.0)
Acne scars	50,162 (17.0)	85,481 (24.3)	135,643 (21.0)
General treatment information	53,675 (18.0)	43,502 (12.4)	97,177 (15.0)
Hormonal acne	43,670 (14.8)	44,158 (12.5)	87,828 (13.6)
Dietary and nutritional strategies	11,780 (4.0)	19,519 (5.5)	31,299 (4.8)
Impact on self-image, self-confidence, and psychological well-being	27,030 (9.2)	NA[a]	27,030 (4.2)
Makeup and acne	6347 (2.2)	16,482 (4.6)	22,829 (3.5)
Discussion about success stories in the treatment navigation	NA	18,643 (5.0)	18,643 (2.9)
Body acne	11,896 (4.0)	2756 (0.8)	14,652 (2.3)
Natural remedies and carcinogens in the treatments	11,083 (2.8)	NA	11,083 (1.7)
Overcoming insecurity beyond the impact on self-image	NA	6747 (1.9)	6747 (1.0)
Total	294,786 (100)	352,023 (100)	646,809 (100)

[a]"NA" indicates cases where topic modeling did not detect enough messages to generate a distinct topic.

Overall, the most frequently discussed topic centered around skincare routines and product recommendations (154,907/646,809, 24.0%), with users mentioning ingredients such as niacinamide, azelaic acid, and salicylic acid. Acne scars were also frequently discussed (135,643/646,809, 21.0%), particularly in the United Kingdom (85,481/135,643, 24.3%) compared to in the United States (50,162/135,643, 17.0%). Users described the treatments they sought for scarring, including microneedling.

General treatment information represented 15.0% (97,177/646,809) of the messages, including discussions about

XSL•FO
RenderX

prescribed treatments such as isotretinoin and birth control and their effects on acne. Hormonal acne was another frequently discussed topic (87,828/646,809, 13.6%), along with dietary and nutritional strategies aimed at improving acne (n=31,299, 4.8%).

Discussions related to the impact of acne on self-image were specific to the United States (27,030/646,809, 4.2%), while topics about success stories in the treatment navigation (n=18,643, 2.9%) and overcoming insecurity beyond self-image (n=6747, 1.0%) were specific to the United Kingdom (Table 2).

Engagement Rate Analysis by Content Topic and Social Media Platform

Engagement rates help indicate user activity levels across platforms, showing where users are most engaged relative to the audience size.

On Instagram, dietary and nutritional strategies (0.1615%, SD 6.36%) showed the highest mean engagement, followed by skincare routines and product recommendations (0.1108%, SD 4.81%). Hormonal acne (0.0863%, SD 1.76%), success stories in treatment navigation (0.0837%, SD 0.76%), and body acne (0.0809%, SD 2.29%) showed similar engagement levels. In contrast, discussions about natural remedies and carcinogens in treatments generated the lowest mean engagement (0.0417%, SD 0.55%). In general, engagement scores were higher in the United Kingdom compared to the United States across all topics (Table 3).

Table . Engagement rate analysis on Instagram. Engagement rate per message=Likes+Comments on that messageFollowers at time of posting×100. Values may exceed one when (1) followers generate multiple interactions (eg, like and comment), (2) followers post more than 1 comment, or (3) content reaches nonfollowers.

Topic	United Kingdom		United States		Total	
	Score (%), mean (SD)	Messages with score available, n	Score (%), mean (SD)	Messages with score available, n	Score (%), mean (SD)	Messages with score available, n
Dietary and nutritional strategy	0.1880 (7.30)	11,350	0.0830 (1.47)	3826	0.1615 (6.36)	15,176
Discussion about skincare routine and product recommendation	0.1300 (5.95)	13,650	0.0760 (1.00)	7519	0.1108 (4.81)	21,169
Hormonal acne	0.1080 (2.08)	1834	0.0350 (0.37)	778	0.0863 (1.76)	2612
Discussion about success stories in the treatment navigation	0.0837 (0.76)	1809	NA[a]	NA	0.0837 (0.76)	1809
Body acne	0.1010 (4.24)	1596	0.0728 (0.37)	3958	0.0809 (2.29)	5554
Overcoming insecurity beyond the impact on self-image	0.0788 (1.28)	2737	0.0540 (0.55)	1019	0.0721 (1.12)	3756
General treatment information	0.0786 (3.07)	15,582	0.0350 (0.34)	3607	0.0704 (2.77)	19,189
Acne scars	0.0667 (1.85)	9084	0.0646 (0.58)	3829	0.0661 (1.58)	12,913
Makeup and acne	0.0638 (0.68)	4392	0.0548 (0.31)	1652	0.0613 (0.60)	6044
Impact on self-image, self-confidence, and psychological well-being	NA	NA	0.0560 (0.55)	1119	0.0560 (0.55)	1119
Natural remedies and carcinogens in the treatments	NA	NA	0.0417 (0.55)	2455	0.0417 (0.55)	2455

[a]NA: not applicable; the topic was not identified in the corresponding subgroup.

According to the 2024 Social Media Industry Benchmark Report by Rival IQ, the median Instagram engagement rate across all industries is 0.43% [39]. As for the health and beauty industry, the "Engagement Rate per Post" is listed as 0.07% [40].

On TikTok, content about makeup and acne had the highest mean engagement score (3.0337%, SD 92.65%). Discussions about overcoming insecurity beyond the impact on self-image also generated strong engagement (1.1801%, SD 12.82%), followed by body acne (0.8977%, SD 16.07%) and general

XSL•FO
RenderX

treatment information (0.8794%, SD 17.13%). Dietary and nutritional strategies (0.8438%, SD 11.22%) and acne scars (0.6198%, SD 7.17%) showed moderate engagement levels, whereas hormonal acne had the lowest mean engagement (0.2570%, SD 2.12%; Table 4).

Table . Engagement rate analysis on TikTok. Engagement rate per Message=Likes+Comments on that messageFollowers at time of posting×100. Values may exceed one when (1) followers generate multiple interactions (eg, like and comment), (2) followers post more than 1 comment, or (3) content reaches nonfollowers.

Topic	United Kingdom		United States		Total	
	Score (%), mean (SD)	Messages with score available, n	Score (%), mean (SD)	Messages with score available, n	Score (%), mean (SD)	Messages with score available, n
Makeup and acne	0.8140 (16.09)	1557	9.4696 (180.85)	537	3.0337 (92.65)	2094
Overcoming insecurity beyond the impact on self-image	1.4120 (12.76)	382	0.5650 (13.02)	144	1.1801 (12.82)	526
Body acne	1.0230 (19.56)	1094	0.7090 (8.41)	726	0.8977 (16.07)	1820
General treatment information	0.5350 (5.07)	4159	1.5450 (28.46)	2152	0.8794 (17.13)	6311
Dietary and nutritional strategy	0.5010 (3.8)	1848	2.3100 (24.55)	432	0.8438 (11.22)	2280
Acne scars	0.5250 (5.6)	2440	0.8890 (10.47)	859	0.6198 (7.17)	3299
Discussion about skincare routine and product recommendation	0.3780 (5.15)	6479	0.8540 (12.91)	3495	0.5448 (8.70)	9974
Impact on self-image, self-confidence, and psychological well-being	NA[a]	NA	0.4850 (2.22)	124	0.4850 (2.22)	124
Natural remedies and carcinogens in the treatments	NA	NA	0.2889 (1.42)	319	0.2889 (1.42)	319
Discussion about success stories in the treatment navigation	0.269 (1.72)	423	NA	NA	0.2690 (1.72)	423
Hormonal acne	0.337 (2.51)	808	0.0660 (0.33)	338	0.2570 (2.12)	1146

[a]NA: not applicable; the topic was not identified in the corresponding subgroup.

According to Rival IQ's TikTok Benchmark Report, the median engagement rate per follower is 2.6% [41].

Impact on QoL

In our study, 52.9% (228,613/432,234) of the users expressed at least 1 impact on their QoL. The most frequently cited domains were signs and symptoms (175,604/228,613, 76.8%), followed by social functioning (n=149,234, 65.3%), mental health (n=107,155, 46.9%), and cost (n=62,008, 27.1%; Multimedia Appendix 3).

Among users describing signs and symptoms, 27.7% (48,671/175,604) mentioned pain and itching, while 21.5% (n=37,755) reported scarring and hyperpigmentation. Pimples and redness were also mentioned (n=16,682, 9.5%).

Regarding social functioning, 48.6% (72,528/149,234) reported adapting their skincare routines to alleviate symptoms, while 48.0% (n=71,632) expressed that acne interfered with their professional or school life. Interference with relationships was also reported by 28.6% (n=42,681).

Mental health impacts included feelings of depression and anxiety linked to negative self-image (55,721/107,155, 52%), as well as fear of recurrence (n=6536, 6.1%) and loneliness (n=2786, 2.6%).

The cost of managing acne was also highlighted as a major concern, particularly the cost of treatments and skincare products (20,835/62,008, 33.6%). Users often searched for more affordable treatments (15,254/62,008, 24.6%) or for ones covered by their insurance (n=5271, 8.5%).

Unmet Needs

A total of 3200 messages were annotated (n=1600 messages per country). Among these, 272 (17.0%) messages in the US dataset and 310 (19.4%) messages in the UK dataset expressed at least 1 unmet need.

The most frequently expressed needs included effective solutions for hormonal acne (111/582, 19.1%), especially treatments considered safe during pregnancy. Users also expressed a need for clarity in identifying acne triggers (84/582, 14.4%). Reported

XSL•FO
RenderX

triggers included diet, makeup products, skincare routines, treatments, sun exposure, weather, stress, hormones, sweat, and fitness-related supplements.

Another commonly expressed need involved support and feedback on treatment experiences (68/582, 11.7%), reflecting a perceived lack of reliable information. Many individuals turned to peers or online "experts" for recommendations and opinions on treatments and skincare products.

Psychological support was another mentioned unmet need (68/582, 11.7%), particularly in relation to negative self-esteem, shame, frustration, and social isolation. Additional unmet needs included improved management strategies for acne scars (n=59, 10.1%), clearer guidance throughout the acne care pathway (n=48, 8.2%), and strategies to manage acne triggered by diet (n=33, 5.7%).

Users also expressed a need for more information about specific treatments, mainly isotretinoin, retinoids, and antibiotics (28/582, 4.8%), as well as general advice on managing acne (n=25, 4.3%). Finally, some users mentioned the need for support in managing the financial burden of treatments (n=16, 2.7%; Table 5).

Table . Top 10 unmet needs related to acne in messages posted in the United States and the United Kingdom in 2024.

Serial number	Needs	United States, n (%)	United Kingdom, n (%)	Total of all messages containing unmet need, n (%)
1	Need for effective solutions for hormonal acne	69 (25.4)	42 (13.5)	111 (19.1)
2	Need for clarity in identifying acne triggers	44 (16.2)	40 (12.9)	84 (14.4)
3	Need for support and feedback on treatment experiences	29 (10.7)	39 (12.6)	68 (11.7)
4	Need for psychological support for acne-related low self-esteem	38 (14)	30 (9.7)	68 (11.7)
5	Need for better management strategies for acne scars	24 (8.8)	35 (11.3)	59 (10.1)
6	Need for clearer guidance through the acne care pathway	17 (6.3)	31 (10.0)	48 (8.2)
7	Need for strategies to manage acne triggered by diet	14 (5.1)	19 (6.1)	33 (5.7)
8	Need for information on specific treatments	8 (2.9)	20 (6.5)	28 (4.8)
9	Need for general advice on acne management	8 (2.9)	17 (5.5)	25 (4.3)
10	Need for support in managing the financial burden of treatments	4 (1.5)	12 (3.9)	16 (2.7)

Discussion

Principal Findings

Our findings highlighted the diverse ways in which individuals with acne engage with social media, from January 1 to December 31, 2024. Over half of the posts expressed the effects of acne on QoL, with users describing their symptoms, mental health, and social functioning. These results are consistent with previous research highlighting the negative impact of the condition [42-44]. Consistent with prior survey work, our study identified psychosocial effects including feelings of depression, anxiety, shame, and social isolation [45-48]. Notably, an impact on QoL was also reported among family members of individuals with acne [49]. This may also reflect that those with acne are often younger and may still be living with their family, and is aligned with similar findings that have been observed among those with atopic dermatitis [50].

Most of the retrieved data came from X, Instagram, and TikTok, highlighting these platforms as key spaces where people discuss acne. Users were predominantly female, which aligns with previous studies [51,52] and could reflect that women are more affected by acne across the lifespan or more likely to engage in discussions about their skin on digital platforms.

Users frequently discussed topics such as scarring and hormonal acne, areas that have been less emphasized in previous research. Discussions also included skincare product recommendations and dietary and lifestyle interventions. Users additionally described important unmet needs in these areas, including the need for better psychological support. This aligns with results from another social media study on dermatology, where discussions centered around exchanging tips and advice [53]. However, as social media content on these topics can often be poor or misleading [27], these findings highlight an opportunity for dermatologists and other health professionals to create

XSL•FO
RenderX

high-quality, evidence-based educational content on these topics for those with acne. They also support the need for additional research on nonpharmacologic management strategies for acne (eg, diet), treatment for sequelae such as scarring and dyspigmentation, and better management of comorbidities, such as depression and anxiety.

Interestingly, the analysis of engagement rates showed that the topics most frequently discussed were not necessarily those that attracted the most interactions. This suggests that users are more interested in seeking solutions than sharing about their condition [26]. The analysis of engagement rates per platform provided new insights into how users interact with acne-related information depending on the social media platform. On Instagram, messages related to nutrition and skincare showed the highest mean engagement scores. This finding was reported in another acne-related social media study [24], although that work relied only on message volume rather than engagement rates. On TikTok, engagement was higher for content about makeup, which shows a strong interest in finding suitable products for sensitive or breakout-prone skin, as cosmetics have been associated with worsening symptoms [54]. Notably, the topic promoting positive body image received a high engagement rate. Despite the well-documented negative impact of acne on body image and self-esteem [43,55,56], our findings reveal that there may be a growing shift toward self-acceptance. A similar theme was reported in previous research analyzing acne content on TikTok [57], although that analysis was based solely on view counts rather than multimetric engagement used in our study.

To our knowledge, this is the first study to combine a multiplatform and multicountry social media analysis of acne discussions in the United States and the United Kingdom. This study enabled the identification of subtle cross-country differences in frequently discussed topics, their engagement rates per platform, and unmet needs related to acne. These findings provide valuable insights, although they should be interpreted in light of the fact that the study population reflects individuals who publicly discuss acne on social media rather than the general acne population. In addition, differences in platform usage, communication styles, and cultural expressions between countries may have influenced the observed patterns.

Limitations

This study includes limitations that are common in social media research. The data were derived from publicly available sources; private sources, such as closed Facebook groups and WhatsApp, were thus excluded. These environments may include individuals with different characteristics, including more vulnerable populations or those less likely to share their experiences publicly, which may introduce selection bias. In addition, social media data are based on user-provided information, which may be subject to recall bias, subjective interpretation, and incomplete context. User characteristics, including demographic and clinical information, could not be verified, which should be considered when interpreting the findings and cross-country comparisons. Additionally, individuals posting on social media may belong to certain socioeconomic backgrounds and possess varying literacy capacities, which may affect the representativeness of our findings.

This study was restricted to the United States and the United Kingdom to ensure language consistency and comparability of social media content, as the methodology relies on language-specific features such as lexical fields, sentiment expressions, and grammatical patterns. However, linguistic variations, cultural norms, and differences in how individuals express emotions, symptoms, and social experiences online may influence how QoL impacts are detected and interpreted across populations. In addition, acne-related terminology, slang, and communication styles may differ between countries and languages. Platform availability and usage may also vary across countries, which may potentially influence both the type of users represented and the nature of the content shared. As a result, the transferability of this approach to other countries may require adaptation and validation in different linguistic and cultural contexts. Moreover, it is worth noting that relevant posts might have been mistakenly eliminated during the filtering process. To reduce background noise, we applied threshold values similar to our previous work [58-60], which may have limited the natural processing analysis. Despite these limitations, this study provides valuable insights into the impact of acne on individuals' lives.

Conclusions

This study found that acne can have diverse and meaningful effects on QoL. Reports of pain, impaired social functioning, depression, anxiety, and the cost of treatments were commonly described in users' messages. Important unmet needs were identified, including nonpharmacologic management strategies for acne (eg, diet), treatment for sequelae such as scarring and dyspigmentation, and management of depression and anxiety. Given the growing role of social media, these findings highlight opportunities for dermatologists and health professionals to educate and engage with the acne community through social media. Future research should explore how digital interventions and education campaigns can better address these unmet needs and reduce the psychosocial burden of acne.

Acknowledgments

Medical writing support was provided by Joelle Malaab (Kap Code).

Funding

This study was funded by Sanofi.

XSL•FO
RenderX

Data Availability

The algorithms used in this study are proprietary intellectual property of Kap Code and, therefore, cannot be disclosed. Furthermore, in accordance with the General Data Protection Regulation (Regulation [EU] 2016/679), these data cannot be shared. Generative artificial intelligence was not used in the preparation of this manuscript.

Authors' Contributions

Conceptualization: JSB, VD, IV, S Samson, EP, JM, S Schück
Data curation: MS, MT
Formal analysis: JM, MS
Investigation: MS
Methodology: MT, S Schück
Supervision: JSB, S Schück
Writing – review and editing: JSB, VD, IV, S Samson, EP, JM
Writing – original draft: JM

Conflicts of Interest

VD, IV, S Samson, and EP are employees of Sanofi. JM, MS, MT, and S Schück are employees of Kap Code. Kap Code was funded by Sanofi to conduct the study. JSB has received personal consulting fees from Sanofi Pasteur, Honeydew Care, Alosa Health, and Twi Biotechnology.

Multimedia Appendix 1
Flowchart of the data cleaning and sample selection processes, showing the number of messages (n) and users discussing acne in 2024.
[DOCX File, 60 KB - derma_v9i1e91126_app1.docx]

Multimedia Appendix 2
Complete list of sources.
[DOCX File, 19 KB - derma_v9i1e91126_app2.docx]

Multimedia Appendix 3
Proportions of US and UK internet users reporting the impact of acne on their quality of life on social media in 2024.
[DOCX File, 18 KB - derma_v9i1e91126_app3.docx]

References

1. Chen H, Zhang TC, Yin XL, Man JY, Yang XR, Lu M. Magnitude and temporal trend of acne vulgaris burden in 204 countries and territories from 1990 to 2019: an analysis from the Global Burden of Disease Study 2019. Br J Dermatol 2022 Apr;186(4):673-683. [doi: 10.1111/bjd.20882] [Medline: 34758111]
2. Skroza N, Tolino E, Mambrin A, et al. Adult acne versus adolescent acne: a retrospective study of 1,167 patients. J Clin Aesthet Dermatol 2018 Jan;11(1):21-25. [Medline: 29410726]
3. Collier CN, Harper JC, Cafardi JA, et al. The prevalence of acne in adults 20 years and older. J Am Acad Dermatol 2008 Jan;58(1):56-59. [doi: 10.1016/j.jaad.2007.06.045] [Medline: 17945383]
4. Holzmann R, Shakery K. Postadolescent acne in females. Skin Pharmacol Physiol 2014;27 Suppl 1:3-8. [doi: 10.1159/000354887] [Medline: 24280643]
5. Khunger N, Kumar C. A clinico-epidemiological study of adult acne: is it different from adolescent acne? Indian J Dermatol Venereol Leprol 2012;78(3):335-341. [doi: 10.4103/0378-6323.95450] [Medline: 22565434]
6. Tanghetti EA, Kawata AK, Daniels SR, Yeomans K, Burk CT, Callender VD. Understanding the burden of adult female acne. J Clin Aesthet Dermatol 2014 Feb;7(2):22-30. [Medline: 24578779]
7. Sharma R, Dogra N, Arora M. Psychosocial impact of acne vulgaris on the quality of life among adolescents versus adults. Clin Med (Lond) 2024 Apr;23(6):35. [doi: 10.7861/clinmed.23-6-s35]
8. Özkesici Kurt B. Comparison of the psychosocial impact of acne in adolescents and adults; body satisfaction, self-esteem, and quality of life. J Cosmet Dermatol 2022 Feb;21(2):836-843. [doi: 10.1111/jocd.14151] [Medline: 33844401]
9. Shields A, Nock MR, Ly S, Manjaly P, Mostaghimi A, Barbieri JS. Evaluation of stigma toward individuals with acne. JAMA Dermatol 2024 Jan 1;160(1):93-98. [doi: 10.1001/jamadermatol.2023.4487] [Medline: 38055249]
10. Karimkhani C, Dellavalle RP, Coffeng LE, et al. Global skin disease morbidity and mortality: an update from the Global Burden of Disease Study 2013. JAMA Dermatol 2017 May 1;153(5):406-412. [doi: 10.1001/jamadermatol.2016.5538] [Medline: 28249066]

XSL•FO
RenderX

11. Vasam M, Korutla S, Bohara RA. Acne vulgaris: a review of the pathophysiology, treatment, and recent nanotechnology based advances. Biochem Biophys Rep 2023 Dec;36:101578. [doi: 10.1016/j.bbrep.2023.101578] [Medline: 38076662]
12. Gollnick H. Current concepts of the pathogenesis of acne: implications for drug treatment. Drugs 2003;63(15):1579-1596. [doi: 10.2165/00003495-200363150-00005] [Medline: 12887264]
13. Tobiasz A, Nowicka D, Szepietowski JC. Acne vulgaris-novel treatment options and factors affecting therapy adherence: a narrative review. J Clin Med 2022 Dec 19;11(24):7535. [doi: 10.3390/jcm11247535] [Medline: 36556150]
14. Kim SK, Lee M, Lee YQ, et al. Genome-scale metabolic modeling and in silico analysis of opportunistic skin pathogen Cutibacterium acnes. Front Cell Infect Microbiol 2023;13:1099314. [doi: 10.3389/fcimb.2023.1099314] [Medline: 37520435]
15. Choe J, Gaurav A, Xia E, Yan A, Mostaghimi A, Barbieri JS. Use and cost of dietary changes, supplements, over-the-counter topicals, and prescriptions for acne: a cross-sectional survey among United States adults. Arch Dermatol Res 2024 Sep 3;316(8):598. [doi: 10.1007/s00403-024-03349-7] [Medline: 39225715]
16. Reynolds RV, Yeung H, Cheng CE, et al. Guidelines of care for the management of acne vulgaris. J Am Acad Dermatol 2024 May;90(5):1006. [doi: 10.1016/j.jaad.2023.12.017] [Medline: 38300170]
17. Number of internet and social media users worldwide as of October 2025. Statista. URL: https://www.statista.com/statistics/617136/digital-population-worldwide/ [accessed 2025-07-10]
18. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
19. Cheston CC, Flickinger TE, Chisolm MS. Social media use in medical education: a systematic review. Acad Med 2013 Jun;88(6):893-901. [doi: 10.1097/ACM.0b013e31828ffc23] [Medline: 23619071]
20. Ahmed W, Bath PA, Sbaffi L, Demartini G. Measuring the effect of public health campaigns on Twitter: the case of World Autism Awareness Day. In: Chowdhury G, McLeod J, Gillet V, Willett P, editors. Transform Digit Worlds (2018): Springer International Publishing; 2018:10-16. [doi: 10.1007/978-3-319-78105-1_2]
21. Chung JE. A smoking cessation campaign on Twitter: understanding the use of Twitter and identifying major players in a health campaign. J Health Commun 2016 May;21(5):517-526. [doi: 10.1080/10810730.2015.1103332] [Medline: 27115179]
22. Gupta P, Khan A, Kumar A. Social media use by patients in health care: a scoping review. Int J Healthc Manag 2022 Apr 3;15(2):121-131. [doi: 10.1080/20479700.2020.1860563]
23. Yousaf A, Hagen R, Delaney E, Davis S, Zinn Z. The influence of social media on acne treatment: a cross-sectional survey. Pediatr Dermatol 2020 Mar;37(2):301-304. [doi: 10.1111/pde.14091] [Medline: 31944359]
24. Ertekin SS, Salici NS, Manav Bas V, et al. Influence of social media and internet on treatment decisions in adult female acne patients: a cross-sectional survey study. Dermatol Pract Concept 2024 Jul 1;14(3):e2024156. [doi: 10.5826/dpc.1403a156] [Medline: 39122512]
25. Wetstone R, Grant-Kels JM. The skincare craze among tweens: ethical and dermatologic implications of social media beauty trends. Clin Dermatol 2025;43(6):897-899. [doi: 10.1016/j.clindermatol.2025.02.009] [Medline: 39971259]
26. Marty T, Khadar M, Autelitano L, et al. Patients' testimonies, feelings, complaints and emotional experiences with dermatoses on open social media: the French infodemiologic patient's free speech study. J Eur Acad Dermatol Venereol 2024 Jul;38(7):1373-1382. [doi: 10.1111/jdv.19781] [Medline: 38314863]
27. Guzman AK, Wang RH, Nazarian RS, Barbieri JS. Evaluation of YouTube as an educational resource for treatment options of common dermatologic conditions. Int J Dermatol 2019 Oct;59(3):e65-e67. [doi: 10.1111/ijd.14693] [Medline: 31631334]
28. Consumer research features. Brandwatch. URL: https://www.brandwatch.com/products/consumer-research/features/ [accessed 2022-07-25]
29. Visibrain [Website in French]. URL: https://www.visibrain.com/fr/ [accessed 2025-07-19]
30. Xgboost: Extreme Gradient Boosting. CRAN R package. URL: https://CRAN.R-project.org/package=xgboost [accessed 2026-06-06]
31. Arquembourg J, Glaser P, Roblot F, et al. Discussions of antibiotic resistance on social media platforms: text mining and mixed methods content analysis study. JMIR Form Res 2025 Apr 25;9(1):e37160. [doi: 10.2196/37160] [Medline: 40289322]
32. Yan X, Guo J, Lan Y, Cheng X. A biterm topic model for short texts. Presented at: WWW '13: Proceedings of the 22nd international conference on World Wide Web; May 13-17, 2013. [doi: 10.1145/2488388.2488514]
33. Renner S, Marty T, Khadhar M, et al. A new method to extract health-related quality of life data from social media testimonies: algorithm development and validation. J Med Internet Res 2022 Jan 28;24(1):e31528. [doi: 10.2196/31528] [Medline: 35089152]
34. Devlin N, Parkin D, Janssen B. An introduction to EQ-5D instruments and their applications. In: Methods for Analysing and Reporting EQ-5D Data: Springer; 2020:1-22. [doi: 10.1007/978-3-030-47622-9_1]
35. Short Form 36—an overview. ScienceDirect. URL: https://www.sciencedirect.com/topics/nursing-and-health-professions/short-form-36 [accessed 2023-10-09]
36. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qual Res 2008 Feb;8(1):137-152. [doi: 10.1177/1468794107085301]
37. Garrett CR, Gask LL, Hays R, et al. Accessing primary health care: a meta-ethnography of the experiences of British South Asian patients with diabetes, coronary heart disease or a mental health problem. Chronic Illn 2012 Jun;8(2):135-155. [doi: 10.1177/1742395312441631] [Medline: 22414446]

XSL•FO
RenderX

38. Lipworth W, Taylor N, Braithwaite J. Can the theoretical domains framework account for the implementation of clinical quality interventions? BMC Health Serv Res 2013 Dec 21;13:530. [doi: 10.1186/1472-6963-13-530] [Medline: 24359085]
39. What is a good engagement rate on Instagram? Rival IQ. URL: https://www.rivaliq.com/blog/good-engagement-rate-instagram/ [accessed 2025-09-30]
40. Live social media benchmarks for health & beauty. Rival IQ. URL: https://www.rivaliq.com/live-benchmarks/health-and-beauty-industry-benchmarks/ [accessed 2025-09-30]
41. What is a good engagement rate on TikTok? Rival IQ. URL: https://www.rivaliq.com/blog/good-engagement-rate-tiktok/ [accessed 2024-06-14]
42. Ding RL, He YZ, Wu QJ, Lin T. Quality of life in patients with acne vulgaris: an observational study. Int J Dermatol Venereol 2025;8(4):205-210. [doi: 10.1097/JD9.0000000000000315]
43. Ali Awad EY, Soliman FES, Maklad SA. The effect of acne vulgaris on body image satisfaction, self esteem and quality of life of adolescents at Tanta city. Tanta Sci Nurs J 2022 Nov 1;27(4):188-206. [doi: 10.21608/tsnj.2022.267578]
44. Alfahl SO, Almehmadi LA, Alamri RS, Almalki DS, Alnakhli F. Impact of acne vulgaris and vitiligo on quality of life and self-esteem in the patient population of Madinah, Saudi Arabia. Cureus 2024 Jan;16(1):e52586. [doi: 10.7759/cureus.52586] [Medline: 38371097]
45. Spitzer C, Buchholz I, Lübke L, et al. Determinants of quality of life in patients with atopic dermatitis and psoriasis: a multivariate approach. Acta Derm Venereol 2025 May 19;105:adv42794. [doi: 10.2340/actadv.v105.42794] [Medline: 40390259]
46. Yamasaki K, Miyachi Y. Perspectives on rosacea patient characteristics and quality of life using baseline data from a phase 3 clinical study conducted in Japan. J Dermatol 2022 Dec;49(12):1221-1227. [doi: 10.1111/1346-8138.16596] [Medline: 36177741]
47. Seetan K, Gablan M, Alnaimi M, et al. Assessment of depressive and anxiety symptoms and health-related quality of life in rosacea patients: a case-control study. Dermatol Res Pract 2024;2024(1):5532532. [doi: 10.1155/drp/5532532] [Medline: 39720781]
48. Spitzer C, Lübke L, Wülfing C, et al. Facets of shame and their impact on quality of life in patients with atopic dermatitis and psoriasis. Sci Rep 2025 Apr 21;15(1):13753. [doi: 10.1038/s41598-025-98353-w] [Medline: 40259009]
49. Tuğrul B, Demirdağ HG, Aslan C, Muştu Koryürek Ö. An overlooked burden of acne in adolescents: the psychosocial well-being of their families [Article in Spanish]. Anales de Pediatría (English Edition) 2023 Jul;99(1):37-43. [doi: 10.1016/j.anpede.2023.06.009]
50. Celestin AK, Sylvain AA, Wilfried.G AK, Pacome GK, Prisca NT, Aziz CS. Quality of life of family of children with atopic dermatitis using the Dermatitis Family Impact Questionnaire in West-Africa. J Pak Assoc Dermatol 2025;35(1):39-44. [doi: 10.66344/jpad.35.1.2025.2860]
51. Aslan Kayıran M, Karadağ AS, Alyamaç G, et al. Social media use in patients with acne vulgaris: what do patients expect from social media? J Cosmet Dermatol 2021 Aug;20(8):2556-2564. [doi: 10.1111/jocd.14272] [Medline: 34087036]
52. Bahaj RK, Alsaggaf ZH, Abduljabbar MH, Hariri JO. The influence of social media on the treatment of acne in Saudi Arabia. Cureus 2022 Mar;14(3):e23169. [doi: 10.7759/cureus.23169] [Medline: 35444887]
53. Reddy A. Skincare in social media: analyzing prominent themes in online dermatologic discussions. Cureus 2021;13(5):e14890. [doi: 10.7759/cureus.14890]
54. Suh DH, Oh H, Lee SJ, Kim HJ, Ryu HJ. Relationship between acne and the use of cosmetics: results of a questionnaire study in 539 Korean individuals. J Cosmet Dermatol 2021 Jul;20(7):2172-2178. [doi: 10.1111/jocd.13853] [Medline: 33206430]
55. Ramadhanty RP, Hamid AYS. Body image perception is related to self-esteem of the adolescents with acne vulgaris. Enferm. Clin 2021 Apr;31:S326-S329. [doi: 10.1016/j.enfcli.2020.09.021]
56. Cecilia MS, Satapathy S, Ramam M. Assessment of body image disturbance, self-esteem and quality of life among adolescents and young adults with acne in a tertiary care facility of India. Indian J Dermatol 2022;67(1):93. [doi: 10.4103/ijd.ijd_342_21] [Medline: 35656260]
57. Iyengar L, Saldanha S, Chong AH. #Acne: a thematic qualitative analysis of acne content on TikTok. Australas J Dermatol 2025 May;66(3):127-134. [doi: 10.1111/ajd.14433] [Medline: 40008491]
58. Grov EK, Fosså SD, Sørebø O, Dahl AA. Primary caregivers of cancer patients in the palliative phase: a path analysis of variables influencing their burden. Soc Sci Med 2006 Nov;63(9):2429-2439. [doi: 10.1016/j.socscimed.2006.06.008] [Medline: 16887247]
59. Abdellaoui R, Foulquié P, Texier N, Faviez C, Burgun A, Schück S. Detection of cases of noncompliance to drug treatment in patient forum posts: topic model approach. J Med Internet Res 2018 Mar 14;20(3):e85. [doi: 10.2196/jmir.9222] [Medline: 29540337]
60. Schück S, Foulquié P, Mebarki A, et al. Concerns discussed on Chinese and French social media during the COVID-19 lockdown: comparative infodemiology study based on topic modeling. JMIR Form Res 2021 Apr 5;5(4):e23593. [doi: 10.2196/23593] [Medline: 33750736]

XSL•FO
RenderX

Abbreviations

BTM: biterm topic model
ML: machine learning
NLP: natural language processing
QoL: quality of life
XGBoost: Extreme Gradient Boosting

Edited by R Dellavalle; submitted 09.Jan.2026; peer-reviewed by A Snyder, X Liang; revised version received 26.May.2026; accepted 27.May.2026; published 23.Jun.2026.

Please cite as:
S Barbieri J, Delore V, Vieillard I, Samson S, Puentes E, Malaab J, Sambrès M, Talmatkadi M, Schück S
Exploring the Lived Experience of Acne in the United States and the United Kingdom: Social Media Analysis
JMIR Dermatol 2026;9:e91126
URL: https://derma.jmir.org/2026/1/e91126
doi:10.2196/91126

© John S Barbieri, Valentine Delore, Isabelle Vieillard, Sandrine Samson, Esteban Puentes, Joelle Malaab, Melanie Sambrès, Manissa Talmatkadi, Stéphane Schück. Originally published in JMIR Dermatology (http://derma.jmir.org), 23.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Promotion of Tanning Behaviors on Instagram: Content Analysis

Tailisha Gonzalez[1], MS, MPH, PhD; Grace C Hillyer[2], EdD; Bailey Gold[3], MA; Corey H Basch[3], MPH, EdD

[1]Department of Health, Administration, Research, and Technology, Lehman College, 250 Bedford Park Blvd W, Bronx, NY, United States

[2]Department of Health Studies and Applied Educational Psychology, Teachers College of Columbia University, New York, NY, United States

[3]Department of Public Health, William Paterson University, Wayne, NJ, United States

Corresponding Author:
Tailisha Gonzalez, MS, MPH, PhD
Department of Health, Administration, Research, and Technology, Lehman College, 250 Bedford Park Blvd W, Bronx, NY, United States

Abstract

Background: Adolescents and young adults may be particularly susceptible to social media content that promotes tanning and appearance-based beauty standards.

Objective: This study examined the characteristics and messaging of pro- and antitanning Instagram posts and the extent to which they promoted health damaging– or health enhancing–behaviors.

Methods: A cross-sectional content analysis of 100 Instagram posts using the hashtag #tanning assessed creator characteristics, tanning practices, commercial promotion, tanning-related messaging, beauty standards, and health risks and benefits. χ^2 and Fisher exact tests were used to determine differences between pro vs anti stances for categorical and dichotomous variables, and the Mann-Whitney *U* test was used for continuous variables.

Results: Most videos were protanning (79/100). Protanning videos were significantly greater than antitanning videos to have been posted in 2026 (77.2% vs 47.6%; *P*=.02), feature teen and young adult creators (87.3% vs 52.4%; *P*<.001), and include shopping tags promoting tanning products (30.4% vs 0%; *P*<.001). Overall, messaging focused on promoting tanning (76%) and advocating for tanned skin as a standard of beauty (63%) and downplayed the use of sun protection (11%) and the risks associated with tanning (20%). Messaging varied significantly by tanning stance, with antitanning videos focused on sun protection and health risk.

Conclusions: Findings indicate an urgent need for clinicians to lead content creation on popular social media platforms to counter the health-damaging effects of unreliable, incomplete, and potentially misdirected information that is posted under popularized hashtags.

(JMIR Dermatol 2026;9:e110157) doi:10.2196/110157

KEYWORDS

Instagram; social media; tanning; skin cancer prevention; content analysis; sun protection; ultraviolet radiation; health communication; adolescent health

Introduction

Social media increasingly plays a significant role in disseminating content with negative health consequences for adolescents and young adults, especially algorithm-amplified content that promotes Western wealth culture such as tanned skin. Dermatology experts in Australia recently reported a resurgence of dangerous tanning practices driven by influencers who glamorize tan lines and promote SPF-free products [1]. A study of Spanish young adults, all of whom experienced a sunburn in the past year, found that women were more likely to prefer tanned skin after reading tanning-related social media messages despite risk warnings [2]. Public health practitioners have attempted to counter tanning misinformation on social media with health-enhancing information and practices. Despite these efforts, an intervention that educated US young adults on risks associated with indoor tanning failed to detect significant change in salon foot traffic [3]. The impact of social media on health-damaging behaviors and the dissemination of incomplete information or misinformation continues to be documented in real time. The American Academy of Dermatology Association (AADA) conducted a survey of Gen Z adults and found that approximately one-third of respondents used sunscreen only when pressed by others, about 1 in 4 prioritized tanning over skin cancer prevention, slightly more than half were unaware of sunburn risk (eg, skin cancer), and more than 1 in 4 reported skin damage from sun exposure [4,5]. More recently, the AADA warned that social media misinformation has increased skin cancer risk for more than 16 million people in the US, with nearly half of survey respondents reporting that they encountered sunscreen misinformation online; an even greater proportion was reported among Gen Z (64%) [6]. The majority of

XSL•FO
RenderX

respondents (83%) believed that tanning contributed to perceptions of good health [6].

Skin cancer is the most prevalent form of cancer, with 70% to 95% of melanomas in fair-skinned people caused by UV radiation [7]. Indoor tanning, which exposes users to UV rays, has been linked to increased incidence of melanoma, basal cell carcinoma, and squamous cell carcinoma, especially among those who started indoor tanning in young adulthood [8,9]. Sunburns have also been linked to increased skin cancer risk [10]. Despite heightened associations with cancer risk, engagement in sun protection behaviors remains low, especially among adults aged 18 to 25 years [11]. Given rising US skin cancer rates, this study examines the characteristics and message content of pro- and antitanning Instagram posts and whether they promote health-damaging or health-enhancing behaviors.

Methods

Instagram was searched on July 20, 2026, using the hashtag #tanning, which was the most popular tanning-related hashtag on the platform at the time of data collection, sorted by recent posts. The first 100 video posts returned under this hashtag were selected for analysis. Posts were excluded if they were duplicates, non-English, or unrelated to tanning; excluded posts were replaced with the next eligible post in the sequence until 100 videos were retained. A codebook was developed a priori by the study team to capture video characteristics and message content. The video characteristics coded were engagement (views, likes, comments, and shares), year posted, perceived creator age (teen, aged 13 - 19 years; young adult, aged 20 - 29 years; adult, aged 30+ years; unclear) and gender, setting (tanning bed, outside in full sun, outside in shade, indoors), use of self-tanner or spray tan, before and after images, product promotion (shopping tag, verbal/visual only, none), humor, and peer influence. Message content was coded as promotion of tanning, tanned skin as a beauty standard, maximizing UV index, claims that tanning beds are safe, use of sun protection, health benefits (mental health, skin health, other), shrugging off risks, and dismissing cancer risk. Videos were classified as pro- or antitanning based on their dominant message.

All 100 videos were coded by one trained coder (BG) using this codebook. To assess interrater reliability, a second coder (CHB) independently coded a randomly selected subsample of 10% of the videos. Interrater reliability on the subsample was assessed using the Cohen kappa (κ=0.88), indicating high agreement; discrepancies between coders were resolved through discussion and consensus. The characteristics and message content were examined using frequency distributions, medians, and IQRs. To compare videos with a pro vs anti stance, we conducted univariable analyses with χ^2 or Fisher exact tests for categorical and dichotomous variables and used the Mann-Whitney *U* test for continuous variables. All tests were 2-tailed, with statistical significance set at $P<.05$. All analyses were performed using SPSS statistical software (version 30; IBM Corp).

Ethical Considerations

This study was reviewed by the William Paterson University Institutional Review Board ([IRB]; protocol number IRB-FY2026-214) and determined to be non–human subjects research. IRB review and approval was not required under 45 CFR 46.102e.

Results

Results for our analyses of video characteristics are shown in Table 1 and messaging in Table 2. The videos in our sample were viewed collectively over 259 million times. The median number of views per video was 602,000 (IQR 91,250 - 2,525,000) and did not differ by stance (Table 1). Overall, videos garnered a median of 79 (IQR 8 - 419) comments, with anti videos gathering nearly 8 times as many comments as pro videos (395 [IQR 34 - 1164] vs 51 [IQR 5 - 269]; P=.006). Videos were predominantly shared by female-identified content creators (83/100, 83%) and teens or young adults (80/100, 80%). All protanning videos (79/79) were posted by consumers. Postings of pro videos in 2026 were significantly greater than anti videos (61/79, 77.2% vs 10/21, 47.6%; P=0.02). Pro videos were also presented more often by teens and young adults (69/79, 87.3% vs 11/21, 52.4%; P<0.001), were more frequently filmed in an indoor setting (29/79, 36.7% vs 1/21, 4.8%; P=.002), and more often promoted a tanning product with a shopping tag (24/79, 30.4% vs 0/21, 0%; P<.001) compared with anti videos. Pro videos also displayed self-tanner or spray tan more than anti videos (31/79, 39.2% vs 1/21, 4.8%; P=.002). Nearly 4 in 10 protanning videos (39.2%) featured self-tanner or spray tan. Of the 21 antitanning videos, 8 (38.1%) were posted by clinicians and 13 (62%) by consumers; none were posted by public health agencies or other health organizations.

XSL•FO
RenderX

Table . Characteristics of Instagram posts (N=100) related to tanning behaviors by stance.

Video characteristics	Total	Pro (n=79)	Anti (n=21)	*P* value
Views				.08
Median (IQR)	602,000 (91,250-2,525,000)	464,000 (72,600-2,000,000)	1,500,000 (215,000-7,850,000)	
Range	723 - 32,600,000	1949 - 32,600,000	723 - 22,300,000	
Likes				.63
Median (IQR)	8983 (508-56,850)	8371 (459-57,000)	11,600 (2810 - 55,450)	
Range	0 - 583,000	0 - 331,000	5 - 583,000	
Comments				.006
Median (IQR)	79 (8-419)	51 (5-269)	395 (34 - 1164)	
Range	0 - 5641	0 - 4178	2 - 5641	
Shares				.53
Median (IQR)	997 (61 - 9617)	971 (57 - 8424)	1022 (142 - 16,238)	
Range	0 - 879,000	0 - 160,000	0 - 879,000	
Year posted, n (%)				.02
2020	1 (1)	0 (0)	1 (4.8)	
2024	6 (6)	5 (6.3)	1 (4.8)	
2025	22 (22)	13 (16.5)	9 (42.9)	
2026	71 (71)	61 (77.2)	10 (47.6)	
Perceived creator age, n (%)				<.001
Teen	23 (23)	22 (27.8)	1 (4.8)	
Young adult	57 (57)	47 (59.5)	10 (47.6)	
Adult	17 (17)	7 (8.9)	10 (47.6)	
Unclear	3 (3)	3 (3.8)	0 (0)	
Perceived gender, n (%)				.06
Male	16 (16)	9 (11.4)	7 (33.3)	
Female	83 (83)	69 (87.3)	14 (66.7)	
Unclear	1 (1)	1 (1.3)	0 (0)	
Setting, n (%)				.002
Tanning bed	15 (15)	13 (16.5)	2 (9.5)	
Outside, full sun	55 (55)	37 (46.8)	18 (85.7)	
Outside, shade	0 (0)	0 (0)	0 (0)	
Indoors	30 (30)	29 (36.7)	1 (4.8)	
Self-tanning/spray tan, n (%)				.002
Self-tanner	21 (21)	20 (25.3)	1 (4.8)	
Spray tanner	11 (11)	11 (13.9)	0 (0)	
Neither	68 (68)	48 (60.8)	20 (95.2)	
Before/after images, n (%)				.49
Yes	30 (30)	25 (31.6)	5 (23.8)	
No	70 (70)	54 (68.4)	16 (76.2)	
Product promotion, n (%)				<.001

XSL•FO
RenderX

Video characteristics	Total	Pro (n=79)	Anti (n=21)	*P* value
Shopping tag	24 (24)	24 (30.4)	0 (0)	
Verbal/visual only	15 (15)	13 (16.4)	2 (9.5)	
None	61 (61)	42 (53.2)	19 (90.5)	
Uses humor, n (%)				.30
Yes	28 (28)	24 (30.4)	4 (19)	
No	72 (72)	55 (69.6)	17 (81)	
Peer influence, n (%)				.55
Yes	39 (39)	32 (40.5)	7 (33.3)	
No	61 (61)	47 (59.5)	14 (66.7)	

XSL•FO
RenderX

Table . Message content of Instagram posts (N=100) related to tanning behaviors by stance.

Messaging	Total	Pro (n=79)	Anti (n=21)	*P* value
Promotes tanning, n (%)				<.001
Yes	76 (76)	75 (94.9)	1 (4.8)	
No	19 (19)	0 (0)	19 (90.5)	
Ambiguous	5 (5)	4 (5.1)	1 (4.8)	
Considers tanned skin a beauty standard, n (%)				<.001
Yes	63 (63)	58 (73.4)	5 (23.8)	
No	37 (37)	21 (26.6)	16 (76.2)	
Maximizing UV index, n (%)				.51
Yes	15 (15)	11 (13.9)	4 (19)	
No	85 (85)	68 (86.1)	17 (81)	
Tanning beds safe, n (%)				.50
Yes	2 (2)	2 (2.5)	0 (0)	
No	10 (10)	7 (8.9)	3 (14.3)	
Unknown	88 (88)	70 (88.6)	18 (85.7)	
Use of sun protection, n (%)				<.001
Yes	11 (11)	3 (3.8)	8 (38.1)	
No	89 (89)	76 (96.2)	13 (61.9)	
Mental health benefit, n (%)				>.99
Yes	6 (6)	5 (6.3)	1 (4.8)	
No	94 (94)	74 (93.7)	20 (95.2)	
Skin health benefit, n (%)				>.99
Yes	2 (2)	2 (2.5)	0 (0)	
No	98 (98)	77 (97.5)	21 (100)	
Other health benefits, n (%)				>.99
Yes	1 (1)	1 (1.3)	0 (0)	
No	99 (99)	78 (98.7)	21 (100)	
Shrugging off risks, n (%)				.006
Yes	20 (20)	20 (25.3)	0 (0)	
No	80 (80)	59 (74.7)	21 (100)	
Dismisses cancer risk, n (%)				.34
Yes	6 (6)	6 (7.6)	0 (0)	
No	94 (94)	73 (92.4)	21 (100)	

Of the 100 posts analyzed, the messaging focused on promoting tanning (76, 76%) and advocating for tanned skin as a standard of beauty (63, 63%) and downplayed the use of sun protection (11, 11%) and the risks associated with tanning (20, 20%) (Table 2). While protanning videos heavily promoted tanning (75/79, 94.9% vs 1/21, 4.8%; *P*<.001) and considered tanned skin a standard of beauty (58/79, 73.4% vs 5/21, 23.8%; *P*<.001), anti videos more often emphasized the use of sun protection (8/21, 38.1% vs 3/79, 3.8%; *P*<.001) and did not shrug off overall risks associated with tanning and UV exposure (21/21, 100% vs 59/79, 74.7%; *P*=.006). Messages promoting mental health, skin health, or other health benefits of tanning did not differ significantly by stance.

Discussion

Within this sample, Instagram content under the #tanning hashtag disproportionately promoted tanning behaviors without including reliable information about health risks (eg, skin cancer, long-term UV radiation exposure). Sunless tanning is considered

XSL•FO
RenderX

a safer alternative to UV exposure, but these videos still promoted tanned skin as a beauty standard and rarely mentioned sun protection. Sunless tans do not protect against UV damage, and spray tans carry their own risks, such as inhaling the product [12]. Research by Seidenberg and colleagues [13] highlights additional risks with sunless tanning. Nearly 1 in 4 women aged 18 to 49 years used both sunless tanning products and indoor tanning, potentially compounding their skin cancer risk [13]. Additional research is needed on co-occurring tanning practices within this population.

A hashtag that garners over 250 million views is considered a high-performing topic; thus, there were multiple lost opportunities to steer users interested in tanning toward health-promoting behaviors. A 2021 review of articles on social media and skin cancer awareness highlighted the positive role dermatologists play in disseminating reliable health information, indicating increased participation by these health care professionals across social media platforms [14]. Unfortunately, the current study found clinician content creators largely absent. Clinicians created 8/21 (38.1%) of the antitanning videos, while all other videos were created by consumers. Clinician videos reached a sizable audience (median 1.25 million views), but the messages with protanning content they were competing with outnumbered them. Most videos promoted tanning as a beauty standard and were dominated by teens and young adults, the majority of whom did not promote any type of sun protection and sometimes actively shrugged off health risks. This disconnect highlights the need for health care professionals to routinely engage in skin cancer prevention, given the understanding that social media runs on platforms that reach new users continuously. Previous research has demonstrated a disconnect between video popularity and the educational value of short-video platform content hosted on sites such as Instagram. The content produced by clinicians may be of greater quality and indicate credibility, but low interaction metrics with audiences indicate a “quality popularity paradox” [15].

Experts previously outlined the ethical implications of social media algorithms that promote high-risk tanning practices and normalize preventable harm [16]. Adolescents and teens are especially vulnerable to appearance-based messaging, risk-taking, and immediate cosmetic reward [16]. More research is required to explore health information posts that emulate successful content creators’ familiarity with social media algorithms [17]. There are many facets that continue to surface, such as greater availability of information in video titles vs descriptions, emotional intensity of videos, and “drama,” or the ability to stimulate conflict among viewers [17,18]. Maximizing these platforms for health communication requires employing hashtags that demonstrate relevance to the audience’s interest instead of the clinician’s (eg, #tanning vs #skincancer) [19].

Experts previously outlined the ethical implications of social media algorithms that promote high-risk tanning practices and normalize preventable harm [16]. Adolescents and teens are especially vulnerable to appearance-based messaging, risk-taking, and immediate cosmetic reward [16]. More research is required to explore health information posts that emulate successful content creators’ familiarity with social media algorithms [17]. There are many facets that continue to surface, such as greater availability of information in video titles vs descriptions, emotional intensity of videos, and “drama,” or the ability to stimulate conflict among viewers [17,18]. Maximizing these platforms for health communication requires employing hashtags that demonstrate relevance to the audience’s interest instead of the clinician’s (eg, #tanning vs #skincancer) [19].

There is an urgent need for clinicians and organizations promoting sun protective behaviors and products to lead content creation on popular social media platforms. Clinicians are positioned to counter the health-damaging effects of unreliable, incomplete, and potentially misdirected information that is posted under popularized hashtags. Commercial enterprises can compound the effects of this messaging by promoting sun-protective clothing, shade devices, hats, sunscreen, UV detectors, and sunglasses. This is a unique and timely opportunity for collaboration across these sectors to counter the damaging impacts of protanning messaging. The need for public-private partnerships may address the null effect some public health practitioners have experienced in implementing tanning interventions via social media platforms [3]. Our study findings support the need for more sustained skin cancer prevention campaigns targeting tanning behaviors [20].

Funding

This research received no external funding.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

CHB and TG designed the study and BG collected the data. GCH analyzed the data. All authors contributed to writing and review of the manuscript.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Gaston J, Daniel BS. Social media and the re-emergence of dangerous tanning practices. Australas J Dermatol 2025 Dec;66(8):497-498. [doi: 10.1111/ajd.14617] [Medline: 41170904]
2. Jiménez-Sánchez L, Fuentes C, Subert A, et al. Sunburns, tanning, and social media: a descriptive study among Spanish university students. Actas Dermosifiliogr 2026 Jun 24;117(10):104715. [doi: 10.1016/j.ad.2026.104715] [Medline: 42342228]
3. de Vere Hunt I, Cai ZR, Nava V, et al. A Social Media-Based Public Health Campaign to Reduce Indoor Tanning in High-Risk Populations. AJPM Focus 2023 Sep;2(3):ne. [doi: 10.1016/j.focus.2023.100123] [Medline: 37662553]
4. Matos TR. Tanning, trending and tumours: can social media and legislation save us from ourselves? Br J Dermatol 2025 Nov 18;193(6):1233-1234. [doi: 10.1093/bjd/ljaf313] [Medline: 40795808]
5. American Academy of Dermatology survey shows Gen Z adults at risk for skin cancer due to increasing rates of tanning and burning. American Academy of Dermatology Association. 2024. URL: https://www.aad.org/news/gen-z-adults-at-risk-skin-cancer [accessed 2026-08-18]
6. Misinformation puts over 16 million Americans at an increased risk for skin cancer. American Academy of Dermatology Association. 2026. URL: https://www.aad.org/news/misinformation-increases-skin-cancer-risk-for-16-million-americans [accessed 2026-10-22]
7. Garbe C, Forsea AM, Amaral T, et al. Skin cancers are the most frequent cancers in fair-skinned populations, but we can prevent them. Eur J Cancer 2024 Jun;204:114074. [doi: 10.1016/j.ejca.2024.114074] [Medline: 38691877]
8. O'Sullivan DE, Brenner DR, Demers PA, et al. Indoor tanning and skin cancer in Canada: a meta-analysis and attributable burden estimation. Cancer Epidemiol 2019 Apr;59:1-7. [doi: 10.1016/j.canep.2019.01.004] [Medline: 30639817]
9. Schulman JM, Fisher DE. Indoor ultraviolet tanning and skin cancer: health risks and opportunities. Curr Opin Oncol 2009 Mar;21(2):144-149. [doi: 10.1097/CCO.0b013e3283252fc5] [Medline: 19532016]
10. Passeron T, Lim HW, Goh C‐, et al. Photoprotection according to skin phototype and dermatoses: practical recommendations from an expert panel. J Eur Acad Dermatol Venereol 2021 Jul;35(7):1460-1469. [doi: 10.1111/jdv.17242]
11. Heckman CJ, Handorf EA, Mitarotondo A, et al. Digital skin cancer risk reduction interventions for young adults: findings from a hybrid type-II effectiveness-implementation trial. Cancer Epidemiol Biomarkers Prev 2025 Jun 3;34(6):962-971. [doi: 10.1158/1055-9965.EPI-24-1636] [Medline: 40131334]
12. Tanning products. US Food and Drug Administration. 2019 Apr 26. URL: https://www.fda.gov/radiation-emitting-products/tanning/tanning-products [accessed 2026-10-05]
13. Seidenberg AB, Julian AK, Hartman AM, Holman DM. Sunless and indoor tanning among U.S. non-Hispanic White women ages 18-49 years. J Dermatol Nurses Assoc 2023;15(3):123-132. [doi: 10.1097/jdn.0000000000000736] [Medline: 38463180]
14. De La Garza H, Maymone MBC, Vashi NA. Impact of social media on skin cancer prevention. Int J Environ Res Public Health 2021 May 9;18(9):5002. [doi: 10.3390/ijerph18095002] [Medline: 34065061]
15. Mao Q, Zhong L, Wang S. Assessment of quality and reliability of Chinese-language myocardial infarction content on Bilibili and TikTok: a cross-sectional study. Front Public Health 2026;14:1748168. [doi: 10.3389/fpubh.2026.1748168] [Medline: 41938936]
16. Fleshner L, Albright C, Klufas T, Grant-Kels JM. Ethical implications of "tanmaxxing" promotion on social media. Clin Dermatol 2026;44(5):862-864. [doi: 10.1016/j.clindermatol.2026.08.001] [Medline: 42567213]
17. Florini S, Watson TX, O'Neil E, Grumbach E, Gray C. Algorithmic folk knowledge: content creators' understandings of TikTok's algorithms. First Monday 2026 Apr 7. [doi: 10.5210/fm.v31i4.14362]
18. Tafesse W. YouTube marketing: how marketers' video optimization practices influence video views. Internet Res 2020 Jul 3;30(6):1689-1707. [doi: 10.1108/INTR-10-2019-0406]
19. Chakrabarti P, Malvi E, Bansal S, Kumar N. Hashtag recommendation for enhancing the popularity of social media posts. Soc Netw Anal Min 2023;13(1):21. [doi: 10.1007/s13278-023-01024-9] [Medline: 36686375]
20. Buller DB, Pagoto S, Henry KL, et al. Persisting effects of a social media campaign to prevent indoor tanning: a randomized trial. Cancer Epidemiol Biomarkers Prev 2022 Apr 1;31(4):885-892. [doi: 10.1158/1055-9965.EPI-21-0059] [Medline: 35064063]

Edited by R Dellavalle; submitted 21.Aug.2026; peer-reviewed by B Walkosz, M Buller; revised version received 07.Oct.2026; accepted 13.Oct.2026; published 30.Oct.2026.

Please cite as:
Gonzalez T, Hillyer GC, Gold B, Basch CH
Promotion of Tanning Behaviors on Instagram: Content Analysis
JMIR Dermatol 2026;9:e110157
URL: https://derma.jmir.org/2026/1/e110157
doi:10.2196/110157

XSL•FO
RenderX

© Tailisha Gonzalez, Grace C Hillyer, Bailey Gold, Corey H Basch. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of TikTok in Education on Hidradenitis Suppurativa in Skin of Color: Cross-Sectional Analysis

Arsema K Zadu*, MS; Jordan Young*, BS; Janyla A Seltzer*, MD; Angel S Byrd*, MD, PhD; Cheri Frey*, MD

Department of Dermatology, College of Medicine, Howard University, 2041 Georgia Ave NW, Washington, DC, United States

*all authors contributed equally

Corresponding Author:
Arsema K Zadu, MS
Department of Dermatology, College of Medicine, Howard University, 2041 Georgia Ave NW, Washington, DC, United States

Abstract

Abstract: This study analyzed 50 TikTok videos returned by a search for "hidradenitis suppurativa in Black skin," revealing that nearly half were patient-created, few had physician involvement (n=10, 20% dermatologists; n=7, 14% plastic surgeons), and few had commission-based (n=7, 14%) or sponsored content (n=2, 4%); they were predominantly patient testimonials on various treatments, highlighting the need for greater physician engagement to address patient needs, hidradenitis suppurativa product safety, and efficacy.

(JMIR Dermatol 2026;9:e71566) doi:10.2196/71566

KEYWORDS

TikTok; skin of color; hidradenitis suppurativa; patient education; social media

Introduction

TikTok has become a prominent platform and search engine for dermatologic information [1]. A majority of videos on the top 20 most commonly diagnosed skin conditions were created by patients [2]. This study aimed to evaluate the primary sources of education on hidradenitis suppurativa (HS) in Black skin on TikTok.

HS is a chronic inflammatory condition marked by sinus tracts, nodules, and abscesses in intertriginous areas [3]. It often affects African American women and is frequently misdiagnosed, leading to delayed identification, accelerated progression, and treatment challenges [4]. Black patients experience a 1.5-year longer diagnostic delay than White patients, and after diagnosis, they typically wait 5 years to see a dermatologist compared to 3 years for White patients [5]. Early intervention is critical to prevent advanced-stage manifestations, which are often refractory to standard therapeutic approaches [6].

A study on dermatologist visibility on TikTok found that most dermatologic education videos are made by individuals without formal medical training [7]. While these platforms allow patients to connect and share experiences with skin conditions, the risk of spreading misinformation is significant, potentially worsening disease severity and reducing treatment effectiveness. This study aimed to describe who produces HS TikTok content featuring Black skin and to quantify the treatments, products, and themes presented.

Methods

TikTok was selected because users often turn to the platform as a general search tool, including for health topics related to HS [1]. The app was used to search for "hidradenitis suppurativa in black skin" in the search bar. A total of 50 videos were viewed one by one by a single reviewer to assess the creator's role, brands or products discussed, key themes, and whether the products were part of paid sponsorships or if the creator received commission from sales.

Ethical Considerations

All data were deidentified prior to analysis. User IDs, screenshots, images, and quoted content contained no personally identifiable information. Data were analyzed and reported in anonymized form only.

Results

Of the 50 analyzed videos, 24 (48%) were created by patients. Board-certified dermatologists produced 10 (20%) videos, while board-certified plastic surgeons produced 7 (14%) videos. One (2%) video was created by a nurse practitioner. Beauty service providers accounted for 2 (4%) videos. The "other" category, which included dietitians, social workers, product creators, and creators with unclear professional titles, accounted for 6 (12%) videos. Board certification status was verified using creator profile biographies and linked professional websites.

Seven videos featured products associated with sales commissions, and 2 videos involved paid sponsorships.

XSL•FO
RenderX

Treatment content represented the largest category, comprising 35 (70%) videos (Table 1). Fourteen (28%) of the 50 videos focused on explaining HS. Of these, 9 (18%) were created by health care professionals, 4 (12%) by patients, and 1 by a product creator. The remaining videos focused on living with HS and dietary approaches to symptom management.

Table . Video themes stratified by creator type across 50 analyzed videos.

	Video type, n				
Creator type	Education	Medical treatment	Over-the-counter treatment	Lifestyle	Mental health
Patient	4	4	14	4	3
Board-certified dermatologist	7	3	5	0	0
Board-certified plastic surgeon	1	2	3	0	0
Nurse practitioner	1	0	0	0	0
Beauty service provider	0	1	1	0	0
Other	1	0	3	0	0

Products appeared across multiple categories, including face masks and exfoliants, body scrubs and cleansers, acne treatments, topical oils, salves, body butters, creams, antiseptics, spot treatments, systemic medications, and hair or scalp care, as shown in Table 2. PanOxyl and Hibiclens washes were the most frequently recommended products, appearing in 5 (10%) and 6 (12%) videos, respectively. Magic Healer body butter was mentioned in 4 (8%) videos, and Humira was mentioned in 3 (6%) videos. CeraVe acne foaming cream cleanser, a turmeric kojic acid soap, and Tend Skin Solution were all separately mentioned in 2 (4%) videos. All other products were mentioned in one video each.

Table . Categories of products mentioned across 50 analyzed hidradenitis suppurativa–related videos.

Category	Products
Face masks and exfoliants	Skinfix Glycolic Renewing Mask, Cocokind Turmeric Mask
Body scrubs and cleansers	Skinfix body scrub, Olay body wash, Naturium Vitamin C body wash, lemon turmeric and kojic acid soap, Dr. Bronner's Soap, turmeric kojic acid soap, Dial bar soap
Acne cleansers and treatments	PanOxyl, PanOxyl acne foaming wash, Inkey List 5 % Benzoyl Peroxide Cleanser, CeraVe acne foaming cream cleanser, Zapzyt, Tend Skin Solution
Topical oils and salves	Relief Natural Company, GuruNanda Tea Tree, Zunda Turmeric, black seed oil, vitamin E, clove water, Magic Healer body butter, Palmer's Body Oil
Body butters, creams, and moisturizers	HS Body Butter, Fenty Skin Cream, Healing Ocean Cream
Antiseptics and antimicrobials	Hibiclens, Magic Healer Product
Specialty treatments and patches	Mighty Patch
Medications	Humira
Hair and scalp care	Head & Shoulders

Discussion

Patients produced the majority of lifestyle-based and experiential content, while physicians concentrated on treatment education and procedural intervention. Commercial messaging remained isolated to product owners and patient testimonials, with minimal clinical creator involvement. This distribution highlights a content gap between medically accurate education and the lived experience narrative dominating public HS discourse on TikTok. Research shows that only about 20% of skin of color videos are created by board-certified dermatologists [8]. These findings highlight an opportunity for dermatologists on TikTok to engage with patients and address the safety and efficacy of these products. Dermatologists, who already have a strong social media presence, can foster patient trust by encouraging open discussions about non–medical-grade treatments. By acknowledging the value of herbal and alternative remedies that patients find helpful, they can assess their safety and efficacy while supporting their continued use when appropriate. Meanwhile, they can provide evidence-based guidance to minimize the risks of harmful or ineffective treatments.

XSL•FO
RenderX

Acknowledgments

AKZ is a recipient of the 2024-2025 Women of Dermatology Society and La-Roche Posey Diversity in Dermatology Fellowship at Howard University.

Conflicts of Interest

ASB is the inaugural recipient of the Skin of Color Society Career Development Award as well as the Society for Investigative Dermatology Freinkel Diversity Fellowship Award, and a recipient of the Robert A. Winn Excellence in Clinical Trials Career Development Award (Winn CDA) funded by Bristol Myers Squibb Foundation (BMSF); consultant for SENTÉ, Inc. and Sonoma Biotherapeutics; Advisory boards for Merck and Novartis. CF is a Primary Investigator for Novartis. All other authors declare no conflicts of interest.

References

1. Zadu AK, Adibe I, Welch D, et al. Social media as a catalyst for health education and early diagnosis of hidradenitis suppurativa. J Drugs Dermatol 2025 Jun 1;24(6):e45. [Medline: 40465496]
2. Villa-Ruiz C, Kassamali B, Mazori DR, Min M, Cobos G, LaChance A. Overview of TikTok's most viewed dermatologic content and assessment of its reliability. J Am Acad Dermatol 2021 Jul;85(1):273-274. [doi: 10.1016/j.jaad.2020.12.028] [Medline: 33359080]
3. van Straalen KR, Prens EP, Gudjonsson JE. Insights into hidradenitis suppurativa. J Allergy Clin Immunol 2022 Apr;149(4):1150-1161. [doi: 10.1016/j.jaci.2022.02.003] [Medline: 35189127]
4. Zouboulis CC, Goyal M, Byrd AS. Hidradenitis suppurativa in skin of colour. Exp Dermatol 2021 Jun;30 Suppl 1(S1):27-30. [doi: 10.1111/exd.14341] [Medline: 34085333]
5. Serrano L, Ulschmid C, Szabo A, Roth G, Sokumbi O. Racial disparities of delay in diagnosis and dermatologic care for hidradenitis suppurativa. J Natl Med Assoc 2022 Dec;114(6):613-616. [doi: 10.1016/j.jnma.2022.08.002] [Medline: 36511276]
6. Andersen RK, Jemec GBE. Treatments for hidradenitis suppurativa. Clin Dermatol 2017;35(2):218-224. [doi: 10.1016/j.clindermatol.2016.10.018] [Medline: 28274363]
7. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
8. Campbell JR, Woolery-Lloyd H, Williams K. TikTok and black skin: is this a missed opportunity for dermatologists? J Drugs Dermatol 2023 Jul 1;22(7):698-700. [doi: 10.36849/JDD.7061] [Medline: 37410044]

Abbreviations

HS: hidradenitis suppurativa

Edited by J Meisenheimer; submitted 26.Jan.2025; peer-reviewed by ED Quincey, G Padilla-Castillo, S Bestavros; revised version received 15.Dec.2025; accepted 15.Dec.2025; published 27.Feb.2026.

Please cite as:
Zadu AK, Young J, Seltzer JA, Byrd AS, Frey C
The Role of TikTok in Education on Hidradenitis Suppurativa in Skin of Color: Cross-Sectional Analysis
JMIR Dermatol 2026;9:e71566
URL: https://derma.jmir.org/2026/1/e71566
doi:10.2196/71566

© Arsema K Zadu, Jordan Young, Janyla A Seltzer, Angel S Byrd, Cheri Frey. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Skin Health Information Seeking on Short Video Platforms in Indonesia: Mixed Methods Approach

Fathi Qushoyyi Ahimsa*, SKom; Kelvin Saputra*, SKom; Michelle Elizabeth Amanda Hutasoit*, SKom; Putu Wuri Handayani*, Prof Dr; Hana Fitriani*, MSc

Faculty of Computer Science, Universitas Indonesia, Jalan Kampus, Depok, Indonesia

*all authors contributed equally

Corresponding Author:
Putu Wuri Handayani, Prof Dr
Faculty of Computer Science, Universitas Indonesia, Jalan Kampus, Depok, Indonesia

Abstract

Background: Many users have now switched to using short video platforms as the main channel in their search for skin health information. With high internet penetration and a large market for the skincare industry, short video platforms play an important role in the "beauty discovery" process and purchase decisions. However, the increasing consumption of skin health content is also accompanied by the risk of misinformation, uneven content quality, and the dominance of creators who are not health professionals. Therefore, it is important to determine what factors affect the use of short video platforms in the search for skin health information.

Objective: By adopting the stimulus-organism-response framework, health belief model, and media richness theory, this study aims to analyze the factors that influence the use of short video platforms in the search for skin health information.

Methods: This study used a mixed methods approach by distributing an online survey to 603 respondents and conducting interviews with 30 interviewees. Survey data were analyzed using the covariance-based structural equation modeling method, and qualitative data were analyzed using the thematic analysis method.

Results: The results of this study found that perceived usefulness (P=.01), attitude (P=.02), perceived severity (P=.009), and perceived susceptibility (P=.02) directly affected the behavior of seeking skin health information on short video platforms. Health content expressiveness (P=.001) and personalized health insights (P<.001) directly affected perceived usefulness. These findings support the media richness theory, which shows that the expressiveness of health content and personalization can increase the perception of the usability of short video platforms. Perceived interactivity has an influence on attitude (P<.001), which then affects skin health information seeking on short video platforms (SHEs; P=.02). Upward skin comparison also had a direct influence on skin stigmatization (P<.001). Moreover, perceived severity (P=.009) and perceived susceptibility (P=.02) have an effect on SHEs. These findings confirm the health belief model's theory that perception of the severity of a skin problem and the perception of a person's likelihood of developing a skin problem can improve skin health seeking behavior. However, no effect of source credibility (P=.17) and skin stigmatization (P=.30) was found on SHEs. This is due to the user's willingness to exchange aspects of trust in information sources such as the functionality of the platform and familiarity with the platform or the existence of other internal cognitive or psychological aspects that can be investigated in the future.

Conclusions: This study can provide guidance for the development of more effective health communication strategies in the digital era using short video platforms.

(JMIR Dermatol 2026;9:e93461) doi:10.2196/93461

KEYWORDS

skin; health information; information seeking; short video platform; social media; Indonesia

Introduction

Background

With the advent of social media, access to health information has increased, with more than 55% of adults using social media to seek health information and advice [1]. Vranica and Kruppa [2] found that about 23% of social media users used short video platforms, specifically TikTok. TikTok handles a high search activity, which reached 1.2 million search requests per minute in 2024 [3]. Liu et al [4] explained that the short video-based social media format is most popular for users conducting health information searches because short videos can help convey complex health information in a short time with a combination of visual, audio, and narrative that is easy to understand.

Short video platforms can present information in various forms of media, such as text (captions), images or photos, audio

XSL•FO
RenderX

(podcasts), and videos, both long-duration videos and short-duration videos. This multimedia format allows users to see visualizations of the disease they are experiencing so that it is easier to recognize early symptoms and compare their condition with visual examples in videos or photos for reference [5]. The format of information media in the search for health information can be explained through media richness theory (MRT). MRT explains that "richer" media have an effective ability to explain ambiguous issues [6]. Short video platforms can combine visual, audio, text, and live demonstration elements in a short duration. This allows complex health information to be conveyed well and improves audience understanding quickly. In addition to the media aspect, several other factors affect the behavior of searching for health information on short video platforms, such as personalization and interactivity.

Han et al [7] found that TikTok's algorithm forms an information space that facilitates the discovery of relevant content that encourages exposure and further searches on the platform. This algorithm serves relevant content based on interactions, such as likes and watch time, so that users can find health content without conducting an explicit search [7]. The behavior of seeking health information on short video platforms can also be understood from a psychological perspective, one of which is through the theory of health belief model (HBM). HBM explains that an individual's decision to engage in the search for health information is influenced by perceived susceptibility (PSU) and perceived severity (PSE) to a health condition [8]. The higher a person's risk perception, the greater their tendency to seek information through digital platforms, including social media [9,10]. Priyambodo et al [11] showed that users tend to seek explanations when they are worried about the potential risks they could face. This pattern is visible in the realm of skin health, which is visual and easy to observe. Concerns about acne, dark spots, eczema, or other changes in skin conditions make short video platforms the main choice because they are able to present visual demonstrations, real examples, and personal experiences in person [12,13].

As of 2025, Indonesia has a population of 284.4 million people, with internet penetration reaching 80.66% [14]. The most widely used short video platforms in the country are TikTok, Instagram Reels, and YouTube Shorts [14]. With this high level of social media penetration, Thang et al [15] found that many users accessed skin health information through short video platforms. Short video platforms are a popular source used to search for or share dermatological or skincare information, with high average views per video of 1.6 million views on TikTok alone [15]. Thus, social media plays a major role in the beauty discovery phenomenon. Beauty discovery is the process of searching for information or definitions related to beauty, such as information about products, personal care techniques, and beauty standards.

Skin Health Information Seeking on Social Media

Although they have many uses, short video platforms also have the risk of presenting disinformation. According to Zhang et al [16], the dissemination of information by lay users can lead to widespread misinformation on social media. Other research has shown that the quality of health videos on these platforms is often low. For example, Ming et al [17] found that the proportion of misinformation in certain disease videos on TikTok could reach 41% - 78% of the total content. In their study, Kirkpatrick and Lawrie [18] found that almost all respondents believed that health misinformation was common on TikTok. The absence of special quality evaluation standards for health information on short video platforms is another obstacle to finding health information on short video platforms. The risk of adopting information without further information verification is further exacerbated by the dominance of beauty content in Indonesia, where the representation of bright and smooth skin is a visual standard widely reproduced by content creators [19]. Ariana et al [20] showed that using TikTok significantly affected users' body image satisfaction. This phenomenon further emphasizes the strong influence of social media on shaping public perception of the definition of an "ideal" body and skin, especially among young audiences who are active in consuming short video content.

Several studies have been conducted on the quality of skin health information content shown on short video platforms. Thang et al [15] discussed acne management information on various short video-based social media platforms and showed that the quality of education varies and does not always correlate with popularity. Khan et al [21] analyzed 120 videos on TikTok related to eczema and found that 65.8% of the content was produced by individuals who did not have a professional background in the health field, while only 15.8% of the content was produced by individuals with a health background. Thang et al [15] found that the quality of content produced by certified dermatologists was significantly higher than that of individual content that did not have a health background. However, videos of better quality made by doctors actually received lower engagement [15]. In addition, several studies have shown misperceptions of information circulating on short video platforms, especially regarding drugs and skin health procedures. Khan et al [21] found many misperceptions about the use of tretinoin, especially exaggerated claims regarding its effectiveness and recommendations for use without medical supervision. A similar pattern can be observed in the content related to derma rolling procedures. For example, Awad et al [22] found that most videos provided recommendations that were not aligned with clinical practice and tended to promise instant results. In general, the most viral videos were those that were demonstrative, emotional, and visually appealing but were not based on strong scientific evidence.

Objectives

Previous studies have examined the behavior of searching for health information on social media in general [9,11,15,23,24]. However, previous studies have only discussed content quality with regard to the credibility of actual information [21,22,25]. Further research exploring the user behavior of social media platforms with short video formats is needed because skin problems are often related to aesthetic factors, social stigma, and comparison of appearance [26]. Thus, the research question of this study is, "What factors affect the use of short video platforms in searching for skin health information?" The findings of this study can be used by the government, academics,

XSL•FO
RenderX

and educational institutions to develop digital literacy and public health programs.

Conceptual Model

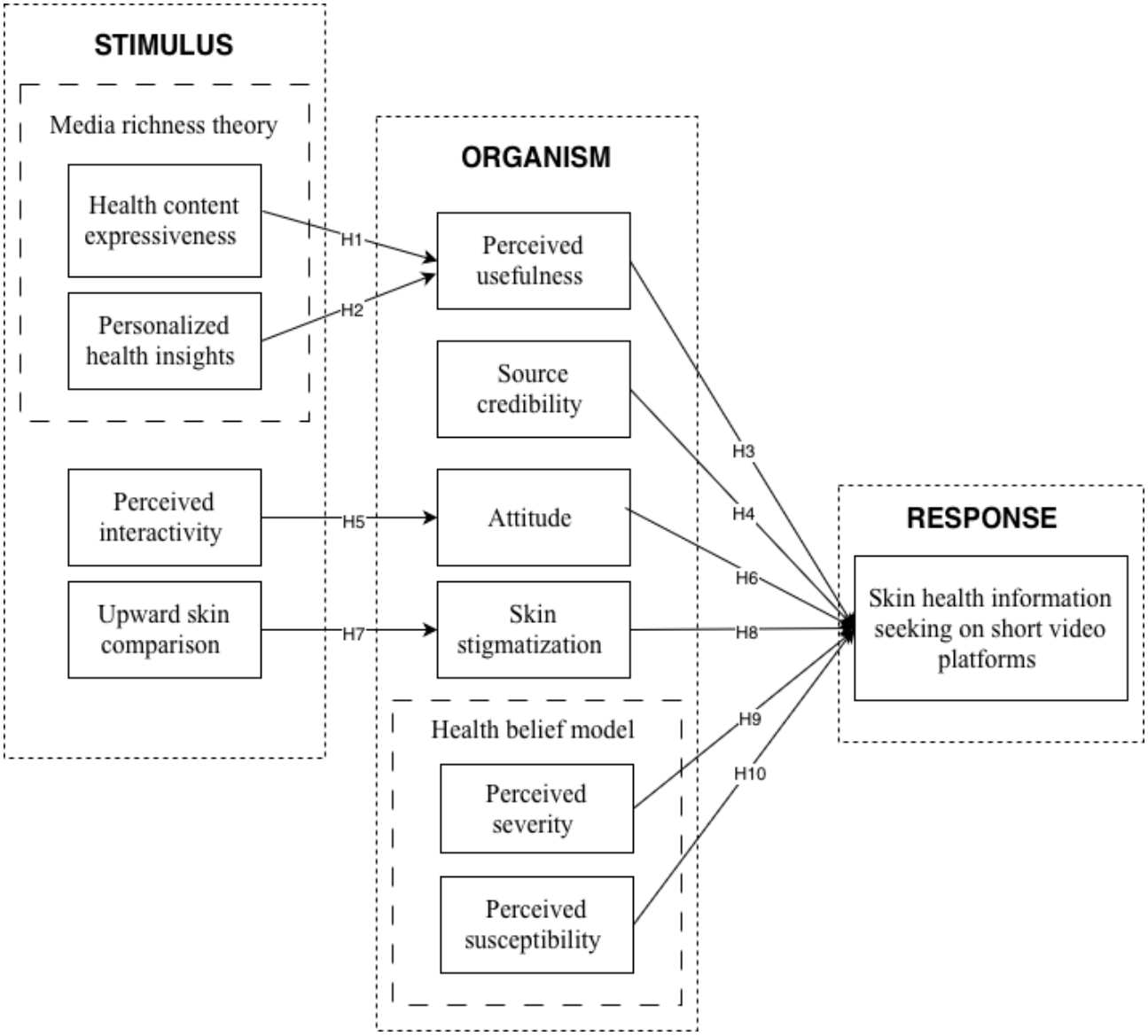
Overview

Brzozowska and Gotlib [27] found that positive engagement develops in response to social media, facilitating users' change behavior on skin health promotion. The stimulus-organism-response (SOR) framework has been used in the health-related studies such as the review study on the impact of social media on skin health promotion [27], analyzing the impact of TikTok on body satisfaction [20], investigating the social media's impact on sugar reduction behavior [28], and the role of beauty influencer in the digital platforms on impulse buying [29]. Thus, this study adopts the SOR framework to develop a deeper understanding of stimuli originating from short video platforms. Cognitive and perceptual constructs are conceptualized as organism-level processes, while skin health information–seeking behavior serves as the response. Within this framework, stimuli refer to external environmental factors that provoke user responses, whereas the organism reflects users' internal cognitive processes [29]. User response is defined as the resulting behavioral reaction that emerges from the interaction between these external stimuli and the organism's internal processes [29]. Therefore, MRT [23] and HBM [9] were adopted in this study to analyze the stimulus and organisms' internal processes from the social media users for skin health information searching on short video platforms. Grounded in SOR framework, we integrated MRT (eg, content expressiveness and personalization) as stimulus to explain the compatibility of the characteristics of short video media with the visual nature of skin health information, while HBM (eg, perceived usefulness [PU], severity, and susceptibility) as organism-level processes to understand the perceptual and motivational factors of individuals that drive the search for skin health information on short video platforms.

Yu et al [23] found that media richness affects the PU of a platform, which increases the adoption of health information on the platform and the continued use of the platform in seeking health information. Health content expressiveness (HCE) and personalized health insights (PHIs) were found to have a significant influence on PU [23]. Niu et al [30] revealed that interactivity positively influenced users' attitudes toward skin health web pages. In addition, Adkins et al [26] found that upward appearance comparisons had a positive influence on feelings of stigmatization in the use of photo-based or visual applications. Malik et al [9] discovered that individual perceptions of a health risk affect users' desire to seek health information. Figure 1 illustrates the research model proposed in this study.

XSL•FO
RenderX

Figure 1. Conceptual model. H: hypothesis.



Hypothesis Development

- Hypothesis 1: HCE positively affects PU

HCE is one of the features that contribute to media richness in MRT, which has an influence on media usefulness [23]. Based on MRT, in the context of media richness, HCE is a component that fulfills the aspect of multiple cues, which refers to the use of various communication channels [23]. With increasing media richness, the ambiguity of information will decrease and encourage the positive behavior of information recipients [31]. According to MRT, through various media, such as visual, audio, and text, content expressiveness can enrich the user experience and improve users' ability to process complex health information [32]. The combination of these various media enables users to have different ways of processing and understanding information [23]. In the context of this study, HCE describes the ability of short video platforms to provide a variety of media, such as visual, audio, and text, which can be used by their users in presenting skin health information. Lenczowski et al [33] found that skin health education in the form of videos was more interesting than written materials. In addition to text, visual and audio elements are important in conveying skin health information online, as skin conditions are visual, and there are many treatment skills, such as self-examination, use of sunscreen, and preparation for actions, that require both visual and verbal demonstrations. If short video platforms can provide all the media needed, then the ability of users to process and adopt the required skin health information will increase as well as increasing user perception of the usefulness of short video platforms in seeking skin health information.

- Hypothesis 2: PHIs positively affect PU

Based on MRT, PHIs are components that meet the aspect of personal focus, which refers to the level of suitability of communication with individual needs [23]. By tailoring recommended health information to the user's preferences, needs, and context, PHIs can increase the practical value of content, as users perceive the platform to address users' unique health challenges [23]. In the context of this study, PHIs describe the ability of a short video platform to display skin health information videos that fit the user's preferences, needs, and

XSL•FO
RenderX

context. In addition, personalization has a significant positive influence on PU in the context of mobile health [34]. Personalization can be conducted by using demographic data, health issues, or user interaction history on a platform [23]. Nittas et al [35] found that personalization on educational platforms regarding the use of sunscreen was well received and considered useful by participants. With a recommendation algorithm that captures and learns every action from a user, such as likes, follows, comments, favorites, and input users against the recommendation results, short video platforms have succeeded in increasing users' use of the platform [36]. Based on MRT, if the short video platform can provide recommendations or display skin health information videos tailored to users' preferences, needs, and contexts, then users will feel that the short video platform can answer their skin health problems [23]. Personalization is considered an added value in the search for skin health information because people's skin conditions vary; thus, common skin diseases, such as acne or psoriasis, can have different treatment responses [37]. The proper treatment of a skin condition can require a long trial-and-error process and can frustrate patients; thus, the personalization achieved through virtual experiments increases users' adoption desires to search skin health information videos on short video platforms, hence, increasing PU of the short video platforms [37].

- Hypothesis 3: PU positively affects skin health information seeking on short video platforms (SHEs)

PU is defined as the level of user confidence that the use of an information system can improve user performance or provide real benefits [23]. Liu et al [4] define PU as a user's subjective assessment of the value of a short video in guiding users regarding health behaviors. Xia et al [38] found that PU had a significant positive influence on users' desire to search for health information online, even exceeding the influence of perceived use. In the context of this study, PU describes the user's perception of the extent to which the short video platform can help users to search and disseminate skin health information. Boyce et al [39] also found that PU had a significant positive influence on health information–seeking behavior in the online health community. In looking for skin health information, one aspect that plays an important role is users' perceptions of the benefits of a platform in finding relevant and applicable information to solve problems or user needs [40]. With features that make the search process more efficient and relevant, the perception of a platform's usability can increase [35]. This increased perception of platform usefulness can then encourage users to use a platform to search for skin health information [39].

- Hypothesis 4: SC positively affects SHE

Source credibility (SC) is defined as the perception of the recipient of information regarding the credibility of the source of knowledge or informant [41]. It is a peripheral cue or precursor used before the user decides to engage further with the information provided [42]. In their study, Claggett et al [42] found that users rated the informant as credible and trustworthy when there was a health degree listed on the informant's name. In seeking health information, the level of credibility of informants was found to have a significant positive influence on the frequency of use of a source to seek information [43]. According to Fan et al [43], sources that are considered credible are more often used by individuals when searching for health information. In the context of short videos such as TikTok, Kirkpatrick and Lawrie [18] found that whistleblowers with professional credentials or markers of affiliation with relevant agencies increased the tendency and frequency of user information searches against those whistleblowers. In the context of skin health, this is because skin health products that are not used appropriately, such as frequency and duration of use, can have side effects on users' skin. Therefore, the credibility of skin health informants is important to ensure that users are educated with credible information so as to avoid adverse or even harmful side effects; thus, the credibility of the information source can increase the user's desire to search for information on the source.

- Hypothesis 5: PI positively affects attitude

Perceived interactivity (PI) is defined as the user's perception of the extent to which the user feels that their experience of using an application resembles interpersonal interaction and provides a sense of social connection [4]. McMillan and Hwang [44] found that PI is influenced by 3 main factors, namely, 2-way communication, user control, and speed or responsiveness. In the context of this study, PI describes the ability of short video platforms to make it easier for their users to interact. Interactions that can be carried out by users do not only include 2-way communication but also include 1-way interactions, such as likes and comments, that allow users to control the videos displayed based on their interactions. Bidirectional communication and user control intersect in active or passive participation factors, user control and speed or responsiveness intersect in interesting or complex activity factors, and speed or responsiveness and bidirectional communication intersect in synchronous or asynchronous communication factors [44]. On TikTok, interactivity was found to be one of the factors contributing to the flow of application use and increasing the desire of users to continue using the application [18]. The interactions made on TikTok, such as likes, comments, and watching content of a longer duration, affect the content recommended by TikTok. Recommended content is tailored to the interaction carried out; thus, user interaction on short video platforms can improve the personalization of the content displayed according to the users' interests. In the context of skin health, personalization is an added value due to different skin conditions [37]. Therefore, when users can have 1-way or 2-way interactions and feel the platform's quick or appropriate response to these interactions, their attitude toward a platform tends to improve [45].

- Hypothesis 6: attitude positively affects SHE

Attitude is defined as an individual's positive or negative feelings toward the use of short videos for health information [4]. Link et al [46] show that attitudes toward seeking health information online, along with risk perception, are the main factors driving individuals' health information–seeking behavior. Mirzaei et al [47] also found that attitudes represented by user experience are one of the significant predictors of health

XSL•FO
RenderX

information search in digital contexts. In the context of this study, attitude describes the user's feelings toward short video platforms to seek skin health information. In this study, users' assessments of their experience were influenced by users' perceptions of the interactivity of short video platforms. Li et al [48] found that users' positive perceptions and attitudes toward health information are associated with increased health information search behavior in digital services. In the context of social media, Neely et al [49] found that user engagement on social media platforms as well as user attitudes are the driving factors for users' intentions and behaviors to search for health information on social media platforms. Specifically, on short video platforms, Kirkpatrick and Lawrie [18] found that the intention and behavior of obtaining health information on TikTok had a high percentage, reaching 65.5% (n=672) for users who intentionally searched for health information and 92.4% (n=948) for those who accidentally searched for it. Kirkpatrick and Lawrie [18] also discovered that users have a positive perception of TikTok in the context of searching for health information. Users' positive attitudes toward online health information searches are associated with higher online health information search behaviors [46]. When users positively rate their experience of searching for health information on short video platforms, users tend to increase the frequency and depth of searching for skin health information through the platform.

- Hypothesis 7: Upward skin comparison (USC) positively affects skin stigmatization (SST)

Adkins et al [26] describe USC as an individual's tendency to compare the appearance of his or her skin to others who have better skin. USC also has a positive correlation with stigmatization in adults with acne-prone skin conditions [26]. USC is a core process that exacerbates the stigma felt by individuals [26]. Tian et al [50] also found that upward social comparisons on social media increase an individual's anxiety about his or her appearance, which is closely related to the individual's stigma and psychosocial burden. In the context of dermatology, Van Beugen et al [51] found that patients with skin diseases that can be seen without aids experienced higher stigmatization than other patients. This suggests that skin conditions that can be seen without aids tend to encourage users to compare themselves, thus creating a psychosocial burden. In the context of this study, USC is defined as an individual's tendency to compare the condition or appearance of their skin with other individuals who are rated as having "better" skin. This tendency is not only limited when individuals use short video platforms only but also the tendency to compare in everyday life. However, the use of social media, especially short video platforms with visual dominance and "ideal skin" content, also contributes by increasing the possibility for individuals to compare their skin conditions with those of other creators or users [26]. Adkins et al [26] found that individuals who compared their skin conditions with other people's skin conditions had higher feelings of stigmatization. Tunca et al [52] found that patients with acne vulgaris disease have significant physical manifestations and emotional and social implications. The disease can increase feelings of guilt, shame, social anxiety, and social isolation [52]. Tunca et al [52] also found that patients with acne often found themselves less attractive, could experience withdrawal behaviors, and avoided social gatherings for fear of being judged.

- Hypothesis 8: SST positively affects SHE

SST is defined as self-perception and anticipation of skin-related social judgments associated with an individual's internalized negative feelings due to certain skin conditions, such as acne [26]. Berger et al [53] revealed that individuals with conditions who experienced stigmatization because of their disease were found to use the internet more often to seek health information than individuals who did not experience stigma. In the context of skin health, Aslan Kayıran et al [54] found that patients with psoriasis routinely use social media and the internet to find information about their disease. This further strengthens the finding that digital platforms are one of the digital channels used by individuals to seek skin health information. Nguyen et al [55] discovered that social media platforms are a relevant and widely used dermatology reference source for patients seeking skin health information. This is due to the accessibility of information, the ease of use of social media platforms, and the social characteristics of the platform [56]. In the context of this study, SST is defined as individuals' negative perceptions of themselves caused by their skin conditions. Negative perceptions or the way individuals judge themselves are not only limited to the clinical severity of the skin condition but also include the anticipation of rejection, sensitivity to the attitudes of others, shyness, and a tendency to hide their skin conditions, which is mapped in the feelings of stigmatization instrument for dermatology patients [57]. In addition to internal factors, these judgments are shaped by social behaviors, such as stereotypes, prejudices, and discrimination [58]. Tunca et al [52] found that skin conditions have physical and emotional effects on patients with certain skin conditions that can be seen without aids, such as acne and psoriasis. The emotional impact experienced can be in the form of feelings of guilt, shame, anxiety, and social isolation [52]. Rowlands et al [59] found that individuals with health conditions who are prone to stigma tend to seek health information through digital platforms. This is because the stigma possessed by individuals encourages them to avoid formal help and choose to look for alternatives, such as social media platforms [58].

- Hypothesis 9: PSE positively affects SHE

Based on HBM, PSE refers to an individual's belief in the intensity of a medical condition and the seriousness of consequences related to a particular health problem [9]. Malik et al [9] found that PSE has a positive influence on individuals' willingness to search for health information through social media. When PSE increases, individuals assess health threats as greater so that the benefits of searching for health information outweigh the effort needed to conduct the search [10]. In addition, in the context of social media, higher PSE encourages users to seek explanations, guidance, and action options that can be done through quick and easily accessible platforms, such as social media platforms [60]. Neely et al [49] also found that during risky public health situations, users with high PSE conducted more searches for health information through social media to reduce uncertainty and guide health decision-making. Moreover, patients with serious skin conditions routinely use

XSL•FO
RenderX

the internet to search for skin care information [54]. Based on HBM, when individuals assess their skin conditions as serious, such as potentially interfering with social interactions, self-confidence, and daily activities, these perceptions can be a driving factor for individuals to seek skin health information as a protective tool that can help them and others avoid or reduce the risks that arise from certain skin health problems [9].

- Hypothesis 10: PSU positively affects SHE

Based on HBM, PSU is defined as an individual's belief about the possibility of being affected by a health problem [9]. PSU is one of the significant predictors of the behavior of seeking online health information [10,46]. When individuals feel vulnerable to a health problem, they have a tendency to seek information on ways to reduce the risk [61]. Therefore, when PSU increases, the expected benefits are considered higher than the effort required, so that individuals are motivated to search for information. In the context of this study, PSU is defined as an individual's perception of the likelihood of being affected by certain skin health problems. This possibility is not only assessed against infectious diseases but also against possibilities caused by genetic factors or family history and nongenetic factors, such as infections, use of certain medications, lifestyle, and the environment. Richard et al [62] found that the location of certain diseases has a different impact on the patient's daily life, and this is due to the burden of the visibility of a disease. Moreover, facial perception and visual attention directed at the face and hands are essential in human social communication; thus, diseases in these parts have a more significant effect [62]. Based on HBM, an increase in PSU strengthens perceived threats, which are predictors of online health information–seeking behavior [9,10]. Perceived threats are also predictors of information insufficiency, which increases the intention to seek information [63]. In the context of social media platforms, users with higher PSU were found to search for health information more actively than users with lower PSU [49].

Methods

Ethical Considerations

This study was approved by the Faculty of Computer Science, University of Indonesia (reference S-20A/UN2.F11.D1.5/PPM.00.00/2025 and dated June 3, 2025). All of the respondents involved in this study agreed in writing to participate in the study, and all respondent data were anonymized. The questionnaire was filled in by the respondents voluntarily, and they could withdraw from participating in the research at any time. We did not provide incentives for the respondents. We used the survey tools of the University of Indonesia and stored the data exclusively on the university's server. All respondents approved the written consent to be involved in this study.

Study Design

This study adopted a mixed methods approach, using an online questionnaire (quantitative) and interviews (qualitative). We designed an online and open questionnaire based on the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist (Checklist 1). All respondents approved the written consent to be involved in this study. The respondents were selected according to the criteria of the research target—that is, Indonesian people who used short video platforms to search for skin health information (purposive sampling). The questionnaire link was shared for 28 days from September 22, 2025, until October 20, 2025, through social media platforms widely used by Indonesian people, such as Facebook, Instagram, and TikTok. In the early stages, quantitative research was used as the primary method to test the hypotheses and the relationships between the variables in the research models. Furthermore, a qualitative approach was carried out to explain and enrich the understanding of the quantitative results through interviews with 30 respondents.

Research Instruments

The research instrument was in the form of a questionnaire, which was divided into 3 pages: respondent validity, respondent demographics, and 49 measurement items (Multimedia Appendix 1). The questionnaire completion time was 10 - 15 minutes, and the respondents could review their answers by clicking the back button for each page. We did not randomize the questionnaire's statements and did not use adaptive questions. For mandatory questions and statements in the questionnaire, we put a red star where the respondents must fill in the questions and statements before they could click the submit button. The respondent validity section consists of questions regarding the knowledge and experience of using one of the short video platforms used by the user, that is—TikTok, Instagram Reels, or YouTube Shorts. In addition, the respondent demographics section consists of questions related to the background of the study respondents. The measurement items were measured using a 5-point Likert scale, in which a value of 1 was interpreted as "strongly disagree," and a value of 5 was interpreted as "strongly agree."

Questionnaire Measurement Item Construction and Pretest

The readability test was conducted to evaluate the ease of the questionnaire questions and to identify aspects that needed improvement from the respondents' perspectives. The readability test was carried out in the form of interviews with 13 respondents. The respondents provided input on the right word selection and the correction of word or punctuation errors. After the readability test stage was completed, the pilot study stage was conducted with 30 respondents to test the reliability of the research indicators that had been made previously. The total Cronbach α (CA) score for the results of the pilot study was 0.928; thus, the questionnaire could be disseminated to a wider range of respondents. Multimedia Appendix 2 presents the demographics of the 603 valid respondents.

Data Cleansing

Of the 889-questionnaire data received, we conducted a manual data completeness check using Microsoft Excel to ensure that all questions had been answered completely. An examination of the results showed that 245 respondents did not complete the questionnaire until the end; thus, the data from these respondents were unable to be used for hypothesis testing. The remaining

XSL•FO
RenderX

644 questionnaires were completely filled out for hypothesis testing.

Analysis Methods

Questionnaire data were processed using the covariance-based structural equation modeling method using AMOS SPSS (version 26; IBM Corp) software, which aims to test the relationships between the variables formulated in the research model. This method was chosen because the study focused on testing hypotheses based on theories that had been tested in previous research. Qualitative data analysis was conducted using thematic analysis. Thematic analysis was chosen because it provides flexibility in identifying, analyzing, and interpreting patterns or themes that systematically emerge from qualitative data [64]. The social, cultural, and psychological contexts behind user behavior can also be found using thematic analysis [64].

Results

Common Method Bias

To detect potential bias, one of the approaches carried out was the confirmatory factor analysis (CFA) single-factor test, which determines one common latent factor that contains all the research indicators [65]. This approach is conducted by estimating the CFA model by pointing all indicators to one common factor as a form of testing whether one common factor can explain most of the variance of the data [65]. Common method bias (CMB) testing using single-factor CFA can be performed by analyzing the feasibility of the model [66]. If the single-factor model does not have a good goodness of fit (GoF), then CMB is not identified in the research model, and the research can continue to the next stage [66]. Table 1 describes the CMB results for measurement and structural model testing.

Table . Common method bias results.

GoF[a] metrics	GoF criteria		Common factor values	Description
	Good fit	Acceptable fit		
CMIN/df[b]	≤2	≤3	7.526	Poor fit
GFI[c]	≥0.95	≥0.9	0.387	Poor fit
RMR[d]	≤0.05	≤0.1	0.109	Poor fit
NFI[e]	≥0.95	≥0.9	0.439	Poor fit
CFI[f]	≥0.97	≥0.9	0.420	Poor fit
TLI[g]	≥0.97	≥0.9	0.449	Poor fit
RMSEA[h]	≤0.05	≤0.08	0.127	Poor fit

[a]GoF: goodness of fit.

[b]CMIN/df: minimum discrepancy of confirmatory factor analysis/degrees of freedom.

[c]GFI: goodness of fit index.

[d]RMR: root mean square residual.

[e]NFI: normed fit index.

[f]CFI: comparative fit index.

[g]TLI: Tucker-Lewis index.

[h]RMSEA: root mean square error of approximation.

Measurement and Structural Model Testing

The convergent validity test was carried out by analyzing the average extracted variance (AVE) value, in which the minimum AVE value for each variable tested must be above 0.5 (Hair et al, 2019). As shown in Table 2, all variables met the AVE value ≥0.5. Therefore, the convergent validity test stage was completed, and the study continued to the next stage. The reliability of the construct was measured using CA and composite reliability. Acceptable CA and composite reliability values are greater than 0.70 [67].

XSL•FO
RenderX

Table . AVE[a], CA[b], and CR[c] values.

Variable	AVE	CA	CR
AT[d]	0.653	0.883	0.882
HCE[e]	0.613	0.864	0.863
PHI[f]	0.593	0.814	0.813
PI[g]	0.685	0.867	0.867
PSE[h]	0.658	0.852	0.852
PSU[i]	0.606	0.885	0.885
PU[j]	0.681	0.914	0.914
SC[k]	0.677	0.913	0.913
SHE[l]	0.727	0.930	0.930
SST[m]	0.702	0.904	0.904
USC[n]	0.641	0.877	0.877

[a]AVE: average extracted variance.
[b]CA: Cronbach α.
[c]CR: composite reliability.
[d]AT: attitude.
[e]HCE: health content expressiveness.
[f]PHI: personalized health insight.
[g]PI: perceived interactivity.
[h]PSE: perceived severity.
[i]PSU: perceived susceptibility.
[j]PU: perceived usefulness.
[k]SC: source credibility.
[l]SHE: skin health information seeking on short video .
[m]SST: skin stigmatization.
[n]USC: upward skin comparison.

The discriminant validity test was performed to ensure that each construct measured in the research model had a significant difference from the other constructs [67]. This stage was conducted to ensure that a construct not only measures the same concept as other constructs but also has unique characteristics so that it has to be separated as a construct in the research model [67]. Table 3 presents the results of the discrimination validity test.

XSL•FO
RenderX

Table . Correlation matrices.

	PSU[a]	PSE[b]	USC[c]	PI[d]	SC[e]	PHI[f]	HCE[g]	AT[h]	SST[i]	PU[j]	SHE[k]
PSU	0.494										
PSE	0.321	0.470									
USC	0.267	0.264	0.495								
PI	0.131	0.144	0.176	0.366							
SC	0.137	0.135	0.134	0.213	0.364						
PHI	0.082	0.125	0.153	0.227	0.189	0.307					
HCE	0.096	0.128	0.111	0.204	0.131	0.148	0.414				
AT	0.096	0.106	0.129	0.269	0.156	0.167	0.150	0.340			
SST	0.198	0.196	0.366	0.130	0.099	0.113	0.082	0.095	0.624		
PU	0.078	0.113	0.122	0.195	0.149	0.219	0.228	0.143	0.091	0.344	
SHE	0.165	0.190	0.168	0.223	0.165	0.182	0.178	0.222	0.126	0.215	0.398

[a]PSU: perceived susceptibility.

[b]PSE: perceived severity.

[c]USC: upward skin comparison.

[d]PI: perceived interactivity.

[e]SC: source credibility.

[f]PHI: personalized health insight.

[g]HCE: health content expressiveness.

[h]AT: attitude.

[i]SST: skin stigmatization.

[j]PU: perceived usefulness.

[k]SHE: skin health information seeking on short video platform.

GoF testing was performed to evaluate the feasibility of the research model by determining whether the GoF values met the specified criteria [67]. The evaluation was carried out by evaluating several GoF values, such as the relative chi-square, GoF index, root mean square residual, normal fit index, comparative fit index, Tucker-Lewis index, and root mean square error of approximation [67]. Table 4 presents the GoF values of the research model.

Table . Goodness of fit (GoF) value.

GoF metrics	GoF criteria	Results
CMIN/df[a]	≤2	1.242
GFI[b]	≥0.9	0.906
RMR[c]	≤0.1	0.032
NFI[d]	≥0.9	0.930
CFI[e]	≥0.9	0.985
TLI[f]	≥0.9	0.982
RMSEA[g]	≤0.08	0.025

[a]CMIN/df: minimum discrepancy of confirmatory factor analysis/degrees of freedom.

[b]GFI: goodness of fit index.

[c]RMR: root mean square residual.

[d]NFI: normed fit index.

[e]CFI: comparative fit index.

[f]TLI: Tucker-Lewis index.

[g]RMSEA: root mean square error of approximation.

XSL•FO
RenderX

Hypothesis Testing

Hypothesis testing was performed using a 1-tailed test to assess the significance of the evaluated hypothesis coefficient. A hypothesis was declared acceptable if the *P* value <.05, while the hypothesis was rejected if the *P* value was ≥.05 [67]. The results of the hypothesis test are summarized in Table 5, which shows that out of the 10 hypotheses submitted, 8 hypotheses were accepted, and 2 hypotheses were rejected because they had a *P* value ≥.05.

Table . Hypothesis (H) test results.

Hypothesis	Estimate	*P* value	Results
H1: PU[a]←HCE[b]	0.425	.001	Accepted
H2: PU←PHI[c]	0.491	<.001	Accepted
H3: SHE[d]←PU	0.327	.01	Accepted
H4: SHE←SC[e]	0.063	.17	Rejected
H5: AT[f]←PI[g]	0.754	<.001	Accepted
H6: SHE←AT	0.302	.02	Accepted
H7: SST[h]←USC[i]	0.627	<.001	Accepted
H8: SHE←SST	−0.041	.30	Rejected
H9: SHE←PSE[j]	0.161	.009	Accepted
H10: SHE←PSU[k]	0.159	.02	Accepted

[a]PU: perceived usefulness.

[b]HCE: health content expressiveness.

[c]PHI: personalized health insight.

[d]SHE: skin health information seeking on short video platform.

[e]SC: source credibility.

[f]AT: attitude.

[g]PI: perceived interactivity.

[h]SST: skin stigmatization.

[i]USC: upward skin comparison.

[j]PSE: perceived severity.

[k]PSU: perceived susceptibility.

Based on the results of the effect size analysis, it can be seen that each construct in the model has a different influence power. The attitude variable showed an effect size of 0.569, which was in the range of 0.50 - 0.79 so it was included in the category of medium effect. These findings show that PI has a strong and substantive influence in shaping users' attitudes. Furthermore, the SST variable has an effect size of 0.392, which is in the category of small effects, so although USC has an effect on increasing skin stigma, the power of influence is relatively weaker than other constructs. The PU variable showed an effect size of 0.562, so it was also included in the medium effect category. This indicates that the combination of the influence of HCE and PHIs makes a significant contribution in increasing the perception of the usefulness of skin health content. Meanwhile, the variable SHE has an effect size of 0.579, which is also in the medium category. This shows that related variables in the research model, such as PU, attitude, PSE, and PSU, have a strong contribution in encouraging skin health information–seeking behavior on short video platforms.

Discussion

Summary of the Findings

The study found that HCE and PHIs affected users' PU of short video platforms in searching for skin health information, thus affecting SHEs. SC had no effect on SHEs because the functionality of the platform and familiarity with the platform were the main points chosen by users. PI had an effect on attitudes, thus affecting SHEs. In the context of short video platforms, interactivity is one of the factors that play a significant role because it is considered to increase personalization, which was previously found to have a positive influence on users' feelings on a platform. USC also had an effect on SST, which had no effect on SHEs. Finally, PSE and PSU affected SHEs.

All qualitative results in each hypothesis strengthen the quantitative results and provide the detailed reason of each hypothesis results according to the Indonesian context. Table 6 describes the mapping summary of the hypothesis testing results with the qualitative results. The unique findings of this study were on the results of SC and SST that has no effect on SHEs. SC only affected an individual's desire to conduct further

XSL•FO
RenderX

searches to validate the information they have just watched. Trusting a figure was also conducted one by one when the content was being watched and not before doing a search and through the dynamics of interactions in the platform ecosystem. Furthermore, in the context of users of short video platforms in Indonesia, SST is not the main factor that drives the behavior of seeking skin health information, especially on short video platforms. The next section will discuss further justification of each hypothesis.

XSL•FO
RenderX

Table . Mapping summary of the hypothesis (H) testing results with the qualitative results.

Hypothesis	Hypothesis testing results	Summary of the qualitative results
H1	Accepted	The expressiveness of the content on the short video platform provides visual evidence, makes it easier to understand, clarifies how to use the product, and presents a direct picture of skin conditions, thus making skin health information more useful and relevant for users.
H2	Accepted	Personalization allows users to receive content relevant to their personal needs, expand insights through additional recommendations, improve search efficiency, and simplify the process of finding skin health information. While some concerns have been raised regarding privacy and information overload, personalization is still perceived as a feature that can improve the usefulness of skin health information on short video platforms.
H3	Accepted	Perceived usefulness is an important factor that encourages users to search for skin health information on short video platforms through ease of access, information relevance, interactivity, and the platform's ability to provide additional insights.
H4	Rejected	The credibility of informants remains important, but perceptions of the level of credibility do not affect an individual's desire not to conduct searches on short video platforms. Rather, they affect an individual's desire to conduct further searches to validate the information they have just watched.
H5	Accepted	Interactions on short video platforms through like, comment, and save features make users feel that they have appreciated content creators who provide useful skin health information. This interaction also positively affects user attitudes because users feel engaged and have control over the videos displayed on the short video platform.
H6	Accepted	In the context of searching for skin health information on short video platforms, attitudes can be judged through users' feelings toward short video platforms to search for skin health information. With the positive feelings of users toward the process of searching for skin health information on short video platforms, users have a desire to search for health information on short video platforms.
H7	Accepted	Users feel motivated to have as good a skin condition as that shown on the short video platform, are more diligent in taking care of their skin, and strive to solve the problems experienced. Users can also understand the content displayed using filters and makeup and promotional content in the form of marketing, which is not actually used by the content creator.

XSL•FO
RenderX

Hypothesis	Hypothesis testing results	Summary of the qualitative results
H8	Rejected	Most of the interviewees did not feel stigmatized and were used to their skin conditions; thus, the stigma did not trigger the search for health information. Even when users' confidence in their skin health increases, some prefer to seek information from other sources (eg, doctors, clinics, or buying products directly) rather than using short video platforms. Thus, in the context of users of short video platforms in Indonesia, skin stigmatization is not the main factor that drives the behavior of seeking skin health information, especially on short video platforms.
H9	Accepted	The higher the perceived severity, the greater the individual's tendency to have skin health information–seeking behavior on short video platforms. Concerns about potential long-term effects and a desire to avoid or minimize skin problems are the main factors that drive users' search for information, both as preventive and early mitigation measures before deciding to seek professional help.
H10	Accepted	The higher the individual's concern about the possibility of skin problems, the greater their tendency to be actively involved in the search, verification, and application of skin health information through short video platforms.

The Effect of HCE on PU

The study found that HCE significantly affected PU. These findings are consistent with Yu et al [23], Armstrong et al [68], and Lenczowski et al [33]. Yu et al [23] found that the use of media elements, such as visual, audio, and text, could enrich the user experience and improve the understanding of complex health information. Armstrong et al [68] revealed that skin health education through online videos was more effective than written media because videos allow for live demonstrations. Lenczowski et al [33] showed that their respondents preferred skin health education through video rather than written materials, confirming that expressive media improves the perception of the usefulness of information.

The findings of the qualitative research also support the results of the quantitative analysis (hypothesis 1). In total, 9 interviewees (9/30, 30%) stated that the expressiveness of content on short video platforms helped them to see first-hand evidence when searching for skin health information. Interviewee 15 said:

> *I can see the results of using the product, my skin texture, and how to use it directly. Of all visual media, photos and videos are important because they provide a real picture and make it easier for me to judge the product or the results.*

In addition, 13 interviewees (13/30, 43.33%) felt that the expressiveness of content on short video platforms made it easier for them to capture information. Interviewee 9 stated:

> *The media is very helpful to see visualizations and make information easier to understand. So, not only do you hear the explanation, but there is also a picture. The most important thing is still videos because you need to see the visuals directly.*

A total of 7 interviewees (7/30, 23.33%) also found this helpful because the expressiveness of short video content on short video platforms can show how the product is used. As Interviewee 22 explained:

> *Visuals are very helpful in giving an overview of the product measurements. If you just write about it, it's usually less imaginative. For example, in the past, I was confused about how much moisturizer to use because each person has a different finger size. The video makes the measurement clearer, so the video is the most important.*

The expressiveness of the content on the short video platform provides visual evidence, makes the content easier to understand, clarifies how to use the product, and presents actual pictures of skin conditions, thus making skin health information more useful and relevant for users.

The Effect of PHIs on PU

This study found that PHIs had a significant effect on PU. The results of this study are in accordance with Yu et al [23], Zhang et al [34], and Nittas et al [35]. Yu et al [23] found that PHIs are one of the features of media richness that significantly affect PU on short video platforms. Information that is appropriate to the users' preferences, needs, and context provides the content with more practical value and can effectively address individual health challenges [23]. Zhang et al [34] showed that PHIs in the context of mobile health had a positive and significant influence on PU because they tailored services to the unique conditions of users in China. Nittas et al [35] found that personalization of skin health education, such as the use of

XSL•FO
RenderX

sunscreen, was well received by participants in Switzerland and improved their perception of the platform's usability.

The results of the interviews support the findings from the questionnaire results (hypothesis 2). In total, 9 interviewees (n=30, 30%) stated that personalization helped them obtain skin health information according to their needs. As Interviewee 8 mentioned, "Personalization makes it easier because the content that appears is already relevant to the question or skin condition that is being experienced." Personalization in the context of skin health on short video platforms is able to display content that is aligned with users' individual dermatological needs based on search history, duration of viewing, interactions, and user preferences for certain skin types or skin concerns. In addition, 7 interviewees (n=30, 23.33%) felt that personalization helped them obtain more information related to skin health. Interviewee 3 said:

> *Personalization makes it easy for me to get information from various sources. I still need varied content to get insights from many sides.*

Personalization allows users to receive content relevant to their personal needs, expand insights through additional recommendations, improve search efficiency, and simplify the process of finding skin health information. While some concerns have been raised regarding privacy and information overload, personalization is still perceived as a feature that can improve the usefulness of skin health information on short video platforms.

The Effect of PU on SHEs

PU had a significant effect on SHEs. The results of the quantitative analysis are in accordance with Yu et al [23], Liu et al [4], Xia et al [38], and Boyce et al [39]. Yu et al [23] showed that PU is one of the main factors driving the search for health information on short video platforms, as users assess the content provided as useful and can improve their ability to understand health information. Xia et al [38] found that PU had a significant positive influence on users' desire to search for health information online, even greater than the effect of perceived ease of use. In addition, Boyce et al [39] emphasized that the PU of a platform drives the behavior of seeking health information in online communities because users feel that the platform helps them find appropriate and relevant information.

In addition, we found that the interview results support the questionnaire result (hypothesis 3). Based on the interviews, 3 interviewees (n=30, 10%) considered short video platforms beneficial because they supported discussion and interactivity, which facilitated the process of finding and understanding skin health information. Interviewee 12 stated, "TikTok remains the main platform for me because its features are interactive and make it easier for me to understand information quickly." Aside from interactivity, the perception of usability was also reflected in the ease of the search process. A total of 9 interviewees (n=30, 30%) stated that short video platforms offer a fast, practical, and easy-to-understand search process. As explained by Interviewee 13, "Because the content has been adjusted to my needs, the information that appears becomes relevant and quickly connects with what I'm looking for." The benefits of the platform were also seen through the platform's ability to provide additional information beyond the users' initial needs. In total, 6 interviewees (n=30, 20%) revealed that short video platforms helped broaden their horizons. As stated by Interviewee 1, "Sometimes I get additional information that I didn't think of before." Aside from adding insights, other interviewees emphasized that the platform also provided a variety of information and the latest content that has been curated. Therefore, PU is an important factor that encourages users to search for skin health information on short video platforms through ease of access, information relevance, interactivity, and the platform's ability to provide additional insights.

The Effect of SC on SHEs

SC did not have a significant effect on SHEs. Moreover, the results of this study are not in accordance with those of Fan et al [43] and Kirkpatrick and Lawrie [18]. Fan et al [43] and Kirkpatrick and Lawrie [18] found that assessments of whistleblower credibility can affect the tendency and frequency of user information searches. However, several studies have found that the credibility of information sources did not have a significant effect on the search for health information through online platforms [69,70].

Denniss et al [70] revealed that the credibility of sources was considered less relevant because users judged the content and social interactions more. In the context of the use of short video platforms in the search for skin health information, SC did not affect the search because the characteristics of the social media ecosystem were more content-oriented and algorithmic. In the interviews, the interviewees who were laymen in the context of skin health watched the videos shown by short video platforms but did not specifically search for a figure. Therefore, trusting a figure was carried out one by one when the content was being watched, and this was not done before conducting a search. This is consistent with the demographic distribution of the study, in which 405 respondents (n=644, 67.1%) stated that skincare products were the type of skin health information they were looking for on short video platforms. This type of information is more general; thus, it can be reviewed by anyone without a formal background in the health field. The type of information that anyone can review emphasizes the role of platforms in building algorithmic ecosystems as a key factor compared with the role of individual content creators in the process of searching for skin health information. Trusting a figure was also conducted one by one when the content was being watched and not before doing a search. To date, users also face an increasing number of hoax content, algorithmic contents, and content generated by artificial intelligence [71,72]; thus, making social media users more careful in looking for informational or educational contents especially for "algorithmic truth" hindering. Figures should align their core values by emphasizing quality and environmental sustainability in their messaging that can strengthen users' trust and emotional connection in the short video platform [73]. Moreover, quantity and quality social proof (ie, likes and comments) could also influence users' behavior in the short video platform [74].

XSL•FO
RenderX

Moreover, interview results also support the quantitative result (hypothesis 4). We found that 20 interviewees (n=30, 67.1%) revealed that skincare products were the type of skin health information they were looking for on short video platforms. This type of information is more general; thus, it can be reviewed by anyone without a formal background in the health field. In addition, based on the results of the interviews, the validation of skin health information on short video platforms is more often done through social interactivity mechanisms, such as comments, the number of likes, and follow-up responses in the form of replies or stitch content. These interactions serve as a form of collective evaluation that helps users assess whether information is trustworthy. In this context, credibility is not built individually through a specific figure but is done every time one watches a video communally through the dynamics of interactions in the platform ecosystem.

Trusting a figure was also conducted one by one when the content was being watched and not before doing a search. Therefore, several exchange rates on the platform could make users continue to use short video platforms to search for skin health information. A total of 12 interviewees (n=30, 40%) stated that their trust in creators' content on short video platforms was quite low, but they still conducted searches on short video platforms. According to Interviewee 27:

> *I don't immediately believe the info from TikTok, anyway. Usually, I dig again and verify it with other content. But if the way of speaking is honest and the form is of an honest review, then it will make me believe more. I also look at the track record, including discussions in the comment column.*

These findings suggest that the credibility of informants remains important, but perceptions of the level of credibility do not affect an individual's desire not to conduct searches on short video platforms. Rather, it affects an individual's desire to conduct further searches to validate the information they have just watched.

The Effect of PI on Users' Attitudes

PI had a significant influence on attitude. This result is in accordance with Liu et al [4], who found that interactivity is one of the factors that contribute to the flow of application use and increases the desire of users to continue using the application. The finding is also consistent with Haykal [37], who found that personalization is an added value in the context of skin health, as everyone's skin condition is different. The qualitative results also support the findings of the questionnaire result (hypothesis 5). A total of 6 interviewees (n=30, 20%) revealed that interactions on short video platforms gave them positive feelings because they helped providing personalized information. Interviewee 2 stated:

> *I usually like or save if the video is useful or interesting. I rarely comment. During that interaction, I felt satisfied because the FYP became more personal and catered to my needs. So, later, similar content will appear more often.*

In addition, 7 interviewees (n=30, 23.33%) had positive feelings about the interactions because they could help retain information. As Interviewee 7 said, "With the save feature, I now have a kind of personal archive that I can open again at any time without having to search from scratch." In total, 3 interviewees (n=30, 10%) also had positive feelings about the interactions because the comments showed communal support. According to Interviewee 16, "If my comment is replied to or someone responds, I feel more confident and don't feel alone." For 3 interviewees (n=30, 10%), the interactions provided positive feelings because they increased their sense of involvement. Interviewee 15 stated, "When I like or save skin health content, I feel more involved with the content." Thus, interactions on short video platforms through the like, comment, and save features make users feel that they have shown appreciation to content creators who provide useful skin health information. This interaction also positively affects user attitudes because users feel engaged and have control over the videos displayed on the short video platform.

The Effect of Users' Attitudes on SHEs

Attitude had a weak effect on SHEs. This result is in accordance with Li et al [48], Neely et al [49], and Link et al [46]. Li et al [48] showed that users' perceptions and positive attitudes in China were associated with increased health information search behavior in digital services. Neely et al [49] found that user engagement on social media platforms and user attitudes were factors driving users' intentions and behaviors to search for health information on social media platforms in the United States. Furthermore, Link et al [46] explained that users' positive attitudes toward online health information searches were associated with increased online health information search behaviors in Germany.

Moreover, interview results support the findings of the quantitative results (hypothesis 6). We found that 10 interviewees (n=30, 33.33%) felt that using short video platforms helped them search for skin health information; thus, they were encouraged to use short video platforms when they wanted to find skin health information. According to Interviewee 9:

> *Yes, I felt that I was given help because my questions were answered. If you need more info, just open the application again.*

Moreover, 8 interviewees (n=30, 26.66%) were satisfied with the experience of searching for skin health information on short video platforms; thus, they were encouraged to use the platform again. As Interviewee 10 stated:

> *I'm satisfied so far because the information is complete and can be seen from many angles. This sense of satisfaction makes me continue to use Instagram to look for information.*

In addition, 3 interviewees (n=30, 10%) felt comfortable using the short video platform when searching for skin health information; thus, they were encouraged to use the platform again. As Interviewee 13 revealed:

> *I found it helpful because the content was diverse and I could search for a lot of things in one place. This positive feeling made me continue to search for information on TikTok, including about skin health.*

XSL•FO
RenderX

> *Because it's comfortable, I keep it as my main platform.*

Thus, in the context of searching for skin health information on short video platforms, attitudes can be judged through users' feelings toward short video platforms when searching for skin health information. With users' positive feelings toward the process of searching for skin health information on short video platforms, they have the desire to search for health information on short video platforms.

The Effect of USC on SST

USC affected SST. This result is in accordance with Tian et al [50], who found that USCs, particularly on social media, increased individuals' anxiety about their appearance in the context of general appearance, not specific to the skin. In the context of searching for skin health information on short video platforms, USC is defined as a situation in which users compare their skin conditions with those on content that displays better skin conditions. Our interview results also support the quantitative finding (hypothesis 7). This study found that 3 interviewees (n=30, 10%) admitted that they had compared their skin conditions with those on the short video platform and that this comparison process gave rise to feelings of insecurity. According to Interviewee 3:

> *Yes, I used to compare a lot. Now, it's not too much because I already understand that the breakout phase comes and goes. In the past, I was also insecure, especially when I saw people with smooth skin. I often think about whether my skin would look good if it wasn't such a mess.*

In addition, 14 interviewees (n=30, 46.66%) were motivated to have a good skin condition similar to that shown on the short video platform, were more diligent in taking care of their skin, and strived to solve the problems they experienced. Interviewee 27 said:

> *If you look at the video of people with really good skin, yes, I compare. Sometimes it's a good idea to take care of your skin more.*

Users can also understand the content displayed using filters and makeup and promotional content in the form of marketing, which is not actually used by the content creator.

The Effect of SST on SHEs

SST had no effect on SHEs. This finding is not consistent with those of Berger et al [53], Aslan Kayıran et al [54], and Rowlands et al [59]. Berger et al [53] showed that stigmatization due to skin health conditions had a significant effect on the behavior of seeking health information online, as individuals who feel stigmatized tend to look for alternative information through the internet as a form of coping with negative social judgments. Conversely, Aslan Kayıran et al [54] found that the perception of stigma in patients with psoriasis had a significant effect on the search for health information through social media and the internet. Rowlands et al [59] also discovered that SST can encourage individuals to search for health information digitally, as they tend to avoid formal help and opt for private channels, such as online platforms.

In accordance with the quantitative results (hypothesis 8), 6 interviewees (n=30, 20%) did not feel insecure about their skin conditions, as explained by Interviewee 12:

> *I've never felt uncomfortable or lacked confidence in my skin condition. So far, I'm pretty confident.*

A total of 4 interviewees (n=30, 13.33%) were used to their skin conditions and thus no longer felt insecure. They conducted information searches more as a preventive measure or practical need. Interviewee 21 stated:

> *I used to lack confidence, but now I feel normal because I already know the pattern. The hormonal acne will go away on its own.*

According to 2 interviewees (n=30, 6.66%), when they felt insecure about their skin conditions, they would conduct an information search through the help of medical professionals, such as doctors or clinics, rather than on short video platforms. Interviewee 18 said, "When I feel uncomfortable or have a breakout, I will immediately stop using the product and look for a solution directly at a physical place or clinic." A total of 5 interviewees (n=30, 16.66%) immediately sought and bought skin health products when they felt insecure about their skin conditions. As Interviewee 5 explained, "If I'm insecure because of severe acne, I usually immediately look for products that can help." Most of the interviewees did not feel stigmatized or were used to their skin conditions; thus, the stigma did not trigger the search for health information. Even when users' confidence in their skin health increases, some prefer to seek information from other sources (doctors, clinics, or buying products directly) rather than using short video platforms. Thus, in the context of users of short video platforms in Indonesia, SST is not the main factor that drives the behavior of seeking skin health information, especially on short video platforms. Hussain et al [75] found that many women relied on filters for social validation despite recognizing their artificiality; thus, to date, beauty standards are increasingly coproduced by platform design and artificial intelligence–driven filters. In addition, beauty standards also influenced by culture, norm, and value [75].

The Effect of PSE on SHEs

PSE had a significant effect on skin health information seeking. This result is in accordance with Malik et al [9], Zhao et al [10], and Neely et al [49]. Zhao et al [10] showed that PSE had a positive effect on individuals' willingness to search for health information through social media in individuals with chronic diseases in China. Increasing PSE was found to increase health information search behavior, as users want to avoid greater health threats [10]. This study is consistent with the study by Neely et al [49], which found that users with high PSE tended to search for more health information to reduce uncertainty in health decisions. Malik et al [9] also asserted that this perception is used to help avoid or reduce the risks that can arise due to certain skin health problems in Pakistan. In the context of searching for skin health information on short video platforms, PSE can be visualized in the content shown on short video platforms.

In addition, the interview result supports the questionnaire result (hypothesis 9). We found that 3 interviewees (n=30, 10%) stated

XSL•FO
RenderX

that concerns about the content and long-term effects of using skincare products prompted them to seek further education on short video platforms. Interviewee 1 said:

> *When I first learned about skincare, I was worried because I layered a lot of products. I was also afraid that my skin would be exposed to chemicals too often and that side effects would appear. So, I looked for info on short video platforms and finally found out about skincare fasting.*

In addition, 9 interviewees (n=30, 30%) were encouraged to actively seek information through short video platforms because of their concerns over the effects of skin problems, such as acne scars, stretch marks, or irritation. As Interviewee 3 stated:

> *Because I'm worried, I'm more diligent when looking for information so that I can avoid the side effects. What I'm most worried about is worsening acne. I once saw a video that said that micellar water used with cotton could make acne worse, so I tried a cleansing balm. But it turned out to be even worse.*

Furthermore, 6 interviewees (n=30, 20%) conducted an information search to minimize the side effects when the initial symptoms began to appear. As Interviewee 23 explained, "I often look for information to reduce the contrast of acne scars when I have a breakout so that acne does not leave marks." The higher the PSE, the greater the individual's tendency to have skin health information–seeking behavior on short video platforms. Concerns about potential long-term effects and the desire to avoid or minimize skin problems are the main factors driving users' search for information, both as preventive and early mitigation measures before deciding to seek professional help.

The Effect of PSU on SHEs

PSU significantly affected SHEs. This result is consistent with Malik et al [9], Zhao et al [10], and Link et al [46]. This indicates that the higher an individual's perception of the likelihood of experiencing skin health problems, the greater the tendency to seek related information through short video platforms. Accordingly, Malik et al [9] showed that an individual's perception of vulnerability to health conditions significantly affected preventive behavior, as individuals tend to take anticipatory steps when feeling vulnerable. Zhao et al [10] confirmed that PSU had a significant effect on health information–seeking behavior driven by the need to understand and control risk. Link et al [46] revealed that the perception of vulnerability to health problems significantly drives the search for information through online platforms.

The qualitative result also supports the findings of hypothesis 10 result. A total of 10 interviewees (n=30, 33.33%) were encouraged to be more active in seeking information as a form of preventive action because of their concerns about the potential appearance of skin problems. As Interviewee 3 stated:

> *Because I'm worried, I'm more diligent when looking for information so that I can avoid the possibility of acne. I'm most afraid that I will have acne, even though my skin is fine, especially because I commute every day and am often exposed to pollution.*

In addition, because of their concerns about skin conditions, 3 interviewees (n=30, 10%) not only looked for solutions but also verified the accuracy of the information they obtained. As Interviewee 2 explained:

> *Because I tend to be paranoid, I'm more diligent when looking for and verifying information so that I can get the most valid information. But yes, sometimes there's no end because everyone has different skin. I'm most worried about having dull skin.*

Thus, the higher the individuals' concern about the possibility of having skin problems, the greater their tendency to be actively involved in the search, verification, and application of skin health information through short video platforms.

Theoretical Implications

This study enriched previous studies that examined the behavior of searching for health information on social media in general [9,11,15,23,24] and discussed content quality with regard to the credibility of actual information [21,22,25]. We found that the role of algorithmic ecosystems in shaping users' skin health behavior, the shift from SC to short video platform credibility, and the importance of visual media richness in aesthetic health contexts could influence users in searching skin health information on short video platforms. The findings of this study also provide theoretical implications for SOR, MRT, and HBM theories by showing that the effectiveness of media in searching for skin health information is highly determined by media wealth, which serve as a stimulus. In the context of MRT, short videos have several media that can be used by their users, such as a combination of video, photos, audio, and caption text, making them more effective in explaining visual skin health topics. This is because the combination of media used can reduce the ambiguity and complexity of information. In the context of HBM, the visualization of skin conditions increases PSE and PSU, which encourage information-seeking behavior represent organisms and responses, respectively. This implication confirms that the effectiveness of media in health communication is not only determined by the form of the media but also by the media's ability to trigger users' beliefs and perceptions of skin health, as described in HBM.

This study found that HCE and PHIs serve as stimuli and had an effect on PU. This finding can expand the MRT, in which the existence of content expressiveness and personalization can make media delivery effective in the context of short videos. In addition, findings related to HCE underscore that expressiveness depends not only on the media format but also on the ways of delivery, demonstration, and visualization of skin health content. These findings are consistent with those of Priyambodo et al [11], who showed that visual appeal affects PU. This study also found that PU is also formed from a combination of content expressiveness and personalization; thus, the research model of health information–seeking behavior needs to consider visual content–based factors as the main variables. This finding is in accordance with the study by Priyambodo et al [11], which revealed that PU is a factor that drives the intention to seek health information. These findings can strengthen the relevance of the technology acceptance model and confirm the research of Malik et al [9] and Jin et al [41]

XSL•FO
RenderX

regarding the relationship between the usefulness of short video platforms and information retrieval behavior.

SC in this study has no effect on SHEs so that it can expand the research of Malik et al [9] by showing that the ecosystem of short video platforms is more influenced by the relevance of content and ease of access compared to the credibility of the source. This is supported by the behavior of users who consume other content before deciding to trust the information received on a short video platform. Based on Joseph [71], most of the short video platforms governed by algorithmic processes, where this platform acts as an intermediary that could shape perceptions of identity and self-representation and amplify pressures on users to conform to idealized norms. Therefore, short video platform users need to engage in active self-construction for developing digital habits that prioritize reflective awareness, diverse media consumption, and the scrutiny of artificial intelligence recommendations to offset the passive exposure to algorithmic feedback [71,72].

The results of this study also show that through interaction features, such as likes, comments, and saves, can have a satisfaction effect on the users' behavior of searching for skin health information. Consistent with Liu et al [4], these findings may provide further support for theories related to interactions in discussing short video platforms. In contrast to Priyambodo et al [11], who showed that user interactivity does not have a significant effect on PU, this study examined the relationship between PI and attitude.

Attitude had a weak effect on SHEs in this study. Although attitude has only a weak effect, these findings suggest that attitude's influence on behavior is not as strong as initial assumptions and therefore requires new theoretical adaptations in the domain of short video platforms. These findings can provide insight that there are other factors that can give users a positive feeling to the users of the short video platform.

Exposure to content with certain beauty standards, such as smooth skin without acne, can trigger a process of upward comparison, which then triggers feelings of dissatisfaction with an individual's skin condition. These findings can strengthen the research of Adkins et al [26] by confirming that USC also drives SST behavior on short video platforms. The insignificance of the effect of SST on SHEs indicates that the relationship between stigma and search behavior may be mediated by other psychological factors that have not been studied in this study. These findings may provide implications that research models related to information-seeking behavior need to integrate psychological and cognitive factors to explain the relationship between stigma and information-seeking behavior.

In the end, the strong influence of PSE and PSU on information-seeking behavior provides validation for HBM. Users who feel that their skin conditions are serious, such as inflamed acne, dark spots, and severe eczema, are likely to be encouraged to seek out skincare information, especially when they feel that the condition they are experiencing has the potential to get worse. In addition, users who have a perception of susceptibility to a disease are also encouraged to seek information to prevent getting the disease. Priyambodo et al [11] have similar results to this study; PSE and PSU have an influence on the intention to seek health information, which then affects the actual online health information–seeking behavior. These findings confirm that PSE and PSU remain the driving factors for information-seeking behavior, especially in health domains that are highly visual, such as skin health. Overall, the findings in this study show that the behavior of seeking skin health information is the result of an interaction between cognitive, psychological, social, and technological factors.

Practical Implications

From the perspective of short video platforms, platform providers need to strengthen their role as responsible information facilitators. This can be done by providing interaction features that allow educational discussions to take place well, such as fact-check prompts. Recommendation algorithms also need to be directed to provide greater space for evidence-based educational content and corrective interaction from professionals, not just viral content. In addition, official information panels and warning labels need to be integrated into content that is vulnerable to misinterpretation so that the interaction network on the platform does not reinforce misinformation. Features that directly connect users with various trusted sources are also needed. Examples of this are displaying verified educational video recommendations, incorporating educational panels on specific dermatology keywords, and providing links to official health institutions. In this way, users can cross-check without having to leave the platform.

As content creators of health care professionals, it is important to stabilize the quality of the information circulating through their active involvement in conversations on the platform. In addition to producing educational content that is clear, visual, and easy to understand, health care workers also need to be present in digital interaction spaces, such as comment columns or video replies, to correct mistakes, answer questions, and strengthen user understanding. The presence of health care workers in the flow of public conversations on short video platforms can increase users' confidence in accurate information. Nonprofessional creators also need to be more careful when presenting skin health information. These creators need to maintain transparency, especially in sponsored content, avoid medical claims that cannot be validated, and open up space for collaboration or clarification with health workers. By leveraging interaction features, such as duets or stitches, creators can support the spread of informative content without misleading users.

Finally, the government needs to play a role in providing official references that can make it easier for users to verify information. Governments can work with platforms to link trusted sources, such as product regulations, dangerous product lists, or clinical dermatology guidelines. In addition, health literacy programs and public education about the risks of illegal products need to be strengthened so that the public is better able to distinguish valid information from false commercial content.

Limitations and Future Work

A limitation of this study is the demographics of the respondents, as most of the respondents were aged 20 - 29 years and were

XSL•FO
RenderX

taking or already had an education level equivalent to a bachelor degree. Thus, this study could not conduct a comparative study of ages. Moreover, the study did not evaluate other cognitive or psychological factors, such as prior knowledge or self-compassion, that could influence users' behavior in searching for health information. This study did not conduct an analysis based on the type of skin health information—that is, primary information on medical and tertiary conditions and aesthetics oriented toward beauty care.

Further studies could explore the effects of SST and SC on more general research objects, such as general health, on short video platforms. Specific content differentiation is also necessary, given that skin health information has diverse types of information, such as product reviews, skin health education, and information related to skin diseases, which are likely to drive different behaviors.

Funding

This study was supported by Direktorat Penelitian dan Pengabdian kepada Masyarakat, Direktorat Jenderal Riset dan Pengembangan, Kementerian Pendidikan Tinggi, Sains, dan Teknologi (contract PKS-145/UN2.RST/HKP.05.00/2026, 228/C3/DT.05.00/PL-BARU/2026, and 122/PKS/WRIII/UI/2026). The funder had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the lack of authority to share data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Measurement items in the questionnaire.
[DOCX File, 26 KB - derma_v9i1e93461_app1.docx]

Multimedia Appendix 2
Questionnaire respondents' demographics.
[DOCX File, 20 KB - derma_v9i1e93461_app2.docx]

Checklist 1
CHERRIES checklist.
[DOCX File, 66 KB - derma_v9i1e93461_app3.docx]

References

1. Schumacher S, Sparks G, Montalvo I, Kirzinger A, Hamel L. KFFtracking poll on health information and trust: health information and advice on social media. KFF. 2025. URL: https://www.kff.org/public-opinion/kff-health-information-and-trust-tracking-poll-health-information-and-advice-on-social-media/ [accessed 2026-05-20]
2. Vranica S, Kruppa M. Google's grip on search slips as TikTok and AI startup mount challenge. WSJ; The Wall Street Journal. 2024. URL: https://www.wsj.com/tech/online-ad-market-google-tiktok-9599d7e8 [accessed 2026-06-04]
3. TikTok what's next 2025 trend report: it's not magic, it's chemistry—introducing "Brand Chem". TikTok. 2025. URL: https://newsroom.tiktok.com/tiktok-whats-next-report-2025?lang=en-ZA [accessed 2026-06-04]
4. Liu J, Huang R, Ren J, Li P, Wang P. The intention to use short videos for health information among Chinese adults: based on the technology acceptance model. Digit Health 2025;11:20552076251335519. [doi: 10.1177/20552076251335519] [Medline: 40297363]
5. Ploderer B, Rezaei Aghdam A, Burns K. Patient-generated health photos and videos across health and well-being contexts: scoping review. J Med Internet Res 2022 Apr 12;24(4):e28867. [doi: 10.2196/28867] [Medline: 35412458]
6. Daft RL, Lengel RH. Organizational information requirements, media richness and structural design. Manage Sci 1986 May;32(5):554-571. [doi: 10.1287/mnsc.32.5.554]
7. Han P, Liu S, Zhang D, Li X, Li X. Research on the factors affecting the adoption of health short videos by the college students in China: unification based on TAM and UTAUT model. Front Psychol 2025;16:1547402. [doi: 10.3389/fpsyg.2025.1547402] [Medline: 40697736]
8. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
9. Malik A, Islam T, Ahmad M, Mahmood K. Health information seeking and sharing behavior of young adults on social media in Pakistan. J Librariansh Inf Sci 2023 Sep;55(3):579-595. [doi: 10.1177/09610006221090228]

XSL•FO
RenderX

10. Zhao YC, Zhao M, Song S. Online health information seeking among patients with chronic conditions: integrating the Health Belief Model and Social Support Theory. J Med Internet Res 2022 Nov 2;24(11):e42447. [doi: 10.2196/42447] [Medline: 36322124]
11. Priyambodo AZ, Zhafirah NA, Shabur R, Handayani PW, Fitriani H. Health information-seeking using short video platforms. Digit Health 2025 May;11:11. [doi: 10.1177/20552076251394619]
12. Hansen S, Jensen TS, Schmidt AM, Strøm J, Vistisen P, Høybye MT. The effectiveness of video animations as a tool to improve health information recall for patients: systematic review. J Med Internet Res 2024 Dec 30;26:e58306. [doi: 10.2196/58306] [Medline: 39753224]
13. Zhu Z, Liu S, Zhang R. Examining the persuasive effects of health communication in short videos: systematic review. J Med Internet Res 2023 Oct 13;25:e48508. [doi: 10.2196/48508] [Medline: 37831488]
14. Survei penetrasi internet dan perilaku penggunaan internet 2025. APJII. 2025. URL: https://survei.apjii.or.id/survei/register/48?type=free [accessed 2026-06-04]
15. Thang CJ, Garate D, Thang J, Lipoff JB, Barbieri JS. Short-form medical media: a multi-platform analysis of acne treatment information in TikTok videos, Instagram reels, and YouTube shorts. JMIR Dermatol 2023 Aug 9;6:e48140. [doi: 10.2196/48140] [Medline: 37624704]
16. Zhang J, Yuan J, Zhang D, et al. Short video platforms as sources of health information about cervical cancer: a content and quality analysis. PLoS One 2024;19(3):e0300180. [doi: 10.1371/journal.pone.0300180]
17. Ming S, Han J, Yao X, Guo X, Guo Q, Lei B. Myopia information on TikTok: analysis factors that impact video quality and audience engagement. BMC Public Health 2024 Apr 29;24(1):1194. [doi: 10.1186/s12889-024-18687-4] [Medline: 38685020]
18. Kirkpatrick CE, Lawrie LL. TikTok as a source of health information and misinformation for young women in the United States: survey study. JMIR Infodemiology 2024 May 21;4:e54663. [doi: 10.2196/54663] [Medline: 38772020]
19. Nagara MRND, Nurhajati L. The construction and adoption of beauty standard by youth female as the consumer of K-beauty products in Indonesia. Jurnal Riset Komunikasi 2022;5(2):258-277. [doi: 10.38194/jurkom.v5i2.543]
20. Ariana H, Almuhtadi I, Natania NJ, Handayani PW, Bressan S, Larasati PD. Influence of TikTok on body satisfaction among generation Z in Indonesia: mixed methods approach. JMIR Hum Factors 2024 Sep 6;11:e58371. [doi: 10.2196/58371] [Medline: 39241225]
21. Khan S, Yee D, Khan S, et al. Biologics to breast milk: a cross-sectional study of popular eczema treatment content on TikTok. Pediatr Dermatol 2022 Nov;39(6):920-922. [doi: 10.1111/pde.15106] [Medline: 35977757]
22. Awad N, Hetzel J, Bhupalam V, Nestor MS. A cross‐sectional content quality analysis of information in TikTok videos on “Dermarolling (Roller Microneedling)”. J Cosmet Dermatol 2024 Sep;23(9):2996-2999. [doi: 10.1111/jocd.16399]
23. Yu M, Guo Y, Wang J, Liu X. Antecedents of health short videos' information adoption and continuous usage intention: evidence from SEM and fsQCA. IMDS 2025 Apr 22;125(5):1869-1891. [doi: 10.1108/IMDS-06-2024-0569]
24. Zhang X, Chen B, Li G, Dong Y. Exploring the health information seeking behavior of social media users under the background of COVID-19 pandemic: an empirical study based on social cognitive theory. Front Psychol 2022;13. [doi: 10.3389/fpsyg.2022.1045914]
25. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
26. Adkins K, Overton PG, Moses J, Thompson A. Investigating the role of upward comparisons and self-compassion on stigma in people with acne: cross-sectional study. JMIR Dermatol 2023 Apr 12;6:e45368. [doi: 10.2196/45368] [Medline: 37632940]
27. Brzozowska JM, Gotlib J. Social media potential and impact on changing behaviors and actions in skin health promotion: systematic review. J Med Internet Res 2025 Jan 6;27:e54241. [doi: 10.2196/54241] [Medline: 39761100]
28. Hu B, Zhu Y, Bao R, Zhao Z, Liu C, Lin A. Social media, health consciousness, and cultural influences on sugar reduction behaviors in Chinese youth: extending the stimulus-organism-response model. J Med Internet Res 2025 Dec 19;27:e68180. [doi: 10.2196/68180] [Medline: 41417575]
29. Ahmad K, Lilani K. From scrolling to buying: role of beauty influencers on impulse buying a stimulus-organism-response perspective. Telemat Inform Rep 2025 Sep;19:100239. [doi: 10.1016/j.teler.2025.100239]
30. Niu Z, Willoughby JF, Coups EJ, Stapleton JL. Effects of website interactivity on skin cancer-related intentions and user experience: factorial randomized experiment. J Med Internet Res 2021 Jan 13;23(1):e18299. [doi: 10.2196/18299] [Medline: 33439131]
31. Tseng FC, Huang TL, Pham TTL, Cheng TCE, Teng CI. How does media richness foster online gamer loyalty? Int J Inf Manage 2022 Feb;62:102439. [doi: 10.1016/j.ijinfomgt.2021.102439]
32. Lin SH, Lin TMY. Demand for online platforms for medical word-of-mouth. J Int Med Res 2018 May;46(5):1910-1918. [doi: 10.1177/0300060518757899] [Medline: 29557232]
33. Lenczowski E, Tung-Hahn E, Higareda J, et al. Video education to improve recognition of common benign and malignant cutaneous lesions and skin cancer prevention in the public. Int J Womens Dermatol 2018 Jun;4(2):80-82. [doi: 10.1016/j.ijwd.2017.10.005] [Medline: 30023424]

XSL•FO
RenderX

34. Zhang X, Guo X, Guo F, Lai KH. Nonlinearities in personalization-privacy paradox in mHealth adoption: the mediating role of perceived usefulness and attitude. Technol Health Care 2014;22(4):515-529. [doi: 10.3233/THC-140811] [Medline: 24763205]
35. Nittas V, Mütsch M, Frey T, Braun J, Puhan MA. Effectiveness of a tailored web app on sun protection intentions and its implications for skin cancer prevention: a randomized controlled trial. PLoS Digit Health 2022 May;1(5):e0000032. [doi: 10.1371/journal.pdig.0000032] [Medline: 36812525]
36. Yang H, Zhang S, Diao Z, Sun D. What motivates users to continue using current short video applications? A dual-path examination of flow experience and cognitive lock-in. Telemat Inform 2023 Nov;85:102050. [doi: 10.1016/j.tele.2023.102050]
37. Haykal D. Digital twins in dermatology: a new era of personalized skin care. Front Digit Health 2025 Feb;7:1534859. [doi: 10.3389/fdgth.2025.1534859]
38. Xia L, Deng S, Liu Y. Seeking health information online: the moderating effects of problematic situations on user intention. J Data Inform Sci 2017 Mar 21;2(2):76-95. [doi: 10.1515/jdis-2017-0009]
39. Boyce L, Harun A, Prybutok G, Prybutok VR. The role of technology in online health communities: a study of information-seeking behavior. Healthcare (Basel) 2024 Jan 29;12(3):336. [doi: 10.3390/healthcare12030336] [Medline: 38338221]
40. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133. [doi: 10.2196/jmir.2143] [Medline: 23026508]
41. Jin XL, Yin M, Zhou Z, Yu X. The differential effects of trusting beliefs on social media users' willingness to adopt and share health knowledge. Inf Process Manag 2021 Jan;58(1):102413. [doi: 10.1016/j.ipm.2020.102413]
42. Claggett JL, Kitchens B, Paino M. Identifying the peripheral cues in the credibility assessment of online health information. Inf Manag 2024 Dec;61(8):104037. [doi: 10.1016/j.im.2024.104037]
43. Fan F, Chan K, Tsang L. Health information-seeking behavior and perceived information source credibility among middle-aged and older adults. Health New Media Res 2024 Jun;8(1):82-88. [doi: 10.22720/hnmr.2024.00073]
44. McMillan SJ, Hwang JS. Measures of perceived interactivity: an exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. J Advert 2002 Oct;31(3):29-42. [doi: 10.1080/00913367.2002.10673674]
45. Liu C, Jiang M, Muhammad ZA. The impact of TikTok short video factors on tourists' behavioral intention among generation Z and millennials: the role of flow experience. PLoS One 2024 Dec 5;19(12):e0315140. [doi: 10.1371/journal.pone.0315140]
46. Link E, Baumann E, Klimmt C. Explaining online information seeking behaviors in people with different health statuses: German representative cross-sectional survey. J Med Internet Res 2021 Dec 10;23(12):e25963. [doi: 10.2196/25963] [Medline: 34890348]
47. Mirzaei A, Aslani P, Luca EJ, Schneider CR. Correction: predictors of health information–seeking behavior: systematic literature review and network analysis. J Med Internet Res 2021 Jul 3;24(6):e39705. [doi: 10.2196/39705]
48. Li H, Li D, Zhai M, Lin L, Cao Z. Associations among online health information seeking behavior, online health information perception, and health service utilization: cross-sectional study. J Med Internet Res 2025 Mar 14;27:e66683. [doi: 10.2196/66683]
49. Neely S, Eldredge C, Sanders R. Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: survey study. J Med Internet Res 2021 Jun 11;23(6):e29802. [doi: 10.2196/29802] [Medline: 34043526]
50. Tian J, Li B, Zhang R. The impact of upward social comparison on social media on appearance anxiety: a moderated mediation model. Behav Sci (Basel) 2024 Dec 26;15(1):8. [doi: 10.3390/bs15010008] [Medline: 39851813]
51. Van Beugen S, Schut C, Kupfer J, et al. Perceived stigmatization among dermatological outpatients compared with controls: an observational multicentre study in 17 European countries. Acta Derm Venereol 2023 Jun 22;103:adv6485. [doi: 10.2340/actadv.v103.6485] [Medline: 37345973]
52. Tunca M, Balik ZB, Uncu HB, Botsali A. Stigmatization and psychosocial burden in acne patients. J Cosmet Dermatol 2025 Feb;24(2):e16678. [doi: 10.1111/jocd.16678] [Medline: 39692178]
53. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
54. Aslan Kayıran M, Karadağ AS, Oğuz Topal İ, et al. Habits of using social media and the internet in psoriasis patients. Dermatol Pract Concept 2022 Jul 1;12(3):e2022143. [doi: 10.5826/dpc.1203a143] [Medline: 36534567]
55. Nguyen SH, Vu GT, Nguyen LH, et al. Understanding social media use and engagement among dermatology patients to inform dermatological prevention and care in Vietnam: cross-sectional study. JMIR Dermatol 2020;3(1):e13424. [doi: 10.2196/13424]
56. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. Health Info Libr J 2017 Dec;34(4):268-283. [doi: 10.1111/hir.12192] [Medline: 29045011]
57. Ginsburg IH, Link BG. Feelings of stigmatization in patients with psoriasis. J Am Acad Dermatol 1989 Jan;20(1):53-63. [doi: 10.1016/s0190-9622(89)70007-4] [Medline: 2913081]

XSL•FO
RenderX

58. Stangl AL, Earnshaw VA, Logie CH, et al. The health stigma and discrimination framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Feb 15;17(1):31. [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]
59. Rowlands IJ, Loxton D, Dobson A, Mishra GD. Seeking health information online: association with young Australian women's physical, mental, and reproductive health. J Med Internet Res 2015 May 18;17(5):e120. [doi: 10.2196/jmir.4048] [Medline: 25986630]
60. Silver RA, Johnson C. Health information seeking behavior on social networking sites and self-treatment: pilot survey study. Online J Public Health Inform 2023;15:e51984. [doi: 10.2196/51984] [Medline: 38179207]
61. Griffin RJ, Dunwoody S, Yang ZJ. Linking risk messages to information seeking and processing. Ann Int Commun Assoc 2013 Jan;36(1):323-362. [doi: 10.1080/23808985.2013.11679138]
62. Richard MA, Saint Aroman M, Baissac C, et al. Burden of visible [face and hands] skin diseases: results from a large international survey. Ann Dermatol Venereol 2023 Jun;150(2):95-100. [doi: 10.1016/j.annder.2022.11.008] [Medline: 36653227]
63. Zhou X, Roberto AJ. An application of the risk information seeking and processing model in understanding college students' COVID-19 vaccination information seeking and behavior. Science Communication 2022 Aug;44(4):446-474. [doi: 10.1177/10755470221120415] [Medline: 38603181]
64. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1-13. [doi: 10.1177/1609406917733847]
65. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
66. Fuller CM, Simmering MJ, Atinc G, Atinc Y, Babin BJ. Common methods variance detection in business research. J Bus Res 2016 Aug;69(8):3192-3198. [doi: 10.1016/j.jbusres.2015.12.008]
67. Hair JF, Babin BJ, Ringle CM, Sarstedt M, Becker JM. Covariance-based structural equation modeling (CB-SEM): a SmartPLS 4 software tutorial. J Market Anal 2025 Sep;13(3):709-724. [doi: 10.1057/s41270-025-00414-6]
68. Armstrong AW, Kim RH, Idriss NZ, Larsen LN, Lio PA. Online video improves clinical outcomes in adults with atopic dermatitis: a randomized controlled trial. J Am Acad Dermatol 2011 Mar;64(3):502-507. [doi: 10.1016/j.jaad.2010.01.051] [Medline: 21236514]
69. Ma T, Atkin D. User generated content and credibility evaluation of online health information: a meta analytic study. Telemat Inform 2017 Aug;34(5):472-486. [doi: 10.1016/j.tele.2016.09.009]
70. Denniss E, Lindberg R, McNaughton SA. Development of principles for health-related information on smedia: Delphi study. J Med Internet Res 2022 Sep 8;24(9):e37337. [doi: 10.2196/37337] [Medline: 36074544]
71. Joseph J. The algorithmic self: how AI is reshaping human identity, introspection, and agency. Front Psychol 2025;16:1645795. [doi: 10.3389/fpsyg.2025.1645795] [Medline: 40718563]
72. Shin D. Automating epistemology: how AI reconfigures truth, authority, and verification. AI Soc 2026 Feb;41(2):1553-1559. [doi: 10.1007/s00146-025-02560-y]
73. Hossain MS, Islam T, Babu M, et al. The influence of celebrity credibility, attractiveness, and social media influence on trustworthiness, perceived quality, and purchase intention for natural beauty care products. Clean Responsible Consum 2025 Jun;17:100277. [doi: 10.1016/j.clrc.2025.100277]
74. Huang W, Wang X, Zhang Q, Han J, Zhang R. Beyond likes and comments: how social proof influences consumer impulse buying on short-form video platforms. J Retail Consum Serv 2025 May;84:104199. [doi: 10.1016/j.jretconser.2024.104199]
75. Hussain B, Aslam S, Imran A. Manufacturing beauty: how AI and social media are redefining aesthetic norms in emerging digital cultures. Acta Psychol (Amst) 2025 Oct;260:105734. [doi: 10.1016/j.actpsy.2025.105734] [Medline: 41075470]

Abbreviations

AVE: average extracted variance
CA: Cronbach α
CFA: confirmatory factor analysis
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CMB: common method bias
GoF: goodness of fit
HBM: health belief model
HCE: health content expressiveness
MRT: media richness theory
PHI: personalized health insight
PI: perceived interactivity
PSE: perceived severity
PSU: perceived susceptibility

XSL•FO
RenderX

PU: perceived usefulness
SC: source credibility
SHE: skin health information seeking on short video platform
SOR: stimulus-organism-response
SST: skin stigmatization
USC: upward skin comparison

Edited by R Dellavalle; submitted 13.Feb.2026; peer-reviewed by K Slaton, M Mardiana; revised version received 08.May.2026; accepted 09.May.2026; published 03.Jul.2026.

Please cite as:
Ahimsa FQ, Saputra K, Hutasoit MEA, Handayani PW, Fitriani H
Skin Health Information Seeking on Short Video Platforms in Indonesia: Mixed Methods Approach
JMIR Dermatol 2026;9:e93461
URL: https://derma.jmir.org/2026/1/e93461
doi:10.2196/93461

© Fathi Qushoyyi Ahimsa, Kelvin Saputra, Michelle Elizabeth Amanda Hutasoit, Putu Wuri Handayani, Hana Fitriani. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Harmonized Dual Deep Learning Architectures for Image-Based Diagnostics of Skin Neglected Tropical Diseases: Benchmark Study via Novel Funnel Framework

Yohannes Minyilu[1*], MSc; Mohammed Abebe Yimer[1*], PhD; Million Meshesha[2*], PhD

[1]Faculty of Computing and Software Engineering, Institute of Technology, Arba Minch University, SE, Arba Minch, Ethiopia

[2]School of Information Science, College of Natural and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia

[*]all authors contributed equally

Corresponding Author:
Yohannes Minyilu, MSc
Faculty of Computing and Software Engineering, Institute of Technology, Arba Minch University, SE, Arba Minch, Ethiopia

Abstract

Background: While deep learning–based methods are the potential technological solutions for the diagnosis of skin Neglected Tropical Diseases (skin NTDs), limited efforts were seen toward the use of such tools in Ethiopia. Data scarcity, methods, and models selection issues created further challenges in an attempt to close the previous gap.

Objective: This study attempts to design a benchmark image-based diagnostic model for skin NTDs through a synergistic combination of feature extraction pretrained models, a custom-designed convolutional neural network (CNN) model trained on the extracted features, and an integrated data augmentation method applied dynamically.

Methods: For this study, a new skin images dataset is created using skin photographs collected by a team of researchers from the NTDs research center of Arba Minch University Medical College. The new dataset contains 1495 images in 3 classes having severe class imbalance. Extensive experiments were conducted to find the optimal deep learning approach by designing a new CNN model, applying transfer learning, and designing the 2-stage approach that uses pretrained models for feature extraction and trains the new CNN model using the extracted features from the pretrained models and applying data augmentation based on the integrated 2-stage approach. For model selection, the study proposed a novel approach, the funnel framework with cascaded selection of methods and models.

Results: After hyperparameter tuning, the model trained using DenseNet121 feature extractor scored the highest accuracy of 96.6%, F_1-score of 95%, and sensitivity of 95%, while the MNv2-based model scored comparable results of 95.6% accuracy, 90% F_1-score, and 90% sensitivity. This study finally selected the DenseNet121 and MNv2 models for feature extraction to build the final model for skin NTDs classification.

Conclusions: The 2-stage approach significantly boosted the models' performance compared with other methods, while the data augmentation method further enhanced the performance of the selected models. Finally, this study suggests further studies using advanced class-balancing methods with more data and a possible integration of other clinical data types.

(JMIR Dermatol 2026;9:e91544) doi:10.2196/91544

KEYWORDS

skin NTDs classification; 2-stage approach; feature extraction; funnel framework; hyperparameter optimization

Introduction

Background

Neglected Tropical Diseases (NTDs) represent 21 different diseases, including podoconiosis, scabies, and tungiasis, affecting more than 1 billion people globally among underserved communities in the tropical areas [1]. As a tropical country, NTDs are highly prevalent in Ethiopia, with the majority of NTDs identified by the World Health Organization present except for Chagas disease and yaws [2], particularly in the remote areas of the country [3,4]. As estimated by the federal ministry of health [5], more than 75 million people are at risk of contracting at least 1 NTD. Of the 21 NTDs recognized by the World Health Organization, more than half (about 18 of them) have skin manifestations and are called skin NTDs [6-8]. Accordingly, the diagnosis of skin NTDs primarily involves examination of patient skin, presenting opportunities for integrated diagnosis [8], involving the use of artificial intelligence–based skin NTDs diagnostic tools [9]. Although it is not fully explored, previous studies [10-12] have shown that deep learning (DL) methods can be used for the diagnosis of skin NTDs.

XSL•FO
RenderX

This study proposes a DL-based diagnostic model for skin NTDs based on skin images of patients using a new skin images dataset we created for this study. However, apart from the limited previous efforts in DL-based skin NTDs diagnostics, dataset-related issues created challenges in building our proposed model, as the dataset used is characterized by small-sized image samples with a severe class imbalance. Generally, as a computer vision task, the development of a DL-based diagnostic model using skin images requires a large-scale higher-quality skin images dataset [11]. In the case of skin NTDs, data scarcity and dataset-related issues are the major challenges in building DL-based intelligent diagnostic tools for skin NTDs [13,14]. These challenges mainly arise from several factors that include poor record keeping, management, and reporting practices [13]. Additionally, model characteristics regarding usability and efficiency issues are also major challenges in building diagnostic models for skin NTDs. Beyond prediction accuracy, efficiency parameters, such as model complexity, inference speed, response time, and deployment platform options (web-based and mobile-based), are also major factors that determine the selection of DL tools, techniques, and models.

Accordingly, given the data-related issues of the dataset we used for this study and the expected operational platforms, which DL method is appropriate to develop an image-based diagnostic model for skin NTDs? While answering this question requires properly devised strategies based on carefully designed experiments, analysis, and interpretation of results, it establishes a foundational benchmarking effort that helps in identifying suitable DL methods to overcome the mentioned challenges. Therefore, this study conducts extensive experimentations to find the optimal DL solution based on the following guiding questions: (1) Which DL approach (baseline model design, transfer learning, or hybrid) would be a feasible strategy to develop the proposed model? (2) Which DL model architecture would help in developing a high-performance diagnostic model for skin NTDs, given the nature of the dataset used? (3) Which approach helps in creating the robust model development pipeline based on the experimental screening of both DL methods and architectures that collectively address high predictive performance with lower architectural and computational complexity?

While addressing these questions, this study develops a benchmark image-based diagnostic DL model for skin NTDs based on experimentally identified suitable methods and approaches that involve designing a new convolutional neural network (CNN) model and applying transfer learning. Evidently, the data scarcity created higher difficulty in capturing relevant features from input images using baseline models, including the custom-designed model, given the diversity and nature of the manifestations of skin NTDs that include “mossy” limbs in podoconiosis [14]. While the use of transfer learning is ultimately the recommended DL strategy, several factors related to pretrained models create challenges that include huge data requirements, domain incompatibility, tendency of capturing noise features, and lack of a standardized robust diagnostic model development pipeline including skin NTDs. To address such challenges, we designed and implemented the 2-stage approach that presents a robust and extensible architectural pipeline integrating the feature mapping (extraction) models and applying domain adaptation. Furthermore, as the 2-stage approach integrates the 10-layer classification head with different regularization methods, it provides deeper feature filtering architectures.

Overall, this study presents several contributions to the problem domain (skin NTDs diagnostics) and to the field through multiple achievements, which include identification of optimal DL methods for skin NTDs that have higher data scarcity problems; establishment of a robust DL model development pipeline, which incorporates designing the 2-stage approach; methodological rigor that includes robust experimental setup and systematic architectural screening by adopting the funnel framework; and, ultimately, development of DL diagnostic models for skin NTDs, which can serve as an architectural benchmark for skin NTDs (skin-related diseases in general).

Related Works

Previous studies showed the potential of the DL-based methods for skin NTDs. Accordingly, Steyve et al [10] proposed an optimized real-time diagnostic approach for 3 skin NTDs (Buruli, leishmaniasis, and leprosy) using a support vector machine classifier optimized by a black hole optimization algorithm. Yotsu et al [15] also presented DL methods for using major CNN architectures (ResNet50 and VGG16 models) for the diagnosis of 5 skin NTDs (Buruli ulcer, leprosy, mycetoma, scabies, and yaws). Another study by Pattnayak et al [16] proposed a DL method for 5 skin NTDs (Buruli ulcer, leprosy, mycetoma, scabies, and yaws). Beesetty et al [17] applied a Siamese-based Few Shot Learning model, trained it on an extremely small dataset with fewer disease classes (368 clinically diagnosed leprosy and 28 nonleprosy skin lesions), and reported higher accuracy.

Challenges Toward Applying DL Methods for Skin NTDs

Multiple factors created challenges toward digitizing the skin NTDs diagnostic procedures using intelligent digital diagnostic tools. Some of the challenges are insufficient infrastructure, data security issues, and limited efforts toward the integration of digital diagnostic tools [18], including the general DL model development challenge, data scarcity, and class imbalance. Multiple large-scale skin image datasets, such as the HAM10000 (Human Against Machine with 10,000 training images) [19] and ISIC (International Skin Imaging Collaboration) [20], are available to train DL models for non-NTD skin diseases. However, it is difficult to find such massive skin image datasets that are publicly available to train DL models for skin NTDs. The other major issues related to data scarcity are the completeness and class distribution imbalance.

Potential DL Solutions for Intelligent Skin NTDs Diagnosis

Data augmentation, mostly for image classification tasks, is the most widely used machine learning operation to address the problem of data scarcity and distribution imbalance [21-23] through artificially generating images. There are 2 major approaches of data augmentation. The traditional augmentation

XSL•FO
RenderX

method uses basic general geometric transformations [21,24], such as cropping, padding, flipping (horizontal or vertical), rotations, permutations, scaling, translations, and addition of noise [23,25]. The cropping, flipping, rotations, and scaling transformations are applied to simulate imaginary variations that might occur in reality. The other data augmentation approach is the class-based conditional augmentation, which is conducted based on predefined conditions either by applying basic geometric transformations or by using generative adversarial (GAN) models [23,26,27].

Gaps Identified

Overall, the vast literature exploration confirmed that very few efforts were seen toward the use of DL-based diagnostic tools for skin NTDs, specifically in the Ethiopian context. Additionally, the data scarcity issues are the subsequent challenges creating the other major research gaps. These major research gaps clearly suggest that further efforts in the area are mandatory, with a clear indication of having an initially established DL-based diagnostic framework that can serve as a benchmark for current and future studies. Therefore, we conducted this foundational study to develop an image-based skin NTDs diagnostic model using a novel skin NTDs image dataset created by using skin photographs of patients collected from one of the remote and highly affected areas in Ethiopia. We conducted this benchmarking study to identify optimal methods and DL model architectures based on properly designed experimental settings, given the nature of the dataset used for the study. Regarding the dataset-related problems, the study demonstrates the dynamic (online) data augmentation method based on the standard general geometric transformations to initially address the data scarcity problem.

Methods

Ethical Considerations

For this study, we created a new dataset using skin photographs of patients with skin NTDs obtained during clinical data collection from a remote affected area in the southwest of Ethiopia. While the data were initially collected by trained health care professionals based on strict ethical procedures using an ethical clearance letter obtained from the institutional review board (IRB) in Arba Minch University (AMU), we acquired the data through institutional research collaboration ("Acknowledgments" section) that warrants full access and use of the firsthand data. To acquire and use the skin images data for this study, proper ethical approval procedures were followed, starting with the acquisition of the ethical clearance letter from the IRB in AMU. Hence, for this study, the authors obtained a specific ethical clearance letter from the concerned institutional review board in AMU (approval protocol number YM23161).

The data collection process was conducted with the full respect of participants' privacy, where data were acquired from each individual participant based on their will confirmed by a written consent. This was achieved using a dedicated checklist prepared and used to systematically monitor the data collection, ensuring that they were collected based on the free will of each participant. This, specifically, involved a mandatory checklist to ensure that each participant has read and signed the consent form using targeted questions, which include the following: "Does the participant read information statement and willing to sign the consent form?" and the consideration of underaged participants—"if the age is below 18, does the guardian sign the consent form?" Overall, all involved participants provided written consent for their participation and authorized the use of their clinical data for academic purposes, including publications.

As the data collection was performed by a team of trained health care professionals, including public health officers, the entire data collection process was carried out in a professional and responsible manner, authorized by the IRB. Regarding this study, we have been authorized by the same institutional ethical review committee before acquiring and using the data based on a series of verification procedures that include careful analysis of the required data for our study, analysis of participant privacy vulnerability issues as a result of using the data, anonymizing all participant-level data in the images to remove all information in the images that could be used to identify patients (study participants), and strictly complying with ethical requirements for using human patient data—which was confirmed by the ethical clearance letter from the IRB. Therefore, before using the images for this study, all images have been anonymized to remove any personal data to ensure that all participant-identifiable features in any images of the manuscript or supplementary material are not visible.

Finally, as the data were collected from one of the highly affected areas during a mass drug administration (MDA) campaign, the ultimate goal of the data collection was to assess the burden of the skin NTDs that will be used for immediate public health decisions. In this process, the patients living in the affected remote community primarily benefited from the MDA-based data collection, with the patients being diagnosed at the MDA site. However, no special compensations were implemented for the participants as a result of using the data for further study. Additionally, the acquired data are used to build a diagnostic model that intends to serve for diagnosing skin NTDs in the same resource-limited areas.

Data Collection and Dataset Description

In this study, we used a new handcrafted dataset containing skin photographs of patients with skin NTDs that were captured to show skin areas affected by the skin NTDs. Initially, the data were collected by a team of researchers from the Collaborative Research and Training Center for Neglected Tropical Disease, College of Medicine Health Sciences of AMU. Data were collected in a project-based research for the assessment of skin NTDs burden through community screening during the scabies MDA campaign from Gacho Baba District, Gamo Zone, southwest of Ethiopia. The dataset contains skin photographs (images) of 3 different skin NTDs, namely, podoconiosis, scabies, and tungiasis. These 3 diseases were included in the dataset since they are identified as the most prevalent skin NTDs identified in the specified affected area. The entire data collection process was conducted in a professional and ethical manner, where the whole process was initiated after all legal and ethical issues were addressed, and an ethical clearance was obtained to collect the data. For this study, we acquired the collected data through institutional research collaboration

XSL•FO
RenderX

between the NTDs research center of the medical college and computing faculty of the technology institute of AMU. Using the acquired data, we created a new skin NTDs image dataset and used it for this study to develop the proposed DL-based skin NTDs diagnostic model.

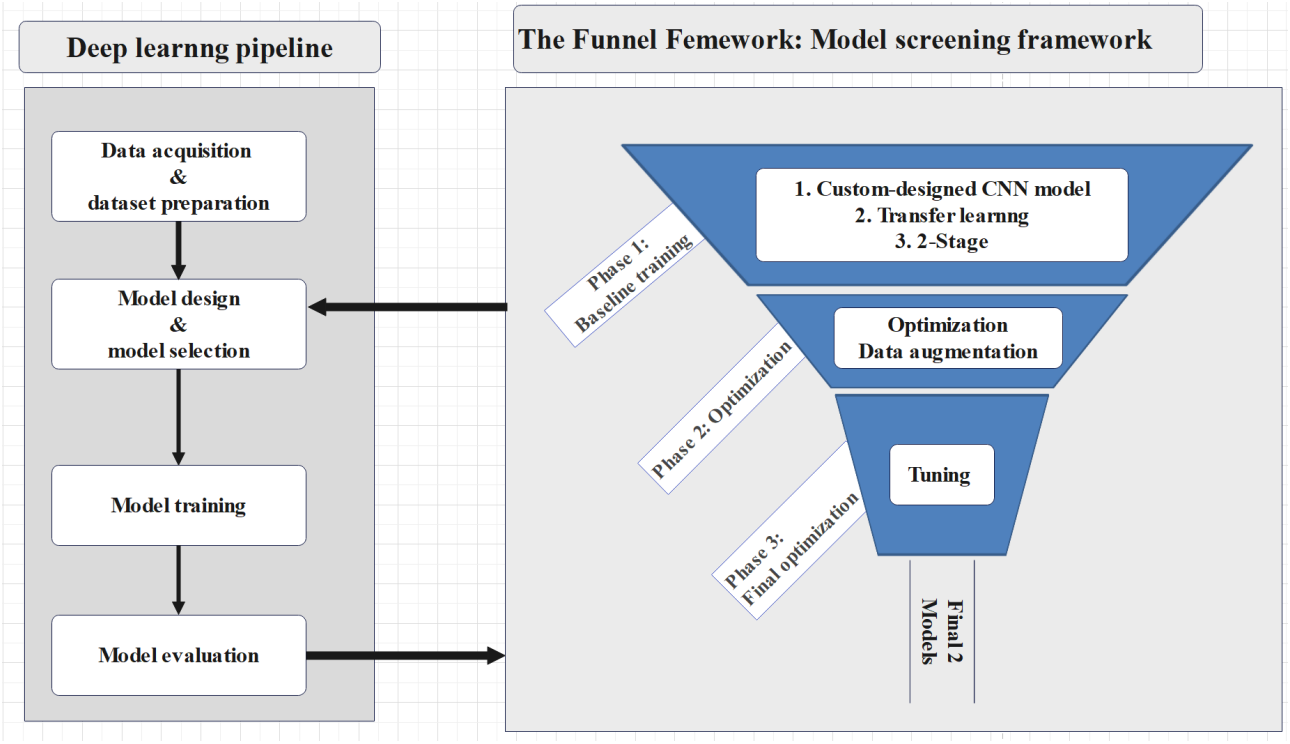
Exploratory Data Analysis

The images were obtained for dermatologist verification and present skin lesions, scratches, excoriations, and other infestations that are typical clinical signs of podoconiosis, scabies, and tungiasis. After acquiring the data, we created 3 separate initial datasets containing the skin images, the unique image IDs, and their corresponding labels for each disease. The final dataset contains 1495 images, as shown in Multimedia Appendix 1. As shown in the statistical distribution, scabies has the largest proportion among the 3 disease classes with a total of 955 instances (955/1495, 63.88%), while tungiasis represents the second largest size with 474 instances (474/1495, 31.71%) and podoconiosis having only 66 instances (66/1495, 4.41%).

Study Design

This study intends to conduct a DL architectural benchmarking research that requires a systematic approach to select optimal DL methods and algorithms, and we proposed a mixed research strategy based on the newly proposed funnel framework we adopted for this study, as summarized in Figure 1. Overall, our study needs a systematic approach to select optimal DL methods and models through multistaged experimental filtering. Specifically, the selection of DL pretrained models requires multiple experimentations with systematically devised screening and filtering methods based on comparative analysis of model performance results. Initially, we propose and experiment with 3 different DL methods: first, train a new custom-designed CNN model; second, transfer learning using the 21 selected pretrained DL models; and third, demonstrate the 2-stage approach.

Figure 1. Architecture of the proposed funnel framework based on the cascaded model selection mechanism, CNN: convolutional neural network.



After the completion of these 3 experiments, the best approach and top 5 pretrained DL models will be selected to apply further enhancements to finally select the best 2 models. All these tasks required a systematically designed approach that can be used as a framework to guide the training, screening, and analysis processes. Therefore, we propose a new approach, the funnel framework with cascaded (phased) selection of models and DL methods, as shown in Figure 1. The funnel framework, adapted from the business-related fields, is used to screen out top-performing models and methods initially identified based on a comparative and phased approach. Accordingly, the experiments are conducted in different training settings to select the optimal DL approach that produces comparatively maximum model performance and screen out the high-performing pretrained models in transfer learning and feature extraction.

Dataset Preparation and Preprocessing

The entire data-splitting process is conducted using the stratified splitting approach with a ratio of 80:20 train-test split, followed by data preprocessing operations. Accordingly, for image resizing, we applied the standard image resolution to all images in the dataset, which required resizing the images to 224 × 224 × 3 pixels for 17 models, while 240 × 240 × 3 pixels and above were used to resize the images for higher EfficientNet models (B1, B3, B5, and V2S). As a next task in data preprocessing, image normalization was programmatically applied to all images.

Model Development: Model Design and Selection

In this study, we use 3 strategies regarding the development of the proposed DL model, which include (1) training a new custom-designed CNN model, (2) applying transfer learning

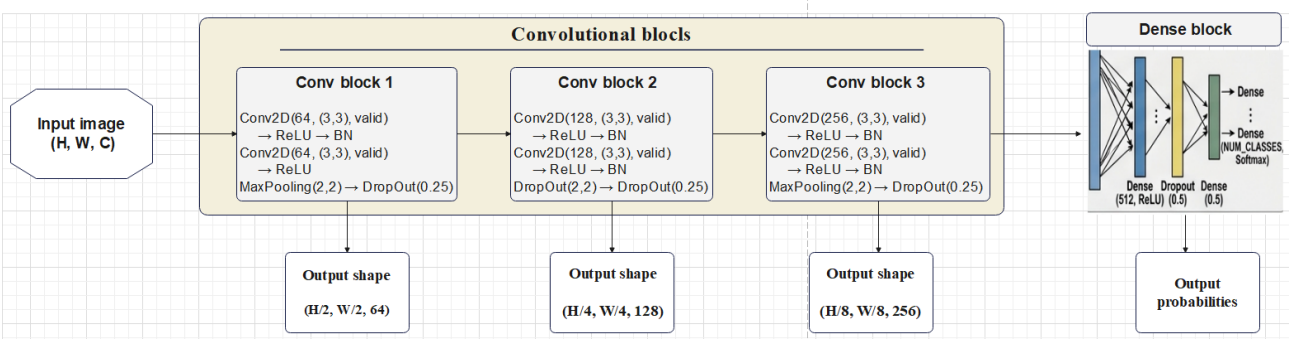
XSL•FO
RenderX

using selected pretrained DL model architectures, and (3) applying the proposed 2-stage approach—a hybrid approach that uses pretrained DL architectures for feature extraction and training the new CNN model for classification. We applied this set of strategies to experimentally demonstrate possible methods to select the optimal strategy resulting in overall higher performance of models.

The New CNN Model

As a first study, we begin our experiments by designing a new custom-designed CNN model that will be used for baseline training and evaluation as well as for the classification of the skin NTDs in the 2-stage approach. Accordingly, we designed a new CNN model consisting of 3 major components (blocks) that represent the 3 different stages of the entire pipeline: feature extraction, the dense layers (including the flattening layer), and the classification head. Based on this general architectural layout, the new model is designed to have 30 layers, containing 8 weight-bearing layers from 6 convolutional and 2 dense layers, 11 regularization layers properly applied across all blocks, 7 activation layers, and 4 spatial refinement layers—3 pooling and 1 flatten layer. Figure 2 visualizes the overall architecture of our custom-designed CNN model.

Figure 2. Architecture of the proposed custom-designed convolutional neural network model.



The model is designed in a consistent manner, implementing a hierarchical structure to capture complex and relevant visual patterns from each input sample image using 3 repeated convolutional blocks having variable filter sizes. Accordingly, the output depth of each block (number of filters) is sequentially doubled across the 3 feature extraction blocks (64 filters in block 1 to 256 filters in block 3). Internally, each feature extraction block is designed to have 2 convolutional layers (Conv2D with 3 × 3 kernels), including properly applied activation, regularization, and dimensionality reduction techniques. Accordingly, the model uses the “relu” activation function, along with robustly implemented regularization methods, which include (1) normalization (BatchNormalization)—applied across all blocks including the dense layers, (2) dropout—applied both in the feature extraction blocks (using Dropout [0.25]) and in the dense block (using Dropout [0.5]), and (3) pooling layers that apply the maximum pooling method—serving dual purposes including regularization, while the maximum pooling method is primarily used for dimensionality reduction [28]. This architectural setup is used to ensure that each feature extraction block extracts and refines features and applies normalization, all before dimensionality reduction (Dropout), to ensure that relevant features and patterns are captured. The overall feature dimension transformations and output shapes of each extraction block are presented in Table 1.

Table . Architectural summary of the new custom-designed CNN[a] model.

Model block	Input shape	Convolutional layers	After pooling (2 × 2)	Output
Input layer	224 × 224 × 3	—[b]	—[b]	3
Block 1 (convolutional block 1)	224 × 224	220 × 220	110 × 110	64
Block 2 (convolutional block 2)	110 × 110	106 × 106	53 × 53	128
Block 3 (convolutional block 3)	53 × 53	49 × 49	24 × 24	256
Dense block	Flattened vector	—[b]	512 units	3: number of classes

[a]CNN: convolutional neural network.

[b]Not available.

On the final block (classification head), the model applies feature map transformation and final skin NTDs classification using 6 layers: flatten—the layer used to transform the final spatial feature map into a 1D feature vector, fully connected (Dense)—the largest layer in the model having 512 units, regularization—2 independent layers applying BatchNormalization and heavier dropout (Dropout [0.5]), activation layer using “relu,” and output layer—the final layer that predicts the probability distribution among the 3 disease classes (podoconiosis, scabies, and tungiasis) using the SoftMax

XSL•FO
RenderX

activation function. Overall, the Adam optimizer and "categorical cross entropy" loss function are used for the final model compilation. All these strategies are properly applied along with a synchronized implementation of early stopping, all to prevent overfitting.

Given the nature of our new dataset, having only 1495 images in the dataset, training a DL model from scratch appeared to be a bit challenging, as the limited size of the feature maps would potentially force the models to learn all the details (including noise pixels) resulting in difficulty to generalize well on new skin NTDs images due to overfitting [28,29]. To overcome this challenge, we conducted further experimental inquiries to identify and use the optimal DL method based on our newly designed CNN model, demonstrating the transfer learning method entirely based on pretrained DL models, followed by the 2-stage approach.

Transfer Learning

To improve the performance of our baseline CNN model and demonstrate the other potential DL methods, we deployed the transfer learning method by using a diverse set of pretrained DL architectures (CNN and transformer-based). To achieve this goal, we applied a systematic DL architectural selection procedure using a predefined set of model selection parameters to validate, comparatively analyze, and finally select the best model and method for the proposed skin NTDs diagnostic model. Hence, we identified 21 pretrained model architectures, selected based on 4 major selection parameters, which include architectural distribution—defining model family and operational principles (CNN and transformer-based models), model complexity—including both architectural (model size) and computational (efficiency) complexity, and novelty (recency) of models. Table 2 presents a summary of the 21 selected pretrained models.

XSL•FO
RenderX

Table . Summary of DL[a] model architectures considered during initial screening.

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
ResNet50	Residual Networks	Skip connections [30]	25.6 M	Large	4.1	Heavy
ResNet18	Residual Networks	Skip connections [30]	11.3 M	Moderate	1.8	Moderate
ConvNext-Small	Modern CNN[c] Architectures	Transformer-like CNN components using standard ConvNet modules [31]	49.7 M	Large	4.5	Heavy
ConvNext-Tiny	Modern CNN[c] Architectures	Transformer-like CNN components using standard ConvNet modules [31]	28M	Moderate	4.5	Heavy
CovNeXtv2-Tiny	Modern Pure CNN Architectures	Fully convolutional masked autoencoder framework with a global response normalization layer [32]	28.1M	Large	4.47	Moderate
CovNeXtv2-Atto	Modern Pure CNN Architectures	Fully convolutional masked autoencoder framework with a global response normalization layer [32]	3.7M	Lightweight	0.55	Lightweight
DenseNet121	Densely Connected Networks	Dense CNN (DenseNet) block architecture with feature reuse [33]	7.3M	Moderate	2.9	Moderate
Xception	Depth-wise Separable CNNs	Depth-wise separable convolutions [34]	21.4	Large	8.4	Heavy
EfficientNetB5	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	29M	Large	9.9	Heavy
EfficientNetB3	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	11.2M	Moderate	1.8	Moderate
EfficientNetB1	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	6.9M	Moderate	0.7	Moderate

XSL•FO
RenderX

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
EfficientNetB0	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	4.4M	Lightweight	0.39	Lightweight
MobileNetV2	Mobile CNNs	Inverted residual blocks with linear bottlenecks [36]	2.6M	Lightweight	0.3	Lightweight
MobileNetV3-Large	Mobile CNNs	Hardware-aware NAS[e] along with the NetAdapt algorithm (platform-aware adaptation [37,38]	3.2M	Lightweight	0.22	Lightweight
MobileNetV3-Small	Mobile CNNs	Hardware-aware NAS[e] along with the NetAdapt algorithm (platform-aware adaptation [37,38]	1.1M	Lightweight	0.06	Lightweight
EfficientNetV2B0	Advanced compound-scaled CNNs	Training-aware NAS and scaling—joint optimization of training speed and parameter efficiency [39]	6.2M	Moderate	0.72	Moderate
EfficientNetV2S	Advanced compound-scaled CNNs	Training-aware NAS and scaling—joint optimization of training speed and parameter efficiency [39]	20.7M	Moderate	2.9	Moderate
Recent CNN and transformer-based models						
RepViT	ViT[f]-Inspired Pure lightweight CNN	Reparameterization convolutions in ViT-like Meta-Former structure [40]	5.1M	Moderate	0.80	Lightweight
FasterViT-0	Hybrid (CNN + ViT)	HAT[g] using window-based self-attention, carrier tokens for local-global representation learning [41]	31.4M	Large	3.34	Moderate
FastViT	Hybrid (CNN + ViT)	RepMixer (structural reparameterization for token mixing) and skip-connection elimination [42]	3.6M	Lightweight	0.70	Lightweight

XSL•FO
RenderX

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
EfficientViTB0	Hybrid (CNN + ViT)	Lightweight multi-scale attention (for context extraction) and MBConv (for local information extraction) [43]	0.7M	Lightweight	0.07	Lightweight

[a]DL: deep learning.

[b]GFLOPs: Giga floating point operations.

[c]CNN: convolutional neural network.

[d]BMConv: mobile-inverted bottleneck convolution.

[e]NAS: neural architecture search.

[f]ViT: vision transformer.

[g]HAT: hierarchical attention.

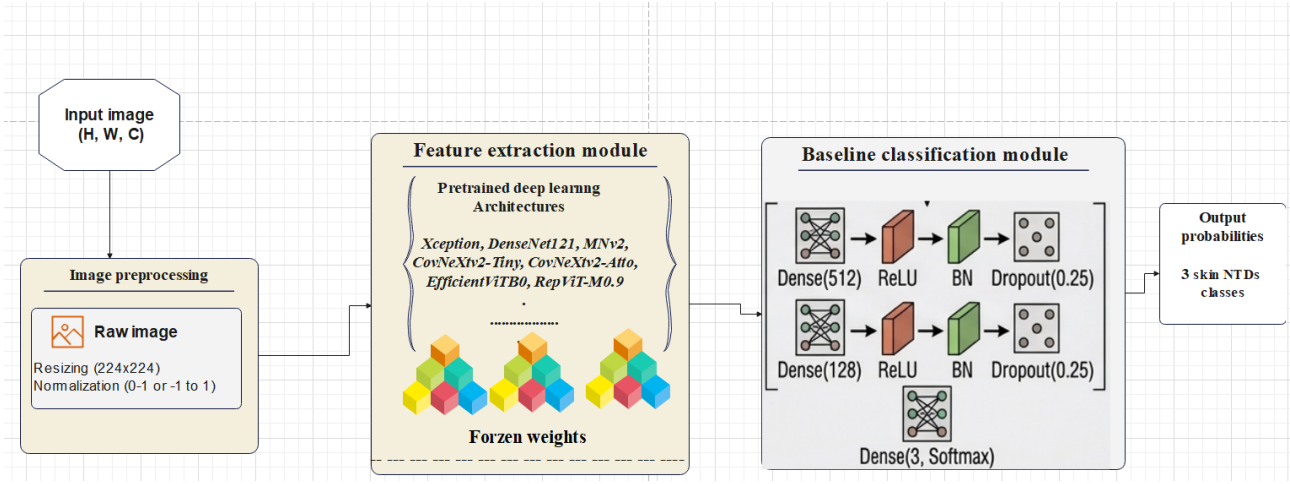
All the 21 selected pretrained models were initially trained on ImageNet-1K dataset, initialized with weights from this standard, large, and diversified dataset having around 3.2 million images [44], helping to create general feature extraction baselines. Furthermore, as shown in Table 1, maintaining architectural distribution, we selected the 21 representative models from nine architectural families, which include (1) models of residual networks (ResNet50 and ResNet18) that apply residual learning (or skip connections) method, (2) modern CNN architectures (ConvNextV1 Small/Tiny and ConvNextV2 Tiny/Atto) that have transformer-like CNN components, (3) model with densely connected CNNs (DenseNet121), (4) model with depth-wise separable CNNs (Xception), (5) the EfficientNet family (B0-B5), (6) mobile CNNs MobileNet (V2 and V3), (7) models that apply advanced compound scaling method (EfficientNetV2 B0 and S), (8) recent state-of-the-art lightweight CNN architecture (RepViT), and (9) hybrid transformer-based architectures (FastViT, FasterViT-0, and EfficientViTB0). Regarding architectural and computational complexities, 6 of these models are heavier models having parameters between 30 and 50 million (M) (30 M < parameters < 50 M), while having computational complexities that range between 4.0 and 9.9 Giga floating point operations (GFLOPs). The other 8 models have moderate levels of complexities (6.2 M < parameters < 28.1 M, and 0.7 billion [B] < FLOP < 4.0 B), while 7 of the 21 models are lightweight models (0.7 M < parameters < 5.1 M, and 0.06 B < FLOP < 0.55 B).

The 2-Stage Approach: Feature Extraction With Integrated CNN Model

On the third experimental setting, our proposed 2-staged approach is demonstrated. In this approach, we crafted a robust hybrid model development pipeline that incorporates 2 different DL model architectures, the selected pretrained and our custom-designed CNN models, integrated to the utility modules (data loading, preprocessing, and evaluation). These 2 groups of models are used independently in 2 phases (stages) subsequently operating one after the other to achieve 2 exclusive DL operations, feature extraction (feature mapping) and disease classification (inference), representing the 2 fundamentally isolated but highly interdependent modules in the pipeline. The selected pretrained models are used only for the purpose of feature extraction. Given these 2 modules are highly fundamental that operate subsequently, the disease classification model operates using the output of the feature extraction model, we named the overall hybrid pipeline as the 2-stage approach. Figure 3 presents the overall architecture of our 2-stage approach, depicting the 2 major stages as modules in the DL architecture.

XSL•FO
RenderX

Figure 3. Architecture of the new 2-stage approach. BN: batch normalization; DL: deep learning; skin NTDs: skin Neglected Tropical Diseases.



In the first stage (feature extractor module), the selected 21 DL architectures are solely used for feature extraction, which we named them as extractor(s) or feature map(s), and we created 21 specific extractors (maps) by freezing all the trainable weights in each of these 21 models. We use this method to effectively extract features using the mapping logic of each pretrained model and derive feature representations within the high-dimensional feature space (R^{2048}) for each input image in our dataset. This method helped us in creating the full feature representations by avoiding pooling layers, as in the case of our custom-designed CNN model, helping us in preserving important spatial representations of features that would potentially be dropped as a result of using pooling method (MaxPooling). After mapping all the input images using the pretrained models, feature matrices are created for each of the separate train and test sets and are prepared to be used by our custom-designed model. This completes the first phase where the feature extraction models completed their only purposes of feature extraction, which are no longer used.

In the second stage (baseline classification module; Figure 3), the feature matrices created using the previously extracted features by each extractor are used to train the new CNN model. However, since our new model was initially designed as a full-fledged CNN architecture, its use in this phase required restructuring. Hence, we restructured our newly designed CNN model by excluding the extraneous feature extraction layers, as these operations are performed by different models in the separate previous module. Guided by our initially designed 30-layer deep and highly regularized CNN model, we specifically redesigned the final 10-layer architecture to enhance the robustness of the classification head, improve classification performance, and add architectural novelty. This approach totally eliminates the 3 feature extraction blocks of the new CNN model, leading to the exclusion of 20 layers. Specifically, this approach directly eliminates the 6 convolutional layers and their corresponding 10 layers (7 activation and 3 maximum pooling), while minimizing the normalization layers to only 3. This creates a final classification model having 10 layers with 3 dense blocks and a final output layer, where each dense block in turn includes regularization methods (normalization and dropout).

Overall, unlike most standard transfer learning pipelines where the pretrained models serve complex tasks with only 1 or 2 classification layers finally added [45], the classification head in our 10-layer customized (redesigned) CNN model has deeper architectural layers. Through its 9 regularization layers, the new CNN head also serves a separate feature filtering purpose, where the final classification layer makes the classification decision using highly relevant skin lesion features.

Model Training and Evaluation

We used a 5-fold cross-validation method to train and evaluate the models. The metrics used to evaluate the models are selected to assist performance analyses from different perspectives. The macro F_1-score and the class-specific metrics (using sensitivity and specificity) are highly used, as accuracy was found to be a misleading metric due to the highly skewed nature of our new dataset used. Hence, we prioritize the macro F_1-score of models on both the train and test sets to effectively evaluate the models' generalizability and learning ability. The overall sensitivity (macro recall), the area under the precision-recall curve (AUPRC), and the area under the receiver operating characteristic (AUROC) scores are also highly used to measure how well each model performs in identifying positive disease classes. Furthermore, visual tools using tables, confusion matrices, ROC curves, including performance plots such as the slope charts and radar plots, are also used to analyze models' performance.

Final Model Selection

Ultimately, this study selects the final best-performing skin NTDs classification models with the highest classification performance. To achieve this, we applied systematic model selection procedures in 2 phases: first-level selection—applied for the initial feature extraction model screening based on baseline performance score (end of phase 1), and final model selection—applied at the end of model screening experiments (end of phase 2). Overall, the selection of top-performing models involved analysis of performance scores that include macro F_1-score, sensitivity, AUPRC, and 4 class-specific performance metrics (podo-recall, tungiasis-recall, scabies-recall, and podo-F_1-score). Hence, at the end of phase 1, using these extended screening parameters, top-performing models that

XSL•FO
RenderX

achieve stable and outperforming scores across the 2 experiments are selected. Accordingly, selected models are used in the next experimental training that applies performance optimization using the dynamic (online) data augmentation method.

Finally, at the end of phase 2, the final 2 best-performing models are selected based on the results achieved during the experiment with the optimization method, and we applied a robust selection procedure based on weighted comparison of performance scores achieved during this last experiment. Accordingly, 6 evaluation metrics (F_1-score [macro], podo-F_1, macro recall [mean], podo-recall, inference speed [samples per second (sps)], and number of parameters) are used, and we applied the weighted scores comparison method with each metric given different weight. To achieve this strategic comparison, we conducted 4 procedures. First, the metrics are categorized as performance (F_1-score [macro], podo-F_1, macro recall [mean], and podo-recall) and efficiency metrics (inference speed and number of parameters). Out of all these metrics, the model complexity (number of parameters) is a metric that is mostly desired to be lower ("lower is good"), while the other 5 metrics represent best model performance when their values are higher ("higher is better"), having opposite directional symmetry with model size. Second, we normalized all values of the 6 comparison metrics using the minimum-maximum normalization method to ensure that all values fall between 0 and 1 and facilitate the combined scores comparison, which includes normalizing the values that already have values between 0 and 1, mainly for (1) avoiding range dilution—minor differences in performance mostly lead to larger ranges that determine comparison; (2) baseline value definition, combined scores mostly perform well with the least score defined as "0.0"; and (3) simplifying directional symmetry. Therefore, to address all these, we applied the minimum-maximum normalization method by defining 2 normalization formulas. First, we normalized the "higher is better" metrics (the metrics that represent best values when their values are higher) using the formula:

$$\text{xnormalized} = \text{x} - \text{xminxmax} - \text{xmin}$$

where, $x_{\text{normalized}}$ is a value in any 1 of the 5 metrics (eg, macro F_1-score) representing a single value for a specific model that is being normalized; x_{min} is the least value in that specific metric; and x_{max} is the maximum score in the same group of metric.

Next, for the "lower is good" metric (model complexity), we applied the normalization using the formula:

$$\text{xnormalize} = \text{xmax} - \text{xminxmax} - \text{xmin}$$

Third, we assigned weights *w* for all evaluation metrics, and as a diagnostic model, disease prediction performance scores are the primary requirements. Hence, we assigned higher weights (0.2) for performance metrics and a relatively lower weight (0.1) for the 2 efficiency metrics (inference speed and number of parameters). Finally, we define a unified objective function that computes the weighted sum of score (WSS) for each model using the formula:

$$\text{WSSm} = \sum \text{i}=16\text{wi} * \text{xi}$$

where WSS_m is the weighted sum of score of any given model evaluated, x_i is a specific evaluation metric, and w_i is the weight of a given metric. After computing the weighted scores, models are ranked accordingly to identify the top 2 models on the overall performance. Accordingly, the first selected model would be a model with higher complexity, with a consideration of using the model as a back-end (web-based) classification model, while the second selected model would be a lightweight model with the potential of being embedded in edge (mobile) devices for the actual and real-time diagnosis of skin NTDs.

Experimental Setup

Overview

As the study is guided by the funnel framework based on extended experiments, the whole training experiments are conducted in a phased approach having 3 phases. The first phase deals with the establishment of baseline skin NTDs classification models—using the new CNN model, transfer learning, and the 2-stage approach; the second phase applies performance optimization (data augmentation); and the third phase applies the final performance optimization through hyperparameter tuning.

Phase 1: Baseline Model Training With Cascaded Model Selection

In this first phase, only baseline skin NTDs classification models are trained in 3 separate experimental setups: the custom-designed CNN model, transfer learning, and the 2-staged approach. This phase is intended for the overall evaluation of baseline models' performance, which includes comparative analysis and first-level model screening.

Custom-Designed CNN Model

The first experimental setting involves training the custom-designed 30-layered CNN model, where the new model is trained and evaluated under 2 different methods. The first training involves the baseline training, where the model is barely trained without applying any advanced machine learning methods for tweaking performance. On the next experiment, the same model is trained by applying data augmentation techniques. These experiments are conducted to evaluate and analyze the performance of this newly crafted model on our new small-sized skin NTDs dataset, as the results from these experiments determine subsequent strategies.

Transfer Learning: Baseline Models Using Pretrained Models

On the second experimental setting, we demonstrated the transfer learning method using the selected pretrained DL architectures to build baseline models. Hence, 21 baseline DL models are fully trained on our new skin NTDs image dataset, with proper evaluation of each model on the classification of the skin NTDs, where the results are used for the overall analysis of models' performance.

Two-Stage Approach: Feature Extraction With Integrated CNN Model

On the third experimental setting, the 2-staged approach is demonstrated. To apply this robust training pipeline, we used

XSL•FO
RenderX

the 21 systematically selected DL architectures for feature extraction and the restructured new 10-layered CNN model architecture. Hence, 21 different skin NTDs classification models are trained using this approach, where each model is evaluated using the predefined appropriate evaluation methods.

Phase 2: Performance Optimization

Data Augmentation

Evidently, DL models are highly data-intensive and require a large amount of data to achieve excellent performance [46]. Hence, using a relatively small-sized dataset, as the case of this study with only 1495 samples, developing a DL-based diagnostic model with higher classification accuracy is a real challenge; it might not be beneficial at all, compared with the use of large-sized datasets [47]. Methodologically, several DL-based data augmentation methods are available, including the advanced standard and conditional augmentation using GAN models, which are also highly suitable for class imbalance handling. However, their execution requires further experiments, analysis, and selection, including higher computational requirements, which in turn extends this study by deviating from the intended benchmarking objective. Therefore, as a benchmarking study, the dynamic or online (on-the-fly) data augmentation approach [29] is used by applying transformations on the training set to increase the number of training samples by a factor of 5 (including the original training images) and create a total of 4785 training images. This method is primarily used to alleviate the data scarcity issue, analyze performance changes, and increase both the size and the diversity of the training set without the permanent creation of the images. Accordingly, we applied selected geometric transformations such as rotation (0.2), scaling (0.2), and horizontal flipping during model training on the input images at the time of model training to mathematically simulate the real-world diagnostics of skin NTDs. Therefore, using this method, 5 different models are trained based on the 2-stage approach using the final 5 selected extractors.

Phase 3: The Final Optimization Methods and Hyperparameter Tuning

As the last experimental training, the hyperparameter tuning operation is applied to the final 2 selected models. To achieve this, the "Hyperband" algorithm was used, which is a faster and resource-efficient hyperparameter optimization algorithm than other hyperparameter searching algorithms. Hyperband is an efficient bandit-based Keras algorithm for hyperparameter optimization that uses early stopping with a successive halving algorithm to quickly find good configurations for models [48,49]. Therefore, the hyperparameter tuning operation is applied by running the Hyperband algorithm to the maximum of 30 epochs, with the maximum number of trials being 60. The optimization was carried out based on the validation accuracy, which was set to be the objective metric. Finally, the final best hyperparameters were saved to use for the final optimized training of the skin NTDs classification model.

Results

This section presents model evaluation results of the 5 experimental settings (baseline training, transfer learning, and the 2-stage approach, including the training with data augmentation and hyperparameter tuning).

Phase 1: Baseline Model Training With Cascaded Screening

The New CNN Model

The new CNN baseline model is trained on the new skin NTDs dataset in 2 different experimental settings, the baseline training that applies no advanced DL method and applying the data augmentation method, as summarized by the overall results in Table 3. Accordingly, on the baseline training, the model achieved an accuracy of 0.674 and F_1-score of 0.42 (with AUROC=0.676 and AUPRC=0.444), with a very high loss (0.978). However, during the second training with data augmentation, the new CNN model achieved improved performance with an F_1-score of 0.446 and an increased loss of 1.458.

However, class-wise, the new model showed the worst sensitivity for podoconiosis (scoring all 0.0 in precision, recall, and F_1-scores), with a macro recall of 0.43 (having a recall of 0.457 and 0.827 for tungiasis and scabies, respectively), as shown in Table 4.

Table . Overall performance scores of the new CNN[a] model across the 2 experimental settings.

Experiment	Model/method	Accuracy	Loss	F_1-score (macro)	AUROC[b] (mean)	AUPRC[c] (mean)
First	Baseline	0.674	0.978	0.422	0.676	0.444
Second	Dynamic data augmentation	0.691	1.458	0.446	0.790	0.608

[a]CNN: convolutional neural network.

[b]AUROC: area under the receiver operating characteristic curve.

[c]AUPRC: area under the precision-recall curve.

XSL•FO
RenderX

Table . Class-specific performance of the new CNN[a] model across the 2 experiments.

Experiment	Model/Method	Recall				F_1-score			
		Podoconiosis	Tungiasis	Scabies	Macro recall	Podoconiosis	Tungiasis	Scabies	Macro F_1-score
First	Baseline	0.0	0.457	0.827	0.428	0.0	0.489	0.776	0.422
Second	Dynamic data augmentation	0.154	0.181	0.979	0.438	0.235	0.296	0.806	0.446

[a]CNN: convolutional neural network.

With data augmentation, however, the model achieved a macro recall of 0.438 and macro F_1-score of 0.446, with improved class-specific scores in recall (podoconiosis=0.154, tungiasis=0.181, and scabies=0.979) and F_1-score (podoconiosis=0.235, tungiasis=0.296, and scabies=0.806). As confirmed by the results, the use of the standard data augmentation method significantly improved the sensitivity and macro F_1-score of the model for the podoconiosis class (SD +0.109) and (SD +0.017).

Transfer Learning: Baseline Performance

On the second and third training settings of the first phase, baseline models are trained using similar pretrained DL model architectures demonstrating (1) the transfer learning method, and (2) the 2-stage approach. Table 5 presents the overall performance of models scored during these 2 experimental settings.

XSL•FO
RenderX

Table . Performance of the models across the 2 experiments of phase 1 model screening experiments[a].

Model	Experiment 1: transfer learning (baseline models)					Experiment 2: 2-stage approach (baseline models)				
	Accuracy	Log-loss	F_1-score (macro)	AUPRC[b] (mean)	Sensitivity (mean)	Accuracy	Log-loss	F_1-score (macro)	AUPRC (mean)	Sensitivity (mean)
ResNet50	0.695	0.733	0.37	0.516	0.39	0.641	1.027	0.26	0.430	0.33
ConvNext-Small	0.641	0.751	0.26	0.552	0.33	0.805	0.532	0.67	0.747	0.62
Xception	*0.973*	*0.106*	*0.94*	*0.975*	*0.91*	*0.94*	*0.219*	*0.9*	*0.937*	*0.88*
EfficientNetB5	0.641	0.785	0.26	0.378	0.33	0.638	0.86	0.34	0.368	0.36
ConvNext-Tiny	0.671	0.71	0.33	0.573	0.37	0.718	0.693	0.43	0.568	0.43
DenseNet121	*0.95*	*0.119*	*0.89*	*0.973*	*0.84*	*0.96*	*0.133*	*0.91*	*0.974*	*0.9*
EfficientNetB3	0.641	0.782	0.26	0.370	0.33	0.668	0.767	0.37	0.429	0.39
MNv3-Large	0.641	0.77	0.26	0.494	0.33	0.681	0.995	0.59	0.625	0.57
MNv3-Small	0.641	0.777	0.26	0.388	0.33	0.708	0.687	0.4	0.640	0.41
EfficientNetB0	0.641	0.786	0.26	0.360	0.33	0.641	1.012	0.26	0.396	0.33
EfficientNetB1	0.641	0.785	0.26	0.363	0.33	0.628	0.828	0.35	0.437	0.37
MNv2	*0.956*	*0.109*	*0.91*	*0.973*	*0.87*	*0.943*	*0.181*	*0.88*	*0.915*	*0.86*
ResNet18	0.641	0.805	0.26	0.352	0.33	0.735	0.737	0.48	0.545	0.49
EfficientNetV2B0	0.641	0.786	0.26	0.349	0.33	0.641	0.793	0.26	0.393	0.33
EfficientNetV2S	0.674	0.694	0.39	0.501	0.4	0.654	0.867	0.44	0.487	0.45
CovNeXtv2-Tiny	*0.95*	*0.124*	*0.919*	*0.983*	*0.886*	*0.94*	*0.222*	*0.876*	*0.935*	*0.862*
CovNeXtv2-Atto	*0.94*	*0.226*	*0.878*	*0.921*	*0.837*	*0.936*	*0.25*	*0.882*	*0.909*	*0.857*
EfficientViTB0	*0.963*	*0.145*	*0.93*	*0.967*	*0.898*	*0.963*	*0.124*	*0.903*	*0.967*	*0.876*
FasterViT0-T8	0.936	0.165	0.883	0.969	0.855	0.671	0.814	0.382	0.469	0.392
FastViT-T8	0.896	0.37	0.602	0.903	0.61	0.93	0.223	0.862	0.939	0.833
RepViT-M0.9	*0.943*	*0.188*	*0.915*	*0.959*	*0.888*	*0.956*	*0.166*	*0.925*	*0.959*	*0.893*

[a]Values presented in italics represent high-performance

[b]AUPRC: area under the precision-recall curve.

On the training with only transfer learning (experiment 1, Table 5), only 8 models scored top results, where Xception outperformed all models with the top accuracy (97.3%), macro F_1-score (0.94), and sensitivity (0.91), followed by EfficientViTB0 (accuracy=0.963, F_1-score=0.930, and sensitivity=0.898) and MNv2 (accuracy=95.6%, F_1-score=0.91, and sensitivity=0.87). DenseNet121 and ConvNeXtV2-Tiny are the next high-performing models scoring the same accuracy (95%) and sensitivity (0.84), with ConvNeXtV2-Tiny scoring better F_1-score (0.89). The RepViT, FasterViT, and ConvNeXtV2-Atto models are the other top-performing models with macro F_1-scores of 0.92, 0.88, and 0.84. FastViT, ResNet50, EfficientNetV2S, and ConvNext-Tiny scored macro F_1-scores of >0.33 and average sensitivity >0.37. However, the

XSL•FO
RenderX

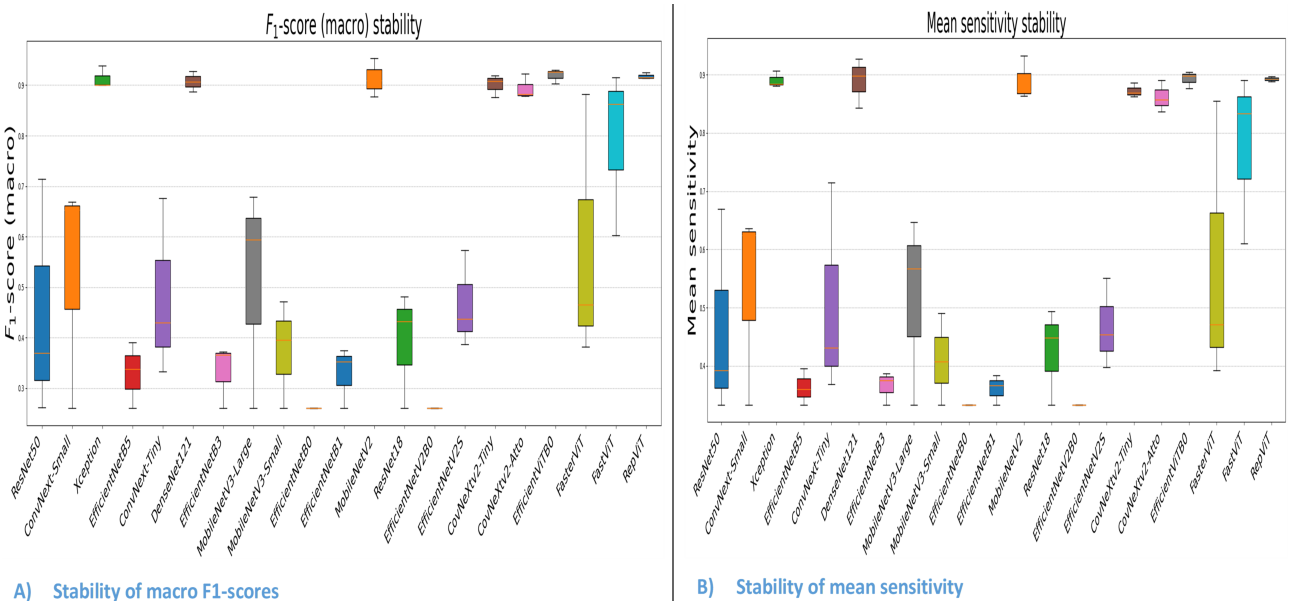
remaining 9 models similarly scored worst class-specific sensitivity (recall=0.0 for podoconiosis and tungiasis classes), leading to the least macro F_1-score (0.26) and average sensitivity (0.33). Using the 2-stage approach as shown in experiment 2 (Table 5), only 4 models trained using RepViT, DenseNet, EfficientViTB0, and Xception extractors helped their corresponding trained classification models to achieve macro F_1-scores of 0.90 and above, while 8 models scored macro F_1-scores of >086. As shown, the model trained on RepViT-extracted features scored the highest macro F_1-score (0.93, mean sensitivity=0.89), where the model using DenseNet-extracted features had the second highest F_1-score (0.91) with the maximum mean sensitivity (0.90). The model using EfficientViTB0-extracted features scored the next top F_1-score (0.90) during this experiment, followed by the models using the Xception feature extractor (F_1-score=0.90 and sensitivity=0.88), CovNeXtv2-Atto feature extractor (F_1-score=0.88 and sensitivity=0.86), MNv2 feature extractor (F_1-score=88 and sensitivity=0.86), CovNeXtv2-Tiny feature extractor (F_1-score=0.88 and sensitivity=0.86), and FastViT feature extractor (F_1-score=0.86 and sensitivity=0.83).

Method-wise, the use of the 2-stage approach improved the overall performance of the majority of the models, where 62% of the models (13 models) exhibited improved macro F_1-scores, with 2 models showing no variations, while the remaining models exhibited minor declines. Class-wise, the 2-stage approach improved the podo sensitivity (podo-recall) of 5 models (using DenseNet121, MNv3-Large, ConvNeXt-Small, ConvNeXtV2-Atto, and FastViT) by the SD of +0.109, +0.218, +0.218, +0.054, and +0.435, respectively, with 3 models (using ConvNeXtV2-Tiny, EfficientViTB0, and FasterVit) showing declining podo-recall, while the other 13 models showing no changes.

Best Models Selection

Overall, the hybrid 2-stage approach applied using the 21 pretrained models (both CNN and transformer-based) yielded superior performance compared with other methods (baseline and transfer learning). Hence, as a DL architectural benchmarking study, we selected this hybrid architecture for further experiments applying optimization methods. However, to screen out short-listed top feature mapping models, we conducted deep and multidimensional analysis that includes performance stability analyses, as shown in Figure 4A and B.

Figure 4. Analysis of models' performance stability across experiments. (A) Stability of the macro F_1-scores of each of the 21 models across the 2 experiments. (B) Stability of the mean sensitivity scores of the 21 models across the 2 experiments.



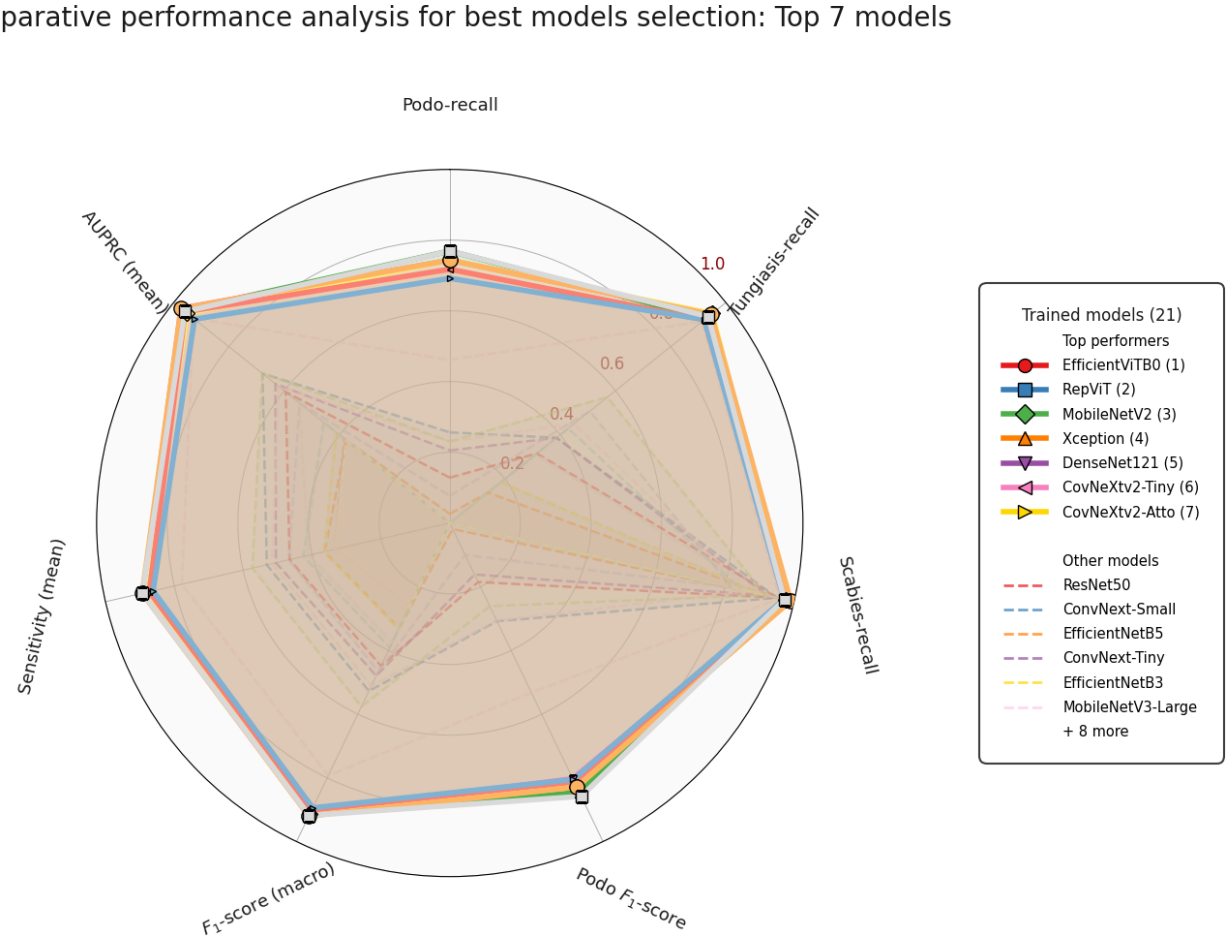
As shown, the box plots clearly depict the stability of the models' performance primarily in macro F_1-score and overall sensitivity (macro recall) across the 2 experiments. Overall, the models fall into three major categories: (1) stable and least-performing models—models trained using features extracted by EfficientNetB0 and EfficientNetV2B0, where the use of these models for feature extraction resulted in the worst macro F_1-score and sensitivity, while the other 6 feature extractors, EfficientNetB5, EfficientNetB3, MNv3-Small, EfficientNetB1, ResNet18, and EfficientNetV2S, resulted in comparatively better performance scores; (2) the unstable models—models trained on ResNet50, ConvNeXt-Small, ConvNeXt-Tiny, MNv3-Large, FasterViT0, and FastViT-T8, for feature extraction, where the models using these extractors scored highly unpredictable performance scores (macro F_1 and sensitivity), as shown by their wide-range scores (tall boxes); and (3) the stable and top-performing models—representing 7 models trained using the features extracted by Xception, DenseNet, MNv2, ConvNeXtV2-Tiny, ConvNeXtV2-Atto, EfficientViTB0, and RepViT. The models using these 7 pretrained models exhibited minimum variability in their macro F_1-score and sensitivity, while achieving the highest median macro F_1-score (0.88) and median mean sensitivity (0.86) across the 2 experiments, as shown in Figure 4A B.

The next analysis includes performance comparison using 7 parameters, as shown by the radar plot in Figure 5. As shown, only 7 models using EfficientViTB0, RepViT, MNv2, Xception,

XSL•FO
RenderX

DenseNet, ConvNeXt-Tiny, and ConvNeXt-Atto feature maps showed exceptionally outperforming performance collectively having a median F_1-score (macro) exceeding 0.95, AUPRC (macro) approaching 1.0, with a comparatively lower median podo-F_1-score approaching 0.90. However, the other 14 models (other than the 7 top-performing models) scored average macro F_1-scores <0.89, average AUPRC<0.92, while showing the worst sensitivity for the minority class with median podo-F_1-score below 0.35.

Figure 5. Overall model performance comparison for best model selection. The radar plot illustrates analysis of the models' performance based on the average scores of the 7 evaluation metrics across the 2 experimental settings, depicting the exceptionally outstanding performance of the 7 top-performing models out of the 21. AUPRC: area under the precision-recall curve.



All these results underlined that the 7 models identified as top-performing models (Figure 5) provided higher feature representation capabilities that led to better overall model performance. Hence, the EfficientViTB0, RepViT, MNv2, Xception, DenseNet, ConvNeXt-Tiny, and ConvNeXt-Atto models are selected for further experiments being used for feature mapping. However, the remaining 14 models were highly challenged to extract useful features that led to poor overall performance and are screened out from being used in further experiments.

Phase 2: Performance Optimization

On the second phase of our experimental framework, the selected 7 models are retrained by applying data augmentation method, and the results in Table 6 are achieved. In this experiment, the model trained using the MNv2 feature mapping significantly outperformed all other models with macro F_1-score of 0.95 and sensitivity of 0.93, while scoring the fourth highest AUPRC (0.97) and the fourth least errors (loss=0.137). As shown, the models using DenseNet121 and EfficientViTB0 extractors scored the next closer performance, where the model trained using DenseNet121 mapping achieved the second highest macro F_1-score and mean sensitivity of 0.928, while the model using EfficientViTB0 for feature mapping scored a macro F_1-score of 0.926 and mean sensitivity of 0.905.

XSL•FO
RenderX

Table . Performance of the 7 screened models in the second phase experiment with data augmentation[a].

Model	Accuracy	Log-loss	F_1-score (macro)	AUPRC[b] (mean)	Sensitivity (mean)
Xception	0.94	0.158	0.901	0.982	0.884
DenseNet121	0.966	0.13	0.928	0.974	0.928
MNv2	*0.973*	*0.137*	*0.953*	*0.967*	*0.933*
CovNeXtv2-Tiny	0.956	0.128	0.908	0.956	0.869
CovNeXtv2-Atto	0.953	0.209	0.922	0.944	0.891
EfficientViTB0	0.97	0.098	0.926	0.985	0.905
RepViT-M0.9	0.956	0.168	0.914	0.959	0.897

[a]Values presented in italics represent high-performance models.

[b]AUPRC: area under the precision-recall curve.

Overall, during this experiment, the 7 feature extraction pretrained models (MNv2, DenseNet121, EfficientViTB0, CovNeXtv2-Atto, RepViT, CovNeXtv2-Tiny, and Xception) helped their corresponding skin NTDs classification models to score macro F_1-scores >0.90, AUPRC >0.94, and mean sensitivity >0.86, with overall loss ≤0.21.

Final Models Selection

For the systematic selection of the final 2 best feature extraction pretrained models that provide the best feature mapping for the skin NTDs classification models, we applied the robust weighted score comparison method defined in the "Methods" section of the study. Hence, applying our predefined WSS formula, we computed the WSS for the scores achieved on the third experiment that applies data augmentation and for the overall average performance of the models across the 3 experiments. For each group of evaluation, we applied two different sets of weights for the metrics: (1) weights given only to the predictive performance metrics—equal weights of 0.25 are given to only the 4 predictive metrics (macro F_1, recall, podo-recall, and podo-F_1), while the speed and number of parameter are not considered to evaluate only predictive performance; and (2) weights given to all the 6 metrics—the predefined weights of 0.2 given for the 4 predictive performance metrics, while the efficiency metrics are given the weights of 0.1. Table 7 summarizes the overall results of the combined weighted evaluation method.

Table . Combined weighted method results for the 7 prescreened models (final model selection)[a].

Model	Performance metrics				Efficiency metrics		Weighted scores (optimization)		Weighted scores (average)	
	F_1-score (macro)	Podo-F1	Macro recall (mean)	Podo-recall	Speed (sps)	Parameters	WSS[b] (predictive)	WSS (overall)	WSS (predictive)	WSS (overall)
RepViT-M0.9	0.05	0.03	0.09	0.10	0.09	0.10	0.335	0.459	*0.974*	*0.975*
DenseNet121	*0.10*	*0.06*	*0.18*	*0.20*	*0.07*	*0.06*	*0.681*	*0.671*	0.511	0.528
EfficientViTB0	0.10	0.03	0.11	0.10	0.00	0.06	0.424	0.399	*0.784*	*0.772*
Xception	0.00	0.03	0.05	0.10	0.03	0.00	0.222	0.211	0.855	0.684
CovNeXtv2-Atto	0.08	0.10	0.07	0.10	0.10	0.10	0.442	0.554	0.009	0.204
MNv2	*0.20*	*0.20*	*0.20*	*0.20*	*0.07*	*0.04*	*1.0*	*0.916*	0.688	0.685
CovNeXtv2-Tiny	0.03	0.00	0.00	0.00	0.05	0.07	0.034	0.155	0.241	0.367

[a]Values presented in italics represent high-performance models.

[b]WSS: weighted sum of scores.

As shown, the model using the MNv2 feature mapping showed an exceptionally highest combined weighted score with an overall WSS of 0.92 on the aggregate weighted scores. This model still showed an exceptionally outperforming combined weighted score on the predictive performance metrics with a WSS of 1.0. The other model using the DenseNet121 mapping has the second highest weighted scores having an overall WSS of 0.67 and a WSS of 0.68 on the predictive performance. On

XSL•FO
RenderX

the same experiment, the models using CovNeXtv2-Atto (0.55) showed the next top overall WSS, followed by the model using RepViT-M0.9 extractors (0.46). On this last experiment, the hybrid transformer-based architecture EffcientViTB0 showed lower weighted scores (0.399 and 0.424) on the overall and predictive WSS. Conversely, using the aggregated performance derived by computing the average scores of each metric for all models across 3 experiments, the combined weighted scores present a completely different set of scores. Based on these results, the model using the RepViT-M0.9 extractors showed an exceptionally higher combined weighted scores both on the overall WSS (0.98) and the predictive WSS (0.97), followed by the model using EfficientViTB0 with an overall WSS (0.98) and predictive WSS (0.97). The models based on MNv2 and DenseNet121 showed the next higher weighted scores (overall WSS of 0.69 and 0.53, respectively) on the aggregated scores of the models across the 3 experiments.

Overall, the data augmentation method is applied to optimize the models' performance, given the study faced the data scarcity problem. However, the models using the RepViT-M0.9 and EfficientViTB0 extractors unexpectedly reacted negatively to this optimization method that expands the sample size and data variance, while the same method significantly boosted the performance of the models based on the MNv2 and DenseNet121 feature mapping. These results reveal a generalization paradox problem, showing (1) overspecialization—the models based on the RepViT-M0.9 and EfficientViTB0 extractors are experiencing performance instability and a potential overfitting due to memorization of the initial small-sized training data, and (2) latent generalization—the models using the MNv2 and DenseNet121 extractors showed their hidden generalization abilities that are exposed as the result of the data augmentation method. Therefore, based on these major driving facts, we selected the MNv2 and DenseNet121 pretrained models for feature mapping in our final skin NTDs diagnostic model.

Phase 3: The Final Optimization

On the final experiment, hyperparameter tuning is applied on the last 2 models selected, which helped the model using the DenseNet121 feature extractor achieve improved performance, as shown in Table 8.

Table . Performance scores of the final 2 selected models after hyperparameter tuning[a].

Model	Accuracy	Macro F_1-score	Loss	AUROC[b] (average)	AUPRC[c] (average)	Recall				F_1-score			
						Podoconiosis	Tungiasis	Scabies	Macro recall	Podoconiosis	Tungiasis	Scabies	Macro F_1-score
DenseNet121	*0.966*	*0.946*	*0.181*	*0.996*	*0.974*	*0.923*	*0.979*	*0.974*	*0.959*	*0.889*	*0.968*	*0.982*	*0.946*
MNv2	0.96	0.906	0.150	0.987	0.955	0.769	0.947	0.979	0.898	0.80	0.937	0.982	0.906

[a]Values presented in italics represent high-performance models.

[b]AUROC: area under the receiver operating characteristic curve.

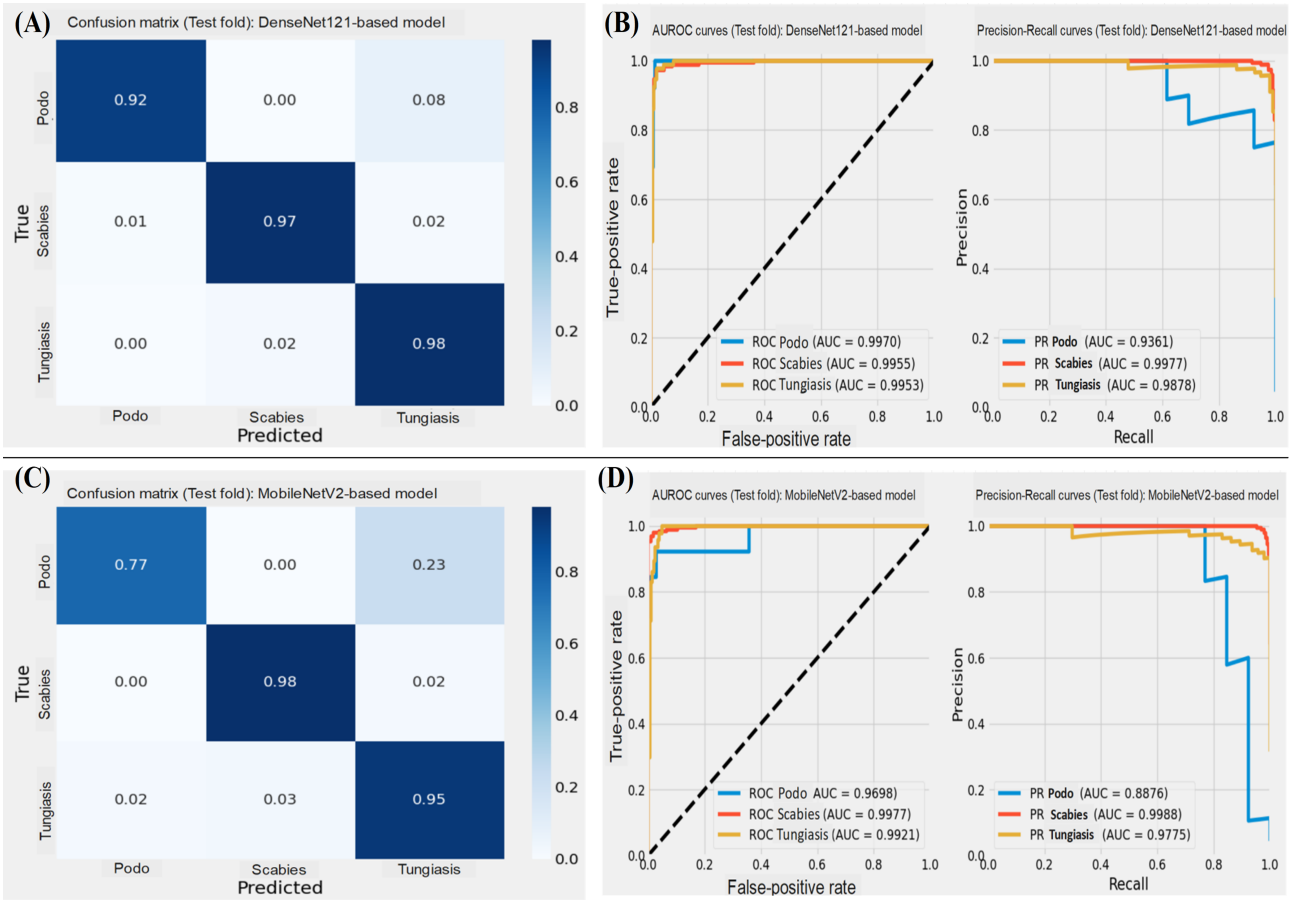
[c]AUPRC: area under the precision-recall curve.

Figures 6A-6D summarizes the overall performance of the final 2 models trained using the DenseNet121 and MNv2 feature extractors after hyperparameter tuning.

After hyperparameter tuning, the DenseNet121-based model scored improved performance, compared with the previous experiment with data augmentation, achieving an accuracy of 0.966, F_1-score of 0.95 (Δ=+0.018), and mean sensitivity of 0.96 (Δ=+0.031). However, the MNv2-based model exhibited declining performance in F_1-score (Δ=−0.047) and macro recall (Δ=−0.035), including podo-specific recall and F_1-score.

XSL•FO
RenderX

Figure 6. Comprehensive classification evaluation of the final 2 models. (A) Confusion matrix showing class-wise results for the model trained using DenseNet121 extractor, (B) AUROC and area under the precision-recall curve (AUPRC) of the model using DenseNet121 feature extractor, (C) confusion matrix showing class-wise results for the model trained using MNv2 extractor, and (D) AUROC and AUPRC of the model using MNv2 feature map. AUC: area under the curve; AUROC: area under the receiver operating characteristic curve; PR: precision-recall; ROC: receiver operating characteristic.



Discussion

Principal Findings

As a DL architectural benchmarking effort, this study developed a diagnostic model for skin NTDs based on the funnel framework with an extensive experiment-based cascaded architectural screening. A significant portion of the experimental design involved identifying robust pretrained DL architectures for feature mapping to be used with our custom-designed CNN model. On average, across the first 2 phases (5 experiments), the custom-designed CNN model and 10 other models trained using the feature mapping architectures yielded poor overall performance with macro F_1-scores below 0.50, 5 models (using MNv3-Large, ConvNext-Small, FasterViT, FastViT, and CovNeXtv2-Atto extractors) showed moderate performance with above 0.50, and 6 models using EfficientViTB0, RepViT, MNv2, Xception, DenseNet121, and ConvNeXtV2-Tiny extractors achieved exceptionally high performance with F_1-scores exceeding 0.90 and sensitivity above 0.87. This exceptionally high performance of the top 6 models is attributed to their robust feature mapping logic that allowed deriving high-dimensional representative features, given our small-sized skin NTDs dataset. Method-wise, the 2-stage approach applying data augmentation resulted in performance improvements for the majority of the models, which underscore the success of the 2-stage approach for skin NTDs classification, given the dataset problems.

The other critical finding of this study highlighted that modern and state-of-the-art DL architectures such as EfficinetNetB5 and EfficinetNetB3, even the highly anticipated lightweight models of EfficinetNetB0 variants, MNv3-Large and MNv3-Small, were unable to extract useful features from the dataset used. All these models proved to be highly data-intensive and extremely sensitive to class imbalance. Conversely, DenseNet121, MNv2, Xception, including the comparatively recent DL architectures of ConvNeXtV2 (Tiny/Atto), and RepViT provided highly robust feature mapping capability, while the transformer-based hybrid architecture EfficientViTB0 showed its potential. Ultimately, the DenseNet121 and MNv2 demonstrated their robustness across the 5 different experimental trainings, including the final optimization. Hence, these 2 models are selected due to their overall feature mapping capability to be integrated with our custom-designed CNN model. Overall, the success of the DenseNet121 model is architecturally linked to its dense connectivity patterns and inductive biases. These factors allowed the trained classifier to reuse feature maps from previous layers through dense blocks, which expands the feature maps into more representative data. Likewise, the success of the MNv2 models is attributed to its internal information bottleneck method. To combat the dataset-related problems such as high variance and possible data memorization, several

XSL•FO
RenderX

regularization strategies were synergistically applied. The strategies included data augmentation, batch normalization, dropout, and early stopping with a synchronized implementation of callbacks. While these methods successfully improved model generalization on unseen test sets, regularization alone was insufficient to fully address the overfitting issues caused by the severely skewed distribution of disease classes. Specifically, for the 12 models excluded after the first phase experiments, all these regularization methods were unable to improve performance even with the application of data augmentation. This clearly proves that the severe class imbalance, especially between scabies and podoconiosis, greatly affected the models' performance. This clearly requires experimental investigation by applying the different methods for class balancing, both at the data and algorithm levels.

The Synergistic Hybrid Approach

Overall, the most effective strategy identified was the 2-stage hybrid approach that combined the high-performing feature extractors with the properly designed CNN classification module. This final harmonized integration involved 3 key strategies: the robust feature extraction architectures, the optimal CNN classification head configuration, and the dynamic data augmentation methods. The harmonized use of all these 3 strategies, along with hyperparameter tuning, provided high-performance classification models by ensuring that each model operated on the optimized output of its preceding module. Ultimately, this study selected the final 2 skin NTD diagnosis models that achieved an F_1-score > 0.95 using MNv2, and an F_1-score > 0.93 using DenseNet121 after data augmentation. All these higher-performance scores are attributed to the combined use of all DL methods, and feature mapping models we use based on our 30-layered very deep CNN model have brought tangible disease-predictive performance improvements across all models.

The integrated 2-stage approach demonstrated its potential for skin NTDs diagnosis, given the study used a dataset with severe class imbalance and small sample images. While most of the DL methods and pretrained models used in the 2-stage approach are preexisting, the harmonization of the feature mapping with our 10-layered classification head created a different DL architecture. This architecture allowed us to establish information bottlenecks using the compact, highly focused, and properly regularized 10-layer classification head that trains on the extracted features. This new architecture added robustness to our model development pipeline, allowing us to build top-performance skin NTDs diagnostic models using the highly constrained dataset, compared with the standard transfer learning method using the same pretrained models. Accordingly, the improved sensitivity of models for the minority class (podo-recall) confirmed the achievements of the architecture, where 5 models showed their sensitivity for podoconiosis, while 62% of the models (13 models) have shown lower overall performance and no variation in podo-recall. These facts clearly underline that the 2-stage approach provided methodological solutions to the severe class imbalance problem. Furthermore, the approach also demonstrated domain adaptation through its two separate modules: (1) feature extraction module—the feature mapping logics of the models initially trained on ImageNet-1K adapted for our skin NTDs diagnostic model, where these diseases have diverse lesion types and textures such as "mossy limbs" that require symmetric analysis of limbs to detect podo based on swellings; and (2) feature refinement module—the highly regularized 10-layered classification head that applies further refinements on the feature matrices extracted by the mapping models.

Clinical Relevance and Further Considerations

This study developed benchmark image-based diagnostic model(s) for skin NTDs using the traditional augmentation method to initially alleviate data scarcity issues (limited sample images with severe class imbalance). We applied this method on our sample input image using our predefined basic geometric transformations at the time of model training to mathematically simulate the real-world diagnostics of skin NTDs. In the actual real-world scenarios, skin NTDs are highly prevalent in remote underserved areas, where the diagnostics of the diseases are mostly undertaken by middle-level health care workers under low-resource settings. Hence, our model design assumed these scenarios that our model is used by middle-level health care workers without the need for highly sophisticated devices, where moderate-quality handheld smartphones can properly serve the diagnostics. Accordingly, we added the basic transformation techniques to our augmentation pipeline that includes rotation (20 degrees), shearing (range=0.1), and zooming (range=0.1) to avoid variations in our models' predictions that can be caused by variable camera angles, inconsistent framing, and varying focal distances in photographs of the skin NTDs. We also added the horizontal flipping transformation to maintain anatomical symmetry, and random brightness adjustments are also applied to simulate the diagnostic practices that can be conducted under different lighting adjustments.

While the study achieved its objective, the data scarcity issues still remained to be the major challenges. However, the top performance scores achieved by the final 2 optimized models (the models using the MN2 and DenseNet121 feature maps) underlined that further experiments using advanced DL-based data augmentations and balancing methods have the potential of further improving the predictive performance of the models. Hence, using each of the selected benchmark models as a basis, advanced DL-based methods such as the use of GAN-generated samples and weighted balancing methods are expected to significantly boost the models' performance, as confirmed by the performance improvements achieved by applying the traditional dynamic data augmentation.

Additionally, as a clinical diagnostic model, the other important aspect expected from such systems is the treatment facility, once the diseases are correctly identified with higher diagnostic accuracy. Generally, the incorporation of treatment recommendations represents the deployment-ready (final) stage of a diagnostic tool, mostly indicating that the model passed several improvement and validation stages. It also requires further system validation and authorization. Therefore, marking the current gaps, incorporation of treatment recommendations, and performance improvements using advanced DL methods are the immediate research tasks.

XSL•FO
RenderX

Conclusions

Ultimately, this study developed a benchmark image-based diagnostic model for skin NTDs through a robust hybrid DL pipeline using a novel skin NTDs dataset collected from a remote area representing an underserved community in the Southwest of Ethiopia. The study developed a new baseline model, applied transfer learning, designed a 2-stage approach, and applied dynamic data augmentation, where all experiments were conducted based on a novel research framework proposed for this study, the funnel framework. Using the funnel framework, optimal methods and high-performing models were selected in a phased approach. With the 2-stage approach being the best model-building approach, the DenseNet121 and MobileNetV2 were top-performing feature extractors. Finally, after the last training applying hyperparameter tuning, the models trained using these extractors still showed performance improvements, DenseNet121 (accuracy=96.6% and F_1-score=0.85) and MobileNetV2 (accuracy=96.6% and F_1-score=0.85). Therefore, the study ultimately selected the DenseNet121 and MobileNetV2 models for feature mapping (extraction) for the final DL-based skin NTDs diagnostic models.

While we developed the intended model for the diagnosis of skin NTDs using the novel skin images dataset, this study also exhibits some downsides that limit the study from contributing to its potential. Data scarcity and severe class imbalance are the primary challenges. The number of diseases represented in this study is also limited to only 3 diseases while there are more than 8 skin NTDs prevalent in Ethiopia. Additionally, the study does not demonstrate data-balancing methods due to resource-related constraints and the extended experiments required. Furthermore, only image data were considered to develop the proposed diagnostic model. Therefore, we recommend further efforts to address the mentioned limitations. First, it is recommended that additional data must be collected from different affected areas (if applicable) with the data being representative of all disease classes. Additionally, DL-based class-balancing methods, such as conditional augmentation based on generative models, are recommended to be experimented. We also suggest the inclusion of multiple types of data, such as text-based patients' data, to expand the dimensionality of the dataset used.

Acknowledgments

The authors of this study gratefully acknowledge Mr Alemayehu Bekele and his team for providing the data used to complete this study, and the technical support he provided us. This study used the data collected from patients with skin Neglected Tropical Diseases (skin NTDs) living in Gacho Baba District of the Gamo Zone, Southwest Ethiopia. The data were initially collected by Mr. Alemayehu Bekele for a project-based research intended to assess the burden of skin NTDs through community screening, led by Mr Alemayehu Bekele from the Collaborative Research and Training Center for Neglected Tropical Disease, College of Medicine Health Sciences of AMU. The authors acquired the dataset for this study through collaboration after having the required ethical clearance letter. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this paper. Finally, the authors would like to declare that no generative tools and generated contents are used for the preparation of any part of this study.

Funding

The authors declared that no financial support was received for this work.

Data Availability

Currently, the new skin NTDs dataset used for this study is not publicly available, as it requires the final verification and consensus regarding dissemination.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Distribution of data samples in the new skin Neglected Tropical Diseases image dataset. The figure illustrates data distribution portraying class imbalance among the 3 disease classes: scabies (dark blue, 63.9%), tungiasis (teal, 31.7%), and podo (light green, 4.4%).

[PNG File, 28 KB - derma_v9i1e91544_app1.png]

References

1. Yohannes Minyilu G, Yimer MA, Meshesha M. Application of deep learning-based multimodal data fusion for the diagnosis of skin neglected tropical diseases: systematic review. JMIR AI 2025 Dec 4;4:e67584. [doi: 10.2196/67584] [Medline: 41344666]

XSL•FO
RenderX

2. Deribe K, Meribo K, Gebre T, et al. The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. Parasit Vectors 2012 Oct 24;5(1):240. [doi: 10.1186/1756-3305-5-240] [Medline: 23095679]
3. The third national neglected tropical diseases strategic plan 2021-2025. : Ministry of Health; 2021 URL: https://espen.afro.who.int/sites/default/files/content/document/Third%20NTD%20national%20Strategic%20Plan%202021-2025_0.pdf [accessed 2026-06-16]
4. Ethiopia: country disease outlook. World Health Organization, Regional Office for Africa, Country Disease Outlook. 2023 Aug. URL: https://www.afro.who.int/sites/default/files/2023-08/Ethiopia.pdf [accessed 2026-06-05]
5. Elimination of neglected tropical diseases (NTDS) in Ethiopia, Woreda level coordination toolkit. : Federal Ministry of Health; 2019 URL: https://www.susana.org/_resources/documents/default/3-3709-7-1571741955.pdf [accessed 2026-06-16]
6. Skin NTDS laboratory network (Skin NTDS LABNET). World Health Organization. URL: https://www.who.int/groups/skin-ntds-laboratory-network [accessed 2026-06-05]
7. Abdela SG, Diro E, Zewdu FT, et al. Looking for NTDs in the skin; an entry door for offering patient centered holistic care. J Infect Dev Ctries 2020 Jun 29;14(6.1):16S-21S. [doi: 10.3855/jidc.11707] [Medline: 32614791]
8. Postigo JAR, Pathak P, Asiedu K. Skin health for all: update on skin neglected tropical diseases with a focus on buruli ulcer and yaws. : Mectizan Donation Program URL: https://iris.who.int/server/api/core/bitstreams/54e59c06-67c7-4c86-8ca9-f210448dab18/content [accessed 2026-06-16]
9. Quilter EEV, Butlin CR, Carrion C, Ruiz-Postigo JA. The WHO Skin NTD mobile application—a paradigm shift in leprosy diagnosis through Artificial Intelligence? Leprosy 2024 Jun;95(2):e2024030. [doi: 10.47276/lr.95.2.2024030]
10. Steyve N, Steve P, Ghislain M, Ndjakomo S, Pierre E. Optimized real-time diagnosis of neglected tropical diseases by automatic recognition of skin lesions. Inform Med Unlocked 2022;33:101078. [doi: 10.1016/j.imu.2022.101078]
11. Brigato L, Barz B, Iocchi L, Denzler J. Image classification with small datasets: overview and benchmark. IEEE Access 2022;10:49233-49250. [doi: 10.1109/ACCESS.2022.3172939]
12. Sales C, Coates SJ. Applications of artificial intelligence for high-burden, underserved skin diseases in global settings: a review. Curr Derm Rep 2025 Jun;14(1):14. [doi: 10.1007/s13671-025-00469-9]
13. Hailu A, Gebre T, Seife F, et al. Challenges and strategies for mainstreaming neglected tropical diseases campaign interventions in Ethiopia. Am J Trop Med Hyg 2025 Feb 5;112(2):467-478. [doi: 10.4269/ajtmh.24-0261] [Medline: 39591648]
14. Tekola Ayele F, Adeyemo A, Finan C, et al. HLA class II locus and susceptibility to podoconiosis. N Engl J Med 2012 Mar 29;366(13):1200-1208. [doi: 10.1056/NEJMoa1108448]
15. Yotsu RR, Ding Z, Hamm J, Blanton RE. Deep learning for AI-based diagnosis of skin-related neglected tropical diseases: a pilot study. PLoS Negl Trop Dis 2023 Aug;17(8):e0011230. [doi: 10.1371/journal.pntd.0011230] [Medline: 37578966]
16. Pattnayak P, Mohanty A, Das T, Patnaik S. Applying artificial intelligence and deep learning to identify neglected tropical skin disorders. Presented at: 2024 3rd International Conference for Innovation in Technology (INOCON); Mar 1-3, 2024; Bangalore, India p. 1-6. [doi: 10.1109/INOCON60754.2024.10511323]
17. Beesetty R, et al. Leprosy skin lesion detection: an AI approach using few shot learning in a small clinical dataset. Indian J Lepr 2023;95:89-102 [FREE Full text]
18. Barnowska EJ, Fastenau A, Penna S, Bonkass AK, Stuetzle S, Janssen R. Diagnosing skin neglected tropical diseases with the aid of digital health tools: a scoping review. PLoS Digit Health 2024 Oct;3(10):e0000629. [doi: 10.1371/journal.pdig.0000629] [Medline: 39374195]
19. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 2018 Aug;5(1):180161. [doi: 10.1038/sdata.2018.161]
20. Cassidy B, Kendrick C, Brodzicki A, Jaworek-Korjakowska J, Yap MH. Analysis of the ISIC image datasets: usage, benchmarks and recommendations. Med Image Anal 2022 Jan;75:102305. [doi: 10.1016/j.media.2021.102305]
21. Johnson JM, Khoshgoftaar TM. Survey on deep learning with class imbalance. J Big Data 2019 Dec;6(1):27. [doi: 10.1186/s40537-019-0192-5]
22. Wang Z, et al. A comprehensive survey on data augmentation. arXiv. Preprint posted online on Oct, 2025. [doi: 10.48550/arXiv.2405.09591]
23. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019 Dec;6(1):60. [doi: 10.1186/s40537-019-0197-0]
24. Kumar T, Brennan R, Mileo A, Bendechache M. Image data augmentation approaches: a comprehensive survey and future directions. IEEE Access 2024;12:187536-187571. [doi: 10.1109/ACCESS.2024.3470122]
25. Islam T, Hafiz M, Jim JR, Kabir M, Mridha MF. A systematic review of deep learning data augmentation in medical imaging: recent advances and future research directions. Healthc Analytics 2024 Jun;5:100340. [doi: 10.1016/j.health.2024.100340]
26. Ribas LC, Casaca W, Fares RT. Conditional generative adversarial networks and deep learning data augmentation: a multi-perspective data-driven survey across multiple application fields and classification architectures. AI 2025 Feb;6(2):32. [doi: 10.3390/ai6020032]
27. Wickramaratne SD, Mahmud M. Conditional-GAN based adversarial networks and deep learning data augmentation: a multi-perspective data-adversarial networks and deep learning data augmentation: a multi-perspective data-data augmentation

XSL•FO
RenderX

for deep learning task classifier improvement using fNIRS data. Front Big Data 2021;4:659146. [doi: 10.3389/fdata.2021.659146] [Medline: 34396092]
28. Wu H, Gu X. Max-pooling dropout for regularization of convolutional neural networks. In: Arik S, Huang T, Lai WK, Liu Q, editors. Neural Information Processing: Springer International Publishing; 2015, Vol. 9489:46-54. [doi: 10.1007/978-3-319-26532-2_6]
29. Cerqueira V, Santos M, Roque L, Baghoussi Y, Soares C. Online data augmentation for forecasting with deep learning. arXiv. Preprint posted online on Jan, 2025. [doi: 10.48550/arXiv.2404.16918]
30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016; Las Vegas, NV, USA p. 770-778. [doi: 10.1109/CVPR.2016.90]
31. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A convnet for the 2020s. Presented at: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-24, 2022. [doi: 10.1109/CVPR52688.2022.01167]
32. Woo S, Debnath S, Hu R, et al. ConvNeXt V2: co-designing and scaling convnets with masked autoencoders. Presented at: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 17-24, 2023; Vancouver, BC, Canada p. 16133-16142. [doi: 10.1109/CVPR52729.2023.01548]
33. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI p. 2261-2269. [doi: 10.1109/CVPR.2017.243]
34. Chollet F. Xception: deep learning with depthwise separable convolutions. Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI p. 1800-1807. [doi: 10.1109/CVPR.2017.195]
35. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. arXiv. Preprint posted online on May 28, 2019. [doi: 10.48550/arXiv.1905.11946]
36. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-23, 2018; Salt Lake City, UT p. 4510-4520. [doi: 10.1109/CVPR.2018.00474]
37. Yang TJ, et al. NetAdapt: platform-aware neural network adaptation for mobile applications. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, editors. Computer Vision ECCV: Springer International Publishing; 2018, Vol. 11214:289-304. [doi: 10.1007/978-3-030-01249-6_18]
38. Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3. Presented at: 2019 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 27 to Nov 2, 2019; Seoul, South Korea p. 1314-1324. [doi: 10.1109/ICCV.2019.00140]
39. Tan M, Le QV. EfficientNetV2: smaller models and faster training. arXiv. Preprint posted online on Apr 1, 2021. [doi: 10.48550/arXiv.2104.00298]
40. Wang A, Chen H, Lin Z, Han J, Ding G. Rep ViT: revisiting mobile CNN from ViT perspective. Presented at: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 16-22, 2024; Seattle, WA, USA p. 15909-15920. [doi: 10.1109/CVPR52733.2024.01506]
41. Hatamizadeh A, et al. FasterViT: fast vision transformers with hierarchical attention. arXiv. Preprint posted online on Jun 9, 2023. [doi: 10.48550/arXiv.2306.06189]
42. Anasosalu Vasu PK, Gabriel J, Zhu J, Tuzel O, Ranjan A. FastViT: a fast hybrid vision transformer using structural reparameterization. Presented at: 2023 IEEE/CVF International Conference on Computer Vision (ICCV); Sep 30-30, 2023; Paris, France p. 5762-5772. [doi: 10.1109/ICCV51070.2023.00532]
43. Cai H, Li J, Hu M, Gan C, Han S. EfficientViT: lightweight multi-scale attention for high-resolution dense prediction. Presented at: 2023 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 1-6, 2023; Paris, France p. 17256-17267. [doi: 10.1109/ICCV51070.2023.01587]
44. Deng J, Dong W, Socher R, Li LJ. ImageNet: a large-scale hierarchical image database. Presented at: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops); Jun 20-25, 2009. [doi: 10.1109/CVPR.2009.5206848]
45. Sahito A, Frank E, Pfahringer B. Transfer of pretrained model weights substantially improves semi-supervised image classification. In: Gallagher M, Moustafa N, Lakshika E, editors. AI 2020: Advances in Artificial Intelligence: Springer International Publishing; 2020, Vol. 12576:433-444. [doi: 10.1007/978-3-030-64984-5_34]
46. Alzubaidi L, Bai J, Al-Sabaawi A, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. J Big Data 2023 Apr;10(1):46. [doi: 10.1186/s40537-023-00727-2]
47. Rukundo O. Effects of image size on deep learning. Electronics (Basel) 2023 Feb;12(4):985. [doi: 10.3390/electronics12040985]
48. Li L, Jamieson K, DeSalvo G, Rostamizadeh A, Talwalkar A. Hyperband: a novel bandit-based approach to hyperparameter optimization. arXiv. Preprint posted online on Jun, 2018. [doi: 10.48550/arXiv.1603.06560]
49. Song Y, Chiangpradit M, Busababodhin P. Hyperband-optimized CNN-BiLSTM with attention mechanism for corporate financial distress prediction. Appl Sci (Basel) 2025 May;15(11):5934. [doi: 10.3390/app15115934]

XSL•FO
RenderX

Abbreviations

AMU: Arba Minch University
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
CNN: convolutional neural network
DL: deep learning
GAN: generative adversarial
GFLOP: Giga floating point operation
IRB: institutional review board
ISIC: International Skin Imaging Collaboration
MDA: mass drug administration
NTD: Neglected Tropical Disease
skin NTD: skin Neglected Tropical Disease
WSS: weighted sum of scores

Edited by S Subhadarshani; submitted 15.Jan.2026; peer-reviewed by R Wu, T Banerjee; revised version received 11.May.2026; accepted 20.May.2026; published 23.Jun.2026.

Please cite as:
Minyilu Y, Yimer MA, Meshesha M
Harmonized Dual Deep Learning Architectures for Image-Based Diagnostics of Skin Neglected Tropical Diseases: Benchmark Study via Novel Funnel Framework
JMIR Dermatol 2026;9:e91544
URL: https://derma.jmir.org/2026/1/e91544
doi:10.2196/91544

© Yohannes Minyilu, Mohammed Abebe Yimer, Million Meshesha. Originally published in JMIR Dermatology (http://derma.jmir.org), 23.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Deep Learning–Based Classification of Melanoma Using Dermoscopic Images: Comparative Evaluation of Convolutional Neural Network and DenseNet Models

Harshal Sanghvi[1], PhD; Argyrios Agritis[2], BS; Isabella Alvarez[2], BS; Mitsu Patel[2], MBBS; Shuwen Wang[3], PhD; Riki Patel[2], PhD; Raphael Banoub[4], MD; Timonthy Makkar[5], MD

[1]Florida Atlantic University, 777 Glades Road, Boca Raton, FL, United States

[2]Department of Technology and Clinical Trials, Advanced Research LLC, Pompano Beach, FL, United States

[3]Clemson University, Clemson, SC, United States

[4]Mayo Clinic in Florida, Jacksonville, FL, United States

[5]Department of Dermatology, Rutgers New Jersey Medical School, Brunswick, NJ, United States

Corresponding Author:
Harshal Sanghvi, PhD
Florida Atlantic University, 777 Glades Road, Boca Raton, FL, United States

Abstract

Background: For melanoma, an aggressive form of skin cancer, early and accurate detection is critical to improving clinical outcomes. Deep learning–based image analysis has emerged as a promising approach for computer-assisted evaluation of dermoscopic lesions. However, model performance is often affected by limited dataset size, class imbalance, overfitting, and differences in network architecture and optimization strategies.

Objectives: This study was aimed at developing and comparatively evaluating convolutional neural network (CNN) and DenseNet-based deep learning models for automated classification of dermoscopic skin lesion images. The effects of data augmentation, network architecture, optimizer selection, and training duration on classification performance were examined.

Methods: Publicly available dermoscopic skin lesion images obtained from Kaggle were evaluated using a conventional CNN, DenseNet169, and DenseNet201. The original dataset contained 2357 images representing multiple benign and malignant skin lesion categories. Images were resized to 250×250 pixels, labeled, and randomly partitioned into training and testing sets at an 80:20 ratio. To address class imbalance, image augmentation using transformations, including ±10° shear, was performed, increasing the dataset size to 6857 images, with approximately 500 additional samples generated per class. Models were trained using adaptive moment estimation (Adam) and root mean square propagation (RMSprop) optimizers, and evaluated after 20, 60, and 100 epochs. Performance was assessed using accuracy, precision, recall (sensitivity), specificity, F_1-score, confusion matrices, and training and validation loss.

Results: Data augmentation substantially improved model performance and reduced evidence of overfitting compared with the original dataset. On the augmented dataset, DenseNet169 and DenseNet201 achieved accuracy and F_1-scores exceeding 94% after 100 epochs, consistently outperforming the conventional CNN. DenseNet169 achieved up to 95% validation accuracy, with only minor performance differences between the Adam and RMSprop optimizers. The CNN showed greater susceptibility to overfitting on the smaller dataset; its F_1-score, calculated over the complete dataset, increased from 31% after 20 epochs to 59% after 60 epochs and 87% after 100 epochs before augmentation. Across the comparative experiments, the DenseNet architectures demonstrated more stable validation performance and better generalization following augmentation.

Conclusions: DenseNet169 and DenseNet201 demonstrated stronger and more stable classification performance than the conventional CNN for dermoscopic skin lesion classification, particularly after augmentation of the training dataset. Increasing the dataset size and improving class balance led to reduced overfitting and improved model performance. These findings support further investigation of DenseNet-based approaches as potential clinical decision support tools for melanoma and skin lesion assessment. However, prospective evaluation using independent, multicenter datasets and diverse patient populations is required before clinical implementation.

(JMIR Dermatol 2026;9:e104086) doi:10.2196/104086

KEYWORDS

deep learning; dermoscopy; medical image analysis; data augmentation; melanoma detection

XSL•FO
RenderX

Introduction

Skin cancer is one of the most prevalent malignancies globally, accounting for approximately 40% - 50% of all cancer diagnoses within the United States. It is broadly categorized into melanoma and nonmelanoma skin cancers (NMSCs) [1]. NMSCs primarily originate from keratinized epithelial cells and include basal cell carcinoma, squamous cell carcinoma, and sebaceous gland carcinoma. Although melanoma is less common, it represents the most aggressive form of skin cancer and accounts for the majority of skin cancer–related deaths [2].

Melanoma arises from melanocytes, which are pigment-producing cells localized in the basal layer of the epidermis. Exposure to UV radiation induces keratinocytes to secrete α-melanocyte-stimulating hormone, which activates melanocytes via the melanocortin 1 receptor, thereby promoting melanin synthesis [3]. This physiological response confers photoprotective properties; however, chronic UV exposure can promote DNA damage and mutagenesis, contributing to malignant transformation and unregulated cellular proliferation. Epidemiological data reveal sex- and age-specific variations in melanoma incidence: incidence is higher in females than in males before the age of 40 years, whereas by the age of 75 years, incidence in males is nearly 3 times that in females. Both genetic predisposition and environmental factors contribute to melanoma pathogenesis, with UV radiation implicated in approximately 60% - 70% of cases. The World Health Organization classifies UVA (315 - 400 nm) and UVB (280 - 315 nm) radiation as complete carcinogens, given their involvement in tumor initiation and progression [4].

Clinically, melanoma is classified into distinct histopathological variants, of which superficial spreading melanoma is the most common, especially in populations with lighter skin phototypes [5]. Despite advancements in understanding melanoma biology, its global incidence continues to rise, emphasizing the importance of early detection in improving patient prognosis. Diagnostic assessment relies predominantly on dermoscopic evaluation, supplemented by heuristic approaches such as the ABCDE (asymmetry, border irregularity, color heterogeneity, diameter, and evolution) criteria. Nevertheless, these methods are inherently subjective, depend on clinical expertise, and are susceptible to interobserver variability [6].

To address these diagnostic challenges, AI and computer-aided diagnosis systems have been developed for automated skin lesion analysis. Deep learning frameworks, particularly convolutional neural networks (CNNs), have demonstrated diagnostic performance comparable to that of experienced dermatologists [7]. Model performance has been improved by incorporating transfer learning and sophisticated feature extraction techniques. Public datasets, such as those from the International Skin Imaging Collaboration and the PH^2 database, have been instrumental in training and validating AI models [8].

Complementary to algorithmic innovations, the use of image preprocessing techniques, such as top-hat filtering and DullRazor for hair artifact removal and noise reduction, has substantially improved lesion segmentation and classification accuracy [9]. Clinical technologies, including digital dermoscopy, total body photography, and mobile health platforms, have enhanced the accessibility of skin surveillance. However, challenges remain regarding cost, patient adherence to follow-up, and variability in diagnostic precision [10].

In this study, we propose a novel deep learning framework that integrates transfer learning with densely connected neural network (DenseNet) architectures (DenseNet169 and DenseNet201) and a conventional CNN model for automated skin lesion classification. The framework iteratively optimizes hyperparameters to enhance diagnostic accuracy and is benchmarked against established machine learning counterparts. By combining rigorous preprocessing protocols with state-of-the-art neural network architectures, this work aims to promote early melanoma detection and support clinical decision-making.

Table 1 summarizes the performance metrics of the proposed approach and relevant methods reported in the literature.

XSL•FO
RenderX

Table . Performance overview of the proposed method and relevant methods reported in the literature.

Reference	Methods	Dataset/validation	Key findings	Sensitivity/accuracy
This study	DenseNet169, DenseNet201, and CNN[a]	Dermoscopic images	Validation and training accuracy greater than 98%	Accuracy 98%
Cheong et al [11], 2021	Feature extraction (entropy, LBP[b], GLCM[c], and ABCDE[d] criteria) with SVM-RBF[e]	Dermoscopic images	Classification accuracy greater than 97.5%	Sensitivity greater than 97.5%
Waheed et al [12], 2023	Deep learning (CNN, hybrid CNN, and transfer learning) with data augmentation	ISIC[f] dataset	Correct classification rate of 81%	Sensitivity 92.46%; specificity 92.23%
Alsaade et al [13], 2021	Hybrid ML[g] and DL[h] (segmentation, CNN, and ANN[i])	ISIC 2018 and PH^2[j]	ANN accuracy of 98% (ISIC) and 97.5% (PH^2)	Sensitivity 83% - 97.5%
Indraswari et al [14], 2022	Transfer learning with MobileNetV2	ISIC Archive	Lightweight model suitable for mobile deployment	Accuracy 85%; sensitivity 85%
Tumpa and Kabir [15], 2021	Image preprocessing (MGI[k] for hair removal), texture and shape features, and ANN	ISIC and PH^2	Benign versus malignant lesions classified with 97.7% accuracy	Sensitivity ~97%
Brinker et al [16], 2019	CNN compared with dermatologists	ISIC Archive	CNN sensitivity comparable to that of dermatologists; higher specificity (68.2% vs 64.4%)	Sensitivity 89.4%
Thaajwer and Ishanka [17], 2020	Image processing (GLCM and color/shape features)+SVM[l]	Dermoscopic images	Classification accuracy of 83%	Not reported
Banerjee et al [18], 2020	Deep learning, fuzzy logic, and DCNN[m] (YOLO[n])	PH^2, ISBI 2017, ISIC 2019	High accuracy across datasets	Sensitivity 97.5% - 98.5%
Bakheet and El-Nagar [19], 2021	Real-time computer-aided diagnosis with DNN[o]	PH^2 dataset	Accuracy of ~97.5%	Sensitivity 100%
Chang et al [20], 2022	CNN, XGB[p] classifier, and SMOTE[q]-based balancing	Stratified 10-fold cross-validation	Accuracy of ~97.5%	Sensitivity 98.1%
Thiyaneswaran et al [21], 2021	GLCM features with ML classifiers	Dermoscopic images	Accuracy of ~90%	Positive prediction 95%; negative prediction 80%

[a]CNN: convolutional neural network.

[b]LBP: local binary pattern.

[c]GLCM: gray level co-occurrence matrix.

[d]ABCDE: asymmetry, border irregularity, color heterogeneity, diameter, and evolution.

[e]SVM-RBF: support vector machine with the radial basis function kernel.

[f]ISIC: International Skin Imaging Collaboration.

[g]ML: machine learning.

[h]DL: deep learning.

[i]ANN: artificial neural network.

[j]PH^2: PH^2 Dermoscopic Image Database.

[k]MGI: morphological gradient image.

[l]SVM: support vector machine.

[m]DCNN: deep convolutional neural network.

[n]YOLO: you only look once.

[o]DNN: deep neural network.

[p]XGB: extreme gradient boosting.

[q]SMOTE: synthetic minority oversampling technique.

XSL•FO
RenderX

Methods

Overview

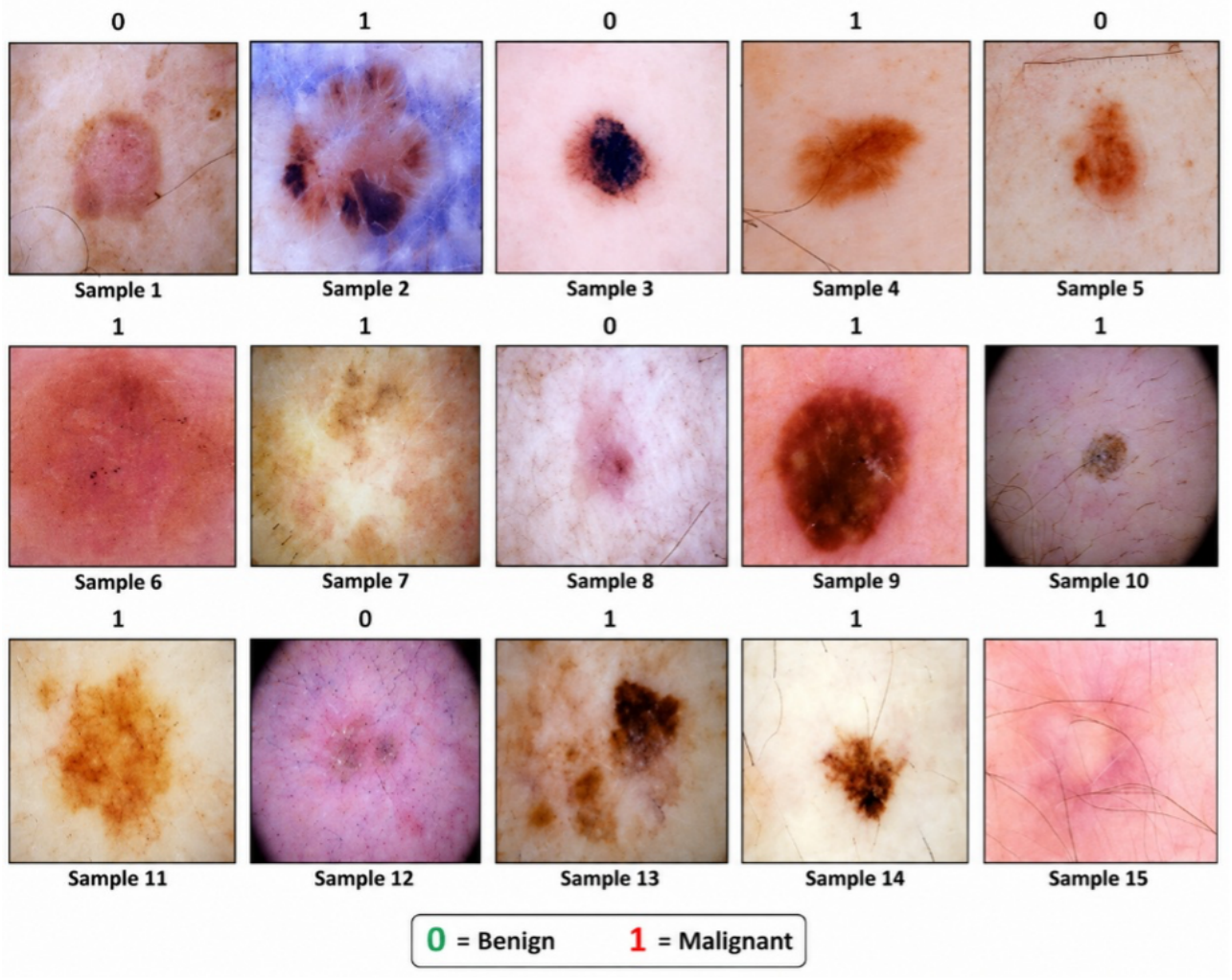
Images of skin lesions were used as inputs for the CNN, DenseNet169, and DenseNet201 models. The images were read using the OpenCV (cv2) library at a size of 250×250 pixels to speed up processing and reduce memory consumption. Subsequently, the images were labeled, and a label binarizer was used to convert the multiclass labels to binary labels. The dataset was randomly partitioned into training (80%) and validation (20%) sets. Each model was trained for 20, 60, and 100 epochs. The CNN model incorporated 3 Conv2D layers with a kernel size of 3×3 and stride of 2×2. The adaptive moment estimation (Adam) and root mean square propagation (RMSprop) optimizers were used to optimize training conditions for the models. Accuracy and loss over the training and validation data were recorded. Loss was calculated as the difference in the actual and predicted class values.

Dataset

The dataset used for classification was obtained from Kaggle. The classification models included a CNN and 2 DenseNet architectures (DenseNet169 and DenseNet201). The dataset included images depicting several forms of skin lesions: actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesions.

We evaluated dataset versions 1 and 2. Both datasets contained images illustrating the same conditions. However, dataset 2 included additional images in each category and exhibited a more balanced distribution of data across conditions. Dataset 2 was used after dataset 1 to help mitigate model overfitting. Several images from both datasets are shown in Figure 1.

Figure 1. Images from the datasets.



Model Architecture

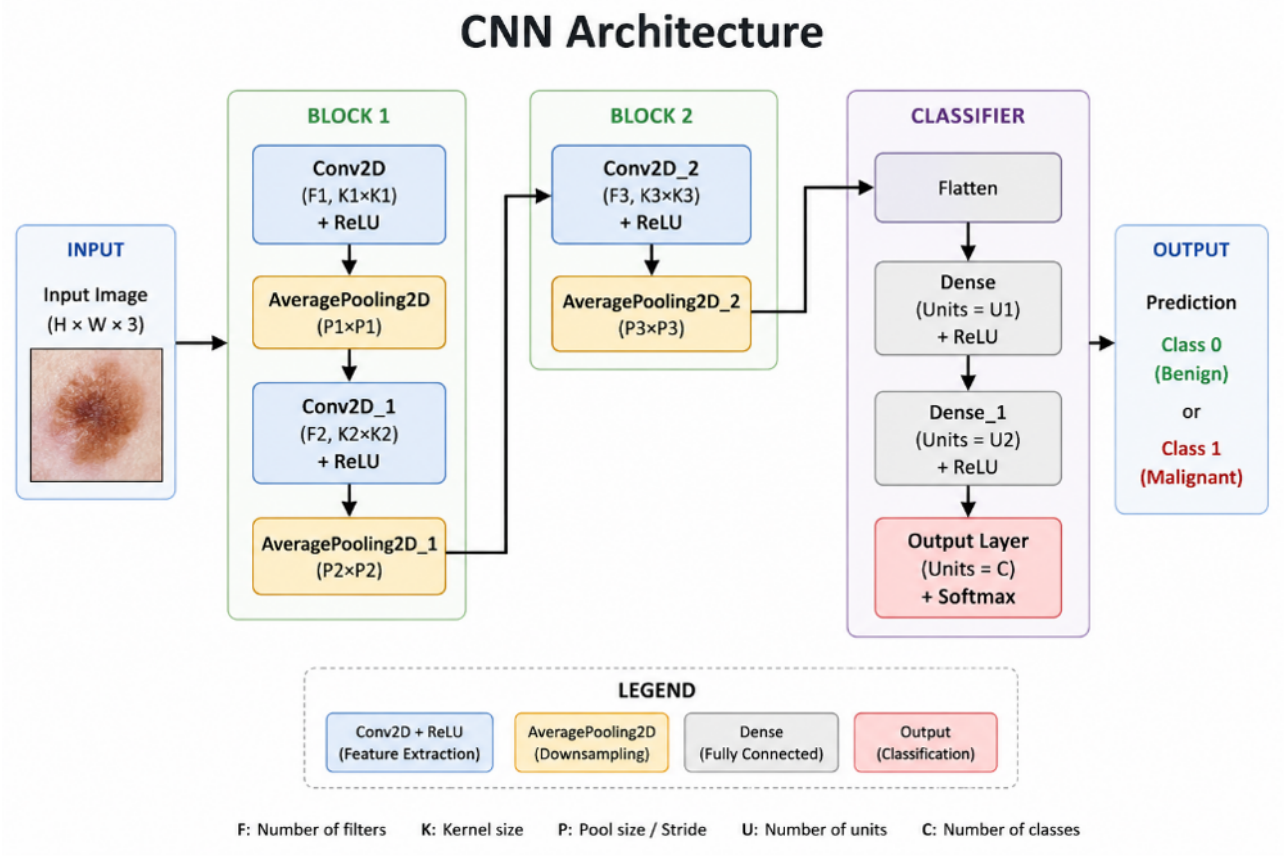
CNN

The CNN architecture is shown in Figure 2. In general, CNNs are composed of convolution layers, pooling layers, and fully connected layers. The workflow involves input, feature extraction, and classification. Feature extraction is performed using convolutional layers with rectified linear unit (ReLu) activation and pooling layers. Classification is performed using a fully connected layer with softmax activation. The first convolution layer processes the input image by convoluting it through a filter that extracts image features, thereby creating a feature map. Notably, convolution layers process only the data within their receptive fields before forwarding the combined

XSL•FO
RenderX

inputs to the subsequent layers. CNN inputs are represented as tensors with specific dimensions; once these tensors pass through the convolutional layer, they are transformed to a feature map. ReLU activation helps reduce errors by improving learning efficiency. Pooling layers help mitigate overfitting by decreasing the spatial dimensions of the feature map [22,23]. Specifically, pooling layers merge the outputs of neuron clusters in one layer into a single neuron in the following layer. The CNN used in this study incorporated average pooling, a technique often used to identify minor changes in images. The fully connected layer lies at the end of the network and connects all neurons of the prior activation layers. Using a flattening process, the resulting maps can be transformed into one-dimensional data [24]. For classification, the flattened matrix is passed to a fully connected layer.

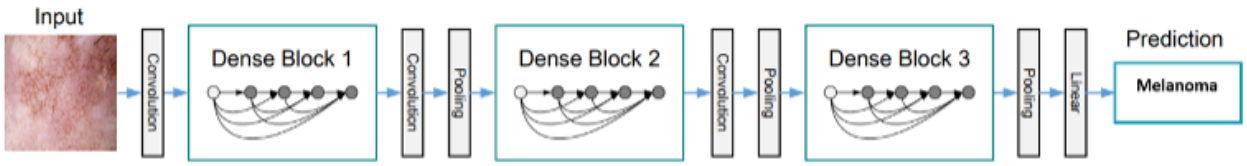
Figure 2. Architecture of the convolutional neural network. CNN: convolutional neural network; ReLU: rectified linear unit.



DenseNet

The general architecture of DenseNet is shown in Figure 3. DenseNets are a class of CNNs featuring dense connections between layers with similar feature-map sizes. Specifically, a feedforward connection exists between a given layer and the remaining layers constituting the network. Unlike other convolutional networks that feature a connection only between a given layer and its following layer, DenseNets connect each layer to its previous and following layers [25]. While a standard convolutional network with L layers typically has L connections, a DenseNet with L layers has L(L + 1)/2 direct connections. Each layer in a dense block receives feature maps from all preceding layers and provides its output to all subsequent layers.

Figure 3. General architecture of DenseNet.



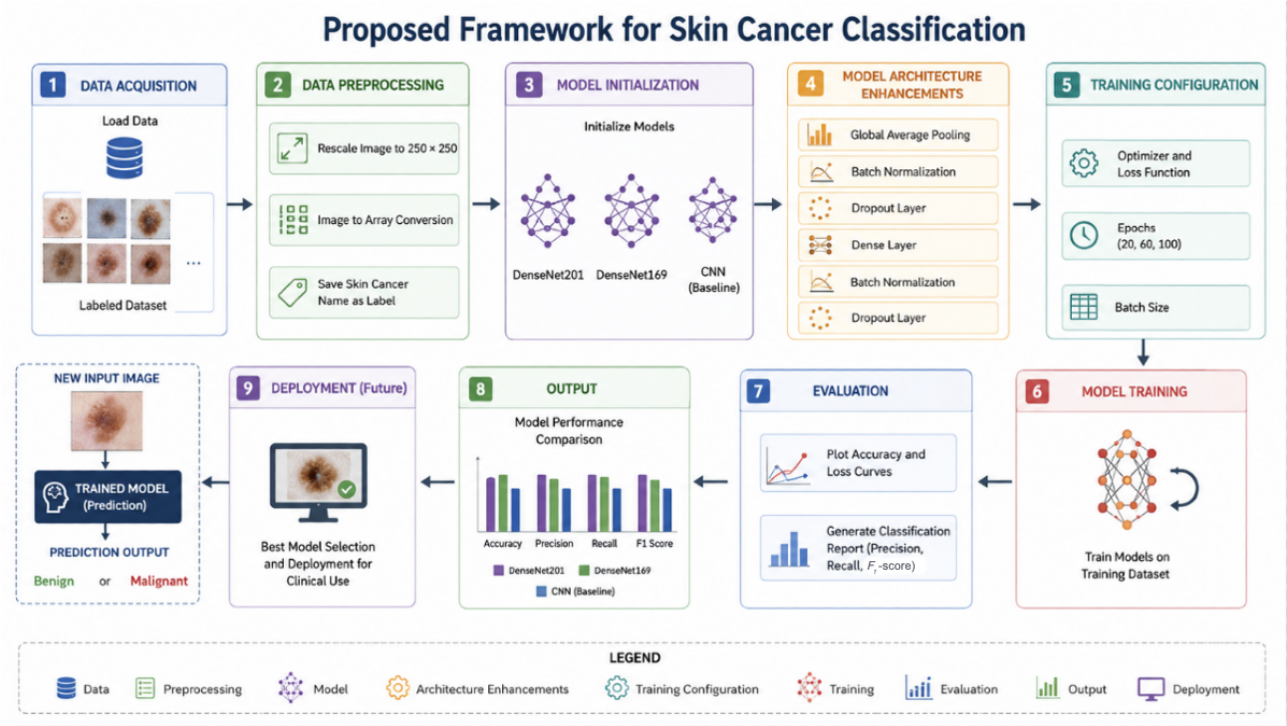
In DenseNets, feature maps obtained from the previous layers are combined through concatenation rather than summing. This architecture offers several advantages, including mitigation of the vanishing gradient problem commonly encountered in CNN frameworks, enhanced strength of connections, more efficient propagation of information, feature reuse, and reduced parameter requirements [26]. The ability of DenseNet to efficiently propagate information facilitates training through implicit deep supervision [27]. DenseNet169 and DenseNet201 include 4 blocks with 6, 12, 32, and 32 layers and 6, 12, 48, and 32 layers, respectively.

XSL•FO
RenderX

Proposed Framework

Figure 4 illustrates the framework used to detect skin lesions in this work.

Figure 4. Flowchart of the proposed framework for skin lesion classification. CNN: convolutional neural network.



Ethical Considerations

This study involved secondary analysis of a publicly available dermoscopic image dataset obtained from Kaggle. No participants were prospectively recruited, and no new patient data or biological specimens were collected. The investigators did not have access to direct patient identifiers or identifiable private information in the dataset used for the analysis. The data were analyzed solely for research purposes in accordance with the terms and conditions of the data source. Because the study was based exclusively on publicly available, deidentified secondary data and involved no direct interaction with or intervention in human participants, informed consent was not obtained for this study. No compensation was provided because no participants were recruited or directly involved.

Results

Overview

The models were trained for 20, 60, and 100 epochs. Loss was calculated by subtracting the predicted class values from the actual class values.

The CNN model was first evaluated using the Adam optimizer. At 20 and 60 epochs, training and validation loss decreased gradually, with negligible differences between the 2 values. This indicates that the model did not exhibit overfitting or underfitting at these durations. However, at 100 epochs, training and validation loss exhibited different trends. The validation loss initially decreased but began to increase at 40 epochs. In contrast, the training loss continued to decrease even after 40 epochs. The difference between training and validation loss at 100 epochs indicated overfitting. Similarly, during training for 100 epochs, the training accuracy continued to increase, whereas validation accuracy remained constant at 50%. The F_1-scores on the validation set were 40%, 50%, and 60% at 20, 60, and 100 epochs, respectively. The F_1-score over the complete dataset increased with training duration from 20 to 100 epochs.

The optimizers assessed in this work were RMSprop and Adam. These optimizers affect model performance by influencing the training process. Both optimizers use moving averages of gradients. Although Adam is similar to RMSprop, it integrates momentum-based quantities. Adam has been reported to be more efficient than other optimizers owing to its reduced memory and time consumption, as well as more effective management of scattered gradients in scenarios involving noise [24].

Experiments were conducted using 3 neural network architectures: a conventional CNN, DenseNet169, and DenseNet201. The models were assessed using 2 optimizers (Adam and RMSprop) and over two dataset versions: the original dataset containing 2357 images (version 1) and an augmented dataset containing 6857 images (version 2). Models were trained for 15 to 100 epochs using GPU acceleration. The runtime per epoch ranged from approximately 33 to 64 seconds.

Model Performance Without Optimizers

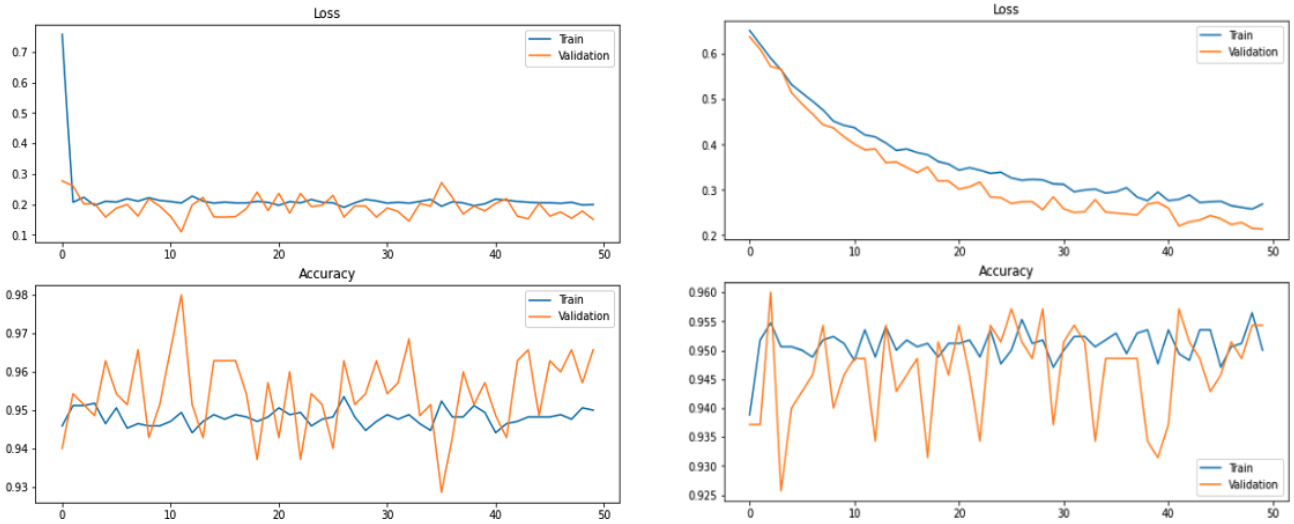
DenseNet169

The model was trained for 50 epochs. The first epoch took 59 seconds, while the remaining epochs took 33 - 39 seconds with GPU acceleration. At the 50th epoch, the model demonstrated strong performance, with a training accuracy of 95.07% and validation accuracy of 94.60%, along with low loss values

XSL•FO
RenderX

(training loss of 0.2014 and validation loss of 0.2101), indicating stable learning. The model achieved a high binary accuracy of 0.9616, along with perfect recall (1.0) and strong precision (0.962). These results highlight that the model identified all true positives with few false positives. Figure 5A shows the training and validation accuracy and loss by epoch for DenseNet169.

Figure 5. Training and validation accuracy and loss by epoch: (A) DenseNet169, (B) DenseNet201.



DenseNet201

The model was trained for 50 epochs. The first epoch took 64 seconds, while the subsequent epochs took 35 seconds with GPU acceleration. At the 50th epoch, the model performed well, achieving a training accuracy of 95% and validation accuracy of 95.43% with low loss values, indicating stable learning. Additionally, the model achieved a binary accuracy of 94.86%, with perfect recall (1.0) and high precision (0.953), indicating it correctly identified all true positives with a low rate of false positives. Figure 5B shows the training and validation accuracy and loss by epoch for DenseNet201.

Model Performance With Optimizers

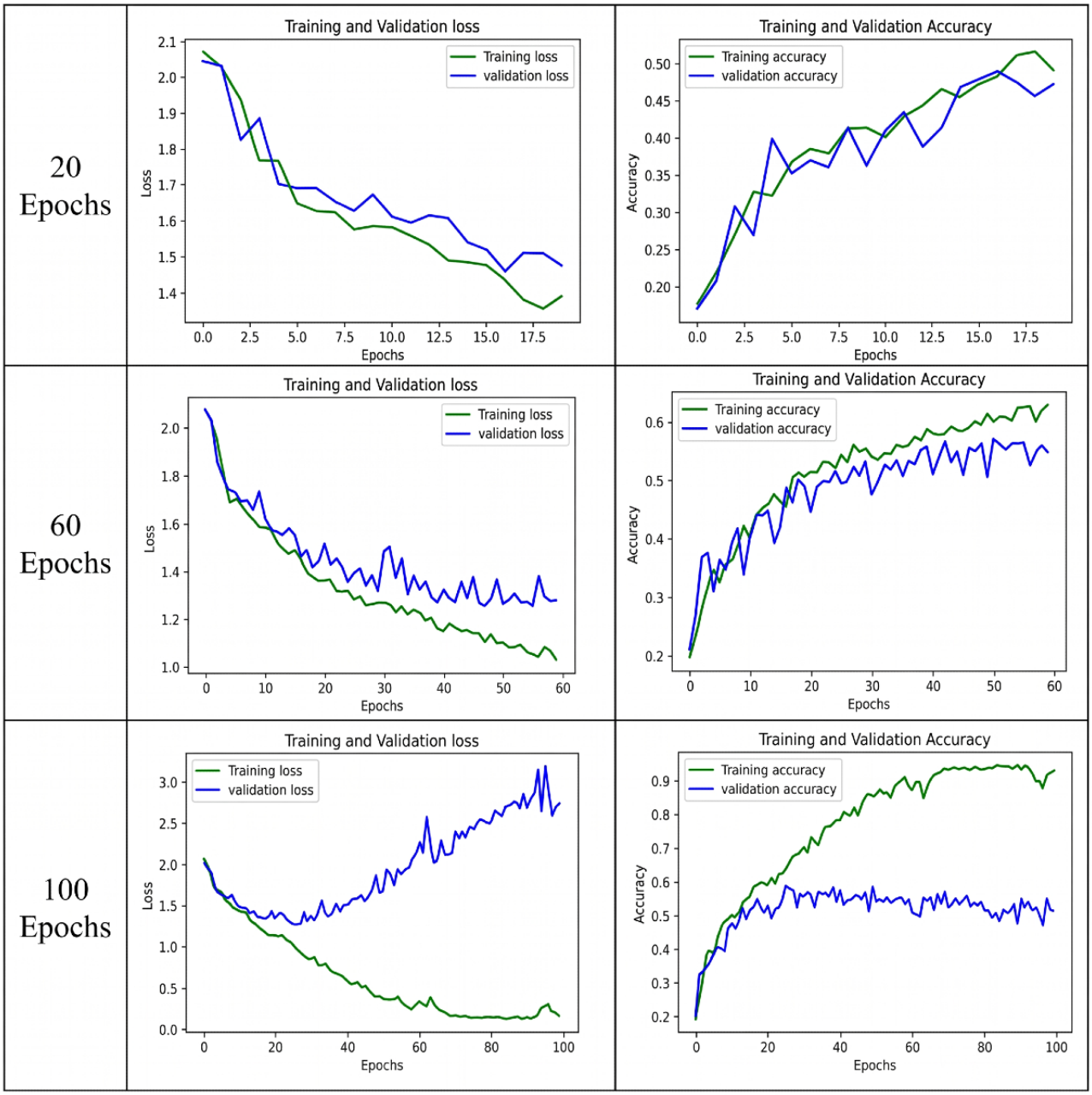
The CNN model showed improvements across epochs and optimizers, as discussed below. Adam is an optimization algorithm that can be used as an alternative to classical stochastic gradient descent, iteratively updating network weights based on training data. With the Adam optimizer, validation accuracy on the augmented dataset increased with the training duration, reaching the maximum value of approximately 85% after 100 epochs. RMSprop yielded comparable accuracy but exhibited fluctuations in loss and validation accuracy at higher epochs, suggesting overfitting.

Comparison of Results Using RMSprop and Adam

Figure 6 shows the training and validation loss and accuracy of the model at 20, 60, and 100 epochs using the Adam optimizer. At 20 and 60 epochs, training and validation loss gradually decreased, with minimal difference between them. These results indicate that the model did not exhibit overfitting or underfitting. However, at 100 epochs, the validation loss initially decreased but began to increase at approximately 40 epochs. In contrast, training loss continued to decrease after 40 epochs. A similar trend was observed in training and validation accuracy. Training accuracy continuously increased with increasing epoch count, but validation accuracy remained approximately 50%. This deviation indicates overfitting. Figure S1 in Multimedia Appendix 1 shows the confusion matrix for the Adam optimizer.

XSL•FO
RenderX

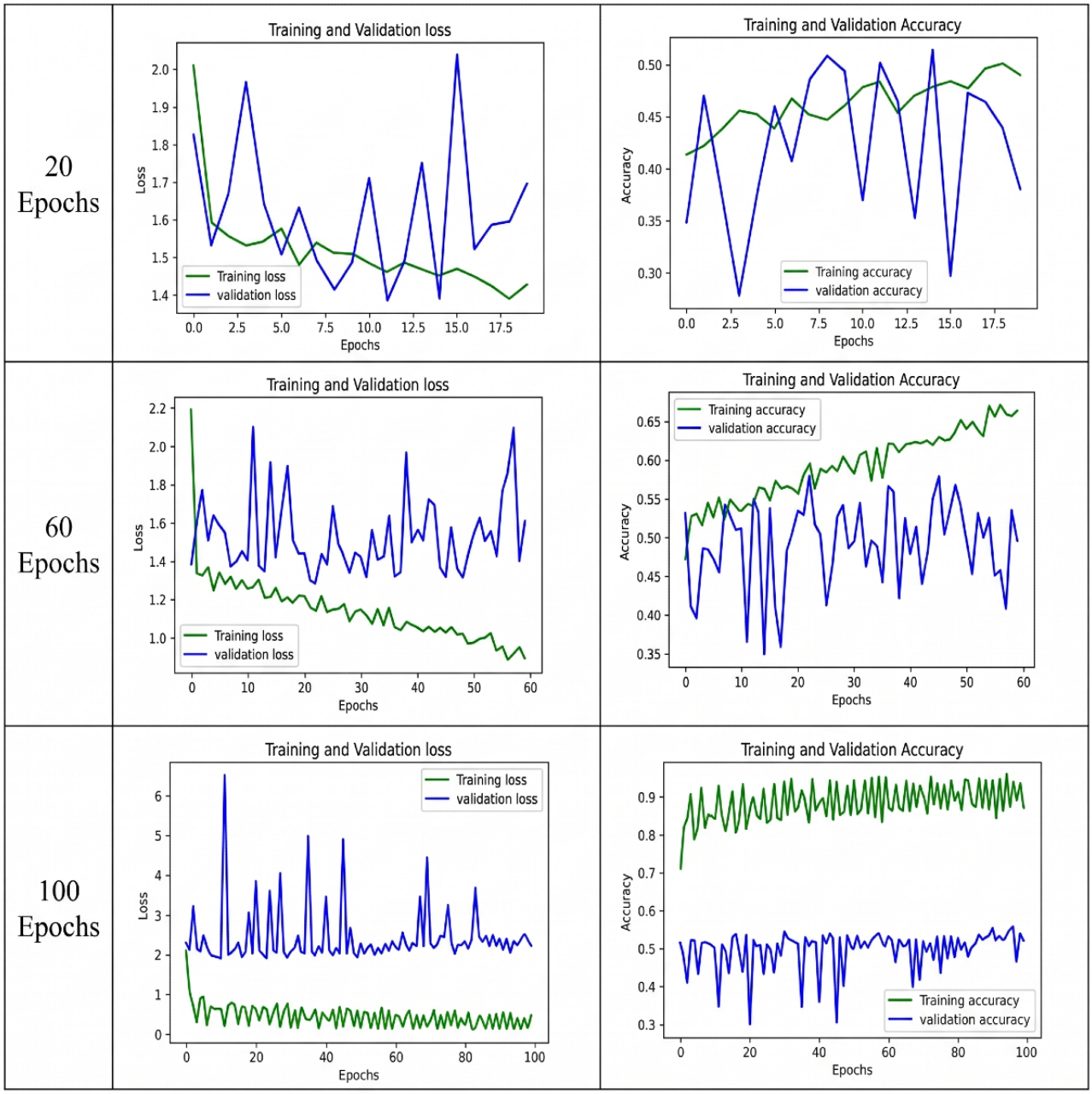
Figure 6. Training and Validation Loss and Accuracy curves for the Adam optimizer. Adam: adaptive moment estimation.



The RMSprop optimizer can be considered similar to the gradient descent algorithm with momentum. This optimizer restricts oscillations in the vertical direction. Therefore, by increasing the learning rate, the model can take larger steps in the horizontal direction, resulting in accelerated convergence. Figure 7 shows the training and validation loss and accuracy of the model at 20, 60, and 100 epochs using the RMSprop Optimizer.

XSL•FO
RenderX

Figure 7. Training and validation loss and accuracy curves for the RMSprop optimizer. RMSprop: root mean square propagation.



The training and validation loss and accuracy curves indicate limited learning efficiency of the model, with sharp fluctuations in both loss and accuracy suggesting overfitting. Figure S2 in Multimedia Appendix 1 shows the confusion matrix for the RMSprop optimizer.

For the complete dataset, the F_1-scores were 31%, 59%, and 87% at 20, 60, and 100 epochs, respectively. Even though the F_1-score was 87% after 100 epochs, the model performance was expected to be unsatisfactory on unseen data (data not in the training and validation sets).

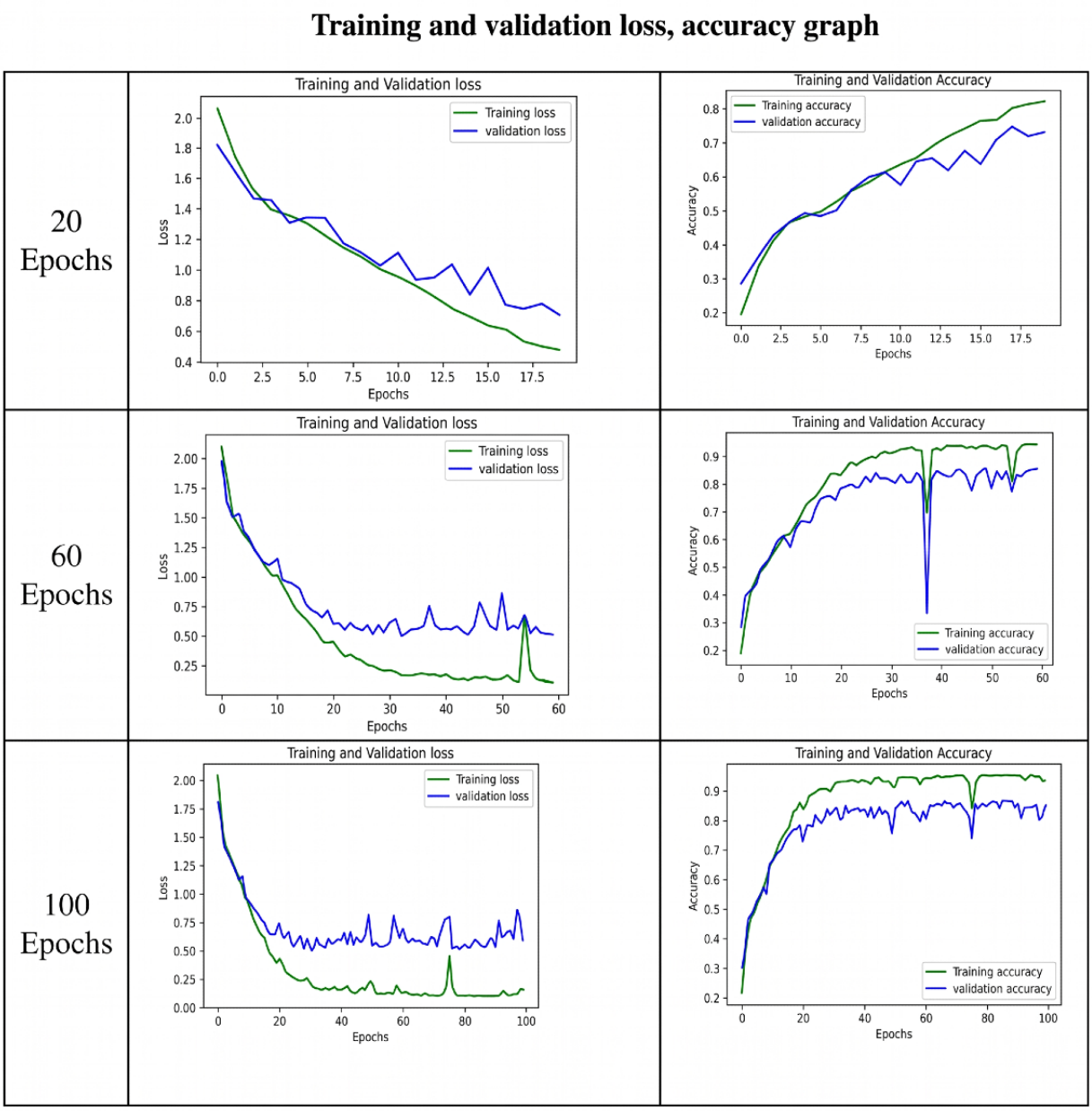
The results indicate that the model overfit after 100 epochs. Lack of sufficient data may contribute to overfitting. Both the Adam and RMSprop optimizers displayed similar F_1-scores after 100 epochs. Although both optimizers exhibit momentum-like behavior (Adam through its bias-corrected moving average and RMSprop with momentum by explicitly adding a fraction of the previous step), the mechanisms are different, which may influence the hyperparameter learning rate.

To enhance model performance, data augmentation was performed using shear rotations of ±10°. Although this transformation led to only minor changes in the existing data, the model treated them as newly input images. Approximately 500 images were added per class to ensure a balanced class distribution.

The CNN was then trained using the augmented dataset. Figures 8 and 9 show the loss and accuracy associated with the augmented training and validation data at 20, 60, and 100 epochs using the Adam and RMSprop optimizers, respectively. Figures S3 and S4 in Multimedia Appendix 1 present the confusion matrices for the augmented validation data using the Adam and RMSprop optimizers, respectively.

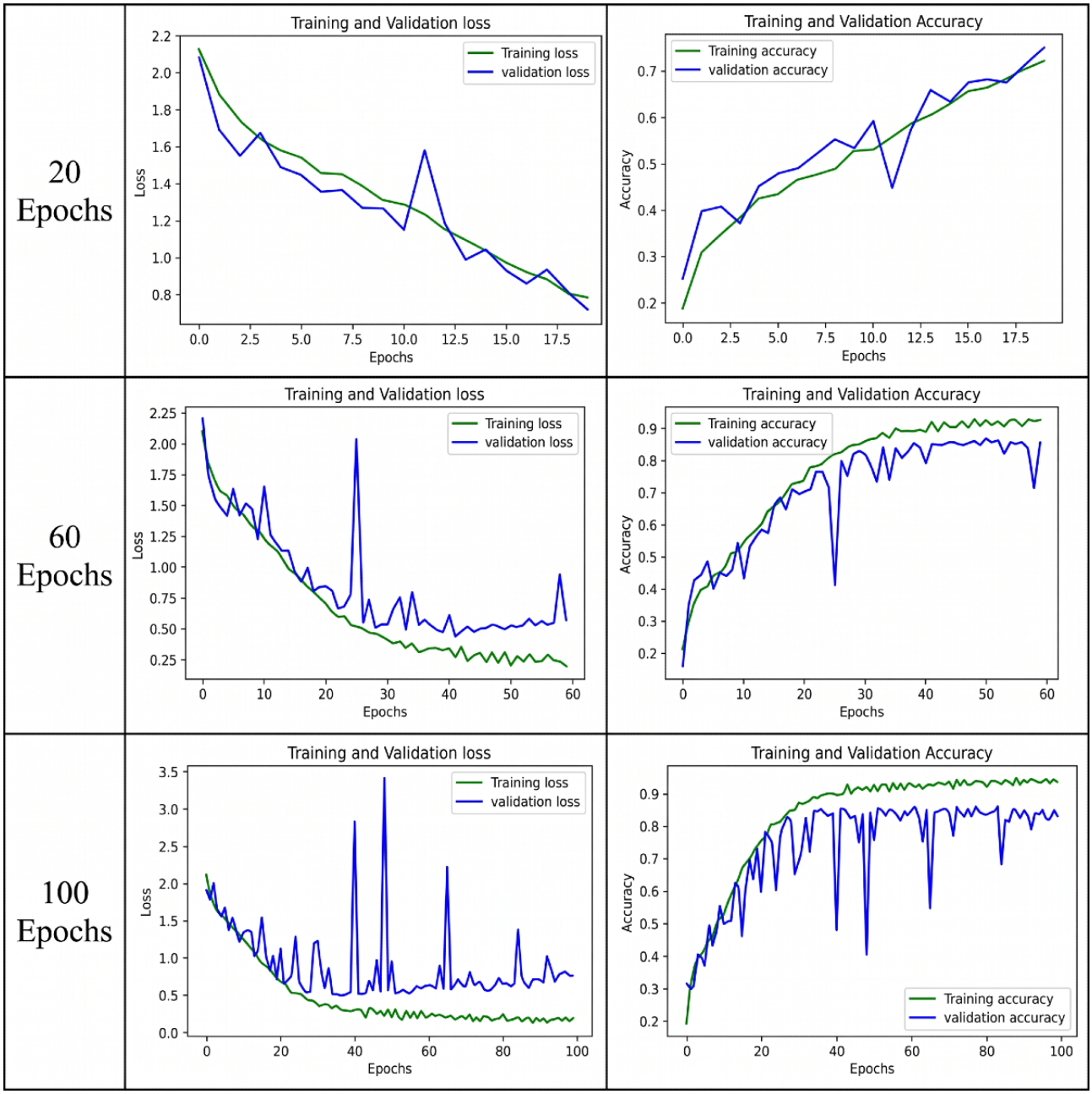
XSL•FO
RenderX

Figure 8. Training and validation loss and accuracy curves for the augmented dataset using the Adam optimizer at 20, 60, and 100 epochs. Adam: adaptive moment estimation.



XSL•FO
RenderX

Figure 9. Training and validation loss and accuracy curves for the augmented dataset using the RMSprop optimizer at 20, 60, and 100 epochs. RMSprop: root mean square propagation.



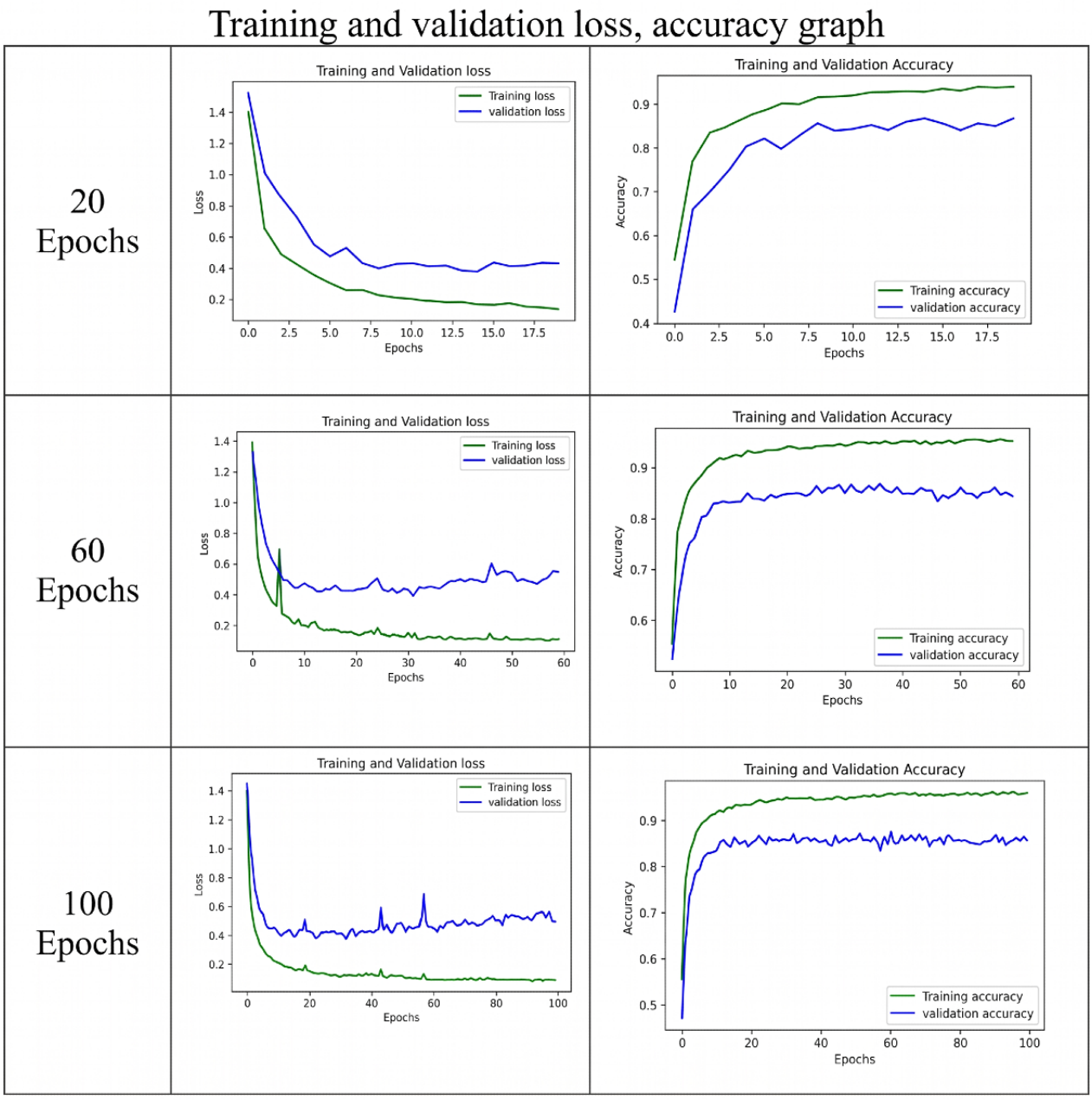
As shown in Figure 8, training and validation loss decreased compared with the results obtained using the original dataset. At 100 epochs, the training and validation accuracy exceeded 90% and 85%, respectively. Although both loss and accuracy improved with the RMSprop optimizer as well (Figure 9), sudden increases in loss and corresponding decreases in accuracy were noted.

DenseNet169

Figures 10 and 11 and Figures S5 and S6 in Multimedia Appendix 1 summarize the performance of DenseNet169 in terms of accuracy, stability, and generalizability. Figures 10 and 11 present the training and validation loss and accuracy curves using the Adam and RMSprop optimizers, respectively. Figures S5 and S6 in Multimedia Appendix 1 present the confusion matrices for the validation set using the Adam and RMSprop optimizers, respectively.

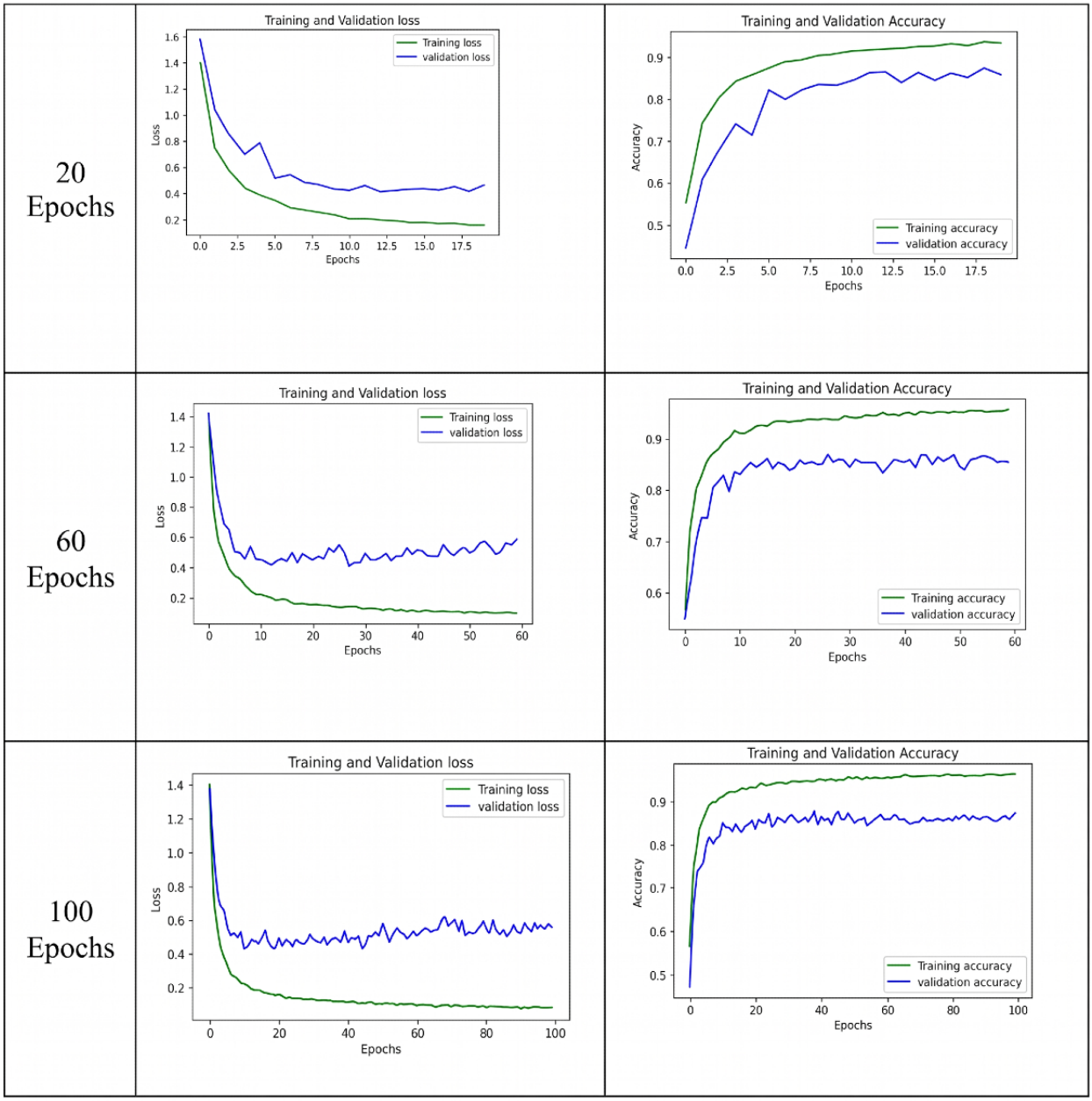
XSL•FO
RenderX

Figure 10. Training and validation loss and accuracy curves for DenseNet169 using the Adam optimizer at 20, 60, and 100 epochs. Adam: adaptive moment estimation.



XSL•FO
RenderX

Figure 11. Training and validation loss and accuracy curves for DenseNet169 using the RMSprop optimizer at 20, 60, and 100 epochs. RMSprop: root mean square propagation.



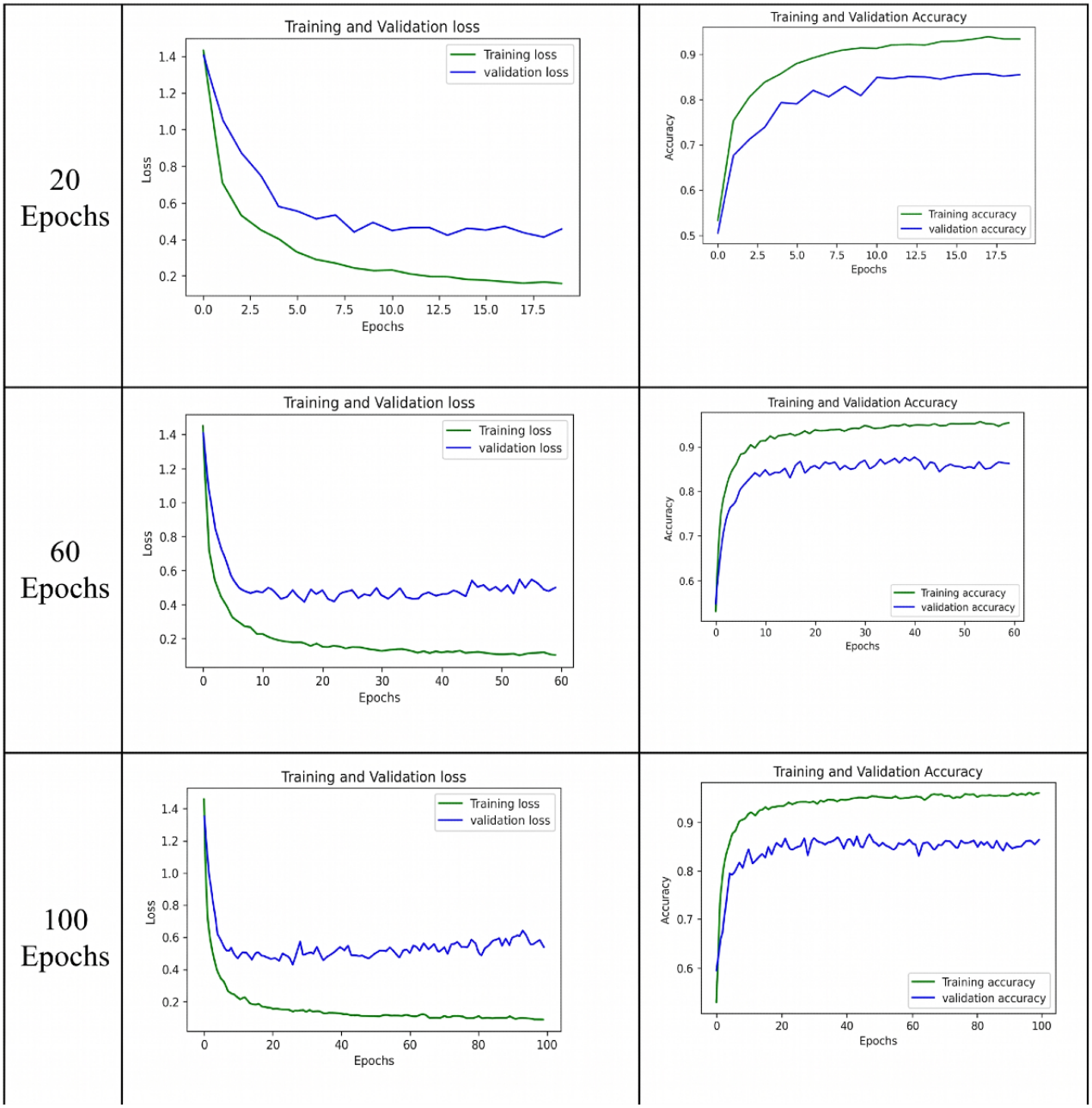
With increase in the dataset size, DenseNet169 achieved higher accuracy and smaller loss than those for the CNN architecture at 20, 40, and 60 epochs. Validation accuracy was 85%

DenseNet201

Figures 12 and 13 and Figures S7 and S8 in Multimedia Appendix 1 summarize the performance of DenseNet201 in terms of accuracy, stability, and generalizability across the two optimization strategies. Figures 12 and 13 present the training and validation loss and accuracy curves using the Adam and RMSprop optimizers, respectively. Figures S7 and S8 in Multimedia Appendix 1 present the confusion matrices for the validation set using the Adam and RMSprop optimizers, respectively.

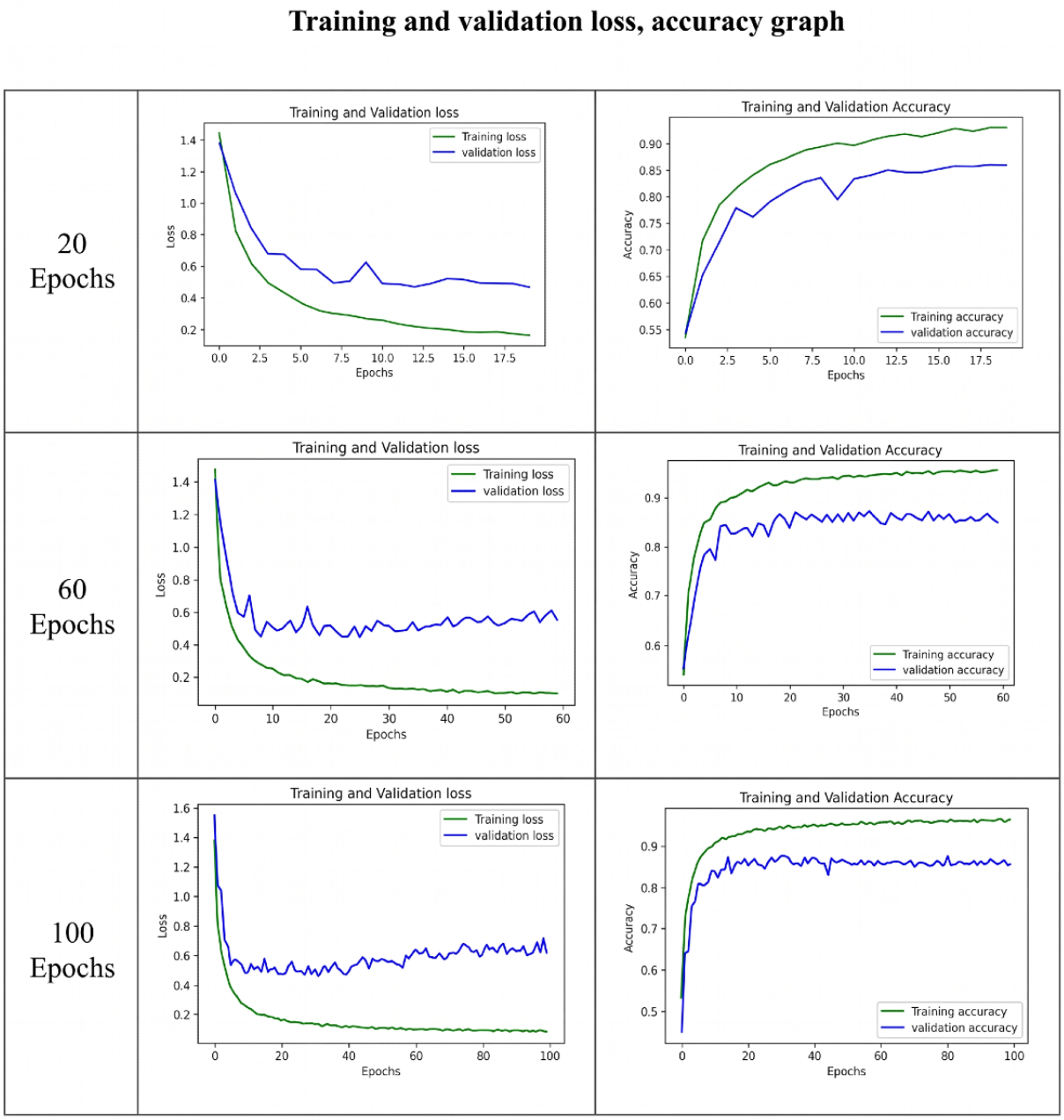
XSL•FO
RenderX

Figure 12. Training and validation loss and accuracy curves for DenseNet201 using the Adam optimizer at 20, 60, and 100 epochs. Adam: adaptive moment estimation.



XSL•FO
RenderX

Figure 13. Training and validation loss and accuracy curves for DenseNet201 using the RMSprop optimizer at 20, 60, and 100 epochs. RMSprop: root mean square propagation.



Comparative Analysis

Both DenseNet variants consistently outperformed the CNN baseline. On the augmented data, both models achieved accuracy and F_1-scores exceeding 94% after 100 epochs, with only minor differences between the Adam and RMSprop optimizers. DenseNet169 achieved a validation accuracy of up to 95% and demonstrated superior generalization benefits compared with both CNN and DenseNet201, attributable to its dense connectivity pattern.

Augmentation involved generating approximately 500 synthetic samples per class using image transformations such as ±10° shear, resulting in improved class balance and model generalizability. The augmented dataset (version 2) was associated with increased accuracy and more stable validation loss curves, suggesting potential reduction in overfitting. Figure 14 shows heatmaps of skin lesion images from the dataset.

XSL•FO
RenderX

Figure 14. Heatmaps of skin lesion images from the dataset.

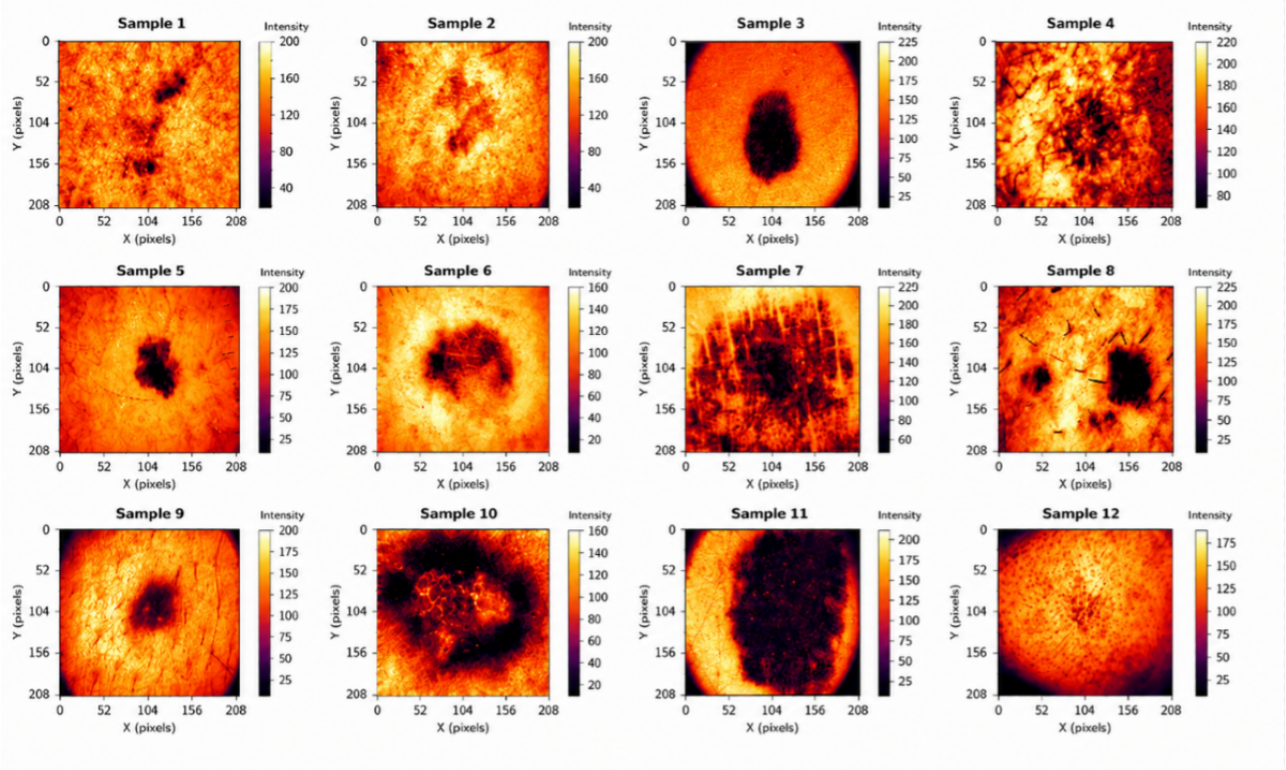


Table 2 presents the accuracy values of the CNN, DenseNet169, and DenseNet201 models. Data were obtained from Kaggle, and the models were also tested on an augmented dataset containing additional samples for each class. The dataset was divided using an 80:20 split. The validation set corresponded to the 20% portion, and the "whole dataset" included both the training and validation sets (80%+20%).

Table . Overall performance (accuracy, %) of the algorithms.

Model and optimizer	Dataset (2357 samples)						Dataset (6857 samples)					
	Validation data			Whole dataset			Validation data			Whole dataset		
	20 epochs	60 epochs	100 epochs	20 epochs	60 epochs	100 epochs	20 epochs	60 epochs	100 epochs	20 epochs	60 epochs	100 epochs
CNN[a]												
Adam[b]	42	52	50	44	59	86	72	86	85	82	94	93
RMSprop[c]	31	48	52	31	59	87	76	86	83	82	94	93
DenseNet169												
Adam	58	58	58	87	88	88	87	84	86	94	94	95
RMSprop	55	57	60	85	87	88	86	86	87	94	95	95
DenseNet201												
Adam	56	58	57	87	88	88	85	86	86	94	95	95
RMSprop	56	56	58	86	88	88	86	84	85	93	94	95

[a]CNN: convolutional neural network.

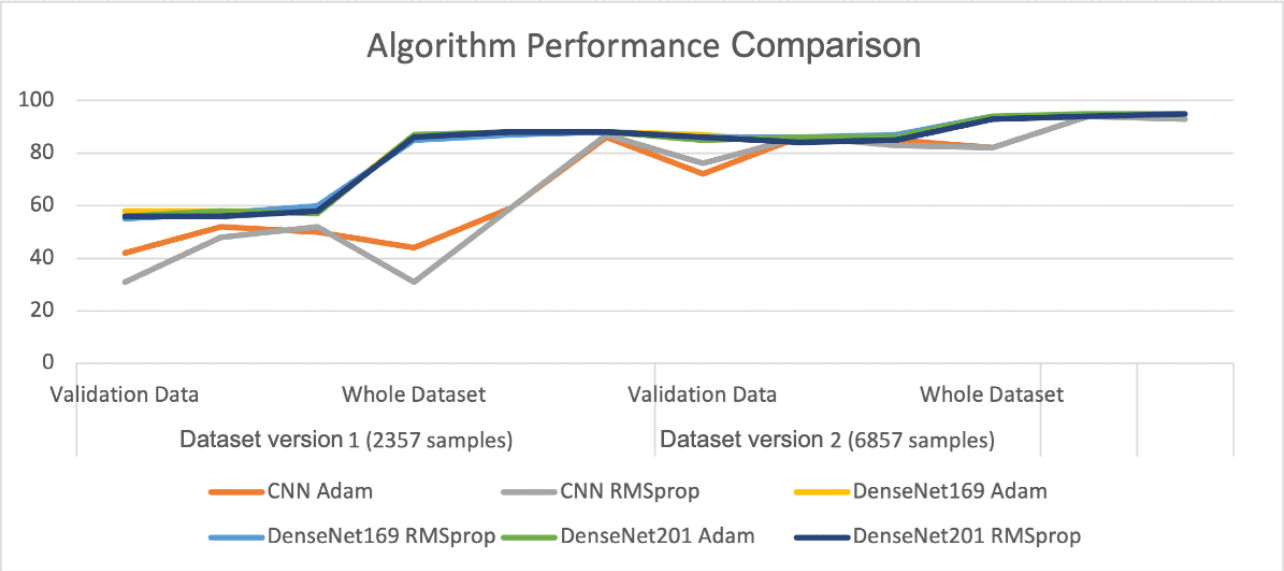
[b]Adam: adaptive moment estimation.

[c]RMSprop: root mean square propagation.

XSL•FO
RenderX

Figure 15 compares the accuracy of the three studied models, highlighting the superior accuracy of DenseNet169 and DenseNet201 compared with the conventional CNN architecture.

Figure 15. Algorithm performance comparison. Adam: adaptive moment estimation; CNN: convolutional neural network; RMSprop: root mean square propagation.



Discussion

Implications

AI is increasingly transforming medicine by enhancing diagnostic accuracy, enabling personalized treatment strategies, and improving overall patient outcomes. Its ability to rapidly process large datasets, identify complex patterns, and support clinical decision-making is reshaping health care delivery. Notably, AI systems can improve diagnostic reliability, reduce human error, and streamline clinical workflows across medical specialties.

In dermatology, AI has garnered considerable attention for its potential to improve the early detection of skin cancer, particularly melanoma. Advanced deep learning models, including CNNs and DenseNet architectures, have demonstrated performance comparable to, and in some cases, exceeding that of dermatologists in the analysis of dermoscopic and clinical images. These approaches can help address inherent challenges, such as subjective interpretation and interobserver variability, offering scalable and consistent solutions for large-scale skin cancer screening.

The integration of CNN and DenseNet architecture into a unified diagnostic framework is especially valuable in clinical practice. When embedded within dermatology workflows as a decision support tool, such systems can aid physicians by automating lesion analysis. Such automation could help reduce diagnostic uncertainty, enhance clinician confidence, and facilitate the efficient classification of multiple lesion types, enabling prioritization of high-risk cases and timely decision-making.

The proposed automated framework incorporates advanced image preprocessing and augmentation techniques to improve model robustness and generalizability. By minimizing reliance on subjective visual assessment and complementing traditional dermoscopy, the system can support earlier and more consistent melanoma detection. Dense connectivity facilitates the extraction of subtle yet clinically relevant features, improving the discrimination of benign and malignant lesions.

AI-assisted analysis can also contribute to improved clinical efficiency. Notably, rapid and accurate lesion classification can accelerate diagnostic workflows, reduce the time spent on image interpretation, and shorten patient visits, potentially yielding an increase in patient throughput while maintaining diagnostic quality. Additionally, AI-driven triage may allow trained technical staff to assist with image acquisition and preliminary screening, optimizing resource utilization within clinical settings.

From an economic perspective, AI implementation in dermatology may have the potential to reduce health care costs. Early and accurate detection of melanoma could decrease unnecessary biopsies and limit the need for costly treatments associated with advanced-stage disease. By identifying high-risk lesions earlier, the proposed framework could support more effective resource allocation, ultimately improving both economic aspects and patient outcomes.

Moreover, automation of routine image analysis may allow clinical staff to devote more time to patient coordination and care delivery. This redistribution of responsibilities would enhance operational efficiency and may help reduce provider burnout.

The compatibility of the proposed framework with dermoscopic imaging may support the use of AI in teledermatology. Patients could remotely submit lesion images for preliminary assessment, enabling prioritization of urgent cases and enhancing access to care in underserved or remote regions. This capability could strengthen continuity of care through remote monitoring and timely intervention.

Future work could be focused on improving model interpretability and performance through the integration of

XSL•FO
RenderX

attention mechanisms into DenseNet architectures. These mechanisms could allow the model to highlight salient image regions, enhancing both diagnostic accuracy and transparency, which may increase clinician trust. Additionally, incorporating multimodal data—such as patient demographics and clinical history—alongside imaging features may further improve diagnostic precision. Validation across diverse populations and real-world clinical settings will be essential to ensure generalizability and support widespread clinical adoption.

Although the findings of this study are encouraging and suggest the potential clinical utility of the proposed framework, further validation is necessary before routine clinical implementation. Specifically, prospective evaluation across multiple institutions using larger and independent patient cohorts will be essential to assess generalizability across patient populations, imaging systems, clinical environments, and real-world workflows. Such multicenter validation may also help identify potential sources of variability and bias that may not be fully captured within the current dataset. Future studies should therefore prioritize prospective external validation to establish the robustness, reproducibility, and clinical reliability of the framework before its integration into routine clinical practice.

Clinical Translation and Future Deployment

Although the proposed CNN and DenseNet architectures demonstrated strong diagnostic performance on the evaluated dataset, clinical deployment requires validation beyond retrospective benchmarking. Before implementation in routine dermatology practice, the models should undergo prospective, multicenter external validation using images acquired from diverse populations, imaging devices, and clinical environments to assess generalizability. Integration into clinical workflows will also require consideration of computational efficiency, interoperability with electronic health record systems, regulatory approval, data privacy, and clinician acceptance. Continuous performance monitoring and periodic model retraining will be essential to mitigate performance degradation resulting from distribution shifts in real-world data. Future work should therefore focus on external validation, explainability, uncertainty estimation, and workflow integration to facilitate the safe and effective adoption of these models as clinical decision support systems.

Conclusion

This study presented a comprehensive comparative evaluation of CNN and DenseNet-based deep learning architectures for automated classification of melanoma using dermoscopic skin lesion images. In experiments conducted under multiple training configurations, the DenseNet models consistently demonstrated superior classification performance compared with the baseline CNN architecture, with higher accuracy, sensitivity, specificity, precision, and F_1-score. These results highlight that dense feature propagation and feature reuse may benefit the identification of malignant skin lesions while reducing false positive and false negative predictions. The presented comparison of commonly used deep learning architectures can guide researchers and clinicians in selecting appropriate models for computer-aided melanoma diagnosis. The proposed framework can serve as a clinical decision support tool, assisting dermatologists in the early detection of melanoma, thereby facilitating timely intervention and improving patient outcomes. Although further prospective, multicenter external validation is required before routine clinical implementation, the findings demonstrate the potential of deep learning–based image analysis in enhancing diagnostic accuracy and supporting precision dermatology. Overall, this work contributes to the growing body of evidence supporting the integration of AI into dermatological practice and establishes a strong foundation for future research on robust, interpretable, and clinically deployable skin cancer screening systems.

Generative AI Use

The authors declare that no generative AI tools were used in the preparation, writing, and revision of this manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data analyzed in this study were obtained from a publicly available dermoscopic skin lesion dataset hosted on Kaggle [28]. The dataset contains images representing multiple categories of benign and malignant skin lesions and was used in accordance with the terms and conditions of the source repository. No new patient-level clinical data were collected specifically for this study.

Authors' Contributions

Conceptualization: HS, TM
Data curation: AA, IA, MP, RP
Formal analysis: AA, IA, MP, SW, RP
Investigation: AA, IA, RB, RP
Methodology: HS, MP, SW, TM
Project administration: HS
Resources: RB
Software: AA, MP

XSL•FO
RenderX

Supervision: HS, TM
Validation: HS, MP, SW, RB, TM
Visualization: AA, IA, SW, RP
Writing – original draft: HS, AA, IA
Writing – review & editing: HS, MP, SW, RB, RP, TM
All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental figures S1 to S8, showing confusion matrices for model validation performance using adaptive moment estimation (Adam) and root mean square propagation (RMSprop) optimizers.
[DOCX File, 1511 KB - derma_v9i1e104086_app1.docx]

References

1. Dildar M, Akram S, Irfan M, et al. Skin cancer detection: a review using deep learning techniques. Int J Environ Res Public Health 2021 May 20;18(10):5479. [doi: 10.3390/ijerph18105479] [Medline: 34065430]
2. Linares MA, Zakaria A, Nizran P. Skin cancer. Prim Care 2015 Dec;42(4):645-659. [doi: 10.1016/j.pop.2015.07.006] [Medline: 26612377]
3. Sample A, He YY. Mechanisms and prevention of UV-induced melanoma. Photodermatol Photoimmunol Photomed 2018 Jan;34(1):13-24. [doi: 10.1111/phpp.12329] [Medline: 28703311]
4. Saginala K, Barsouk A, Aluru JS, Rawla P, Barsouk A. Epidemiology of melanoma. Med Sci (Basel) 2021 Oct 20;9(4):63. [doi: 10.3390/medsci9040063] [Medline: 34698235]
5. Davis LE, Shalin SC, Tackett AJ. Current state of melanoma diagnosis and treatment. Cancer Biol Ther 2019;20(11):1366-1379. [doi: 10.1080/15384047.2019.1640032] [Medline: 31366280]
6. Rastrelli M, Tropea S, Rossi CR, Alaibac M. Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. In Vivo 2014;28(6):1005-1011. [Medline: 25398793]
7. Dzwierzynski WW. Melanoma risk factors and prevention. Clin Plast Surg 2021 Oct;48(4):543-550. [doi: 10.1016/j.cps.2021.05.001] [Medline: 34503715]
8. D'Orazio J, Jarrett S, Amaro-Ortiz A, Scott T. UV radiation and the skin. Int J Mol Sci 2013 Jun 7;14(6):12222-12248. [doi: 10.3390/ijms140612222] [Medline: 23749111]
9. Smaoui N, Bessassi S. A developed system for melanoma diagnosis. Int J Comput Vis Signal Process 2013;3:10-17 [FREE Full text]
10. Marchetti MA, Liopyris K, Dusza SW, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: results of the International Skin Imaging Collaboration 2017. J Am Acad Dermatol 2020 Mar;82(3):622-627. [doi: 10.1016/j.jaad.2019.07.016] [Medline: 31306724]
11. Cheong KH, Tang KJW, Zhao X, et al. An automated skin melanoma detection system with melanoma-index based on entropy features. Biocybern Biomed Eng 2021 Jul;41(3):997-1012. [doi: 10.1016/j.bbe.2021.05.010]
12. Waheed SR, Saadi SM, Mohd Rahim MS, et al. Melanoma skin cancer classification based on CNN deep learning algorithms. Mal J Fund Appl Sci 2023;19(3):299-305. [doi: 10.11113/mjfas.v19n3.2900]
13. Alsaade FW, Aldhyani THH, Al-Adhaileh MH. Developing a recognition system for diagnosing melanoma skin lesions using artificial intelligence algorithms. Comput Math Methods Med 2021;2021:9998379. [doi: 10.1155/2021/9998379] [Medline: 34055044]
14. Indraswari R, Rokhana R, Herulambang W. Melanoma image classification based on MobileNetV2 network. Procedia Comput Sci 2022;197:198-207. [doi: 10.1016/j.procs.2021.12.132]
15. Tumpa PP, Kabir MA. An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features. Sens Int 2021;2:100128. [doi: 10.1016/j.sintl.2021.100128]
16. Brinker TJ, Hekler A, Enk AH, et al. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer 2019 Apr;111:148-154. [doi: 10.1016/j.ejca.2019.02.005] [Medline: 30852421]
17. Thaajwer MAA, Ishanka UAP. Melanoma skin cancer detection using image processing and machine learning techniques. Presented at: 2020 2nd International Conference on Advancements in Computing (ICAC); Dec 10-11, 2020; Malabe, Sri Lanka p. 363-368. [doi: 10.1109/ICAC51239.2020.9357309]
18. Banerjee S, Singh SK, Chakraborty A, Das A, Bag R. Melanoma diagnosis using deep learning and fuzzy logic. Diagnostics (Basel) 2020 Aug 9;10(8):577. [doi: 10.3390/diagnostics10080577] [Medline: 32784837]
19. Bakheet S, El-Nagar A. A deep neural approach for real-time malignant melanoma detection. Appl Math Inf Sci 2021 Jan 1;15(1):89-96. [doi: 10.18576/amis/150111]

XSL•FO
RenderX

20. Chang CC, Li YZ, Wu HC, Tseng MH. Melanoma detection using XGB classifier combined with feature extraction and k-means SMOTE techniques. Diagnostics (Basel) 2022 Jul 19;12(7):1747. [doi: 10.3390/diagnostics12071747] [Medline: 35885650]
21. Thiyaneswaran B, Anguraj K, Kumarganesh S, Thangaraj K. Early detection of melanoma images using gray level co - occurrence matrix features and machine learning techniques for effective clinical diagnosis. Int J Imaging Syst Technol 2021 Jun;31(2):682-694. [doi: 10.1002/ima.22514]
22. Young AT, Vora NB, Cortez J, et al. The role of technology in melanoma screening and diagnosis. Pigment Cell Melanoma Res 2021 Mar;34(2):288-300. [doi: 10.1111/pcmr.12907] [Medline: 32558281]
23. Kim P. MATLAB Deep Learning: Apress; 2017. [doi: 10.1007/978-1-4842-2845-6]
24. Fu'adah YN, Pratiwi NC, Pramudito MA, Ibrahim N. Convolutional neural network (CNN) for automatic skin cancer classification system. IOP Conf Ser: Mater Sci Eng 2020 Dec 1;982(1):012005. [doi: 10.1088/1757-899X/982/1/012005]
25. Gajera HK, Nayak DR, Zaveri MA. A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. Biomed Signal Process Control 2023 Jan;79:104186. [doi: 10.1016/j.bspc.2022.104186]
26. Huang G, Liu Z, Maaten LVD, Weinberger KQ. Densely connected convolutional networks. arXiv. Preprint posted online on Jan 28, 2018. [doi: 10.48550/arXiv.1608.06993]
27. Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z. Deeply-supervised nets. arXiv. Preprint posted online on Sep 25, 2014. [doi: 10.48550/arXiv.1409.5185]
28. Hussain SI. ISIC 2019 skin cancer detection dataset. Kaggle. URL: https://www.kaggle.com/dsv/16086084 [accessed 2026-10-26]

Abbreviations

ABCDE: asymmetry, border irregularity, color heterogeneity, diameter, and evolution
Adam: adaptive moment estimation
CNN: convolutional neural network
DenseNet: densely connected neural network
NMSC: nonmelanoma skin cancer
ReLu: rectified linear unit
RMSprop: root mean square propagation

Edited by R Dellavalle; submitted 08.Jun.2026; peer-reviewed by J Meisenheimer, X Liang; revised version received 02.Sep.2026; accepted 23.Sep.2026; published 27.Oct.2026.

Please cite as:
Sanghvi H, Agritis A, Alvarez I, Patel M, Wang S, Patel R, Banoub R, Makkar T
Deep Learning–Based Classification of Melanoma Using Dermoscopic Images: Comparative Evaluation of Convolutional Neural Network and DenseNet Models
JMIR Dermatol 2026;9:e104086
URL: https://derma.jmir.org/2026/1/e104086
doi:10.2196/104086

© Harshal Sanghvi, Argyrios Agritis, Isabella Alvarez, Mitsu Patel, Shuwen Wang, Riki Patel, Raphael Banoub, Timonthy Makkar. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Rise in Homemade Sunscreen Trends and Future Impacts on Skin Cancer Risk: Systematic Review

Gabi Kaftan[1], PharmD; Mia Panillo[2], BA; Lydia Maxon[3], BS; Emily Esposito[1], PhD

[1]College of Pharmacy and Health Sciences, Sullivan University, 2100 Gardiner Ln, Louisville, KY, United States

[2]College of Osteopathic Medicine, Rocky Vista University, Englewood, CO, United States

[3]St. Olaf College, Northfield, MN, United States

Corresponding Author:
Gabi Kaftan, PharmD
College of Pharmacy and Health Sciences, Sullivan University, 2100 Gardiner Ln, Louisville, KY, United States

Abstract

Background: The trend of homemade sunscreen recipes has rapidly gained popularity over the last decade, being largely fueled by social media influencers, natural health blogs, and the growing mistrust of large health organizations like the US Food and Drug Administration (FDA). Consumers are increasingly drawn to products labeled "*natural*," "*organic*," "*vegan*," and "*cruelty-free*," often conflating these terms with safety and effectiveness. However, when applied to sun protection, these assumptions can be dangerously misleading. According to the FDA, there is no verified mathematical formula that can be used to determine an accurate sun protection factor (SPF) rating for the amount of zinc oxide used in many of the recipes found online.

Objective: The objective of this review is to examine the rising popularity and efficacy of homemade sunscreens compared to commercial sunscreens and highlight the potential public health implications of skin cancer risk related to homemade sunscreen use.

Methods: Multiple databases were searched to find current and relevant literature analyzing the social media impact relating to homemade sunscreens and the efficacy of homemade sunscreens compared to commercially available products. Digital content that included or discussed homemade sunscreen recipes was included as well to provide examples of information trending online.

Results: One study analyzing social media trends found that many strong, unverified claims relating to homemade sunscreen versus commercial sunscreen use are being presented in the media, potentially influencing public knowledge. Additional studies analyzing the components of homemade sunscreen recipes compared to components in commercially available products found that homemade sunscreens offer little to no sun protection, therefore increasing the risk of UV damage for individuals using homemade products.

Conclusions: Based on the presented evidence, the adoption of unregulated homemade sunscreens presents a substantial threat to public health. Misinformation spread through social media and influencer culture may unintentionally contribute to an increase in UV-related skin cancers. Greater efforts are needed from health care professionals, regulatory bodies, and online platforms to educate the public and promote the use of scientifically validated sun protection methods.

(JMIR Dermatol 2026;9:e84694) doi:10.2196/84694

KEYWORDS

sunscreen; SPF; homemade; commercial; ultraviolet; ultraviolet radiation; sun protection; social media; UV

Introduction

Skin cancer is the most commonly diagnosed cancer in the United States. Incidence rates for skin cancers continue to rise, especially for basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), which together account for the majority of cases. In the United States, approximately 5.4 million cases of BCC and SCC are diagnosed annually in about 3.3 million individuals, as many people develop multiple tumors [1]. Melanoma is much less common (about 1% of skin cancers), but it is responsible for most skin cancer deaths.

Within the last decade, the popularity of do-it-yourself (DIY) skin care products has overtaken social media due to the rise in fear-mongering over synthetic ingredients deemed "bad" or "harmful." A post on the popular blog *Wellness Mama* [2] mentions that many sunscreen products contain ingredients such as oxybenzone, which are "endocrine-disrupting chemicals" that "harm ocean life, especially coral." Statements such as these lead consumers to perceive DIY, "natural," and/or "organic" sunscreens as safer and more eco-friendly alternatives to commercial products. However, their effectiveness in protecting against harmful UV radiation remains largely unverified. The US Food and Drug Administration (FDA) does not define or regulate "organic" as it applies to skincare, and

XSL•FO
RenderX

without FDA regulation, “organic” can be placed on any product without further verification [3].

Sunscreen works to protect from UV radiation by blocking, reflecting, or refracting UV rays from the skin [4]. UVA rays penetrate deeper into the dermis, which induces photoaging, generates reactive oxygen species, and causes DNA damage [5]. UVB primarily penetrates the epidermis and is directly linked to skin cancers and DNA damage [6]. There are two main types of compounds used in sunscreens to block UV radiation, which are physical filters, such as the minerals titanium dioxide and zinc oxide, and chemical filters, such as oxybenzone and avobenzone [4]. Chemical filters are absorbed systemically and absorb high-intensity UV radiation, which excites the chemical compounds to a higher energy state. When the compounds eventually return to their ground state, they then convert the absorbed energy into lower-energy wavelengths, such as infrared radiation. In contrast, physical filters are not absorbed systemically and act as a physical UV barrier that sits on top of the skin.

A study sponsored by the FDA in 2019 found that the blood level of 4 organic compounds surpassed FDA guidelines when sunscreen was applied under maximum use conditions [4]. However, consumers typically do not use this amount of sunscreen, and the study still promoted the use of sunscreen due to its protection from UV rays. The systemic absorption of chemical filters has raised concerns about sunscreen use, especially since animal studies have suggested potential endocrine, reproductive, and neurological toxicities [7]. A study published in 2001 found endocrine-disrupting effects in rats after being administered oral oxybenzone [8]. However, the doses used in this study are essentially unrealistic and unattainable in humans. At this time, there is no strong evidence to suggest negative biological effects from sunscreen.

The homemade sunscreen recipes found throughout the internet all share a common theme: they include different amounts of UV-blocking agents that equate to specific SPF values. For example, *Modern Hippie, Inc* [9] is a health and wellness blog that promotes the use of homemade sunscreen recipes and advises against wearing commercial sunscreens. The blog author claims that 0.27 ounces of zinc oxide equate to an SPF of 2 - 5, following the statement that SPF values are not exact, as they are not tested in a laboratory setting. While this conversion seems helpful to users who intend to make their own sunscreen, it is not approved by any formal governing bodies, such as the FDA, and is therefore unreliable in determining safety and efficacy [10].

In contrast to these posts, a 2024 blog post from The Skin Cancer Foundation [11] states, “Homemade formulations that are being shared on social media have no science behind them, and can be dangerous to use.” Commercial sunscreens are considered over-the-counter drugs instead of cosmetic products, which means they must undergo rigorous testing to meet the FDA’s strict criteria. Because of this, the United States only has 16 UV filters approved for use, whereas the European Union has approved 29 filters since sunscreens are labeled as cosmetics [5]. In 2019, the FDA developed generally recognized as safe and effective (GRASE) categories for UV filters, with the 3 categories being category 1: GRASE; category 2: not GRASE; and category 3: insufficient data [7]. Currently, only zinc oxide and titanium dioxide are categorized as GRASE, indicating that these UV filters have undergone a series of clinical investigations that proved them to be noncarcinogenic and without reproductive or developmental risks. These FDA guidelines demonstrate the importance of using FDA-approved sunscreen products for consumer safety.

The clean beauty movement has traded in safe and effective ingredients for an increased risk of skin cancer and contact dermatitis [12]. While it has been proven that the use of a broad-spectrum SPF 30+ decreases the risk for skin cancer, this demographic believes they will develop cancer from FDA-approved sunscreens, not from unprotected sun exposure. This rejection of modern science raises important public health concerns, especially regarding the potential increase in skin cancer risks due to insufficient sun protection. This systematic review aims to analyze the rising popularity, ingredients, and efficacy of DIY sunscreen and highlight the potential risks associated with these non–FDA-approved, homemade sunscreen products.

Methods

Articles were selected from peer-reviewed journals based on article content, quality, and relevance to the topic of interest. PubMed and Google Scholar were used to find the selected articles using the search terms “homemade sunscreen,” “DIY sunscreen,” and “homemade sunscreen and social media.” Articles were included if they mentioned relevant information regarding the components and/or effectiveness of homemade or commercial sunscreens, or if the article presented information relating to the impacts of homemade sunscreen social media trends. If the article was not available in English, it was excluded. Popular search engines, social media sites, blogs, and other online forums that discussed the production of homemade sunscreens were also searched to provide examples of information relating to homemade sunscreen products being presented in nonscientific or non–peer-reviewed forms. A summary of the inclusion and exclusion criteria is provided in Textbox 1.

The systematic review was conducted based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) systematic review methodology. All researchers participated in screening and evaluated the suitability of the articles based on the inclusion and exclusion criteria. Titles and abstracts were screened initially by dividing the number of articles in half, with two researchers screening one half of the articles, and two screening the other half, and each pair discussing any discrepancies. For the full-text screening, all authors participated in screening all articles and discussing any discrepancies as a group. Articles were only selected if all authors agreed on inclusion. The completed PRISMA checklist can be found in Checklist 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Article content: mentioned relevant information regarding the components and/or effectiveness of homemade or commercial sunscreens. Presented information relating to the impacts of homemade sunscreen social media trends.
- Language: available in English.

Exclusion criteria

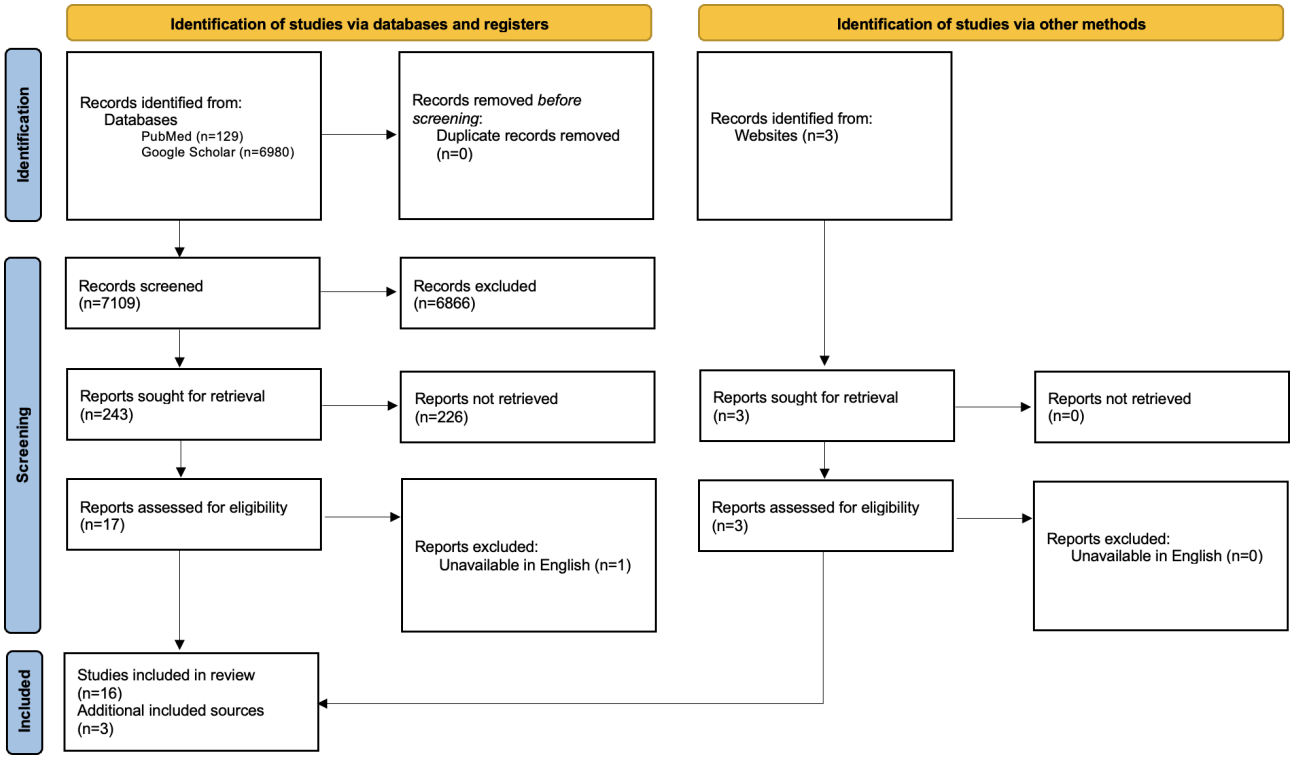
- Language: not available in English or an English translation.

Results

Article Selection

From the search using the specified search terms, 129 articles were identified from PubMed and 6980 were identified from Google Scholar, for a total of 7109 articles (Figure 1). No duplicates were removed in the initial screening. After initial title and abstract screening, the articles were narrowed down to 243 articles, and then to 17 after full article screening based on the inclusion and exclusion criteria. Of the 17 articles selected, one article was excluded because it was not available in English, resulting in 16 peer-reviewed articles being chosen for inclusion. Three online blogs were chosen for additional context relating to our topic.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Information Being Portrayed by the Media

In a study published in 2020, two researchers coded Pinterest pins via a pilot-tested codebook that had an interrater reliability of 90% [13]. Of the 189 pins included in the study, 95.2% described the effectiveness of homemade sunscreens in a positive manner, and 68.3% of pins recommended recipes for homemade sunscreens that offered insufficient UV protection. In addition, 33.3% of pins made SPF claims ranging from SPF 2 to SPF 50; however, these claims were not backed by evidence-based literature or scientifically verified. This study demonstrates how strong, unverified claims can be presented in the media, potentially influencing public knowledge in a way that may contrast with evidence-based science.

Comparing the Ingredients Between Homemade Sunscreens and Commercial Sunscreens

Homemade sunscreens are typically made of 4 components: an emollient (oil) base, a photoprotective agent, a water-resistant agent, and essential oils [14]. The emollient base provides a smooth, spreadable component for other ingredients to be mixed into so that the resulting product is easily spreadable on one's skin. The photoprotective agent is typically an inorganic mineral compound, such as zinc oxide or titanium dioxide, which is the primary UV-blocking component of the recipe that also reflects UV rays as well. The third component usually is a water-resistant agent, such as beeswax, which is included to prevent water loss from the skin and provide environmental protection. Lastly, some recipes include essential oils, which

XSL•FO
RenderX

are an optional ingredient that can be added for fragrance or color. A summary of the mentioned homemade sunscreen components, as well as common examples for each, can be found in Table 1.

Table . Main components of homemade sunscreens.

Homemade sunscreen component	Purpose	Examples
Emollient (oil) base	Creates a smooth base to make the sunscreen easily spreadable onto the skin	Olive oil, coconut oil, almond oil, jojoba oil, carrot seed oil, rose hip seed oil, avocado oil, argan oil, sunflower seed oil
Photoprotective agent	Inorganic minerals added to absorb UV radiation	Zinc oxide, titanium dioxide
Water resistant agent	Prevents dermal water loss and protects against environmental irritants	Beeswax (Cera flava: yellow; Cera alba: white)
Optional: essential oil(s)	Added for fragrance or color	Lavender oil, peppermint oil, tea tree oil, eucalyptus oil, geranium oil

In contrast to homemade sunscreens, FDA-approved sunscreens are typically made of the following components: active ingredients, stabilizing agents, sensory enhancers, and optional ingredients [15]. The active ingredients in the product are the primary UV-filtering agents, and the stabilizing agents are added to the formulation to prevent active ingredient degradation and promote overall shelf stability. More specifically, stabilizing agents such as photostabilizers help prevent degradation of the UV-filtering active ingredients, and emulsifiers help create a stable emulsion between the oil and water phases of the sunscreen, ensuring the ingredients are evenly dispersed throughout the product [15]. Sensory enhancers are compounds that improve the feel or scent of the formula, such as fragrances, moisturizers, and emollients, and they help make application of the product a pleasant experience. The purpose and common examples of the commercial sunscreen components mentioned are summarized in Table 2.

Table . Main components of commercial sunscreens.

Commercial sunscreen component	Purpose	Examples
Active ingredients	Compounds that filter the sun	• Chemical: avobenzone, octinoxate, octisalate, octocrylene, homosalate, oxybenzone • Mineral: zinc oxide, titanium dioxide
Stabilizing agents	Photostabilizers, preservatives, thickening agents, and emulsifiers aimed at preventing degradation so that the active ingredients maintain their efficacy and the overall product can remain shelf-stable for long periods of time	• Photostabilizers: ethylhexyl methoxycrylene, polyester-7, neopentyl glycol diheptanoate, chelating agents, antioxidants • Preservatives: phenoxyethanol, parabens, and formaldehyde-releasing preservatives such as DMDM hydantoin and imidazolidinyl urea • Thickening agents: carbomer, cellulose derivatives, acrylates, xanthan gum, agar, arrowroot powder, tapioca starch • Emulsifiers: Lecithin, cetearyl alcohol, glyceryl stearate, polysorbates, sorbitan esters, xanthan gum, and beeswax
Sensory enhancers	Ingredients that change the way the product feels or smells	• Caprylic/capric triglyceride, squalane, silicones, aloe vera, green tea extract, hyaluronic acid
Optional ingredients	Compounds can be added to provide additional benefits to the sunscreen's basic sun protection	• Aloe, blue-light protectors (ie, iron oxides, zinc oxide, titanium dioxide)

Efficacy of Homemade Sunscreens

While the recipes for homemade sunscreens seem rather straightforward, consumers fail to understand that when compounded incorrectly, they can degrade the UV component or produce a volatile substance. Certain compounds, such as furocoumarins, bergapten, and psoralens, are phototoxic and contact-sensitizing compounds commonly present in natural oils used for DIY sunscreens [14]. For example, essential oils from *Seseli libanotis* contain furocoumarins that, upon exposure to UV light, generate bergapten—the primary phototoxic compound responsible for photosensitization reactions.

A toxicological study published in 2024 investigated the phototoxicological risks of plant-derived ingredients used in various homemade sunscreen formulations under simulated sunlight [16]. When testing the formulations, many, if not all, of the homemade recipes showed instability and changes in properties such as odor, color, and phase separations when

XSL•FO
RenderX

compared to pharmacy-grade or commercially available sunscreens. This ingredient instability can contribute to the development of irritant compounds within the resulting sunscreen, which then can put individuals at a heightened risk for the development of contact dermatitis. The study also found that certain plant-derived components in these DIY recipes did not specify the exact source of herbal ingredients or the need for ingredient standardization in the recipe, and having this standardization of herbal components from correctly identified plant sources is crucial for consistent and reliable effects when using these homemade recipes.

The efficacy of photoprotective compounds, such as zinc oxide, also depends on the particle size and suspension homogeneity of compound particles within the resulting sunscreen [17]. Incorrectly prepared emulsions often create particle aggregates because truly homogeneous suspensions are difficult to create. It is also important to note that these DIY recipes do not use accurate measuring devices that would otherwise be used when compounding in a lab, which can impact the true SPF levels these products are aiming to achieve [18]. Two of the studies selected for this review specifically investigated the efficacy of homemade sunscreens in meeting minimum SPF levels and their ability to protect the skin from UV-induced sun damage.

One study followed the protocols for 15 homemade sunscreen recipes found online, and then analyzed the SPF, the production factor in the UVA domain (PF-UVA), and the critical wavelength of each homemade product in a laboratory setting [18]. The study found that 3 of the products did not create any form of sunscreen with any photoprotective effect whatsoever, and thus carried a major risk of sun damage for users of these recipes. The other 12 recipes resulted in a product that had measured SPF levels of 6 or less, which fails to meet the dermatologist-recommended minimum SPF of 30, therefore carrying a major risk of sun damage for users of these recipes as well.

An additional study conducted in Australia in 2025 compared the efficacy of another natural homemade sunscreen recipe found on a wellness blog to a commercially available SPF 50+ sunscreen [19]. The natural homemade sunscreen contained almond oil, coconut oil, shea butter, beeswax, red raspberry seed oil, carrot seed oil, and zinc oxide. Skin explants were treated with either the homemade product, the commercial SPF 50+ sunscreen, or a base lotion prior to UV exposure, and then assessed via immunohistochemistry for the levels of UV-induced DNA damage and the generation of sunburn cells. The study found that the natural homemade sunscreen was comparable to the commercial sunscreen in reducing UV-induced DNA damage, but it did not effectively protect against the generation of sunburn cells. In comparison, the commercial SPF 50+ sunscreen reduced both UV-induced DNA damage and the generation of sunburn cells. The base lotion provided little to no protection, showing the most UV-induced DNA damage under histological examination and the highest proportion of sunburn cells generated per linear mm of skin.

Discussion

Principal Findings

This study found that due to the virality of social media, many claims regarding homemade sunscreen products, which are not backed by scientific evidence, have supported the use of non–FDA-approved products, contributing to misinformation surrounding sun protection. In addition to validating previous studies demonstrating that homemade sunscreen trends promoted online frequently rely on unverified SPF claims and formulations lacking stability, standardization, and regulations, our findings also expand upon prior work by analyzing available social media data and laboratory evidence relating to homemade sunscreen use.

Prior research relating to this topic has established a strong evidence base that, when adequately formulated products are applied and reapplied correctly, both chemical or mineral sunscreen use can significantly reduce an individual's risk of UV-induced sun damage and sun cancer. Despite this established evidence base, information from unverified blogs and other platforms claiming how commercially available sunscreens contain toxic chemicals that are harmful to both the user and the environment tends to overshadow scientific research [4,7,8]. Available evidence relating to this claim is minimal, as related studies in this review found that unrealistic amounts of commercial sunscreen must be used for any level of toxicity to occur.

Rubin and Brod's article [12] discusses how claims such as these have contributed to the "clean beauty" movement and led to consumer mistrust of commercially available products. As a result, many consumers have begun to look for more "natural" or "organic" alternatives to commercially available products. As mentioned previously, the FDA does not formally define or regulate the "organic" label for skincare products, which has led to many products claiming to be "organic" without any verification processing. The rise in the use of DIY or homemade sunscreens as a result is largely unregulated, and many homemade recipes involve the use of substances that have not undergone adequate laboratory standardization testing that would ensure the resultant homemade product has adequate product safety and efficacy. Use of homemade recipes can then lead to increased risk of skin cancer and contact dermatitis since the product has not been tested under standardized processes.

When analyzing the alleged SPF against UV radiation in homemade sunscreen, articles included in this review supported that homemade sunscreen products provide minimal to zero efficacy, therefore putting users at a greater risk for developing UV-related skin cancers. The low efficacy of homemade sunscreens may be due to different factors, such as a lack of stabilizing agents to prevent product degradation and improper compounding of ingredients. DIY recipes also lack accurate measuring devices, with influencers equating different amounts of UV-blocking agents to a specific SPF. Without a regulated formulation, the amount of actual UV protection in homemade sunscreens is unknown. Given this inconsistency, individuals who wear homemade sunscreen have unreliable and ineffective protection from the sun. Based on the evidence presented,

XSL•FO
RenderX

homemade sunscreens may have an SPF as high as 6, with some providing no SPF at all. Dermatologists recommend sunscreens with an SPF of 30 for adequate sun protection, and if homemade sunscreens do not meet this minimum, then users are at higher risk of experiencing UV damage and developing skin cancers.

Commercially available sunscreens provide significantly higher protection against the generation of sunburn cells in comparison to homemade sunscreens [19]. Commercially available, FDA-approved sunscreens thus offer more reliable and effective sun protection. Additionally, FDA-approved sunscreens undergo rigorous testing and are regulated to ensure safety and UV protection. In comparison, homemade sunscreens lack evidence supporting their efficacy and do not follow regulatory standards, which can lead to inconsistent products. People should use commercially available sunscreens that are scientifically backed by evidence for their overall safety and to ensure they are using a product with verifiable efficacy.

Limitations

Several limitations are present in this review. For article selection, only articles available in English were selected. This may have excluded key articles that are only available in other languages. There were only a few studies that specifically analyzed the efficacy of homemade sunscreens in a small-scale laboratory setting, therefore limiting the generalizability of our review. Additionally, our research is limited due to narrow search terms. By searching under phrases such as “homemade sunscreen,” “DIY sunscreen,” and “homemade sunscreen and social media,” relevant articles that exclude this terminology may have been missed. The majority of the studies presented in this review examine the chemical properties of homemade sunscreens due to there being a lack of clinical investigations on the efficacy of homemade sunscreens. This further limits the generalizability of the review.

Future Research

Due to the minimal number of studies that have been published, larger-scale experimental trials would be beneficial in creating a more detailed and generalizable knowledge base of what provides the best sun protection. These studies should include a histological analysis comparing skin reactions after UV exposure on a cellular level between homemade sunscreens and commercial sunscreens, as well as a targeted SPF comparison of UV blocking agents. Additionally, measuring SPF levels among more different homemade sunscreen recipes and seeing how long the homemade products last on the skin compared to commercial products would be helpful as well. Conducting these additional studies could help make the public more well-informed when they are selecting what kind of sunscreen to use and provide more insight into ingredient effectiveness.

Conclusion

With the rise in DIY sunscreen trends on social media, people need to be wary of whether the information readily available is truly backed by scientific evidence. For the most accurate information, people should seek out peer-reviewed journals and instruction from reputable medical professionals. Social media consumers also need to be willing to verify the information they find online and not blindly trust influencer claims. The increase in misinformation on social media should also raise concerns for physicians and public health officials. Physicians should work to increase education on the dangers of UV damage and preventive measures, such as appropriate sun protection and UV protective clothing, to help combat the influence of social media.

Acknowledgments

Artificial intelligence was not used for any portion of the conception, writing, editing, or preparation of this manuscript.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Checklist 1
PRISMA checklist.
[DOCX File, 274 KB - derma_v9i1e84694_app1.docx]

References

1. Wehner MR. Keratinocyte carcinoma: a review. JAMA 2026 Jan 6;335(1):70-81. [doi: 10.1001/jama.2025.18749] [Medline: 41165676]
2. Wells K. Natural homemade sunscreen recipe. Wellness Mama. 2024. URL: https://wellnessmama.com/beauty/homemade-sunscreen/ [accessed 2025-07-01]
3. “Organic” cosmetics. US Food and Drug Administration.: FDA URL: https://www.fda.gov/cosmetics/cosmetics-labeling-claims/organic-cosmetics [accessed 2026-06-26]
4. Sander M, Sander M, Burbidge T, Beecker J. The efficacy and safety of sunscreen use for the prevention of skin cancer. CMAJ 2020 Dec 14;192(50):E1802-E1808. [doi: 10.1503/cmaj.201085] [Medline: 33318091]

XSL•FO
RenderX

5. Ziglar J, Mohammad TF, Gilaberte Y, Lim HW. Sunscreens: updates on sunscreen filters and formulations. Photoderm Photoimm Photomed 2025 May;41(3). [doi: 10.1111/phpp.70026]
6. de Gruijl FR. Photocarcinogenesis: UVA vs UVB. Methods Enzymol 2000;319:359-366. [doi: 10.1016/s0076-6879(00)19035-4] [Medline: 10907526]
7. Abdel Azim S, Bainvoll L, Vecerek N, DeLeo VA, Adler BL. Sunscreens part 2: regulation and safety. J Am Acad Dermatol 2025 Apr;92(4):689-698. [doi: 10.1016/j.jaad.2024.02.066] [Medline: 38777185]
8. Wang SQ, Burnett ME, Lim HW. Safety of oxybenzone: putting numbers into perspective. Arch Dermatol 2011 Jul;147(7):865-866. [doi: 10.1001/archdermatol.2011.173] [Medline: 21768493]
9. Bevan-Cormier C. Tinted sunblock with zinc oxide. Modern Hippie Inc. URL: https://web.archive.org/web/20260115133138/https://www.modernhippiehw.com/blog/tinted-sunblock-zinc-oxide/ [accessed 2025-07-01]
10. Kanellis VG. Sharing recipes and creating potentially dangerous homemade sunscreens. Australas J Dermatol 2020 May;61(2):161. [doi: 10.1111/ajd.13237] [Medline: 31976549]
11. Skin Cancer Foundation. Ask the expert: is homemade sunscreen safe? Skin Cancer Foundation. 2024. URL: https://www.skincancer.org/blog/is-homemade-sunscreen-safe/ [accessed 2025-07-01]
12. Rubin CB, Brod B. Natural does not mean safe-the dirt on clean beauty products. JAMA Dermatol 2019 Dec 1;155(12):1344-1345. [doi: 10.1001/jamadermatol.2019.2724] [Medline: 31553406]
13. Merten JW, Roberts KJ, King JL, McKenzie LB. Pinterest homemade sunscreens: a recipe for sunburn. Health Commun 2020 Aug;35(9):1123-1128. [doi: 10.1080/10410236.2019.1616442] [Medline: 31111735]
14. Arkhipenko Z, Shah SC, Yi RC, Feldman SR. A review on homemade sunscreen. J Integr Dermatol 2025 [FREE Full text]
15. Osterwalder U, Sohn M, Herzog B. Global state of sunscreens. Photodermatol Photoimmunol Photomed 2014;30(2-3):62-80. [doi: 10.1111/phpp.12112] [Medline: 24734281]
16. Ulutaş OK, Karadogan F, Akgül S, Yildiz ZG. Assessing the safety and efficacy of DIY-homemade sunscreens with natural, aromatic and herbal ingredients. Phytochem Lett 2024 Feb;59:101-107. [doi: 10.1016/j.phytol.2023.12.008]
17. Guedens WJ, Reynders M, Van den Rul H, Elen K, Hardy A, Van Bael MK. ZnO-based sunscreen: the perfect example to introduce nanoparticles in an undergraduate or high school chemistry lab. J Chem Educ 2014 Feb 11;91(2):259-263. [doi: 10.1021/ed300851a]
18. Couteau C, Dupont C, Paparis E, Coiffard LJM. Demonstration of the dangerous nature of “homemade” sunscreen recipes. J Cosmet Dermatol 2021 Jun;20(6):1788-1794. [doi: 10.1111/jocd.13783] [Medline: 33047478]
19. Ince FA, Nayar JC, Wong AL, et al. Uncovering the efficacy of a natural homemade sunscreen in protection from ultraviolet radiation. Photochem Photobiol Sci 2025 Jun;24(6):1017-1029. [doi: 10.1007/s43630-025-00742-7] [Medline: 40481984]

Abbreviations

BCC: basal cell carcinoma
DIY: do-it-yourself
FDA: Food and Drug Administration
GRASE: generally recognized as safe and effective
PF-UVA: production factor in the UVA domain
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SCC: squamous cell carcinoma
SPF: sun protection factor

Edited by A Kristensen-Cabrera; submitted 23.Sep.2025; peer-reviewed by M Pissavini, OK Ulutas; revised version received 13.Feb.2026; accepted 13.Feb.2026; published 03.Jul.2026.

Please cite as:
Kaftan G, Panillo M, Maxon L, Esposito E
The Rise in Homemade Sunscreen Trends and Future Impacts on Skin Cancer Risk: Systematic Review
JMIR Dermatol 2026;9:e84694
URL: https://derma.jmir.org/2026/1/e84694
doi:10.2196/84694

© Gabi Kaftan, Mia Panillo, Lydia Maxon, Emily Esposito. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Navigating the Intersection of Radiofrequency Microneedling and Surgical Facelifts: Scoping Review

Mia Panlilio[1], BA; Rebecca Bolen[1], BA; Olnita Martini[1], MS; Alexa Bonk[1], BS; John Tedesco[2], DO

[1]Department of Biomedical Sciences, College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States
[2]College of Osteopathic Medicine, Oklahoma State University, Tulsa, OK, United States

Corresponding Author:
Mia Panlilio, BA
Department of Biomedical Sciences, College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

Abstract

Background: Optimal management of facial skin laxity requires a nuanced approach by health care providers working in aesthetics. Radiofrequency microneedling (RFMN) devices have emerged as a popular noninvasive treatment for facial rejuvenation and improving skin laxity. While RFMN has demonstrated efficacy in enhancing skin tightening and complementing aesthetic procedures, its long-term impact on subsequent surgical facelifts remains uncertain.

Objective: The objective of this scoping review is to explore the interplay between RFMN and surgical facelift outcomes, with a focus on potential complications such as excessive skin tightening, dermal scarring, and altered tissue planes that may pose surgical challenges.

Methods: A search using PubMed and Google Scholar was conducted, and articles were selected from peer-reviewed journals based on specific inclusion and exclusion criteria. Only articles available in English were selected. In total, 21 articles were included in this scoping review.

Results: Papers included in this review discussed the mechanisms of action involved with RFMN, RFMN-related tissue changes, and how these changes could impact future facelift procedures. Most of the papers found that RFMN may drastically alter multiple tissue planes involved in facelift procedures due to collagen deposition through multiple tissue layers and increased tissue fibrosis. Patient factors influencing the effectiveness of RFMN and its role in facial rejuvenation were also examined, emphasizing the importance of navigating patient-specific demographics as a future consideration when creating an individualized treatment plan for each patient.

Conclusions: Patients should be informed that RFMN may lead to dermal fibrosis, tissue adhesions, and altered superficial musculoaponeurotic system composition, which could interfere with future facelift procedures and the patient's desired treatment goals. This emphasizes the importance of detailed discussion between the patient and health care provider to improve pretreatment consultation, increase patient education, and set realistic expectations. Further research is needed to determine optimal timing and treatment strategies for patients considering both RFMN and surgical facelifts to achieve the best aesthetic outcomes.

(JMIR Dermatol 2026;9:e78385) doi:10.2196/78385

KEYWORDS

radiofrequency; microneedling; facial; laxity; facelift

Introduction

Radiofrequency (RFMN) microneedling is a minimally invasive procedure that combines mechanical injury and thermal stimulation via tiny microneedles with radiofrequency (RF) energy to induce collagen remodeling [1]. The microneedling component creates controlled microtraumas, triggering a postinflammatory cascade that promotes neocollagenesis, elastin production, and angiogenesis. RF energy delivered through the microneedles then generates thermal coagulation within the dermis and hypodermis to induce collagen denaturation and subsequent contraction of tissue for skin tightening benefits [2]. This fractional approach allows for targeted treatment while preserving surrounding tissue, reducing recovery time [3]. Compared to traditional microneedling, evidence suggests that RFMN brings about greater improvements in aged skin, likely by eliminating senescent fibroblasts and increasing the number of nonsenescent fibroblasts [4].

A systematic review done in 2021 by Tan et al [5] analyzed 42 studies evaluating RFMN use across various conditions, with the largest evidence base for skin rejuvenation, followed by acne scars, acne vulgaris, striae, and axillary hyperhidrosis. A smaller number of studies were available supporting RFMN use for melasma, rosacea, cellulite, and androgenetic alopecia.

XSL•FO
RenderX

Based on the large and growing body of evidence for skin rejuvenation, RFMN devices have gained immense popularity for addressing skin laxity in patients seeking noninvasive alternatives to surgical facelifts. These devices offer treatment options for individuals outside the average age range for a facelift, those who have previously undergone a facelift, or patients desiring minimally invasive interventions [5]. Advances in RFMN technology, such as interchangeable tips with various microneedle pin configurations and dual treatment modes, allow for targeted treatments in delicate anatomical areas like the periorbital region [5]. While these technological refinements enhance customization, the process may impact deeper dermal structures critical to surgical outcomes, setting the stage for potential interference with future facelift procedures.

Long-term effects of RFMN before surgical facelifts remain unclear, raising concerns about potential complications. RFMN treatment prior to an elective facelift may have the potential to interfere with optimal facelift results due to excessive skin tightening, scarring, and damage to the dermis. This emphasizes the importance of pretreatment discussion about expectations and the adverse effects of RFMN if it is being used with a patient who is considering a facelift in the future. Given the interplay between RFMN and surgical facelifts, what are the long-term implications of RFMN for patients who may eventually pursue surgical facelift procedures? Could the very technology we're using to delay surgery unintentionally complicate it later? This scoping review explores the anatomical and clinical intersections between RFMN and surgical facelifts for patients realistically considering either a facelift or RFMN. It draws on current literature, evolving device technology, and real-world considerations to guide thoughtful treatment planning for optimal patient outcomes.

Methods

Searches done on PubMed and Google Scholar using the terms "skin laxity and radiofrequency microneedling," "skin laxity and microneedling," "skin laxity and facelift," "facelift and radiofrequency microneedling," and "facelift and microneedling" were conducted on January 4, 2024, and again on June 20, 2025, to account for any newly published or updated literature since the original search. The second search did not yield any new articles. Articles from peer-reviewed journals were included if they provided information on the mechanism of action of RFMN, described the techniques involved when performing a facelift procedure, or examined the effects of RFMN and/or facelift procedures on the skin. Only articles available in English were selected, and articles were excluded if they provided information on the use of RFMN on parts of the body other than the face and neck, or if they described RFMN treatment or surgical intervention unrelated to improving facial skin laxity. A summary of the inclusion and exclusion criteria are highlighted in Textbox 1. The review was conducted based on the 2005 methodology of Arksey and O'Malley [6]. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) reporting guidelines were followed, and the completed PRISMA-ScR checklist for this review can be found as Checklist 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- From peer-reviewed journals
- Content of article: mechanism of action of radiofrequency microneedling (RFMN); facelift procedures; effects of RFMN and/or facelift procedures on the skin
- Available in English

Exclusion criteria

- Content of article: RFMN being used on parts of the body other than the face and neck; RFMN treatment or surgical intervention unrelated to improving facial skin laxity
- Not available in English or an English translation

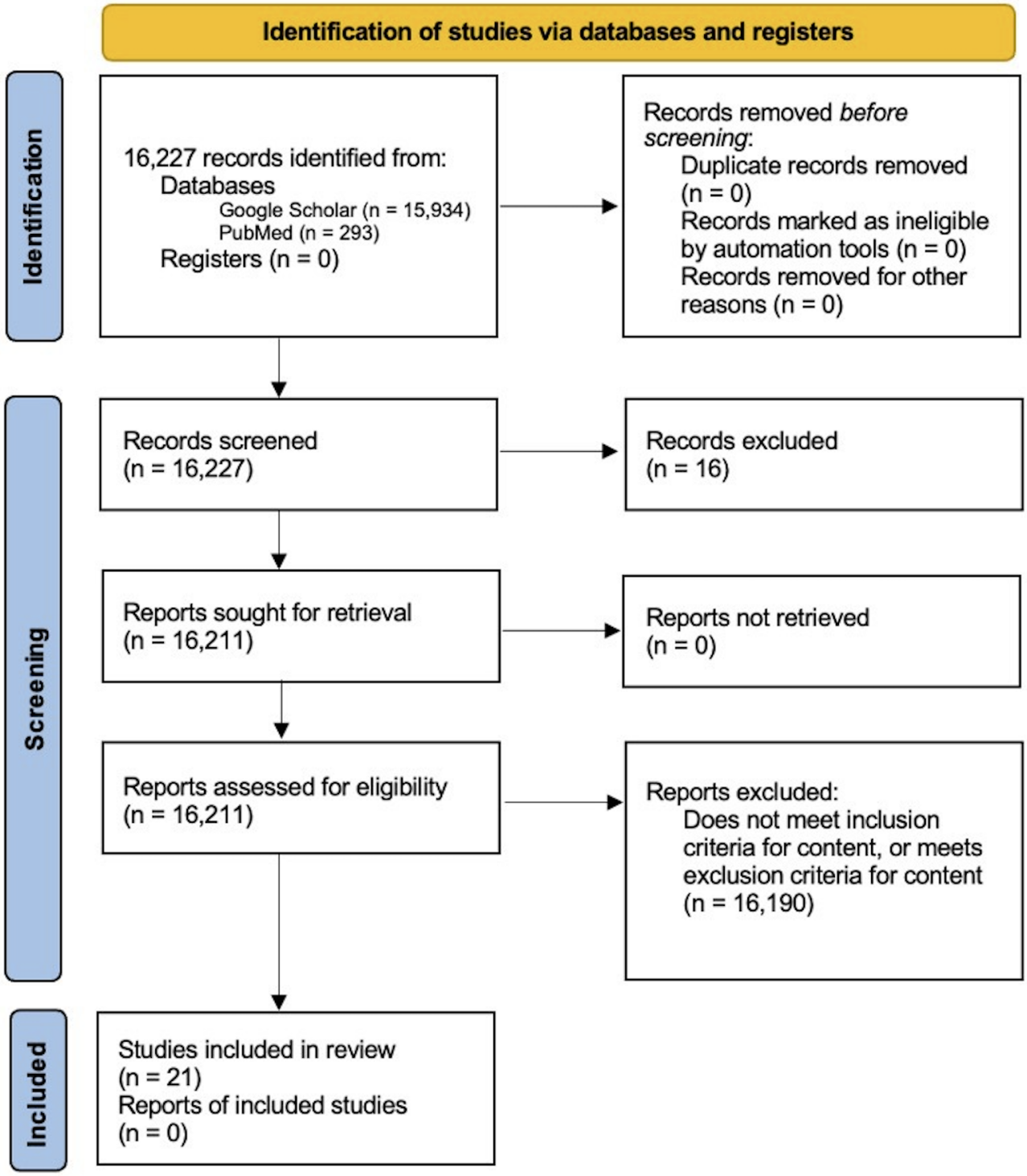
Results

Article Selection

From the initial search using the selected search terms, 15,934 and 293 articles were respectively identified from Google Scholar and PubMed, for a total of 16,277 articles (Figure 1). Out of the 16,277 articles screened, 16 PubMed articles were excluded because they were not available in English. Google Scholar does not have a language screening filter, so all 15,934 articles from the initial search were still considered. After all authors screened the remaining articles for content based on the inclusion and exclusion criteria, 21 articles in total were selected for this review. A PRISMA diagram of the article selection is available below in Figure 1.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of article selection.



Summary of Included Articles

After article screening was completed and the 21 articles were selected, 12 articles were identified as original research, 5 as narrative literature reviews, 2 as systematic reviews, and 1 as an educational reference article. One article, by Arksey and O'Malley [6], was used as a framework to guide writing of this review, and did not necessarily fit the inclusion and exclusion criteria for articles specifically related to the topic of this scoping review. Table 1 displays a summary of the articles mentioned, as well as their article types and study designs.

XSL•FO
RenderX

Table . Summary of included articles.

Authors (year)	Journal	Article type	Study design
Devgan et al (2019) [1]	*Otolaryngology Clinics of North America*	Literature review	Narrative review
Spataro et al (2022) [2]	*Facial Plastic Surgery Clinics of North America*	Literature review	Narrative review
Hendricks and Farhang (2022) [3]	*Journal of Cosmetic Dermatology*	Literature review	Narrative review
Hwang et al (2025) [4]	*Scientific Reports*	Original research	Split-face comparative clinical trial
Tan et al (2021) [5]	*Dermatologic Surgery*	Literature review	Narrative review
Arksey and O'Malley (2005) [6]	*International Journal of Social Research Methodology*	Original research	Not applicable
Dayan et al (2020) [7]	*Plastic and Reconstructive Surgery—Global Open*	Original research	Prospective case series (single-arm clinical study)
Arnaoutakis et al (2022) [8]	*Facial Plastic Surgery & Aesthetic Medicine*	Literature review	Narrative review
Ramaut et al (2018) [9]	*Journal of Plastic, Reconstructive & Aesthetic Surgery*	Systematic review	Systematic review
Huang et al (2014) [10]	*Biochemistry*	Basic science research	In vitro experimental study
Nguyen et al (2025) [11]	*Lasers in Medical Science*	Original article	Clinical and histologic study (prospective cohort)
Xu et al (2025) [12]	*Lasers in Surgery and Medicine*	Original research	Animal study (porcine model)
Zheng et al (2014) [13]	*Dermatologic Surgery*	Original research	Experimental histologic study
Wang et al (2025) [14]	*Lasers in Medical Science*	Original research	Pilot clinical study
Wang et al (2024) [15]	*Lasers in Surgery and Medicine*	Original research	Animal study (porcine model)
Cho et al (2024) [16]	*Skin Research & Technology*	Original research	Animal study (minipig model)
Hohman et al (2023) [17]	*StatPearls*	Reference article	Narrative review (educational)
Ghassemi et al (2003) [18]	*Aesthetic Plastic Surgery*	Original research	Anatomical cadaveric study
Demesh et al (2021) [19]	*Journal of Cosmetic Dermatology*	Original research	Clinical case series
Seo et al (2012) [20]	*Lasers in Surgery and Medicine*	Original research	Clinical and histologic study (prospective cohort)
Austin et al (2022) [21]	*Lasers in Surgery and Medicine*	Systematic review	Systematic review

RFMN-Related Tissue Changes

While RFMN effectively improves skin laxity and wrinkle reduction, its impact on future facelift procedures remains uncertain. An article published in 2022 found that a single session of noninvasive fractional bipolar RFMN achieved approximately 37% of the skin laxity improvement seen with a surgical facelift, suggesting multiple treatments may be required for significant results [8]. However, a 2018 systematic review from the *Journal of Plastic, Reconstructive & Aesthetic Surgery* states that repeated sessions risked dermal fibrosis, particularly in the papillary dermis, potentially complicating future surgical interventions [9].

As RFMN devices introduce repeated and organized microtraumas into the skin, the depth of penetration directly impacts targeted tissue layers. Microneedling from such devices introduces a targeted mechanism of repair that excludes highly inflammatory cellular cascades, such as transforming growth factor β-1 (TGF-β1) and transforming growth factor β-2 (TGF-β2), and instead is driven through a less-inflammatory cascade via transforming growth factor β-3 (TGF-β3), a protein known to lead to fibroblast migration and collagen matrix remodeling [10]. Platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), and epidermal growth factor (EGF) are all released locally in response to the microtrauma, allowing natural skin tightening via angiogenesis and collagen deposition [2]. Repeated treatments thus increase collagen deposition, which increases the risk for the development of dermal fibrosis.

A better understanding of the timeline of collagenesis following RFMN is critical when considering the interplay with surgical facelifts. Acute inflammation and early collagen deposition dominate the first week after treatment, followed by organized collagen remodeling and maturation over the ensuing 1 to 3 months [2,10]. Persistent changes in dermal structure, including fibrosis or altered tensile strength, may interfere with surgical flap elevation, tissue pliability, and healing after a facelift.

RFMN energy settings can be optimized to balance skin tightening with control of fibrosis by titrating energy per needle,

XSL•FO
RenderX

pulse duration, and depth to achieve sufficient dermal coagulation for neocollagenesis and elastogenesis while avoiding excessive thermal injury that may promote fibrotic remodeling. Data supports targeting moderate energy settings (eg, energy per needle 20 - 60 mJ, pulse durations 100-300 ms) and limiting the number of passes to induce controlled dermal coagulation, maximizing skin tightening while minimizing the risk of fibrosis [11,12]. Adjusting needle depth to target the reticular dermis and using insulated needles can further localize RFMN thermal effects, thus reducing epidermal damage and unwanted fibrosis [13,14]. Sequential or pulsed energy delivery, as well as energy feedback systems, can help regulate tissue response and prevent overtreatment [15,16].

Effects on Future Elective Facelift Procedures

Surgical facelift procedures rely on the manipulation of the superficial musculoaponeurotic system (SMAS), a fibrofatty connective tissue layer continuous with the superficial cervical fascia, connected to the platysma muscle inferiorly and the galea superiorly [17]. It plays an integral role in the anatomic relation of the superficial dermis to the underlying facial muscles. There are two distinct SMAS compositions given anatomic regions, and the abrupt junction of differing compositions resides at the nasolabial fold region, where medially, there are fewer fat lobules and a more direct connection of the SMAS to the superficial dermis as muscle fibers are seen to extend superficially into the dermis [17]. However, the other regions of the SMAS lateral to the nasolabial fold still carry the same properties of communication of facial muscle to skin by muscle tendon fibers connecting both regions via the SMAS [18]. Beneath this layer, the SMAS has a complex relation with deep ligaments and connections that limits the mobility of superficial structures. These connections are crucial to release to generate the most optimal movement for desired facelift outcomes [17].

Due to the sophisticated relationship of neighboring structures, RFMN, particularly at greater depths, may alter these structural relationships. Traditional surgical facelifts target a single plane of tissue in a primary horizontal plane [17]. On the contrary, RFMN targets a small treatment area in a vertical configuration through multiple planes of tissue, potentially leading to increased tissue adhesions, difficult surgical dissection, impaired flap mobility, and suboptimal facelift outcomes [10].

Although undergoing RFMN treatments prior to a facelift could potentially induce a level of fibrosis that may help delay the timeline when a patient would be a candidate for a facelift, evidence suggests RFMN may be more beneficial postoperatively. Following a facelift, RFMN could enhance skin tightening and improve aesthetic outcomes by stimulating additional collagen production either as an immediate adjunctive therapy or as a method to combat recurrent long-term skin laxity [19]. Careful planning is necessary to determine the appropriate time frame between treatments to avoid excessive fibrosis and impaired wound healing.

Patient Considerations for RFMN Procedures

When considering therapy using RFMN, specific patient populations should be considered when discussing treatment options, as certain age groups, as well as patients with jowl laxity, have been shown to experience better outcomes with facial surgery [19]. Taking into consideration that each treatment has varying mechanisms and different anatomical targets, a comparative study found that surgical facelifts improved skin laxity by 46% relative to baseline, whereas RFMN alone achieved only a 16% improvement [20]. These findings underscore the importance of setting realistic patient expectations regarding treatment efficacy. Additionally, patient age should be considered, as older individuals (≥55 years) experience more pronounced skin tightening with RFMN compared to younger patients [2]. Younger patients typically have a higher collagen content in their skin compared to older patients, who undergo collagen loss due to age; thus, the relative resulting decrease in skin laxity is much more noticeable in older patients versus younger patients. This knowledge prompts early discussion of RFMN treatment to address its potential effectiveness or lack thereof, especially at a younger age.

In addition to considering a patient's age, premature neck and jowl laxity are common concerns among patients seeking skin-tightening treatments. However, RFMN does not effectively target subplatysmal fat, necessitating careful patient selection—individuals with significant subplatysmal fat may achieve superior results with surgical interventions such as liposuction [19]. For patients with pronounced skin laxity, RFMN alone may be insufficient and could exacerbate sagging if incidental heat-induced fat loss occurs without concurrent skin excision [19]. Other complications of RFMN reported include hyper- or hypopigmentation of treated skin, thermal burns, blistering, and scarring; these can often be mitigated with proper technique and equipment settings [8].

Discussion

Main Findings

The nuances of RFMN in facial rejuvenation necessitate a deeper understanding of its implications for future surgical facelifts. This calls for detailed discussion between the patient and health care provider to improve pretreatment consultation, patient education, and results. Patients should be informed that RFMN may lead to dermal fibrosis, tissue adhesions, subcutaneous adipose denaturation, and altered SMAS composition, which could complicate facelift procedures and their desired outcomes. Additionally, providers should address the limitations of RF to effectively target jowl laxity and set realistic expectations regarding RF results in patients younger than 55 years. Understanding these points would allow for individualized treatment planning, ensuring patients receive the most appropriate interventions based on their anatomical considerations and aesthetic goals. Furthermore, current evidence suggests RFMN may be better positioned as a postoperative adjunct rather than a presurgical intervention, especially in patients known to be surgical candidates in the future.

It is also important to note that RFMN has primarily been used for skin rejuvenation, mild laxity, and conditions such as acne scars, rather than as a substitute for surgical facelift procedures [5]. Additional systematic reviews and clinical trials, such as those by Austin et al [21] and Nguyen et al [11], further support

XSL•FO
RenderX

that early RFMN and radio frequency protocols are designed for modest rejuvenation, targeting mild-to-moderate laxity and stimulating collagen production. Among these studies, reported patient satisfaction was highest among those seeking subtle improvements rather than facelift-level results.

Limitations

This scoping review is limited by the minimal availability of long-term studies specifically examining the effects of RFMN on subsequent surgical facelift procedures. The available literature mainly consists of case reports, small-scale studies, and expert opinion, which restricts the generalizability of conclusions. In addition, the large amount of variation between RFMN device settings, treatment protocols, and patient demographics across studies further limits the ability to standardize findings or establish definitive treatment guidelines.

Conclusion

Cosmetic surgical providers trained in RFMN and/or facelift procedures should give careful consideration to this new technology when discussing different options for facial rejuvenation. Factors to weigh in these considerations should include, but are not limited to, age-related expectations, area of treatment, and the potential impact of subsequent facelifts. Rather than viewing RFMN and surgical facelifts as isolated interventions, providers should consider how early noninvasive treatments may influence future surgical options. This has important implications for clinical decision-making, patient education, and informed consent.

Ultimately, optimizing aesthetic outcomes will require a more integrated strategy that aligns patient goals with both immediate and long-term treatment trajectories. Future studies are needed to evaluate the cumulative impact of RFMN on facial anatomy and to guide safe, evidence-based treatment planning for patients considering both noninvasive and surgical facial rejuvenation options. Additional research should also focus on establishing guidelines for the optimal timing of RFMN relative to surgical facelifts and identifying strategies to minimize adverse effects of RFMN. As RFMN technology evolves, ongoing studies will be critical in refining its role in facial rejuvenation and improving patient outcomes.

Acknowledgments

All authors declare no use of generative artificial intelligence in drafting any portion of the manuscript (this includes any and all tools, including ChatGPT, Gemini, Claude, or any other comparable tool not explicitly listed).

Funding

The authors declare no financial support was received for this work.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 19 KB - derma_v9i1e78385_app1.docx]

References

1. Devgan L, Singh P, Durairaj K. Minimally invasive facial cosmetic procedures. Otolaryngol Clin North Am 2019 Jun;52(3):443-459. [doi: 10.1016/j.otc.2019.02.013] [Medline: 30954270]
2. Spataro EA, Dierks K, Carniol PJ. Microneedling-associated procedures to enhance facial rejuvenation. Facial Plast Surg Clin North Am 2022 Aug;30(3):389-397. [doi: 10.1016/j.fsc.2022.03.012] [Medline: 35934440]
3. Hendricks AJ, Farhang SZ. Dermatologic facial applications of Morpheus8 fractional radiofrequency microneedling. J Cosmet Dermatol 2022 Oct;21 Suppl 1:S11-S19. [doi: 10.1111/jocd.15231] [Medline: 35916259]
4. Hwang JM, Lee SH, Baek EJ, et al. Comparison of the effects of fractional microneedle radiofrequency and microneedling on modulating the senescent fibroblast milieu in aged skin. Sci Rep 2025 May 26;15(1):18296. [doi: 10.1038/s41598-025-02545-3] [Medline: 40419589]
5. Tan MG, Jo CE, Chapas A, Khetarpal S, Dover JS. Radiofrequency microneedling: a comprehensive and critical review. Dermatol Surg 2021 Jun 1;47(6):755-761. [doi: 10.1097/DSS.0000000000002972] [Medline: 33577211]
6. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
7. Dayan E, Rovatti P, Aston S, Chia CT, Rohrich R, Theodorou S. Multimodal radiofrequency application for lower face and neck laxity. Plast Reconstr Surg Glob Open 2020 Aug;8(8):e2862. [doi: 10.1097/GOX.0000000000002862] [Medline: 32983756]
8. Arnaoutakis D, Kochhar A, Magro I, Karimi K. The role of subcutaneous radiofrequency-assisted liposculpture in the facial plastic surgeon's practice. Facial Plast Surg Aesthet Med 2022;24(S1):S11-S16. [doi: 10.1089/fpsam.2022.0225] [Medline: 36103269]

XSL•FO
RenderX

9. Ramaut L, Hoeksema H, Pirayesh A, Stillaert F, Monstrey S. Microneedling: where do we stand now? A systematic review of the literature. J Plast Reconstr Aesthet Surg 2018 Jan;71(1):1-14. [doi: 10.1016/j.bjps.2017.06.006] [Medline: 28690124]
10. Huang T, Schor SL, Hinck AP. Biological activity differences between TGF-β1 and TGF-β3 correlate with differences in the rigidity and arrangement of their component monomers. Biochemistry 2014 Sep 16;53(36):5737-5749. [doi: 10.1021/bi500647d] [Medline: 25153513]
11. Nguyen L, Bartholomeusz J, Schneider SW, Herberger K. Histological and clinical dose-response analysis of radiofrequency microneedling treatment for skin rejuvenation. Lasers Med Sci 2025 Feb 7;40(1):75. [doi: 10.1007/s10103-025-04335-9] [Medline: 39915343]
12. Xu Y, Zhang Y, Wang H, Tang H, Zeng W, Wen X. Microneedle radiofrequency induces extracellular matrix remodeling through fibroblast activation: a histological study in a porcine model. Lasers Surg Med 2025 Aug;57(6):528-543. [doi: 10.1002/lsm.70033] [Medline: 40444432]
13. Zheng Z, Goo B, Kim DY, Kang JS, Cho SB. Histometric analysis of skin-radiofrequency interaction using a fractionated microneedle delivery system. Dermatol Surg 2014 Feb;40(2):134-141. [doi: 10.1111/dsu.12411] [Medline: 24373135]
14. Wang M, Li Y, Lai X, Shi C, Yan Y. The clinical effects and skin histological changes induced by a novel insulated radiofrequency microneedle: a pilot study. Lasers Med Sci 2025 Feb 17;40(1):98. [doi: 10.1007/s10103-025-04354-6] [Medline: 39960528]
15. Wang H, Hamblin MR, Zhang Y, Xu Y, Wen X. Histological evaluation of monopolar and bipolar radiofrequency microneedling treatment in a porcine model. Lasers Surg Med 2024 Mar;56(3):288-297. [doi: 10.1002/lsm.23768] [Medline: 38334177]
16. Cho SB, Kang SY, Lee YJ, Choi M, Kim B, Ahn JC. Effect of sequential delivery of 1- and 2-MHz bipolar microneedling radiofrequency energy on thermal tissue reactions in a minipig model. Skin Res Technol 2024 Sep;30(9):e13898. [doi: 10.1111/srt.13898] [Medline: 39213160]
17. Hohman MH, Raggio BS, Patel BC. Deep plane facelift. In: StatPearls: StatPearls Publishing; 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK545277/ [accessed 2024-01-22]
18. Ghassemi A, Prescher A, Riediger D, Axer H. Anatomy of the SMAS revisited. Aesthetic Plast Surg 2003;27(4):258-264. [doi: 10.1007/s00266-003-3065-3] [Medline: 15058546]
19. Demesh D, Cristel RT, Gandhi ND, Kola E, Dayan SH. The use of radiofrequency-assisted lipolysis with radiofrequency microneedling in premature jowl and neck laxity following facialplasty. J Cosmet Dermatol 2021 Jan;20(1):93-98. [doi: 10.1111/jocd.13824] [Medline: 33128284]
20. Seo KY, Yoon MS, Kim DH, Lee HJ. Skin rejuvenation by microneedle fractional radiofrequency treatment in Asian skin; clinical and histological analysis. Lasers Surg Med 2012 Oct;44(8):631-636. [doi: 10.1002/lsm.22071] [Medline: 22936274]
21. Austin GK, Struble SL, Quatela VC. Evaluating the effectiveness and safety of radiofrequency for face and neck rejuvenation: a systematic review. Lasers Surg Med 2022 Jan;54(1):27-45. [doi: 10.1002/lsm.23506] [Medline: 34923652]

Abbreviations

EGF: epidermal growth factor
FGF: fibroblast growth factor
PDGF: platelet-derived growth factor
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RF: radio frequency
RFMN: radio frequency microneedling
SMAS: superficial musculoaponeurotic system
TGF-β1: transforming growth factor β-1
TGF-β2: transforming growth factor β-2
TGF-β3: transforming growth factor β-3

Edited by R Dellavalle; submitted 01.Jun.2025; peer-reviewed by A Edalatpour, P Huang; revised version received 08.Dec.2025; accepted 06.Jan.2026; published 27.Apr.2026.

Please cite as:
Panlilio M, Bolen R, Martini O, Bonk A, Tedesco J
Navigating the Intersection of Radiofrequency Microneedling and Surgical Facelifts: Scoping Review
JMIR Dermatol 2026;9:e78385
URL: https://derma.jmir.org/2026/1/e78385
doi:10.2196/78385

XSL•FO
RenderX

© Mia Panlilio, Rebecca Bolen, Olnita Martini, Alexa Bonk, John Tedesco. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Informatics-Based Psychotherapeutic and Psychiatric Interventions in Dermatology: Scoping Review of Impacts on Skin Disease Severity and Mental Health Outcomes

Caroline Lamarre[1], BSc; Jeffrey Chivinski[2], BPharm, MD; Alexandre Hudon[3,4,5,6,7], BEng, MSc(Ed), MD, PhD

[1]Department of Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada

[2]Beacon Dermatology, Calgary, AB, Canada

[3]Department of Psychiatry and Addictology, Faculty of Medicine, Université de Montréal, 2900 Boulevard Édouard-Montpetit, Montréal, QC, Canada

[4]Department of Psychiatry, Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada

[5]Department of Psychiatry, Institut national de psychiatrie légale Philippe-Pinel, Montréal, QC, Canada

[6]Centre de recherche de l'Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada

[7]Centre de recherche en pédagogie de la santé, Montréal, QC, Canada

Corresponding Author:
Alexandre Hudon, BEng, MSc(Ed), MD, PhD
Department of Psychiatry and Addictology, Faculty of Medicine, Université de Montréal, 2900 Boulevard Édouard-Montpetit, Montréal, QC, Canada

Abstract

Background: Chronic dermatologic conditions such as psoriasis, atopic dermatitis, and hidradenitis suppurativa are associated with a high burden of psychiatric comorbidities, including depression, anxiety, and suicidality. Despite growing awareness of the psychosocial impact of skin diseases, mental health needs remain underaddressed in dermatologic care. Digital technologies (including teledermatology, mobile health apps, and internet-delivered psychotherapies) offer promising avenues for integrating psychotherapeutic and psychiatric interventions into dermatology. However, the scope, effectiveness, and implementation of such informatics-based approaches remain poorly mapped in the literature.

Objective: This scoping review aimed to systematically identify, categorize, and synthesize studies on digital psychotherapeutic and psychiatric interventions targeting patients with dermatological conditions, with a focus on clinical, mental health, and implementation outcomes.

Methods: Following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines, we conducted a comprehensive search across 5 databases (MEDLINE, Embase, Web of Science, PsycINFO, and Google Scholar) for articles published up to March 2025. Studies were included if they involved patients with dermatologic conditions and assessed interventions that combined a digital informatics component (eg, telehealth, apps, artificial intelligence, virtual platforms) with a psychotherapeutic or psychiatric element (eg, cognitive behavioral therapy [CBT], mindfulness, consult-liaison psychiatry). Eligible study designs included clinical trials, observational studies, and mixed methods research. Data were extracted systematically, and methodological quality was assessed using JBI tools.

Results: Out of 15,176 records identified, 11 studies met the inclusion criteria. Most interventions targeted psoriasis (9/11) and used asynchronous digital platforms such as internet-based CBT and mobile apps. Across studies, dropout rates ranged from 10% to 76%. Improvements in dermatologic quality of life were reported in 6 of 11 studies, with statistically significant reductions in depression and anxiety observed in multiple trials (eg, internet-based CBT and mindfulness-based interventions), alongside reductions in psoriasis severity (Psoriasis Area and Severity Index) and itch intensity in randomized controlled trials. Intervention duration ranged from single-session virtual reality exposure to 8 - to 12-week structured programs. However, long-term outcomes beyond 3 to 12 months were rarely assessed, and reporting of sociodemographic variables and equity-related factors was limited.

Conclusions: Informatics-based psychotherapeutic and psychiatric interventions represent a promising frontier in psychodermatology, with early evidence suggesting feasibility and potential clinical benefit. Digital platforms may expand access to mental health support and improve holistic care for patients with dermatologic conditions. However, significant gaps remain in terms of equity, long-term effectiveness, integration into clinical workflows, and adaptation for diverse populations. Future research should focus on rigorous, inclusive trials and the development of hybrid models that blend digital and face-to-face care to ensure sustainable and equitable impact.

(JMIR Dermatol 2026;9:e82096) doi:10.2196/82096

XSL•FO
RenderX

KEYWORDS

psychodermatology; digital health; cognitive behavioral therapy; teledermatology; mental health; dermatologic conditions; mobile health apps; digital therapeutics; psychiatric comorbidity; scoping review

Introduction

Dermatologic Issues and Mental Health

Chronic dermatologic conditions, such as psoriasis, atopic dermatitis (AD), acne, vitiligo, and hidradenitis suppurativa (HS), are among the most psychologically burdensome medical conditions and impose a substantial, multifaceted impact on mental health. Not only do these cutaneous diseases cause visible disfigurement, discomfort, and functional and social impairment, but they are also associated with a high prevalence of various psychiatric comorbidities. Chronic symptoms like pain and pruritus associated with the relapsing nature of many skin conditions often mediate these psychological effects [1]. Large-scale studies have found that around 10% of patients with dermatologic conditions meet the criteria for clinical depression and around 17% for clinical anxiety, with about 12% to 13% reporting suicidal ideation [1]. These rates are roughly 2 to 3 times higher than in the general population, underscoring how common serious psychological comorbidities (eg, depression, anxiety, social anxiety or phobia, and suicidal thoughts) significantly add to patients' overall disease burden [1]. For instance, psoriasis and AD are each associated with notably elevated risks of new-onset depression and anxiety. A recent UK cohort study reported that adults diagnosed with psoriasis or AD had a 10% to 20% greater hazard of developing depression or anxiety within the first year, compared to matched individuals without skin disease [2]. In particular, patients with more severe skin involvement (such as extensive psoriasis) show even higher relative risks of depression and other psychiatric sequelae [2].

Psychiatric morbidity appears especially pronounced in HS, which research indicates is accompanied by a high prevalence of depression and anxiety, as well as increased rates of substance use disorders and even serious mental illnesses such as psychosis or bipolar disorder [3]. This higher prevalence is due to a combination of biopsychosocial factors, such as chronic pain and physical discomfort secondary to malodor and constant drainage. Many patients develop social withdrawal and avoidance behaviors out of fear of stigmatization. As the disease often affects intertriginous and genital areas, patients experience an impact on their sexual health and relationship difficulties. A lack of cure or repeated ineffective treatments leads to hopelessness and despair over time. Most alarmingly, patients with HS face a dramatically elevated risk of suicidality. For example, epidemiologic studies have found that suicide is roughly twice as likely in HS sufferers compared to those without HS [3]. Such findings illustrate that chronic inflammatory skin diseases carry not only physical burdens but also profound psychological and social challenges for patients. This underlines the need for integrated, multidisciplinary approaches to care for these patients.

Despite this heavy psychosocial burden, patients' mental health needs are often overlooked in dermatologic care. In 1 large survey of over 7200 patients with inflammatory dermatoses, two-thirds reported high stress levels, yet fewer than 15% had ever been offered any psychological support [4]. Dermatology clinics traditionally focus on cutaneous findings, and many patients never receive screening or referral for underlying depression, anxiety, or other psychological issues related to their skin condition [1]. These unmet needs can be due to under-recognition by care providers, limited access to mental health care like psychology or psychiatry, or simply the stigma associated with mental illness. Dermatologists often lack the time or referral pathways to manage the psychologic comorbidities. This gap in care means that a significant portion of patients with dermatologic disorders endure unaddressed emotional distress, which can, in turn, exacerbate disease outcomes. Furthermore, this can lead to poor adherence to medical treatments and a worsening quality of life. Consequently, there is a growing consensus that integrating mental health screening and psychotherapeutic support into routine dermatologic practice is critical for the optimal management of chronic skin diseases. Experts advocate for the regular use of mental health questionnaires (eg, Patient Health Questionnaire-2 for depression and Generalized Anxiety Disorder-7 for anxiety) in dermatology settings and prompt referral to psychiatry or psychology when needed [5]. Studies suggest that comprehensive care, combining dermatologic treatment with interventions, such as cognitive behavioral therapy (CBT), stress management techniques, or support groups, can improve patients' quality of life and potentially even their skin disease severity [6]. By recognizing and addressing the psychological comorbidities of chronic dermatoses, clinicians can substantially reduce the overall disease burden and improve long-term outcomes for these patients.

Psychotherapeutic Interventions and the Emerging Role of Digital Technologies

Digital platforms (including teledermatology, mobile health [mHealth] apps, artificial intelligence [AI]–supported therapeutics, and virtual assistants) are reshaping dermatologic care and have emerged as innovative ways to provide and deliver therapeutic interventions to patients with chronic medical conditions. Teledermatology (via store-and-forward images or live video) enables remote diagnosis and management of chronic inflammatory skin diseases with high reliability. It can significantly shorten waiting times for patients while maintaining diagnostic accuracy comparable to in-person visits [7]. For example, 1 study on patients with chronic skin conditions found that live-video teledermatology achieved 96% diagnostic concordance with face-to-face consultations and reduced consultation turnaround time by 90% [8]. This translates to faster treatment and greatly improved access for patients who might otherwise face long delays or travel barriers due to living in remote areas. In addition to bypassing geographic barriers, teledermatology can help reduce stigma or anxiety associated with in-person therapy and is more flexible. mHealth apps offer

XSL•FO
RenderX

accessible tools for patient self-monitoring, education, and behavior change. Dermatology apps can help patients track symptoms (for instance, logging eczema or psoriasis flares), adhere to treatment plans, and learn about skincare, thereby empowering them in day-to-day disease management [9]. Multiple scoring systems self-reported by patients can be integrated into these apps, for example: the Numeric Rating Scale for evaluating itch, or Skindex, which measures the emotional, symptomatic, and functional burden of skin diseases. Recent evidence confirms that such mHealth interventions improve patient outcomes: a 2024 meta-analysis showed that smartphone apps and telemonitoring significantly enhanced patients' quality of life and self-management of AD, while also increasing patient satisfaction and engagement in care [10]. Meanwhile, AI-driven digital therapeutics and virtual assistants are being integrated into dermatology workflows to support both providers and patients. In a teledermatology program for psoriasis, adding a virtual assistant (chatbot) improved medication adherence and gave patients greater peace of mind by connecting them to their dermatologist; notably, participants' Dermatology Life Quality Index (DLQI) scores improved with the chatbot intervention [11]. Such digital assistants can streamline triage, answer patient questions, and automate follow-ups, augmenting the efficiency of teledermatology services [11].

Importantly, these technologies do not just address skin health in isolation. They also impact mental health outcomes. Many chronic dermatoses carry a significant psychological burden (stress, low self-esteem, anxiety), so bridging dermatologic care with psychosocial support is crucial. Digital platforms are beginning to fill this need. For instance, an internet-based cognitive behavioral therapy (iCBT) program for patients with atopic eczema significantly reduced itching and insomnia while improving quality of life, with outcomes comparable to traditional therapist-led care [12]. This kind of online psychodermatologic intervention is valuable because access to in-person psychological services for patients with skin diseases is often limited [12]. In sum, teledermatology and related digital innovations are enhancing chronic skin disease management by improving convenience and access, supporting patient self-management, and addressing the psychosocial dimensions of dermatologic care in ways that were not previously possible. These interventions enhance the resilience and quality of life of patients with chronic skin diseases.

Gaps in the Literature

Despite the growing promise of digital platforms in dermatologic and psychodermatologic care, their real-world implementation remains fraught with challenges. The presence of digital psychodermatology tools continues to be fragmented and sparse in clinical practice. Clinicians cite significant barriers related to workflow integration, including the time required for image capture and documentation; limited interoperability with electronic medical records; and medicolegal concerns surrounding data privacy, reimbursement models, and clinical liability [13,14]. From the patient perspective, digital literacy and technology acceptance remain key hurdles (particularly among older adults), whose comfort with telehealth interfaces is shaped by varying levels of support, familiarity, and personal values [15]. While digital tools such as teledermatology and online psychotherapeutic interventions have demonstrated improved access and high patient satisfaction, their clinical effectiveness remains inconsistent. Some models yield optimal outcomes only when used in conjunction with traditional in-person care [13,16]. Furthermore, many digital mental health interventions in dermatology, such as app-based CBT or stress management platforms, face high attrition rates and low long-term engagement, limiting their sustained impact on mental health outcomes [17]. These implementation and engagement challenges highlight the need for better-designed, clinically integrated, and user-centered digital solutions in psychodermatology.

Objectives and Hypotheses

This scoping review aims to systematically map, categorize, and synthesize the scientific literature on informatics-enabled psychotherapeutic and psychiatric interventions designed to improve both dermatologic conditions and their associated mental health outcomes. It also seeks to address key gaps identified in the existing literature. Specifically, the objectives are to: (1) chart the landscape of digital platforms (such as teledermatology, mHealth apps, AI-based decision-support tools, iCBT modules, mindfulness-based digital programs, and e-psychiatry consult-liaison services) used therapeutically in dermatologic care; (2) profile the dermatologic diagnoses most frequently targeted and characterize the demographic and clinical features of the patient populations involved; (3) synthesize reported clinical, psychological, and quality-of-life outcomes, including dermatologic severity indices, itch and pain scores, validated mental health scales, and patient-reported outcomes; (4) analyze implementation models, including modes of delivery, duration, provider roles, integration into clinical care pathways, and equity-related considerations (eg, digital literacy, socioeconomic factors, and geographic access); (5) assess the methodological quality of included studies using validated critical appraisal tools to identify strengths, limitations, and risks of bias; and (6) identify knowledge gaps and research priorities to guide the future development, evaluation, and equitable implementation of digital psychodermatologic interventions.

Methods

Search Strategies

To comprehensively capture the current state of research on informatics-based psychotherapeutic and psychiatric interventions in dermatology, we conducted a structured scoping review of studies published up to March 2025. Searches were performed across 5 major electronic databases: MEDLINE (via PubMed), Web of Science, Embase, PsycINFO (via Ovid), and Google Scholar. The review followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Checklist 1), which provides methodological guidance for mapping emerging fields with conceptual or methodological diversity [18].

The search strategy was developed iteratively, combining free-text keywords and controlled vocabulary (Medical Subject

XSL•FO
RenderX

Headings and Emtree terms) to ensure broad yet relevant coverage. Terms related to dermatologic diagnoses (eg, "psoriasis," "acne," "atopic dermatitis," "hidradenitis suppurativa") were combined with those describing digital modalities (eg, "teledermatology," "mobile health," "artificial intelligence," "digital therapeutics," "virtual assistants") and psychosocial or psychiatric constructs (eg, "depression," "anxiety," "psychotherapy," "CBT," and "mental health"). The objective was to capture literature at the intersection of digital tools and mental health outcomes in populations with dermatologic disorders. Search strategies were peer-reviewed internally and piloted for sensitivity before full execution. The scoping review was not registered.

All searches were conducted by the lead reviewer (CL), with results independently verified by a second reviewer (AH) to ensure consistency, accuracy, and completeness. No restrictions on study geography, clinical setting, or publication language were applied at the search stage, although inclusion criteria were applied during screening. Complete search algorithms for each database are provided in Multimedia Appendix 1.

Study Eligibility

Studies were eligible for inclusion if they met all of the following criteria: (1) the population under investigation comprised children, adolescents, or adults with a clinically diagnosed chronic dermatologic disorder (eg, psoriasis, AD, acne, HS, vitiligo, alopecia areata, chronic urticaria, or other specified inflammatory or pigmentary skin diseases); (2) the study evaluated a therapeutic approach that incorporated both an informatics component (eg, teledermatology, mobile or e-health apps, web-based platforms, AI-driven tools, virtual reality [VR], chatbot technologies) and a psychotherapeutic or psychiatric element aimed at improving mental health or dermatologic outcomes. Acceptable psychological components included, but were not limited to, CBT, mindfulness-based interventions, psychoeducation, liaison psychiatry, or psychotropic medication management. Studies were eligible if they reported at least 1 skin-related outcome (eg, disease severity scores, itch, pain, or DLQI) or 1 mental health outcome (eg, anxiety, depression, psychological distress, or health-related quality of life). Eligible study designs encompassed randomized and nonrandomized controlled trials, quasi-experimental studies, cohort or cross-sectional analyses, mixed methods evaluations, and implementation studies with quantitative or qualitative outcome reporting.

Studies were excluded if they focused exclusively on diagnostic tools, image recognition systems, or data-capture technologies without a therapeutic or psychiatric objective. Similarly, studies assessing only pharmacological, surgical, or cosmetic interventions without an integrated digital or psychological component were not considered. Single-case reports with fewer than 5 participants, editorials, opinion articles, conference abstracts without full text, narrative reviews, and protocol papers lacking original data were also excluded. Additionally, studies involving only healthy volunteers, individuals without a dermatologic diagnosis (eg, cosmetic users), or interventions targeting sun protection, skin cancer screening, or esthetic counseling were not included. Only articles published in English or French were considered for review; gray literature and unpublished manuscripts were excluded.

Data Extraction

A structured data extraction process was conducted using a predesigned coding framework in Microsoft Excel to ensure systematic and reproducible capture of study details. One reviewer (CL) performed the initial data extraction, and a second reviewer (AH) independently reviewed all entries to verify accuracy and consistency across studies. Discrepancies in interpretation or classification were resolved through discussion until consensus was reached.

Variables extracted encompassed bibliographic details (eg, author names, publication year, country of origin, and publication type), participant characteristics (eg, dermatologic diagnosis, sample size, age distribution, and gender), and intervention attributes (eg, program title, type of informatics platform, psychotherapeutic or psychiatric components, mode of delivery, and intervention duration). Study design parameters were also recorded, including methodological approach, setting, comparator groups, and factors related to health equity. Outcome measures included both clinical (eg, skin severity scores, itch or pain levels, DLQI) and psychological indicators (eg, symptoms of depression, anxiety, stress, or overall mental well-being), as well as implementation-related variables, such as adherence, usability, accessibility, and equity considerations. Where applicable, methodological quality and potential sources of bias were assessed using validated critical appraisal tools.

To enhance interpretability and align with digital health implementation frameworks, interventions were further categorized according to interaction modality and level of human involvement into three predefined groups: (1) *synchronous interventions*, defined as real-time interactions between patients and providers (eg, telepsychiatry, video consultations, live e-consults); (2) *asynchronous interventions*, defined as time-independent, user-initiated, or platform-delivered content without real-time clinician interaction (eg, mobile apps, web-based iCBT modules, educational platforms); and (3) *automated interventions*, defined as systems delivering therapeutic or support functions with minimal or no human involvement, often leveraging algorithmic or AI components (eg, chatbots, AI-assisted triage tools, VR environments). Classification was performed independently by 2 reviewers, and discrepancies were resolved by consensus.

Quality Assessment

The methodological quality of all included studies was assessed using design-specific critical appraisal tools developed by the Joanna Briggs Institute (JBI) [19]. Two reviewers (CL and AH) independently conducted the quality assessment, with any discrepancies resolved through consensus.

The appropriate JBI checklist was applied according to the study design: randomized controlled trials, quasi-experimental studies, cohort studies, cross-sectional analyses, case series (with ≥5 participants), and qualitative research were all appraised using the corresponding JBI instruments. For mixed methods studies, both quantitative and qualitative components were assessed separately using the relevant JBI tools.

XSL•FO
RenderX

Each checklist evaluated key domains of methodological rigor, including selection bias, measurement validity, confounding, and clarity of outcome reporting. Studies were not excluded based solely on quality ratings; however, appraisal results were documented to guide the interpretation of the strength and credibility of the evidence. Findings from the quality assessments are presented in tabular format in Multimedia Appendix 2 and summarized narratively to highlight common methodological strengths and limitations across the included literature.

Results

Description of Studies

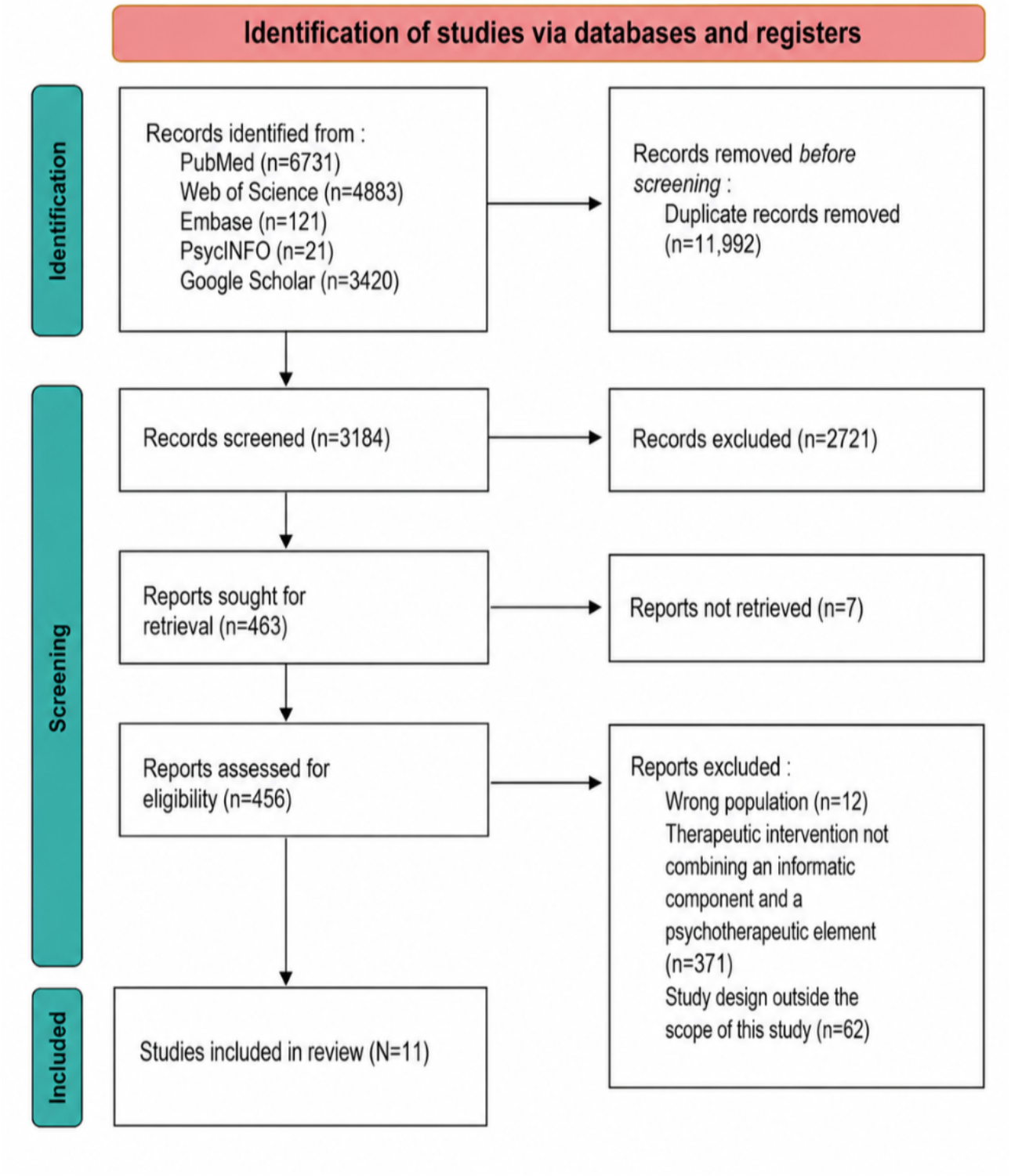
A total of 15,176 records were identified through electronic database searches. After removing 11,992 duplicates, 3184 unique titles and abstracts were screened for relevance. Of these, 2721 were excluded based on eligibility criteria during the initial title and abstract review. Seven additional full-text reports could not be retrieved, leaving 456 articles for full-text review.

During this stage, 445 reports were excluded: 12 due to an ineligible population, 371 for lacking a combined informatics and psychotherapeutic intervention, and 62 for using study designs outside the scope of the review.

Ultimately, 11 studies met all inclusion criteria and were included in the final synthesis. A detailed flowchart of the screening and selection process is presented in Figure 1, and key study characteristics are summarized in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart illustrating the inclusion of studies.



Main Outcomes

This scoping review synthesized findings from 11 peer-reviewed studies evaluating digital interventions with psychotherapeutic or psychiatric components aimed at improving both dermatologic and mental health outcomes. Most studies focused on chronic inflammatory skin diseases, especially psoriasis (n=9), with fewer addressing AD, HS, or vitiligo.

Digital platforms included iCBT, mindfulness-based self-help modules, mobile apps (eg, *DermaScope*, *Allay*), VR interfaces, and AI-assisted screening tools. Psychotherapeutic elements comprised guided iCBT, compassion-focused therapy, acceptance-based therapy, and self-guided mindfulness recordings, while psychiatric components involved consult-liaison psychiatry or psychotropic medication management within digitally integrated pathways.

Many studies reported statistically significant improvements in both dermatologic severity (eg, Psoriasis Area and Severity Index [PASI] scores, itch scales, DLQI) and mental health outcomes (eg, depression, anxiety, illness perceptions), though

XSL•FO
RenderX

effect sizes and adherence rates varied. Dropout rates ranged from 10% to 46%, with generally high acceptability among users who completed the interventions. Most interventions were delivered as adjuncts to standard dermatologic care, but only a minority were integrated within structured care pathways.

Risk of bias was moderate in most studies, with common limitations including small sample sizes, nonblinded designs, and short follow-up periods. Despite these constraints, the evidence suggests that digital psychodermatologic interventions are both feasible and potentially beneficial, warranting further development and refinement. The main outcomes are summarized in Table 1.

Table . Main outcomes across the 11 identified peer-reviewed studies.

Study, year	Population with condition	Intervention summary	Key outcomes	Interaction modality
van Beugen et al, 2016 [20]	Adults with psoriasis	Tailored therapist-guided (iCBT[a]) targeting coping and disease adjustment	Improved physical functioning and daily life impact; mixed psychological effects	Hybrid (asynchronous + synchronous)
Hedman-Lagerlöf et al, 2021 [21]	Adults with AD[b]	Therapist-guided iCBT (12 wk)	Reduced AD severity, itch, stress, depression; sustained at 12 months	Hybrid (asynchronous + synchronous)
Kern et al, 2023 [22]	Adults with AD	Self-guided digital CBT[c]-based intervention (8 wk)	Improvements in QoL[d], itch, depression, stress; feasibility demonstrated	Asynchronous
Kern et al, 2025 [12]	Adults with AD	Self-guided vs clinician-guided online CBT (RCT[e])	Self-guided noninferior to clinician-guided; reduced symptoms (POEM[f])	Hybrid (asynchronous + synchronous)
Fortune et al, 2022 [23]	Adults with psoriasis	Mobile app (*Allay*) delivering digital therapeutic content (CBT-informed)	Improved QoL, resilience, depression, illness perceptions	Asynchronous
Domogalla et al, 2021 [24]	Adults with psoriasis	Smartphone app + educational program integrated into care	Reduced depressive symptoms and QoL improvements (long-term)	Hybrid (asynchronous + synchronous)
Muftin et al, 2022 [25]	Adults with psoriasis	Online compassion-based and mindfulness-based self-help (RCT)	Reduced shame; improved QoL	Asynchronous
Clarke et al, 2024 [26]	Adults with skin conditions and depression	Online compassion-based guided self-help intervention	Improved depression, QoL, self-compassion; feasibility established	Hybrid (asynchronous + synchronous)
Sengupta and Wagani, 2025 [27]	Adults with various skin conditions	Online mindful self-compassion intervention (structured sessions)	Reduced depression, anxiety, stress; improved QoL and well-being	Synchronous (scheduled sessions)
Zhou et al, 2024 [28]	Adults with psoriasis	Online MBCT[g] + usual care	Improved PASI[h], depression, QoL, itch	Asynchronous
Leibovici et al, 2009 [29]	Adults with pruritus (AD/psoriasis)	VR[i] immersion and audiovisual distraction	Reduced pruritus intensity (short-term)	Automated

[a]iCBT: internet-based cognitive behavioral therapy.

[b]AD: atopic dermatitis.

[c]CBT: cognitive behavioral therapy.

[d]QoL: quality of life.

[e]RCT: randomized controlled trial.

[f]POEM: Patient-Oriented Eczema Measure.

[g]MBCT: mindfulness-based cognitive therapy.

[h]PASI: Psoriasis Area and Severity Index.

[i]VR: virtual reality.

Landscape of Digital Interventions

The digital interventions identified in this review encompassed a broad range of informatics modalities designed to support psychotherapeutic or psychiatric care in populations with dermatologic disorders.

Web-based iCBT was the most extensively studied approach. In the trial by van Beugen et al [20], a 12-week secure online platform delivered guided modules targeting stress,

XSL•FO
RenderX

dysfunctional beliefs, and disease coping in patients with psoriasis, producing moderate improvements in quality of life and psychological flexibility compared to care-as-usual. Similarly, Fortune et al [23] embedded a therapist-supported iCBT program within the *Allay* mobile app, providing structured psychoeducation and skill-building for emotional distress in psoriasis, with early evidence of improved illness perceptions and emotional adjustment.

Compassion and mindfulness-based audio programs were examined in 2 studies. They were delivered via MP3 or digital links. Muftin et al [25] compared self-directed compassion-focused therapy with mindfulness-based therapy for psoriasis, with both interventions showing moderate reductions in distress and improved skin-related attitudes. Clarke et al [26] evaluated a guided audio intervention using kindness and acceptance-based recordings for patients with AD and HS, noting favorable user engagement and emotional improvements.

mHealth apps such as *DermaScope* and *Allay* were also common platforms [23,24]. *DermaScope* offered educational content and emotional support prompts, achieving high user satisfaction and reducing depressive symptoms in a German psoriasis cohort. *Allay*, implemented in the United States and Ireland, enabled daily symptom monitoring, psychoeducational content, and therapist communication, with most users requesting continued access after the study.

AI-driven tools were more exploratory. Zhou et al [28] developed a machine learning model to predict psychiatric comorbidity risk in HS, demonstrating the feasibility of AI-assisted mental health triage within dermatology.

Similarly, VR was used in Leibovici et al [29] as a distraction technique for patients experiencing severe itching, with immersive 3D environments reducing perceived discomfort during acute episodes.

Emerging modalities included Sengupta et al [27] conversational chatbot integrated with skin-tracking functions to provide CBT-based support to adolescents with vitiligo, which showed positive user attitudes and symptom control over time. Kern et al [12,22] piloted a novel e-psychiatry consult-liaison model for dermatology referrals, enabling synchronous psychiatric evaluation via teleplatforms, thereby improving continuity of care for unmet psychiatric needs.

Target Populations and Diagnoses

Psoriasis was the primary focus in most included studies, with 9 out of 11 trials targeting individuals with moderate-to-severe plaque psoriasis. Participants were typically adults aged from 35 to 53 years, with a predominance of female respondents and relatively high levels of education. Several protocols assumed or prescreened for digital literacy, reflecting a bias toward digitally engaged populations. Few studies reported detailed breakdowns of socioeconomic status, race, or rural versus urban residence, limiting the generalizability to marginalized groups.

HS was the sole focus in Zhou et al [28], which applied AI-based stratification to improve psychiatric detection and triage. Sengupta et al [27] included a broader diagnostic mix, engaging adolescents and young adults with vitiligo, AD, and other pigmentary disorders via a chatbot-enabled support system. AD and chronic itch were secondarily mentioned in Clarke et al [26] and Leibovici et al [29], respectively, though without significant subgroup analyses.

No study targeted pediatric or geriatric patients exclusively, and few interventions were adapted for linguistic or cognitive accessibility. Clinical samples were largely restricted to mild-to-moderate symptom severity, and most trials excluded patients with active suicidality, substance use disorders, or psychosis. This suggests a significant gap in psychodermatologic research involving more complex or underserved patient populations.

Outcomes

All included studies assessed at least 1 validated mental health outcome, most commonly using the Hospital Anxiety and Depression Scale, Patient Health Questionnaire-9, Brief Illness Perception Questionnaire, and visual analog scales for stress and emotional burden. Dermatologic outcomes were also consistently reported, with 8 studies using standardized measures such as the PASI, the Dermatology Life Quality Index (DLQI), and Numeric Rating Scale for itch and pain.

Several studies have demonstrated statistically significant, though modest, reductions in emotional distress. Domogalla et al [24] reported a mean 3.2-point decrease on the Patient Health Questionnaire-9 following a mobile app intervention, while van Beugen et al [20] found reduced self-reported helplessness and avoidance behavior using a digital CBT protocol. Fortune et al [23] observed improvements in illness beliefs (particularly in perceived control and emotional representation) following therapist-supported iCBT via the *Allay* app.

Multicomponent programs incorporating mindfulness or compassion elements yielded improvements not only in depression and anxiety scores but also in broader functional outcomes, such as sleep quality, social participation, and perceived stigmatization [25,26]. DLQI scores improved in 6 studies, with reductions of 3 to 7 points from baseline, indicating clinically meaningful gains in daily functioning and disease coping.

Long-term effects were less certain, as few studies included follow-up beyond 8 to 12 weeks. Notable exceptions included Hedman-Lagerlöf et al [21] who conducted a randomized controlled trial of exposure-based iCBT for AD, demonstrating sustained reductions in skin-related avoidance, anxiety, and depressive symptoms at a 12-month follow-up.

The study by Zhou et al [28] was the only one to focus on predicting psychiatric morbidity algorithmically, rather than evaluating therapeutic outcomes directly, underscoring a gap between predictive and interventional research.

Overall, while effect sizes varied, the evidence consistently pointed to positive impacts on both dermatologic and psychological well-being when digital tools incorporated structured therapeutic content. A summary of the findings is presented in Table 2.

XSL•FO
RenderX

Table . Summary of the findings.

Category	Summary
Most studied condition	Psoriasis (n=9 studies), with some inclusion of HS[a], AD[b], and vitiligo
Age range of participants	Adults aged 18 - 65 y; mean range across studies: 35 - 53 y
Gender representation	Majority female, digitally literate, higher education level
Digital modalities used	Web platforms, mobile apps, chatbots, VR[c], AI[d] triage tools, e-consults
Therapeutic approaches	CBT[e], ACT[f], mindfulness, compassion-based therapy, liaison psychiatry
Mental health outcomes measured	HADS[g], PHQ-9[h], B-IPQ[i], stress and mood scales, QoL[j] indicators
Dermatologic outcomes measured	PASI[k], DLQI[l], itch or pain scales, visual symptom tracking
Key mental health results	Small-to-moderate reductions in anxiety, depression, and distress; improved illness beliefs
Key dermatologic results	DLQI improved in 6 studies; itch and pain modestly reduced in short-term
Duration of interventions	From 10-min VR to 8- to 12-wk programs; mostly 4 - 8 wk
Provider involvement	Ranged from self-guided applications to full psychiatric consults; minimal hybrid models
Barriers identified	Digital fatigue, limited literacy, emotional burden of disease reminders
Facilitators identified	Autonomy, ease of access, therapist support, personalization of content
Equity considerations	Minimal focus on older adults, underserved groups, or multilingual adaptation
Gaps in research	Lack of long-term outcomes, economic data, implementation strategies, equity, and digital alliance measures

[a]HS: hidradenitis suppurativa.
[b]AD: atopic dermatitis.
[c]VR: virtual reality.
[d]AI: artificial intelligence.
[e]CBT: cognitive behavioral therapy.
[f]ACT: acceptance and commitment therapy.
[g]HADS: Hospital Anxiety and Depression Scale.
[h]PHQ-9: Patient Health Questionnaire-9.
[i]B-IPQ: Brief Illness Perception Questionnaire.
[j]QoL: quality of life.
[k]PASI: Psoriasis Area and Severity Index.
[l]DLQI: Dermatology Life Quality Index.

Implementation and Sustainability

The majority of interventions were designed as adjunctive supports rather than standalone treatments, aiming to complement rather than replace dermatologic care. Delivery was predominantly asynchronous, allowing users to access modules, audio content, or app-based features on their own schedules without real-time clinician involvement. Intervention duration varied widely—from a single 10-minute VR session for itch relief to 8- to 12-week therapist-supported iCBT or self-guided compassion-based therapy protocols [20,25,26,29]. Clinical involvement ranged from fully automated, self-guided platforms (eg, Sengupta 2025 chatbot, Domogalla mobile app) to integrated psychiatry consultations delivered via telehealth, as seen in Kern et al [12], where dermatologists referred patients to e-psychiatry services for diagnostic clarification and psychotropic management [22].

Reported barriers to implementation included digital fatigue (with daily app reminders or disease-focused content perceived as burdensome) [23,25], technological literacy challenges, and data privacy concerns, particularly among older or digitally underserved populations [28]. Dropout rates ranged from 10% to 46%, with reasons for attrition often unreported or multifactorial. Conversely, facilitators of engagement included content personalization, 24/7 accessibility, and a sense of autonomy and empowerment, especially in self-help platforms [24,26].

Despite these insights, few studies examined long-term sustainability, cost-effectiveness, or integration into routine dermatology workflows. Only Kern et al [12,22] assessed provider workflow impact, and no study conducted formal equity analyses by race, income, rurality, or language. Overall, while the evidence supports feasibility, substantial gaps remain in

XSL•FO
RenderX

implementation science and health equity evaluation for digital psychodermatology.

Guidance Level and Adherence

Across the included studies, interventions varied substantially in the degree of human support provided, ranging from fully self-guided (unguided) digital interventions to clinician- or therapist-supported (guided) models, with several hybrid approaches combining asynchronous content with limited human interaction. Guided interventions, including therapist-supported iCBT and structured clinician feedback embedded within digital platforms, were reported in studies such as Hedman-Lagerlöf et al [21] and van Beugen et al [20], where participants received periodic support, feedback, or monitoring alongside digital modules. In contrast, unguided interventions—including self-help mobile apps and fully automated online programs (eg, Fortune et al [23], Muftin et al [25], and Kern et al [12] self-guided arm)—relied primarily on user-driven engagement without real-time or personalized clinician input [20,21]. Hybrid models, such as those incorporating minimal email support or periodic check-ins (eg, [24,26]), occupied an intermediate position in terms of resource intensity. When examined in relation to adherence, a pattern emerged whereby studies reporting higher levels of human support tended to demonstrate lower attrition and more sustained engagement, whereas fully unguided interventions were more frequently associated with higher dropout rates [24,26]. For example, therapist-guided iCBT trials reported relatively stable completion rates despite minimal clinician time investment, while several self-guided interventions noted usability challenges or attrition over time. Overall dropout rates varied across studies with variability partly attributable to differences in intervention intensity, duration, and user support. Although the heterogeneity of study designs and reporting precludes formal quantitative comparison, these findings suggest a potential association between the degree of human support and intervention adherence, with guided or hybrid models appearing to mitigate disengagement relative to fully self-directed approaches.

Quality Appraisal

Using JBI appraisal tools, most studies demonstrated a moderate risk of bias. Randomized controlled designs were generally of higher methodological quality but remained susceptible to open-label effects and attrition [20,25]. Quasi-experimental and observational studies provided useful insights but were limited by small sample sizes and short follow-up periods [24,27]. Overall, quality ratings across studies ranged from moderate to high based on the JBI checklists. Detailed assessments are presented in Multimedia Appendix 2.

Knowledge Gaps

A number of critical knowledge gaps were identified across the reviewed literature. First, the demographic reach of interventions was limited: most samples comprised middle-aged, digitally literate adults, with minimal inclusion of older adults, individuals with low health or digital literacy, or patients from rural or socioeconomically disadvantaged backgrounds. No study provided stratified analyses by race, ethnicity, or income, and none were explicitly designed to address the needs of underserved populations.

While several interventions demonstrated short-term clinical or psychological benefits, none assessed long-term psychiatric outcomes, such as suicidal ideation, substance use disorders, or psychiatric hospitalization, despite the elevated prevalence of these conditions in chronic dermatologic illnesses. Furthermore, sustainability and cost-effectiveness were rarely addressed. Only Kern et al [12] investigated system-level integration into dermatologic workflows, and no study conducted a formal economic evaluation or budget impact analysis.

There was also a notable absence of key implementation metrics, such as intervention fidelity, training burden, and reimbursement feasibility, limiting the applicability of findings to real-world clinical settings. Although several interventions were delivered via mobile apps or web-based portals, blended care models that combine digital and in-person therapeutic components remain largely unexplored.

Additionally, the digital therapeutic alliance, a key factor in mental health treatment efficacy, was not systematically assessed in any study. Equity-focused implementation strategies, including culturally adapted content or multilingual delivery formats, were entirely absent.

Addressing these gaps is essential to ensure that psychodermatological digital tools are scalable, inclusive, and clinically meaningful across diverse patient populations and care systems.

Discussion

Principal Results

This scoping review synthesized findings from 11 studies evaluating informatics-enabled psychotherapeutic or psychiatric interventions in dermatology. Most interventions were delivered via digital platforms, including web-based cognitive CBT, mHealth apps, mindfulness audio modules, VR tools, and AI-supported systems. The observed effects of digital psychodermatologic interventions can be interpreted through the lens of the brain-skin axis, a bidirectional neuroimmunological pathway linking psychological stress and cutaneous inflammation. By reducing stress and improving coping and behavioral responses (eg, scratching), interventions such as iCBT may attenuate activation of the hypothalamic-pituitary-adrenal axis and downstream inflammatory processes, contributing to improvements in both mental health and skin severity (eg, PASI). This mechanistic framework supports the integration of digital mental health tools as adjunctive strategies in the management of chronic inflammatory skin diseases.

Psoriasis emerged as the most frequently studied dermatologic condition, while fewer studies addressed AD, HS, or mixed skin disorders. All studies assessed mental health outcomes, most commonly targeting depression, anxiety, stress, or illness perception. Several interventions demonstrated moderate improvements in emotional distress, self-efficacy, and

XSL•FO
RenderX

dermatologic quality of life, particularly those integrating CBT or self-compassion frameworks.

However, digital equity considerations, long-term adherence, and integration into routine dermatologic care were rarely assessed. Methodological quality was generally moderate to high, but many studies were underpowered, lacked control groups, or experienced high attrition.

Comparison With Prior Work

Psychological therapies have demonstrated measurable benefits in chronic skin conditions such as psoriasis [30]. A recent systematic review of randomized trials found that interventions such as cognitive behavioral therapy (CBT), mindfulness-based interventions, and structured patient education can improve both mental health outcomes and dermatologic severity in psoriasis [31]. Similarly, Qureshi et al [30] reported that adjunctive psychological programs (including stress management and biofeedback) were associated with better skin clearance and quality-of-life improvements in psoriatic disease. However, most interventions in these studies were delivered in person, with only a minority using digital or telehealth formats [31]. This gap highlights an underexplored opportunity in dermatology, given the logistical advantages of digital platforms for chronic care.

Our review highlights the unique value of informatics platforms (such as web portals and mobile apps) in psychodermatology. Digital tools can reduce access barriers and extend care beyond clinic visits through on-demand support and self-management resources. Unlike weekly face-to-face sessions, online programs can provide asynchronous, low-cost coaching that patients can access between dermatology appointments. Evidence from general mental health research suggests this convenience does not necessarily compromise efficacy: multiple meta-analyses have shown that guided iCBT produces outcomes equivalent to those of in-person therapy for anxiety and depression [32]. Andersson and colleagues [32] further demonstrated that across conditions, iCBT achieves symptom reductions comparable to face-to-face care, although direct comparisons remain limited.

Despite these findings, dermatology still lacks head-to-head trials explicitly comparing digital and traditional psychotherapeutic delivery. This research underscores the importance of formally evaluating whether internet-delivered psychotherapies can match the clinical outcomes of office-based care in populations with dermatologic disorders.

Digital CBT has also proven effective in other chronic medical conditions, such as improving depression in diabetes and reducing pain catastrophizing in arthritis [33]. However, adaptation of these interventions to dermatology-specific needs remains limited. Content must be tailored to address skin-related stressors, such as visible lesions, itch-scratch cycles, and social stigma. Yet, personalized psychodermatology programs remain rare, with few adapted to specific clinical phenotypes (eg, itch-dominant psoriasis or high-stigma AD), which reinforces the call for more personalized digital therapeutics. As Kazdin [34] has argued in his framework for mental health innovation, digital platforms offer a unique opportunity to personalize therapy content to individual patient profiles at scale, which is an important direction for future research.

A striking finding of this review is the disproportionate focus on psoriasis, which accounted for 9 of the 11 included studies, with comparatively minimal representation of other high-burden dermatologic conditions such as AD and HS. This imbalance likely reflects several structural drivers within the field. First, psoriasis benefits from a long-standing research infrastructure, including the widespread use of standardized and validated outcome measures such as the PASI and the DLQI, which facilitate trial design, comparability, and regulatory alignment. In contrast, while AD has measures such as the Patient-Oriented Eczema Measure and Eczema Area and Severity Index, and HS has emerging indices such as Hidradenitis Suppurativa Clinical Response, these tools have been less consistently integrated into psychodermatology and digital intervention research, potentially limiting study development and cross-study synthesis. Second, psoriasis has been a primary target of substantial pharmaceutical investment, particularly with the expansion of the biologics market, which has historically driven broader innovation ecosystems, including adjunctive behavioral and digital interventions. This contrasts with conditions such as HS, which, despite significant psychosocial burden and high rates of psychiatric comorbidity, has received comparatively less research funding and fewer large-scale intervention studies. Third, psoriasis populations are often easier to recruit within structured clinical pathways and registries, enabling digital trial implementation, whereas patients with other conditions may be more fragmented across care settings. The consequence is a literature that may overrepresent conditions with strong biomedical and commercial ecosystems while underrepresenting those with equally or greater psychosocial need. This raises important concerns regarding the generalizability of current findings and highlights a critical gap in the development of digital psychodermatologic interventions for underserved patient populations. Future research should prioritize conditions such as HS and AD not only to ensure equity in innovation but also to better align digital intervention development with the true distribution of psychological burden across dermatologic diseases.

The present findings should be interpreted in light of the mixed methods systematic review by Hewitt et al [35], which represents the most closely related synthesis of digital psychological interventions in dermatology and identified 23 studies, including 15 randomized controlled trials, concluding that such interventions may improve psychological outcomes such as mood, quality of life, and illness-related knowledge despite substantial heterogeneity limiting definitive conclusions. While both reviews examine the intersection of digital technologies and psychological support in populations with dermatologic disorders, this study extends and refines this evidence base through several key distinctions. Our inclusion criteria were deliberately more restrictive, requiring interventions to integrate both an informatics component and a structured psychotherapeutic or psychiatric element, and to report dermatologic and/or mental health outcomes; this approach excludes purely educational or communication-based tools without explicit therapeutic intent, thereby enhancing

XSL•FO
RenderX

clinical interpretability despite yielding a smaller number of included studies (n=11). The review also extends the literature temporally and conceptually by incorporating studies published up to March 2025, capturing the emergence of newer modalities such as AI-assisted tools, virtual assistants, and hybrid care models that were largely absent from earlier syntheses and signal a shift toward more interactive, personalized, and scalable psychodermatologic interventions. In contrast to prior work, this synthesis explicitly treats dermatologic and psychological outcomes as coprimary domains, integrating clinical severity indices (eg, PASI, DLQI) with validated mental health measures to adopt a more comprehensive biopsychosocial perspective that reflects the bidirectional relationship between skin disease and psychological distress. Additionally, greater emphasis is placed on implementation considerations, including adherence, integration into clinical workflows, and equity-related barriers; although prior work identified issues of acceptability and usability, this analysis further examines how such interventions are operationalized in practice and highlights persistent gaps, particularly regarding long-term engagement and access among underserved populations. Taken together, this review does not contradict prior findings but rather refines and extends them through a more clinically focused definition of intervention, the inclusion of emerging technologies, and the positioning of outcomes within an integrated dermatologic-psychiatric framework, thereby contributing to the advancement of scalable and clinically embedded models of psychodermatologic care.

Limitations

This scoping review has several limitations. Despite a comprehensive multidatabase search strategy, relevant studies may have been missed, particularly those published in nonindexed journals, in languages other than English or French, or under alternative terminologies not captured by our search terms.

The heterogeneity in study designs, populations, interventions, and outcomes limited our ability to perform quantitative synthesis or direct cross-study comparisons. Many included studies had small sample sizes, high attrition rates, and limited follow-up, reducing the generalizability and robustness of their findings. Most trials also lacked standardized measures of adherence, usability, or cost-effectiveness, which are essential for assessing scalability. Also, a key finding of this review is the widespread lack of reporting on sociodemographic variables such as race or ethnicity, education, and socioeconomic status, despite frequent references to barriers like digital literacy. This gap limits the interpretability and generalizability of findings and raises concerns that current evidence may disproportionately reflect more digitally literate and advantaged populations, underscoring the need for standardized reporting and more inclusive study designs.

This review did not systematically include gray literature, which may have excluded unpublished or program-level innovations from clinical settings. Finally, although we aimed to capture both dermatological and mental health outcomes, few studies reported comprehensively on both, and even fewer examined long-term effects or implementation into routine dermatological care.

Conclusions

This scoping review provides a comprehensive synthesis of the emerging field of informatics-based psychotherapeutic and psychiatric interventions in dermatology. The reviewed studies highlight the feasibility and preliminary efficacy of digital tools in addressing both psychological distress and dermatologic symptom burden across a range of chronic skin conditions. These interventions offer accessible and scalable, patient-centered models of care that can complement traditional dermatology services, particularly in resource-limited or remote settings.

However, this evidence base is constrained by methodological heterogeneity, small sample sizes, high attrition, and insufficient reporting on equity, usability, and long-term impact. The lack of integration into dermatologic workflows and the absence of head-to-head comparative studies with face-to-face psychotherapy also represent additional knowledge gaps.

Future research should prioritize designing trials that evaluate digital interventions in diverse populations, incorporate equity-informed frameworks, and test hybrid models of care that blend digital tools with human support. As the fields of psychodermatology and digital health continue to converge, the development of personalized, adaptable, and clinically integrated digital mental health interventions will be essential to improving patient outcomes and reducing disparities in dermatologic care.

Funding

This study received indirect funded from the La Fondation de l’Institut universitaire en santé mentale de Montréal and operating funds from IVADO for AH. We did not use generative AI tools for this manuscript.

Authors' Contributions

Conceptualization: CL, AH.
Formal analysis: CL, AH.
Methodology and validation: CL, AH.
Funding acquisition: AH.
Investigation: AH.
Project administration: AH.
Supervision: AH, JC.

XSL•FO
RenderX

Resources: AH.
Data curation: CL, AH, JC.
Writing — original draft: CL, AH, JC.
Writing — review and editing: CL, AH, JC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Electronic search strategy for the scoping review conducted.
[DOCX File, 17 KB - derma_v9i1e82096_app1.docx]

Multimedia Appendix 2
Systematic review study selection detailed results.
[DOCX File, 33 KB - derma_v9i1e82096_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 111 KB - derma_v9i1e82096_app3.docx]

References

1. Dalgard FJ, Gieler U, Tomas-Aragones L, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015 Apr;135(4):984-991. [doi: 10.1038/jid.2014.530] [Medline: 25521458]
2. Henderson AD, Adesanya E, Mulick A, et al. Common mental health disorders in adults with inflammatory skin conditions: nationwide population–based matched cohort studies in the UK. BMC Med 2023 Aug 4;21(1):285. [doi: 10.1186/s12916-023-02948-x] [Medline: 37542272]
3. Caccavale S, Tancredi V, Boccellino MP, Babino G, Fulgione E, Argenziano G. Hidradenitis suppurativa burdens on mental health: a literature review of associated psychiatric disorders and their pathogenesis. Life (Basel) 2023 Jan 9;13(1):189. [doi: 10.3390/life13010189] [Medline: 36676137]
4. Misery L, Chesnais M, Merhand S, et al. Perceived stress in four inflammatory skin diseases: an analysis of data taken from 7273 adult subjects with acne, atopic dermatitis, psoriasis or hidradenitis suppurativa. J Eur Acad Dermatol Venereol 2022 Aug;36(8):e623-e626. [doi: 10.1111/jdv.18016] [Medline: 35181931]
5. Mostaghimi L, Noughani H. “Psychodermatology”: the present standing and a path forward. J Eur Acad Dermatol Venereol 2022 Aug;36(8):1162-1163. [doi: 10.1111/jdv.18278] [Medline: 35848144]
6. Christensen RE, Jafferany M. Psychiatric and psychologic aspects of chronic skin diseases. Clin Dermatol 2023;41(1):75-81. [doi: 10.1016/j.clindermatol.2023.03.006] [Medline: 36878453]
7. Chow A, Smith HE, Car LT, et al. Teledermatology: an evidence map of systematic reviews. Syst Rev 2024 Oct 12;13(1):258. [doi: 10.1186/s13643-024-02655-5] [Medline: 39396040]
8. Seghers AC, Seng KH, Chio MTW, Chia E, Ng SK, Tang MBY. A prospective study on the use of teledermatology in psychiatric patients with chronic skin diseases. Australas J Dermatol 2015 Aug;56(3):170-174. [doi: 10.1111/ajd.12297] [Medline: 25754857]
9. Kaliyadan F, Ashique KT. Use of mobile applications in dermatology. Indian J Dermatol 2020;65(5):371-376. [doi: 10.4103/ijd.IJD_422_20] [Medline: 33165393]
10. Cherrez-Ojeda I, Robles-Velasco K, Osorio MF, et al. A systematic review and meta-analysis of mobile health applications and telemonitoring in atopic dermatitis self-management. Dermatol Ther (Heidelb) 2024 Jul;14(7):1787-1798. [doi: 10.1007/s13555-024-01213-0] [Medline: 38909171]
11. Roca S, Almenara M, Gilaberte Y, et al. When virtual assistants meet teledermatology: validation of a virtual assistant to improve the quality of life of psoriatic patients. Int J Environ Res Public Health 2022 Nov 5;19(21):14527. [doi: 10.3390/ijerph192114527] [Medline: 36361408]
12. Kern D, Ljótsson B, Lönndahl L, et al. Self-guided vs clinician-guided online cognitive behavioral therapy for atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2025 Feb 1;161(2):183-190. [doi: 10.1001/jamadermatol.2024.5044] [Medline: 39693097]
13. Peracca SB, Fonseca A, Hines A, et al. Implementation of mobile teledermatology: challenges and opportunities. Telemed J E Health 2021 Dec;27(12):1416-1422. [doi: 10.1089/tmj.2020.0500] [Medline: 33691074]
14. Dovigi E, Kwok EYL, English 3rd JC. A framework-driven systematic review of the barriers and facilitators to teledermatology implementation. Curr Dermatol Rep 2020;9(4):353-361. [doi: 10.1007/s13671-020-00323-0] [Medline: 33200042]

XSL•FO
RenderX

15. Johnson A, Shukla N, Halley M, et al. Barriers and facilitators to mobile health and active surveillance use among older adults with skin disease. Health Expect 2021 Oct;24(5):1582-1592. [doi: 10.1111/hex.13229] [Medline: 34190397]
16. Maddukuri S, Patel J, Lipoff JB. Teledermatology addressing disparities in health care access: a review. Curr Dermatol Rep 2021;10(2):40-47. [doi: 10.1007/s13671-021-00329-2] [Medline: 33747638]
17. Ghani M, Adler C, Yeung H. Patient factors associated with interest in teledermatology: cross-sectional survey. JMIR Dermatol 2021 May 10;4(1):e21555. [doi: 10.2196/21555] [Medline: 37625162]
18. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
19. Munn Z, Barker TH, Moola S, et al. Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. JBI Evid Synth 2020 Oct;18(10):2127-2133. [doi: 10.11124/JBISRIR-D-19-00099] [Medline: 33038125]
20. van Beugen S, Ferwerda M, Spillekom-van Koulil S, et al. Tailored therapist-guided internet-based cognitive behavioral treatment for psoriasis: a randomized controlled trial. Psychother Psychosom 2016;85(5):297-307. [doi: 10.1159/000447267] [Medline: 27508937]
21. Hedman-Lagerlöf E, Fust J, Axelsson E, et al. Internet-delivered cognitive behavior therapy for atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2021 Jul 1;157(7):796-804. [doi: 10.1001/jamadermatol.2021.1450] [Medline: 34009282]
22. Kern D, Ljótsson B, Lönndahl L, et al. A digital self-help intervention for atopic dermatitis: analysis of secondary outcomes from a feasibility study. JMIR Dermatol 2023 Mar 20;6:e42360. [doi: 10.2196/42360] [Medline: 37632924]
23. Fortune DG, Ravnkilde V, Ryan S, Ramsay B, Clough S, Richards HL. A digital therapeutic for management of psychosocial aspects of psoriasis: a pre-post proof of concept study. Skin Health Dis 2022 Jun;2(2):e103. [doi: 10.1002/ski2.103] [Medline: 35677915]
24. Domogalla L, Beck A, Schulze-Hagen T, Herr R, Benecke J, Schmieder A. Impact of an eHealth smartphone app on the mental health of patients with psoriasis: prospective randomized controlled intervention study. JMIR mHealth uHealth 2021 Oct 25;9(10):e28149. [doi: 10.2196/28149] [Medline: 34431478]
25. Muftin Z, Gilbert P, Thompson AR. A randomized controlled feasibility trial of online compassion-focused self-help for psoriasis. Br J Dermatol 2022 Jun;186(6):955-962. [doi: 10.1111/bjd.21020] [Medline: 35041766]
26. Clarke EN, Norman P, Thompson AR. Online compassion-based self-help for depression in people with skin conditions: a feasibility study. Pilot Feasibility Stud 2024 Apr 16;10(1):63. [doi: 10.1186/s40814-024-01486-4] [Medline: 38627850]
27. Sengupta A, Wagani R. Mindful self-compassion for psychological distress associated with skin conditions: an online intervention study. Indian J Dermatol Venereol Leprol 2025;91(1):71-75. [doi: 10.25259/IJDVL_451_2023] [Medline: 38314971]
28. Zhou Y, Zhao Y, Li Y, et al. Online mindfulness-based cognitive therapy as a treatment for patients with psoriasis: a randomized controlled trial. SSRN. Preprint posted online on Aug 30, 2024. [doi: 10.2139/ssrn.4940246]
29. Leibovici V, Magora F, Cohen S, Ingber A. Effects of virtual reality immersion and audiovisual distraction techniques for patients with pruritus. Pain Res Manag 2009;14(4):283-286. [doi: 10.1155/2009/178751] [Medline: 19714267]
30. Qureshi AA, Awosika O, Baruffi F, Rengifo-Pardo M, Ehrlich A. Psychological therapies in management of psoriatic skin disease: a systematic review. Am J Clin Dermatol 2019 Oct;20(5):607-624. [doi: 10.1007/s40257-019-00437-7] [Medline: 30937679]
31. Xiao Y, Zhang X, Luo D, et al. The efficacy of psychological interventions on psoriasis treatment: a systematic review and meta-analysis of randomized controlled trials. Psychol Res Behav Manag 2019;12:97-106. [doi: 10.2147/PRBM.S195181] [Medline: 30799963]
32. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295. [doi: 10.1002/wps.20151] [Medline: 25273302]
33. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
34. Kazdin AE. Annual research review: expanding mental health services through novel models of intervention delivery. J Child Psychol Psychiatry 2019 Apr;60(4):455-472. [doi: 10.1111/jcpp.12937] [Medline: 29900543]
35. Hewitt RM, Ploszajski M, Purcell C, et al. A mixed methods systematic review of digital interventions to support the psychological health and well-being of people living with dermatological conditions. Front Med (Lausanne) 2022;9:1024879. [doi: 10.3389/fmed.2022.1024879] [Medline: 36405626]

Abbreviations

AD: atopic dermatitis
AI: artificial intelligence
CBT: cognitive behavioral therapy
DLQI: Dermatology Life Quality Index
HS: hidradenitis suppurativa

XSL•FO
RenderX

iCBT: internet-based cognitive behavioral therapy
JBI: Joanna Briggs Institute
PASI: Psoriasis Area and Severity Index
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
VR: virtual reality
mHealth: mobile health

Edited by R Dellavalle; submitted 08.Aug.2025; peer-reviewed by H Sadeghsalehi; revised version received 22.May.2026; accepted 25.May.2026; published 10.Jun.2026.

Please cite as:
Lamarre C, Chivinski J, Hudon A
Informatics-Based Psychotherapeutic and Psychiatric Interventions in Dermatology: Scoping Review of Impacts on Skin Disease Severity and Mental Health Outcomes
JMIR Dermatol 2026;9:e82096
URL: https://derma.jmir.org/2026/1/e82096
doi:10.2196/82096

© Caroline Lamarre, Jeffrey Chivinski, Alexandre Hudon. Originally published in JMIR Dermatology (http://derma.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Perceptions of Climate Change Impacts on Atopic Dermatitis: Cross-Sectional Survey Study

Gunnar Mattson[1], MPH, MD; Sarah Coates[2], MD; Amanda R Twigg[2], MD

[1]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]Department of Dermatology, University of California, San Francisco, 1701 Divisadero Street, Third Floor, San Francisco, CA, United States

Corresponding Author:
Amanda R Twigg, MD
Department of Dermatology, University of California, San Francisco, 1701 Divisadero Street, Third Floor, San Francisco, CA, United States

Abstract

This cross-sectional survey study (63.5% response rate) characterized how patients with atopic dermatitis (AD) perceive and experience the effects of climate change on their AD. Most participants reported that environmental factors such as heat and air pollution worsened their AD and expressed a desire for climate-health education, yet few had discussed these concerns with their dermatologist. These findings reveal a gap in patient-centered dermatologic care and support the development of tools to integrate environmental health into atopic dermatitis management.

(JMIR Dermatol 2026;9:e80679) doi:10.2196/80679

KEYWORDS

atopic dermatitis; eczema; climate change dermatology; climate change; environmental health; climate health impacts; patient experience; survey study; health communication; patient education

Introduction

Climate change is recognized as the foremost global health threat of the 21st century [1]. Environmental shifts (rising temperatures, air pollution, and extreme weather) can impair the skin barrier, alter the microbiome, and induce inflammation, increasing the prevalence and severity of atopic dermatitis (AD) among other skin conditions [2,3]. Among dermatologists, 79.6% agree it affects their patients [4]. Yet, few routinely discuss this with patients, and limited research explores how patients perceive and experience these impacts. To address these gaps, this cross-sectional study surveyed patients with AD to assess how they perceive climate change's impact on their condition and whether these concerns are addressed in dermatologic care.

Methods

Survey Instrument Development

The survey was informed by climate-health literature, dermatologic environmental impacts, and health communication frameworks (eg, message framing, perceived susceptibility, and severity from the Health Belief Model) [5]. Five UCSF (University of California, San Francisco) dermatologists reviewed the instrument for clinical relevance and clarity. Ten adult AD patients piloted it, and feedback informed wording and usability.

Study Population & Recruitment

Eligible participants were English-speaking adults with AD seen at UCSF dermatology clinics between August 2023 and August 2024. A total of 2164 patients were identified by the electronic health record (EHR) query. To reduce selection bias, patients were contacted via EHR messaging or mailed letters to account for differences in digital health access; 326 patients expressed interest and became the study population. These patients were sent the study description and a secure Qualtrics link to the online survey.

Statistical Analysis

Descriptive statistics using Microsoft Excel were used to summarize participant demographics and survey responses. Frequencies were calculated for categorical variables. No inferential or hypothesis testing was conducted, as the study aimed to characterize trends and patient-reported experiences rather than test associations or determine causality.

Ethical Considerations

This study received exempt certification from the UCSF medical ethical review committee (IRB 21 - 33538). All participants provided consent to participate in the study, and their responses were deidentified.

Results

Of 326 individuals, 207 completed the survey (63.5% response rate). A majority of individuals (n=166/207, 80.2%, 95% CI

XSL•FO
RenderX

74.8% - 85.6%) reported that environmental-climate factors impact their AD, particularly extreme heat (n=157, 75.8%, 95% CI 70.0% - 81.7%) and poor air quality (n=81, 39.1%, 95% CI 32.5% - 45.8%). Commonly reported effects included increased medication use (n=168, 81.2%, 95% CI 75.8% - 86.5%), more symptomatic flares (n=167, 80.7%, 95% CI 75.3% - 86.1%), more skin affected (n=139, 67.1%; 95% CI 60.8% - 73.5%), and changes to daily behaviors (n=130, 62.8%; 95% CI 56.2% - 69.4%). Most participants (n=179, 86.5%; 95% CI 81.8% - 91.1%) expressed interest in understanding how environmental-climate factors affect their AD, yet only 76 participants (36.7%; 95% CI 30.1% - 43.3%) said their dermatologist addressed these concerns. The most valued strategies for addressing climate-health impacts included more information (n=164, 79.2%; 95% CI 73.7% - 84.8%), dedicated time during visits to plan for exposures (n=105, 50.7%; 95% CI 43.9% - 57.5%), and more in-person visits (n=101, 48.8%; 95% CI 42.0% - 55.6%). Table 1 shows participant characteristics, and Table 2 shows survey response data.

Table . Participant demographics and background information.

Demographics	Participants (N=207)
Age in years (mean, SD)	46.4 (18.6)
Sex, n (%)	
Male	75 (36.2)
Female	129 (62.3)
Nonbinary	3 (1.4)
Race/Ethnicity, n (%)	
American Indian or Alaskan Native	2 (1.0)
Asian or Asian American	82 (39.6)
Black or African American	12 (5.8)
Hispanic or Latino	12 (5.8)
Native Hawaiian or Pacific Islander	1 (0.5)
White	107 (51.7)
Other	5 (2.4)
Years living with atopic dermatitis (mean, SD)	21.6 (18.3)
Treatments used for atopic dermatitis, n (%)	
Topical steroid	193 (93.7)
Topical medication other than a steroid	145 (70.4)
Topical over the counter product (does not require a prescription)	139 (67.4)
Pill medication (eg, methotrexate, cellcept, tofacitinib, upadacitinib)	47 (22.8)
Injection medication (eg, dupilumab, tralokinumab)	94 (45.6)
Phototherapy	41 (19.9)

XSL•FO
RenderX

Table . Responses to survey questions using the 5-point Likert scale, where 1 indicates "Strongly disagree," 2 "Somewhat disagree," 3 "Neutral," 4 "Somewhat agree," and 5 "Strongly agree." A reported mean greater than 3 indicates agreement and less than 3 indicates disagreement.

Statement, agreement ranked using the 5-point Likert scale	Score, mean (SD)
Climate and environmental factors have impacted your experience with eczema	4.2 (1.0)
The following factor has impacted your experience with eczema:	
Extreme Heat	4.2 (1.1)
Wildfires	3.3 (1.1)
Poor Air Quality	3.4 (1.1)
Drought	3.2 (1.1)
Extreme Rainfall	3.0 (1.3)
Sea Level Rise	2.4 (1.0)
Flooding	2.6 (1.1)
Climate and environmental factors' impact on your eczema include:	
More symptomatic with exacerbations or flares	4.2 (1.0)
More skin affected	3.9 (1.2)
Need for extra appointments with healthcare team	3.1 (1.2)
Sending additional messages to dermatologist or calling their office	3.0 (1.2)
Using medication more often	4.1 (1.0)
Change to your medication	3.2 (1.3)
Change to lifestyle or daily behaviors	3.8 (1.1)
You want to know how the climate and environment impact your eczema	4.2 (1.0)
Your dermatologist has talked about how the climate and environment affect your eczema	2.9 (1.3)
This strategy would be helpful in managing changes to your eczema from the climate and environment:	
More visits in person	3.4 (1.1)
More telehealth visits	3.2 (1.1)
Time during visits to make plans for climate or environmental problems	3.5 (1.1)
More information on the topic	4.1 (0.9)
Support groups	2.9 (1.1)

Discussion

Principal Findings

While this study does not evaluate clinical causality, it provides novel insight into how patients perceive and experience the effects of environmental-climate factors on their AD. Most participants perceived climate-related changes in their AD and desired clinical guidance, yet few reported receiving it. These findings suggest that dermatologists should initiate brief conversations about common triggers, particularly heat and air pollution, and provide anticipatory guidance and resources. This insight underscores previously reported low self-efficacy among dermatologists in discussing climate change with patients [4]. Understanding these patient insights is vital to providing patient-centered care and forming effective partnerships with patients about their skin health. These efforts align with the American Academy of Dermatology's commitment to "educate our patients about the effects of climate change on the health of their skin." [6]

Limitations and Future Direction

Limitations include a single-center design limiting generalizability, reliance on self-reported data with potential recall bias, and possible self-selection bias, as patients more affected by climate change may have been more likely to participate. Future research should validate these findings in broader populations, explore climate-health experiences in other skin conditions, and develop educational and clinical strategies to help navigate these climate-health conversations with patients. Even in short visits, dermatologists can explore patient experiences with climate change using supportive prompts (eg, "Would it be helpful to discuss how environmental factors might relate to your flares?") to validate patient concerns and provide opportunities for personalized climate-health conversations to be continued in subsequent visits.

Conclusions

This study highlights a disconnect between how patients with AD experience climate-related triggers and how often these

XSL•FO
RenderX

concerns are addressed in clinical care. Findings underscore the need for tools and strategies to support climate-health conversations in dermatology. Integrating environmental health into AD management can enhance patient-centered care, improve outcomes, and reinforce dermatology's role at the intersection of clinical care, public health, and patient advocacy.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Watts N, Amann M, Arnell N, et al. The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. Lancet 2021 Jan 9;397(10269):129-170. [doi: 10.1016/S0140-6736(20)32290-X] [Medline: 33278353]
2. Mieczkowska K, Stringer T, Barbieri JS, Williams M, Rosenbach M. Surveying the attitudes of dermatologists regarding climate change. Br J Dermatol 2022 Apr;186(4):748-750. [doi: 10.1111/bjd.20900] [Medline: 34787308]
3. Ahn K. The role of air pollutants in atopic dermatitis. J Allergy Clin Immunol 2014 Nov;134(5):993-999. [doi: 10.1016/j.jaci.2014.09.023] [Medline: 25439225]
4. Celebi Sozener Z, Ozdel Ozturk B, Cerci P, et al. Epithelial barrier hypothesis: effect of the external exposome on the microbiome and epithelial barriers in allergic disease. Allergy 2022 May;77(5):1418-1449. [doi: 10.1111/all.15240] [Medline: 35108405]
5. Alyafei A, Easton-Carr R. The health belief model of behavior change. In: StatPearls [Internet]: StatPearls Publishing; 2025. URL: https://www.ncbi.nlm.nih.gov/books/NBK606120 [accessed 2025-03-30]
6. Position Statement on Climate and Health. American Academy of Dermatology. 2018. URL: https://server.aad.org/forms/policies/Uploads/PS/PS%20-%20Climate%20and%20Health.pdf [accessed 2024-03-31]

Abbreviations

AD: atopic dermatitis
EHR: electronic health record
UCSF: University of California, San Francisco

Edited by E Parker; submitted 15.Jul.2025; peer-reviewed by C Nnajiofor, J Volkman; revised version received 01.Oct.2025; accepted 01.Jan.2026; published 20.Jan.2026.

Please cite as:
Mattson G, Coates S, Twigg AR
Patient Perceptions of Climate Change Impacts on Atopic Dermatitis: Cross-Sectional Survey Study
JMIR Dermatol 2026;9:e80679
URL: https://derma.jmir.org/2026/1/e80679
doi:10.2196/80679

© Gunnar Mattson, Sarah Coates, Amanda R Twigg. Originally published in JMIR Dermatology (http://derma.jmir.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Shifting Perceptions of Cosmetic Procedures: Rise of Medical Spas in Online Search Trends

Hong Ye[1*], MS; Sophia Yanjia Zhu[2], BS; Britta Verhey[3], BS; Joy Xu[4], BHSc; Gordon Hyeonjin Bae[5*], MD

[1]Virginia Tech Carilion School of Medicine, Virginia Polytechnic Institute, Roanoke, VA, United States

[2]Department of Molecular, Cellular, and Developmental Biology, College of Letters and Sciences, University of California, Santa Barbara, Santa Barbara, CA, United States

[3]School of Medicine, University of Nevada, Reno, Reno, NV, United States

[4]David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

[5]Department of Dermatology, Stanford School of Medicine, Stanford University, 450 Broadway Street, Redwood City, CA, United States

[*]these authors contributed equally

Corresponding Author:
Gordon Hyeonjin Bae, MD
Department of Dermatology, Stanford School of Medicine, Stanford University, 450 Broadway Street, Redwood City, CA, United States

Abstract

Our analysis of online search trends suggests that over the past decade, public search interest has increasingly associated certain aesthetic procedures, particularly injectables and laser hair removal, with medical spas rather than dermatologists.

(JMIR Dermatol 2026;9:e83509) doi:10.2196/83509

KEYWORDS

patient safety; cosmetic procedures; aesthetic procedures; med spa; medical spa; nonphysician provider; dermatologist; online search trends; Google Trends

Introduction

In recent years, the rise of medical spas (med spas) has reshaped aesthetic dermatology. As cosmetic procedures are increasingly framed as wellness and self-care procedures, patients may view them as routine, driving demand for services performed by nonphysician providers. This raises concerns about patient safety and clinical oversight. Demand for cosmetic procedures continues to increase globally, with preferences varying by age, ethnicity, and geography [1]. Social media engagement has paralleled rising search interest in popular cosmetic terms, further shaping patient decision-making. To assess evolving public perceptions, we analyzed online search trends, comparing queries pairing common cosmetic procedures with either “med spa” or “dermatologist.”

Methods

Study Design

We used Google Trends (GT) to examine US relative search volumes (RSVs) from June 2015 to June 2025 for 7 aesthetic procedures—Botox injection, laser hair removal, chemical peels, laser skin resurfacing, microdermabrasion, microneedling, and dermal fillers. These procedures were selected to represent cosmetic treatments ranging from minimally invasive (eg, chemical peels) to more complex interventions (eg, laser skin resurfacing), spanning varying levels of procedural risk and required expertise. Search terms were identified via GT’s “Related queries” feature (filters: United States, all categories, web search) and supplemented by researcher consensus, capturing both generic descriptors and brand names. Each term was queried in combination with “med spa” or “dermatologist” to create matched pairs (Multimedia Appendix 1). Terms were excluded if GT returned RSVs of 0 for the majority of the study period.

For procedures with multiple search terms, monthly RSVs were averaged within each framing to produce a single composite time series. Paired 1-tailed *t* tests compared mean monthly RSVs between framings, with each month as a matched time point. Effect size was defined as the mean RSV difference (“med spa” minus “dermatologist”); positive values indicated higher RSVs for “med spa.” We reported 95% CIs. The normality of paired differences was assessed via the Shapiro-Wilk test; 3 procedures (Botox injection, microneedling, and lip fillers) showed departures from normality, and Wilcoxon signed rank tests confirmed all results. Statistical significance was set at $P<.05$.

Ethical Considerations

This retrospective study is exempt from informed consent. It was a secondary analysis of existing data, it did not involve intervention or interaction with human participants, and the reviewed data were deidentified per the deidentification standard

XSL•FO
RenderX

defined in Section §164.514(a) of the HIPAA (Health Insurance Portability and Accountability Act) Privacy Rule. The process by which the data were deidentified was attested to through a formal determination by a qualified expert, as defined in Section §164.514(b)(1) of the HIPAA Privacy Rule. This formal determination refreshed on December 2020.

Results

Significant differences in search interest emerged between "med spa" and "dermatologist" queries for all 7 procedures (Figure 1). "Med spa" framing was associated with significantly higher RSVs for lip fillers ("med spa"=35.7; "dermatologist"=9.1; effect size=26.6), Botox injection ("med spa"=49.4; "dermatologist"=29.1; effect size=20.3), laser hair removal ("med spa"=71.1; "dermatologist"=63.7; effect size=7.4), and microneedling ("med spa"=32.6; "dermatologist"=28.5; effect size=4.1) (Table 1). "Dermatologist" framing showed higher RSVs for chemical peels ("dermatologist"=71.6; "med spa"=22.8; effect size=−48.8), laser skin resurfacing ("dermatologist"=48.0; "med spa"=8.9; effect size=−39.1), and microdermabrasion ("dermatologist"=51.7; "med spa"=18.9; effect size=−32.8). All comparisons were statistically significant (all *P* values≤.004).

RSV is a normalized metric ranging from 0 to 100, where 100 represents peak search interest for a given term within the specified time period and geography. Lower values reflect proportionally lower interest relative to that peak. RSVs are not directly comparable across different search queries.

Figure 1. Relative search volumes (RSVs) over time for cosmetic procedures paired with "med spa" versus "dermatologist" (2015 - 2025).

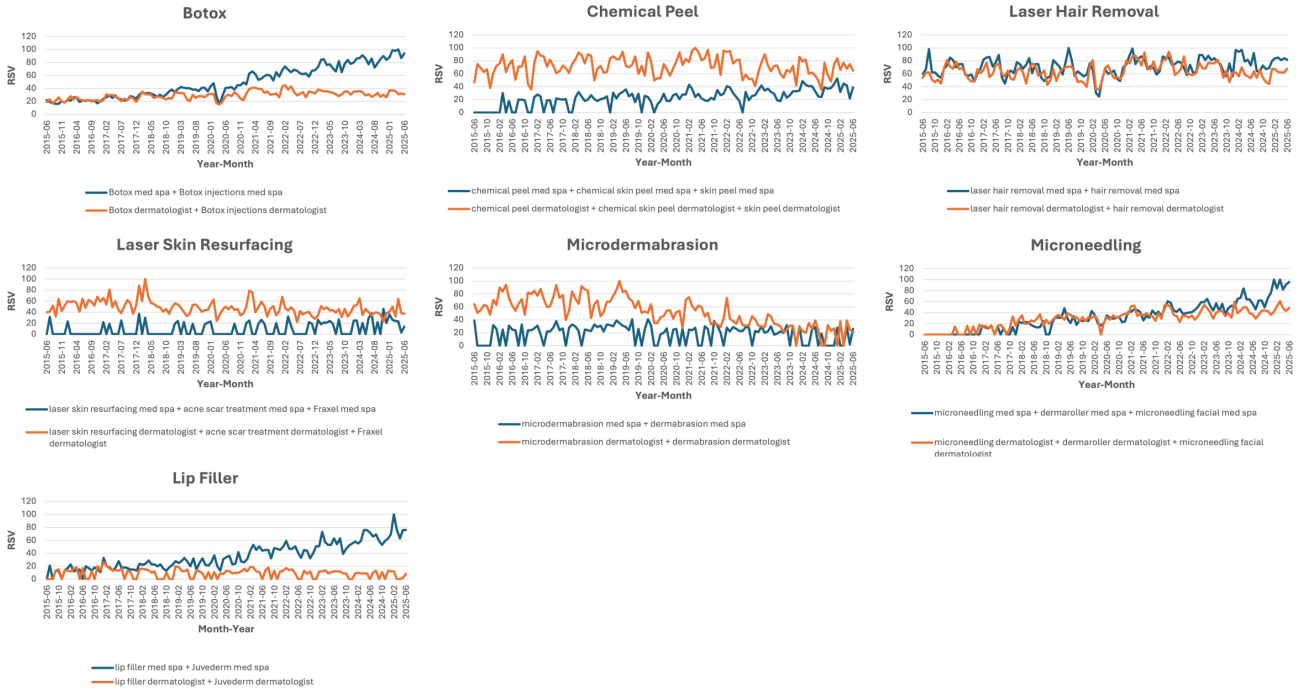


Table . Relative search volumes (RSVs) for cosmetic procedures paired with "med spa" versus "dermatologist" queries.

Procedure type	"Med spa" RSV, mean (SD)	"Dermatologist" RSV, mean (SD)	Effect size (95% CI)[a]	*P* value	Query with greater RSV
Laser hair removal	71.1 (13.7)	63.7 (12.0)	7.39 (5.19 to 9.58)	<.001	Med spa
Botox injection	49.4 (24.5)	29.1 (6.2)	20.3 (16.5 to 24.0)	<.001	Med spa
Lip filler	35.7 (20.4)	9.10 (6.6)	26.6 (22.7 to 30.4)	<.001	Med spa
Microneedling	32.6 (25.3)	28.5 (15.8)	4.05 (1.32 to 6.78)	.004	Med spa
Laser skin resurfacing	8.89 (12.2)	48.0 (12.8)	−39.1 (−42.2 to −35.9)	<.001	Dermatologist
Microdermabrasion	18.9 (13.3)	51.7 (22.9)	−32.8 (−37.0 to −28.7)	<.001	Dermatologist
Chemical peel	22.8 (13.2)	71.6 (14.8)	−48.8 (−51.8 to −45.7)	<.001	Dermatologist

[a]Effect size = mean RSV ("med spa") – mean RSV ("dermatologist"); a negative value indicates that "dermatologist" queries had higher RSVs than "med spa" queries.

Discussion

Our findings suggest a substantial shift in public interest toward med spas over board-certified dermatologists for certain popular cosmetic procedures. Although increased RSVs do not necessarily equate to higher procedure volumes at med spas, prior analyses demonstrated reliable correlations between search trends and real-world demand for botulinum toxin injections

XSL•FO
RenderX

and hyaluronic acid fillers [2]. Seasonality in "sunscreen" search trends has further demonstrated GT's utility as a surrogate marker for population-level behaviors relevant to public health campaigns [3]. Together, these examples support the utility of search data as early signals of shifting patient preferences and potential emerging risks. Notably, even microneedling, which showed the smallest effect size among med spa–favoring procedures, reached statistical significance, suggesting broad rather than selective shifts in public preferences.

These trends are especially concerning, given persistent gaps in med spa regulatory oversight. Only 6.5% of surveyed US med spas reported on-site medical directorship, and fewer than half of listed directors had training in dermatology or plastic surgery [4]. State regulations governing med spa ownership, physician oversight, and nonphysicians' scope of practice remain highly variable [5]. Cosmetic procedures performed by nonphysician operators have been associated with higher rates of adverse events when compared to physician-performed treatments (73.3% vs 51.8%), yet over 70% of patients believe nonphysician providers are adequately qualified [5]. Earlier surveys similarly linked complications from laser and resurfacing procedures to nonphysician operators, with dermatologists frequently managing the resulting adverse events [6]. As access to cosmetic services expands outside traditional medical settings, the risk of harm may increase.

This study has several limitations inherent to GT data. RSVs are normalized within regions and reflect relative Google search interest rather than absolute public demand, precluding direct comparisons across countries. GT may also underrepresent low-volume but clinically meaningful terms, and sampled results can vary across repeated queries. In a systematic review, only 7% of GT-based health studies reported sufficient methodological detail for replication [7]. RSVs may also be influenced by seasonality, as dermatologic search topics have been shown to peak during warm months, and by media-driven events producing short-lived spikes in search volumes [8,9]. Such variation could affect the apparent magnitude of differences between "med spa" and "dermatologist" queries. GT data do not account for demographic differences in internet use, may overrepresent younger populations, and do not capture data from other search engines [6,9,10].

Public search interest in med spas now outpaces that in dermatologists for several in-demand cosmetic procedures, highlighting a growing disconnect between patient perceptions and professional concerns about training and oversight. The increasing online visibility of med spas underscores the need for stronger patient and policy education regarding provider qualifications and procedural safety. Dermatologists and professional societies should prioritize accessible, evidence-based resources that help patients distinguish board-certified providers from nonphysician practitioners and understand the risks of inadequate oversight. Our findings also support ongoing advocacy for standardized credentialing requirements, physician oversight of injectable and laser-based procedures, and transparent adverse event reporting. Future research should integrate clinical outcome data, including complication rates from dermatology practices and emergency departments, with search trend analyses to more fully quantify the impact of these shifting patterns on patient safety and to guide evidence-based health policy.

Acknowledgments

Artificial intelligence was not used in manuscript composition.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

Conceptualization: HY (lead), GHB (equal)
Data curation: HY
Formal analysis: HY (lead), GHB (supporting), SYZ (supporting), BV (supporting)
Visualization: HY (lead)
Writing – original draft: HY (lead), GHB (supporting)
Writing – review & editing: HY (lead), GHB (supporting), SYZ (supporting), BV (supporting), JX (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms used in analyses.
[DOCX File, 28 KB - derma_v9i1e83509_app1.docx]

References

XSL•FO
RenderX

1. Maisel A, Waldman A, Poon E, Alam M. Types of cosmetic procedures requested by different types of patients and the reasons for these preferences. Dermatol Surg 2020 Dec;46(12):1728-1732. [doi: 10.1097/DSS.0000000000002632] [Medline: 32852422]
2. Motosko C, Zakhem G, Ho R, Saadeh P, Hazen A. Using Google to trend patient interest in botulinum toxin and hyaluronic acid fillers. J Drugs Dermatol 2018 Nov 1;17(11):1245-1246. [Medline: 30500150]
3. Hopkins ZH, Secrest AM. Public health implications of Google searches for sunscreen, sunburn, skin cancer, and melanoma in the United States. Am J Health Promot 2019 May;33(4):611-615. [doi: 10.1177/0890117118811754] [Medline: 30428681]
4. Gibson JF, Srivastava D, Nijhawan RI. Medical oversight and scope of practice of medical spas (med-spas). Dermatol Surg 2019 Apr;45(4):581-587. [doi: 10.1097/DSS.0000000000001719] [Medline: 30829773]
5. Rabin DA, Chaudhry AS, Sumra H, et al. Emerging legal risks in medical spa procedures: insights from 20 malpractice cases. Ann Plast Surg 2026 Jul 1;97(1):7-16. [doi: 10.1097/SAP.0000000000004757] [Medline: 42048538]
6. Brody HJ, Geronemus RG, Farris PK. Beauty versus medicine: the nonphysician practice of dermatologic surgery. Dermatol Surg 2003 Apr;29(4):319-324. [doi: 10.1046/j.1524-4725.2003.29088.x] [Medline: 12656807]
7. Kamiński M, Tizek L, Zink A. 'Dr. Google, what is that on my skin?'-internet searches related to skin problems: Google Trends data from 2004 to 2019. Int J Environ Res Public Health 2021 Mar 4;18(5):2541. [doi: 10.3390/ijerph18052541] [Medline: 33806391]
8. Tijerina JD, Morrison SD, Nolan IT, Parham MJ, Richardson MT, Nazerali R. Celebrity influence affecting public interest in plastic surgery procedures: Google Trends analysis. Aesthetic Plast Surg 2019 Dec;43(6):1669-1680. [doi: 10.1007/s00266-019-01466-7] [Medline: 31392394]
9. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the internet as a resource for health information: results from a California population-based study. Int J Med Inform 2017 Jul;103:49-54. [doi: 10.1016/j.ijmedinf.2017.04.008] [Medline: 28551001]
10. Milinovich GJ, Williams GM, Clements ACA, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. Lancet Infect Dis 2014 Feb;14(2):160-168. [doi: 10.1016/S1473-3099(13)70244-5] [Medline: 24290841]

Abbreviations

GT: Google Trends
HIPAA: Health Insurance Portability and Accountability Act
med spa: medical spa
RSV: relative search volume

Edited by R Dellavalle; submitted 03.Sep.2025; peer-reviewed by FA Dhabbari, K Kesty, MA Zoubi; revised version received 14.May.2026; accepted 22.May.2026; published 15.Jul.2026.

Please cite as:
Ye H, Zhu SY, Verhey B, Xu J, Bae GH
Shifting Perceptions of Cosmetic Procedures: Rise of Medical Spas in Online Search Trends
JMIR Dermatol 2026;9:e83509
URL: https://derma.jmir.org/2026/1/e83509
doi:10.2196/83509

© Hong Ye, Sophia Yanjia Zhu, Britta Verhey, Joy Xu, Gordon Hyeonjin Bae. Originally published in JMIR Dermatology (http://derma.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Interest in Sun Protection and Misinformation: A Google Trends and TikTok Analysis

Rajyk Bhala, BA; Sriram Ambadi, BS; Reid Oldenburg, MD, PhD

Department of Dermatology, School of Medicine, University of California San Diego, 8899 University Center Lane, Suite 350, San Diego, CA, United States

Corresponding Author:
Rajyk Bhala, BA
Department of Dermatology, School of Medicine, University of California San Diego, 8899 University Center Lane, Suite 350, San Diego, CA, United States

Abstract

In this analysis of Google Trends and TikTok content, public interest in anti-sun protection topics increased in recent years, while non–physician-created content demonstrated similar engagement but lower quality and viewer experience than physician-created content, highlighting a mismatch between public interest and information accuracy.

(JMIR Dermatol 2026;9:e99635) doi:10.2196/99635

KEYWORDS

social media; health information; misinformation; TikTok; sunscreen; Google Trends; sun protection; dermatology

Introduction

Social media platforms such as TikTok have become influential sources of dermatologic health information and misinformation [1]. Increasing health misinformation has raised concerns about the reliability of online health information [2]. Google search volume reflects information-seeking behavior, whereas TikTok provides algorithm-driven health information. Together, these platforms provide complementary perspectives on public engagement with sun protection topics and the quality of related information on TikTok.

Methods

Overview

We conducted a cross-sectional analysis by choosing the top 30 TikTok videos listed under each of the 3 terms, “sun exposure,” “sunscreen toxic,” and “sunblock,” categorized as neutral, anti-sun protection, and pro-sun protection, respectively. A naïve TikTok account with no previous viewing history was used to minimize the algorithm’s personalization of search results and better reflect content encountered by users. Videos were included if they were relevant to the search term, in English, contained commentary, and were uploaded between May 25, 2021, and January 10, 2026. Google search trends were quantified using the Glimpse extension to obtain worldwide and United States absolute search volume data for these terms from 2020 to 2025 [3,4]. Mean annual search volume was calculated to reduce seasonal variation in search trends, and percent change was determined using 2020 as baseline and 2025 as endpoint [4]. Public interest was operationalized as Google search volume for the chosen terms [4]. Viewer engagement scores were calculated as (likes + comments + shares)/views, and creators were classified as physicians or non-physicians based off creators’ self-classification as an MD or DO in their profile [1]. Physician status was not used as a proxy for video quality or viewer experience. Video quality was determined using DISCERN (1=extensive shortcomings to 5=minimal shortcomings) and viewer experience using the Armstrong Viewer Assessment (AVA; 0=poor to 4=very good) by 2 independent evaluators [1,5]. Both evaluators independently assessed each video using predefined criteria. Discrepancies were resolved through discussion and consensus prior to final analysis. DISCERN assesses the quality and reliability of consumer health information and has been widely applied to social media content [5]. The AVA assesses viewer experience and educational value of dermatological online content [1]. Two-tailed *t* tests compared mean DISCERN and AVA scores between physician and non-physician creators, with statistical significance set at .05.

Ethical Considerations

This study used publicly available, deidentified data and was exempt from institutional board review.

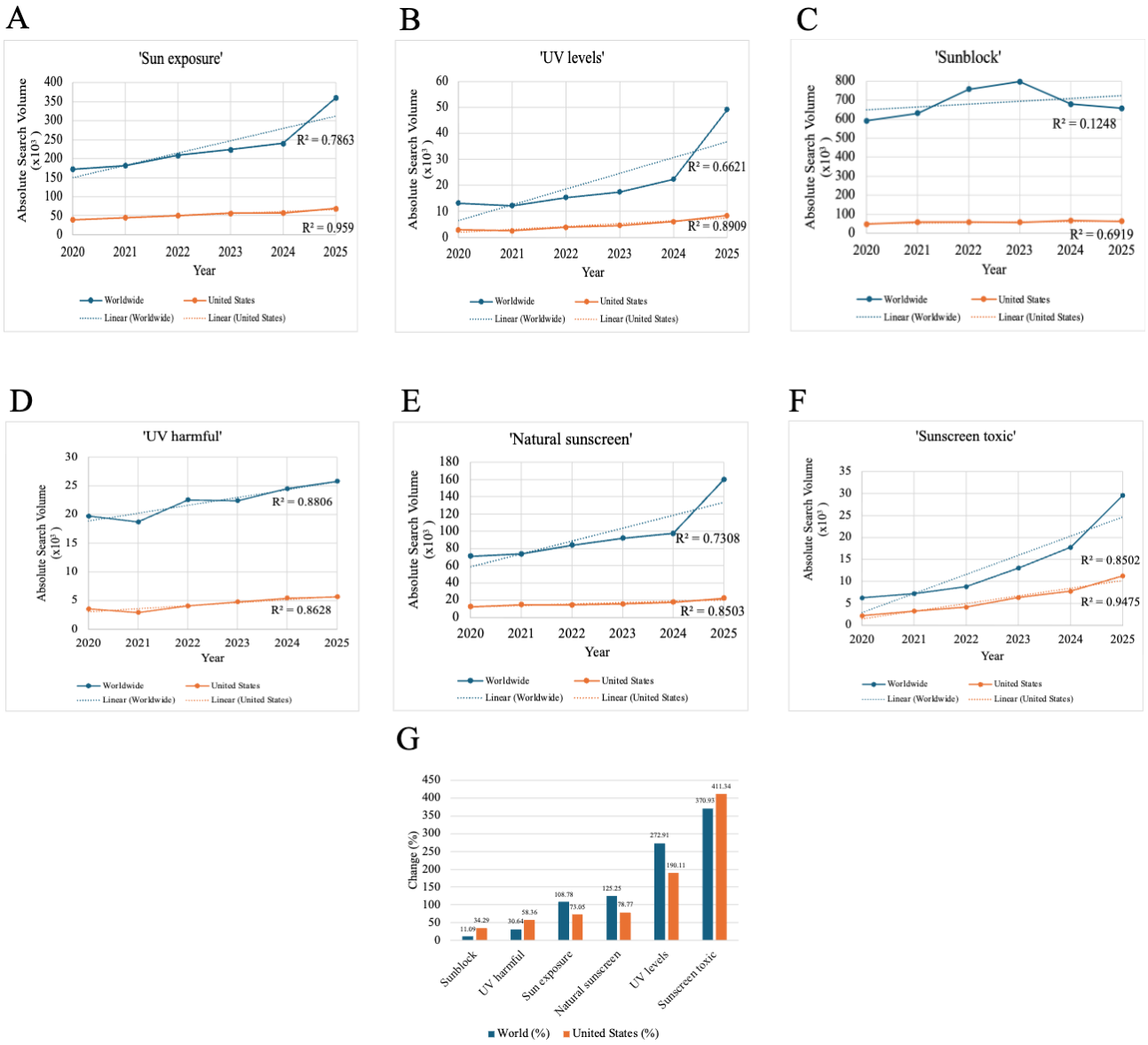
Results

Google search volume for anti-sun protection terms like “natural sunscreen” (United States: 1.78-fold) and “sunscreen toxic” (United States: 5.1-fold) showed the greatest growth (Figure 1). Neutral search terms like “sun exposure” (United States: 1.73-fold) and “UV levels” (United States: 2.9-fold) revealed moderate growth, whereas pro-sun protection terms reflected the smallest increases by comparison, including “sunblock” (United States: 1.34-fold) and “UV harmful” (United States:

XSL•FO
RenderX

1.6-fold) (Figure 1). Calculated viewer engagement scores did not differ significantly across all 3 terms (sun exposure: 0.12. SD 0.23 vs 0.06. SD 0.02, *P*=.22; sunscreen toxic: 0.06, SD 0.04 vs 0.07. SD 0.04, *P*=.70; sunblock: 0.06, SD 0.04 vs 0.04, SD 0.03, *P*=.10; Table 1). Physician-produced videos achieved higher DISCERN scores across all terms, indicating higher quality (sun exposure: 2.68, SD 2.39 vs 4.10, SD 1.07, *P*<.001; sunscreen toxic: 1.25, SD 0.29 vs 2.75, SD 1.29, *P*=.04; sunblock: 2.68, SD 1.31 vs 4.25, SD 0.38, *P*<.001; Table 1). Physician-produced videos achieved higher AVA scores, indicating better viewer experience (sun exposure: 1.38, SD 1.13 vs 3.10, SD 1.07, *P*<.001; sunscreen toxic: 0.25, SD 0.29 vs 1.75, SD 1.29, *P*=.04; sunblock: 1.68, SD 1.31 vs 3.25, SD 0.38, *P*<.001; Table 1).

Figure 1. Trends in public interest in sun protection search terms from 2020 to 2025. Panels A–F show annual absolute search volume for the terms "sun exposure," "UV levels," "sunblock," "UV harmful," "natural sunscreen," and "sunscreen toxic" using Glimpse extension data (accessed February 8, 2025). Panel G summarizes the percent (%) change in absolute search volume between 2020 and 2025. Values are shown for worldwide (blue) and United States (orange) search interest. Absolute search volumes were averaged by year to reduce seasonal variation and obtained using the Glimpse extension.



XSL•FO
RenderX

Table . Engagement and quality metrics across sun protection search terms. Viewer engagement, DISCERN quality scores, and Armstrong Viewer Assessment (AVA) scores are reported as mean (SD) for videos associated with the search terms “sun exposure,” “sunscreen toxic,” and “sunblock,” stratified by physician status.

	Videos, n	Engagement, mean (SD)	DISCERN, mean (SD)	AVA, mean (SD)
Sun exposure (n=30)				
Non-physician	20	0.12 (0.23)	2.68 (2.39)	1.38 (1.13)
Physician	10	0.06 (0.02)	4.10 (1.07)	3.10 (1.07)
Sunscreen toxic (n=30)				
Non-physician	24	0.06 (0.04)	1.25 (0.29)	0.25 (0.29)
Physician	6	0.07 (0.04)	2.75 (1.29)	1.75 (1.29)
Sunblock (n=30)				
Non-physician	22	0.06 (0.04)	2.68 (1.31)	1.68 (1.31)
Physician	8	0.04 (0.03)	4.25 (0.38)	3.25 (0.38)

Discussion

This study uses Google search volume data to measure public interest in sun protection topics [6]. Concurrently, content related to these topics on TikTok may not be accurate or high-quality. Lower DISCERN scores do not necessarily indicate misinformation, but they may reflect incomplete, poorly sourced, or less reliable health information.

These findings reveal a gap between rising public interest in sun protection topics and the quality of readily accessible information. These results are limited by the subjective nature of the DISCERN and AVA methods and the limited sample size. Dermatologists and health care professionals may benefit from addressing misconceptions during clinical encounters, given widespread engagement with sun protection content on social media. Higher-quality, evidence-based content on social media may support informed health decision-making.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

RB and SA performed data analysis.
RB and RO contributed to manuscript writing.
RB, SA, and RO revised the manuscript and approved the final version.

References

1. Khan S, Reddy R, Maynard N, et al. Tips, trends, and truths: a study of psoriasis treatment content on TikTok. J Drugs Dermatol 2024 Feb 1;23(2):e67-e69. [doi: 10.36849/JDD.7050] [Medline: 38306133]
2. Rhodes HG, Taylor TE, Viswanath K. Understanding and Addressing Misinformation about Science, 1st edition: National Academies Press; 2025.
3. Glimpse. URL: https://meetglimpse.com [accessed 2026-03-03]
4. Correia E, Mandel J, Cullison SRJ. Analyzing social media trends in cosmeceuticals: insights from Google Trends and TikTok analytics. J Cosmet Dermatol 2025 Apr;24(4):e70172. [doi: 10.1111/jocd.70172] [Medline: 40198513]
5. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
6. Southwell BG, Dolina S, Jimenez-Magdaleno K, Squiers LB, Kelly BJ. Zika virus-related news coverage and online behavior, United States, Guatemala, and Brazil. Emerg Infect Dis 2016 Jul;22(7):1320-1321. [doi: 10.3201/eid2207.160415] [Medline: 27100826]

Abbreviations

AVA: Armstrong Viewer Assessment

XSL•FO
RenderX

Edited by R Dellavalle; submitted 27.Apr.2026; peer-reviewed by A Moreno, B Southwell, N Morena; revised version received 08.Jul.2026; accepted 10.Jul.2026; published 23.Jul.2026.

Please cite as:
Bhala R, Ambadi S, Oldenburg R
Public Interest in Sun Protection and Misinformation: A Google Trends and TikTok Analysis
JMIR Dermatol 2026;9:e99635
URL: https://derma.jmir.org/2026/1/e99635
doi:10.2196/99635

© Rajyk Bhala, Sriram Ambadi, Reid Oldenburg. Originally published in JMIR Dermatology (http://derma.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Differences in Electronic Consultation Conversion Rates Between Advanced Practice Providers and Board-Certified Dermatologists

Dakota Hitchcock[1*], MD; Sabrina Newman[2*], MD

[1]Department of Internal Medicine, Saint Joseph Hospital, 1960 Ogden Street Suite 400, Denver, CO, United States

[2]Department of Dermatology, University of Colorado School of Medicine, Department of Dermatology, Aurora, CO, United States

[*]all authors contributed equally

Corresponding Author:
Dakota Hitchcock, MD
Department of Internal Medicine, Saint Joseph Hospital, 1960 Ogden Street Suite 400, Denver, CO, United States

Abstract

In this analysis of dermatology e-consults at a large academic health system, advanced practice providers had nearly threefold higher conversion rates to in-person visits compared to board-certified dermatologists, with potential implications for access and resource utilization.

(JMIR Dermatol 2026;9:e83922) doi:10.2196/83922

KEYWORDS

e-consult; telehealth; dermatology; advanced practice provider; teledermatology

Introduction

Electronic consultations (e-consults) have become an increasingly valuable tool in improving access to specialty care, reducing unnecessary in-person referrals, and supporting timely management of patients by primary care providers [1,2]. By allowing clinicians to consult with specialists asynchronously through the electronic health record, e-consults can help streamline workflows, decrease wait times, and conserve specialist resources [2,3]. Dermatology services receive a high number of e-consult requests, likely due to the visual diagnostic nature of the specialty [3,4]. As the use of e-consults expands across health care systems, understanding how different provider types use this tool, particularly in high-demand specialties such as dermatology, is critical to optimizing efficiency and effectiveness. Furthermore, identifying whether variations in conversion patterns reflect provider-level practice differences or system-level routing processes is essential for ensuring that e-consults function as intended.

Methods

We conducted a retrospective analysis to evaluate whether e-consult conversion rates differed by provider type, specifically comparing advanced practice providers (APPs), including nurse practitioners and physician assistants, to board-certified dermatologists. e-consult data specific to dermatology were extracted from the University of Colorado Hospital electronic health record system for the period of January 2020 to April 2025. An e-consult was considered “converted” if it resulted in a subsequent in-person specialist visit or full referral, rather than being resolved entirely through asynchronous communication.

In this system, e-consults are routed to APPs versus dermatologists primarily based on provider availability rather than consult content or patient acuity. As a result, patients evaluated by APPs and physicians likely represent comparable clinical populations, reducing the likelihood that differences in conversion rates were driven by systematic triage of more complex cases to one provider group.

Results

A total of 2572 dermatology e-consults were submitted during the study period. Of these, 1205 were addressed by APPs, with 321 (26.6%) resulting in conversion to an in-person visit (Table 1). In contrast, only 125 of the 1367 e-consults addressed by physicians (9.1%) were converted (Table 2). e-consults managed by APPs were nearly three times more likely to lead to an in-person referral compared to those managed by physicians.

XSL•FO
RenderX

Table . Total number and percentage of e-consults converted from advanced practice professionals.

e-consult converted	N (%)
No	884 (73.4)
Yes	321 (26.6)
Total	1205 (100.0)

Table . Total number and percentage of e-consults converted from dermatologists.

e-consult converted	N (%)
No	1242 (90.9)
Yes	125 (9.1)
Total	1367 (100.0)

Discussion

This analysis reveals a notable difference in e-consult conversion rates between APPs and physicians. This disparity suggests potential differences in how each provider group approaches triage and decision-making in specialty care. If APP-handled e-consults were converted at the same rate as physician-handled e-consults, over 200 additional dermatology clinic appointments during the study period may have been available for patients with higher-acuity needs. Despite this variation in appointment conversion, it is important to note that the majority of e-consults from both groups were resolved without the need for in-person follow-up. This reinforces the broader value of e-consults in improving efficiency and reducing unnecessary specialist visits and aligns with current literature [2,3].

The higher conversion rate observed among APPs may reflect a range of underlying factors. One possibility is that APPs may be more likely to convert e-consults conservatively due to comparatively less specialty-specific training or comfort managing complex cases. Importantly, in our system, APPs and dermatologists receive e-consults based largely on provider availability rather than clinical complexity. This reduces the likelihood that differences in patient or case characteristics explain the observed variation. Existing literature on provider-level differences in e-consult use and impact have shown mixed results. For example, one study comparing e-consults submitted by nurse practitioners and family physicians found that nurse practitioners were more likely to report that the consultation led to new clinical guidance and less likely to say it avoided an in-person referral [5]. In contrast, a systematic review of referral practices found no significant difference in overall referral rates between nurse practitioners and family physicians [6]. However, these studies largely examine differences among referring providers rather than responding providers. Because our study investigates variation among the providers performing the e-consults themselves, it represents a novel contribution to the literature. To our knowledge, no published studies have specifically examined provider-level variation in dermatology e-consult outcomes from the specialist side, underscoring the importance of our findings.

While our findings shed light on differences in provider behavior, they also raise questions about the clinical appropriateness of these conversions. Without detailed outcome data, it remains unclear whether the higher conversion rate among APPs were clinically necessary or reflective of a lower threshold for referral. Future research should explore the clinical drivers and downstream outcomes of converted e-consults, considering patient complexity, consult content, and specialty-specific considerations.

In addition to clinical impact, the higher conversion rate among APPs may have broader implications for system efficiency and resource use. Given the higher conversion rate, APP-managed e-consults may increase health care utilization, with potential cost implications for patients and health systems. Assuming a standard new patient visit billed at a level 3 or level 4 (estimated reimbursement US $120–$180 per visit), the additional ~200 appointments potentially consumed due to higher APP conversion rates translates to an estimated US $24,000–$36,000 in additional health care costs during the study period. Future work could further investigate whether these conversions lead to improved outcomes or represent avoidable costs.

This study contributes to the growing body of literature on e-consult optimization and provider practice variation. As health systems increasingly adopt team-based models of care and integrate APPs more fully into specialty workflows, ensuring consistent and effective use of e-consults across provider types will be essential. Implementing structured guidance, standardized triage protocols, and targeted training modules, particularly for APPs, may help promote more consistent decision-making and appropriate referral thresholds. Additionally, health systems may consider establishing limitations or clinical guidelines regarding the types of dermatologic conditions appropriate for independent APP e-consult management to ensure high-quality care, reduce unnecessary referrals, and minimize avoidable health care costs. By equipping all members of the care team with the tools and guidance needed to manage e-consults effectively, we can improve access, preserve specialist capacity, and enhance the overall efficiency of care delivery.

XSL•FO
RenderX

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

Conceptualization: DH, SN
Data curation: SN
Formal analysis: DH, SN
Investigation: DH
Methodology: DH, SN
Supervision: SN
Validation: DH, SN
Writing – original draft: DH, SN
Writing – review & editing: DH, SN

Conflicts of Interest

None declared.

References

1. Vimalananda VG, Orlander JD, Afable MK, et al. Electronic consultations (E-consults) and their outcomes: a systematic review. J Am Med Inform Assoc 2020 Mar 1;27(3):471-479. [doi: 10.1093/jamia/ocz185] [Medline: 31621847]
2. Vimalananda VG, Gupte G, Seraj SM, et al. Electronic consultations (e-consults) to improve access to specialty care: a systematic review and narrative synthesis. J Telemed Telecare 2015 Sep;21(6):323-330. [doi: 10.1177/1357633X15582108] [Medline: 25995331]
3. Keely E, Liddy C, Afkham A. Utilization, benefits, and impact of an e-consultation service across diverse specialties and primary care providers. Telemed J E Health 2013 Oct;19(10):733-738. [doi: 10.1089/tmj.2013.0007] [Medline: 23980939]
4. Bradley C, Smith L, Youens K, White BAA, Couchman G. Formalizing the curbside: digitally enhancing access to specialty care. Proc (Bayl Univ Med Cent) 2023;36(6):716-720. [doi: 10.1080/08998280.2023.2240364] [Medline: 37829223]
5. Liddy C, Deri Armstrong C, McKellips F, Keely E. A comparison of referral patterns to a multispecialty eConsultation service between nurse practitioners and family physicians: The case for eConsult. J Am Assoc Nurse Pract 2016 Mar;28(3):144-150. [doi: 10.1002/2327-6924.12266] [Medline: 25965249]
6. Laurant M, Reeves D, Hermens R, Braspenning J, Grol R, Sibbald B. Substitution of doctors by nurses in primary care. Cochrane Database Syst Rev 2005 Apr 18(2):CD001271. [doi: 10.1002/14651858.CD001271.pub2] [Medline: 15846614]

Abbreviations

APP: advanced practice provider
E-consults: electronic consultations

Edited by R Dellavalle; submitted 10.Sep.2025; peer-reviewed by ES Temin, P Aggarwal; revised version received 30.Dec.2025; accepted 06.Jan.2026; published 30.Jan.2026.

Please cite as:
Hitchcock D, Newman S
Differences in Electronic Consultation Conversion Rates Between Advanced Practice Providers and Board-Certified Dermatologists
JMIR Dermatol 2026;9:e83922
URL: https://derma.jmir.org/2026/1/e83922
doi:10.2196/83922

© Dakota Hitchcock, Sabrina Newman. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analysis of Melanoma Educational Content and Skin of Color Representation on TikTok: Cross-Sectional Study

Daniella Jaguan[1,2], BS; Zoya Siddiqui[1,3], BA; Jennifer Nam Choi[1], MD

[1]Department of Dermatology, Feinberg School of Medicine, Northwestern University, 676 N St Clair St #1600, Chicago, IL, United States

[2]School of Medicine, Georgetown University, Washington, DC, United States

[3]School of Medicine, University of Missouri–Kansas City, Kansas City, MO, United States

Corresponding Author:
Jennifer Nam Choi, MD
Department of Dermatology, Feinberg School of Medicine, Northwestern University, 676 N St Clair St #1600, Chicago, IL, United States

Abstract

Melanoma-related TikTok content lacks adequate skin of color representation and comprehensive educational content. This study found that engagement differed by content type, with response-style videos performing best and educational videos performing worst; in multivariate analysis, patient-created videos were also associated with higher engagement, while educational depth was not associated with engagement, suggesting that format, relatability, and presentation style may influence audience engagement more than educational depth, highlighting potential opportunities to improve the reach and inclusivity of melanoma-related health communication.

(JMIR Dermatol 2026;9:e87064) doi:10.2196/87064

KEYWORDS

melanoma; health education; skin of color; social media; health communication

Introduction

Background

The global incidence of melanoma remains high, underscoring the need for improved public health messaging and early detection [1]. Social media has become a major source of health-related information, with 82.4% of individuals seeking medical information online or through social media, and 16.4% reporting it as their primary source [2]. TikTok, one of the fastest-growing platforms, hosts over 1 billion monthly users who spend roughly 95 minutes daily on the platform [3]. Studies show that most dermatology content on TikTok is created by nondermatologists, raising concerns about accuracy [4,5]. Furthermore, there is a persistent gap in addressing skin of color (SOC) populations, a group disproportionately affected by delayed diagnosis and poorer outcomes [6-10].

Objective

Social media represents an opportunity to deliver comprehensive and accurate education to wide audiences with the potential to improve public health outcomes [11]. We evaluate the content, educational topic coverage, and SOC representation of highly engaged melanoma-related TikTok posts to identify opportunities to enhance equitable skin cancer education on social media.

Methods

Study Design and Data Collection

In this cross-sectional analysis, we searched the top 100 most popular public posts using the keyword “melanoma” across multiple time points between August 5 and 15, 2025, to capture highly visible content and reduce algorithm variability. Duplicate posts were removed. Posts were included if they were in English, publicly accessible, and contained a mention of melanoma diagnosis or signs, risks, prevention, treatment, or a personal story with a health message. All posts were reviewed independently by 2 reviewers (DJ and ZS).

Content Classification

Content was categorized as personal experience, patient experience, educational, research, or response. Author type was classified as dermatologist, other physician health care provider (HCP), nonphysician HCP, patient, influencer, or health information page or brand. Detailed category definitions are provided in Multimedia Appendix 1.

Assessment of Educational Topic Coverage

Educational topic coverage was assessed using a 6-point structured scoring framework (“educational depth”), awarding 1 point for each melanoma-related domain addressed: ABCDE (asymmetry, border irregularity, color variation, diameter, or evolution) criteria, diagnostic process, melanoma subtypes,

XSL•FO
RenderX

treatment, prevention, and risk factors. This exploratory measure was designed to quantify the number of predefined melanoma-related topics covered within each post.

Posts were also evaluated for misinformation, citations from credible sources, and inclusion of call-to-action language. Definitions for these variables are provided in Multimedia Appendix 1.

SOC Representation

SOC representation was assessed by reviewing Fitzpatrick skin types (FSTs) depicted in melanoma images and identifying any mention of SOC-related disparities in melanoma outcomes.

Engagement Metrics and Statistical Analysis

Total engagement was calculated as the sum of likes and comments for each post. Group comparisons were performed using chi-square and Kruskal-Wallis tests. Associations between educational depth and engagement were analyzed using Spearman correlation. Multivariable linear regression was used to evaluate associations of engagement with educational depth, content type, and author type. Additional methodological details are included in Multimedia Appendix 1.

Results

Author and Content Characteristics

Of the 100 most popular TikTok posts analyzed, the content was most frequently created by patients (n=41) and dermatologists (n=41). Nearly half were educational (n=47) or personal experiences (n=41). Post characteristics are shown in Table S1 in Multimedia Appendix 2.

Content and Educational Depth

Overall, educational depth was low, with most posts scoring 0 to 2 on the 6-point scale. Depth did not differ significantly between medical professionals (dermatologists, physicians, nonphysician HCPs; n=54) and nonmedical creators (patients, influencers, health information page or brands; n=46; χ^2_5=7.28; *P*=.20; Table S2 in Multimedia Appendix 2).

Patients primarily posted personal experiences (40/41, 97.6%), while dermatologists primarily produced educational content (35/41, 85.4%) and responses (4/41, 9.8%).

No overt misinformation was identified. No posts cited external references or sources. A call-to-action appeared in 59 posts.

Engagement

Total engagement differed significantly by content category (Kruskal-Wallis χ^2_4=11.65, *P*=.02). Educational posts had the lowest median total engagement, whereas response videos achieved the highest median engagement (Table 1). Response videos remained associated with higher engagement after adjusting for educational depth and author type (β=3.54; *P*<.001).

Engagement did not differ significantly by author type (Kruskal-Wallis χ^2_5=6.53; *P*=.37); however, patient-created videos were associated with significantly higher engagement after adjusting for educational depth and content type (β=6.81; *P*=.009).

Educational depth demonstrated a nonsignificant inverse correlation with engagement (Spearman ρ=–0.13; *P*=.19) and remained nonsignificant after adjusting for content and author type (β=–0.28; *P*=.11). Video duration was also not significantly associated with engagement (Spearman ρ=0.15; *P*=.15).

XSL•FO
RenderX

Table . Engagement metrics by author and content type among melanoma-related TikTok posts.

Author/content type		Posts, n	Likes, n	Comments, n	Engagement[a], n	Engagement, median (IQR)
Author type						
	Patient	41	2,416,972	18,759	2,435,731	6028 (2261 - 17456)
	Dermatologist	41	1,888,594	26,448	1,915,042	3124 (652 - 33,603)
	Other physician HCP[b]	10	317,074	4644	321,718	17,854 (3572.3 - 22,458)
	Influencer	2	237,400	5277	242,677	121,338.5 (69,382.3 - 173,294.8)
	Non-physician HCP[b]	3	6807	249	7056	1343 (789.5 - 3410)
	Brand/health information page	3	2053	335	2388	1776 (1058 - 10,374)
Content type						
	Educational	47	1,669,778	14,916	1,684,694	2598 (649.5 - 18,969)
	Personal experience	41	1,823,735	17,943	1,841,678	5748 (1970 - 15,286)
	Patient experience	7	361,487	7508	368,995	17,426 (4802.5 - 57,468.5)
	Response	4	902,000	11,307	913,307	227,215 (67,625.8 - 387,915.5)
	Research	1	130,500	4038	134,538	134,538 (—[c])

[a]Calculated as the sum of likes and comments for each post

[b]HCP: health care provider.

[c]Not applicable; IQR was not calculated as this category contained only a single post (n=1).

SOC Representation

FSTs 1 to 3 predominated (n=51 posts); only 3 posts showed types 4 to 6, and only 3 addressed SOC disparities. Dermatologists included FSTs 4 to 6 in only 2.4% of posts (1/41) and other physicians in 20% (2/10). No nonmedical creators featured skin types above 3. The 3 posts with SOC content made by dermatologists and other physicians had a median engagement of 18,000 (IQR 9097-41000) likes and a median 276 (148-613) comments.

All posts showing SOC were educational, with educational depth scores of 2 (2/3) and 4 (1/3). Topics included misconceptions that darker skin cannot develop melanoma or does not need sun protection, delayed diagnoses in darker skin tones, and melanoma subtypes common in SOC.

Discussion

Principal Findings

In this cross-sectional analysis of highly engaged-with melanoma-related TikTok posts, patients and dermatologists were the most frequent contributors; patients mainly shared personal experiences, while dermatologists primarily shared educational content. Coverage of melanoma-related educational domains was generally low.

Engagement varied significantly by content type, with educational posts having the lowest total median engagement and response videos the highest. In an adjusted analysis, patient-created videos and response-style content were associated with higher engagement, whereas educational depth was not. These findings suggest that among the highly engaged melanoma-related TikTok posts included in our sample, characteristics such as personal storytelling, relatability, and interactive presentation styles may attract more engagement than traditional fact-focused didactic content. However, these findings should be interpreted cautiously given the limited sample size of top-performing posts and the presence of potential confounders like creator popularity.

Only 3% (3/100) of highly engaged posts featured SOC representation. Although based on a very small sample, this observation is consistent with prior reports describing underrepresentation of darker skin tones in dermatologic educational resources [12]. Platform algorithms may also reinforce existing visibility gaps by amplifying higher-performing content.

XSL•FO
RenderX

Implications for Dermatology Communication on Social Media

Consistent with prior studies demonstrating limited visibility of board-certified dermatologists on TikTok, dermatologists in our cohort primarily produced educational videos, which generated lower engagement than other content types [4]. Response-style videos demonstrated the highest engagement in our cohort, and patient-authored videos were independently associated with higher engagement in a multivariable analysis. These findings suggest that interactive presentation styles and relatability can influence engagement. Further study is needed to investigate whether collaborations with patient creators or interactive features like video "stitching" can improve the reach of educational content while maintaining educational value. Moreover, given the lack of SOC representation observed in our sample, future content should incorporate diverse imagery and address disparities in melanoma diagnosis and outcomes in underrepresented populations.

Limitations

This study is limited by its cross-sectional design, small sample size, and potential selection bias related to inclusion of only the most highly engaging English-language posts identified through TikTok's algorithm. As a result, lower-engagement content may be underrepresented, limiting generalizability to the broader ecosystem of melanoma-related TikTok content. Lastly, we note that the health information page and brand category encompassed a heterogeneous group of nonindividual accounts, and grouping these accounts together may have obscured differences in content strategies and engagement patterns.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

DJ contributed to conceptualization, data curation, formal analysis, methodology, and visualization, and wrote the original draft. DJ, ZS, and JNC contributed to writing—review and editing. JNC contributed to conceptualization and methodology and provided supervision. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed study methodology.
[DOCX File, 20 KB - derma_v9i1e87064_app1.docx]

Multimedia Appendix 2
Supplementary tables including overall post characteristics and educational depth scores by author type.
[DOCX File, 20 KB - derma_v9i1e87064_app2.docx]

References

1. Zhou L, Zhong Y, Han L, Xie Y, Wan M. Global, regional, and national trends in the burden of melanoma and non-melanoma skin cancer: insights from the Global Burden of Disease Study 1990–2021. Sci Rep 2025;15(1):5996. [doi: 10.1038/s41598-025-90485-3]
2. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
3. Time spent on TikTok per day - 95 min on average in 2026. Blank Spaces. 2026. URL: https://www.blankspaces.app/blog/tiktok-screen-time-statistics [accessed 2026-09-03]
4. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
5. Nguyen M, Youssef R, Kwon A, Chen R, Park JH. Dermatology on TikTok: analysis of content and creators. Int J Womens Dermatol 2021 Sep;7(4):488-489. [doi: 10.1016/j.ijwd.2021.04.006] [Medline: 34621965]
6. Campbell J, Williams K, Woolery-Lloyd H. DermTok: how TikTok is changing the landscape of dermatology patient education. J Drugs Dermatol 2023 Mar 1;22(3):302-304. [doi: 10.36849/JDD.6676] [Medline: 36877875]
7. Fritsche M, King T, Hollins L. Skin of color in dermatology: analysis of content and engagement on TikTok. J Drugs Dermatol 2022 Jul 1;21(7):718-722. [doi: 10.36849/JDD.6784] [Medline: 35816063]
8. Villa-Ruiz C, Kassamali B, Mazori DR, Min M, Cobos G, LaChance A. Overview of TikTok's most viewed dermatologic content and assessment of its reliability. J Am Acad Dermatol 2021 Jul;85(1):273-274. [doi: 10.1016/j.jaad.2020.12.028] [Medline: 33359080]

XSL•FO
RenderX

9. Swearingen A, Olagun-Samuel C, Sikora M, Alomary S, Adotama P. Exploring the scope of melanoma educational content on Instagram. J Drugs Dermatol 2025 May 1;24(5):e34. [doi: 10.36849/JDD.8751] [Medline: 40327571]
10. Dawes SM, Tsai S, Gittleman H, Barnholtz-Sloan JS, Bordeaux JS. Racial disparities in melanoma survival. J Am Acad Dermatol 2016 Nov;75(5):983-991. [doi: 10.1016/j.jaad.2016.06.006] [Medline: 27476974]
11. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85. [doi: 10.2196/jmir.1933] [Medline: 23615206]
12. Syder NC, Omar D, McKenzie S, Brown-Korsah JB, Taylor SC, Elbuluk N. Gaps in medical education curricula on skin of color in medical school, residency, and beyond: part 1. J Am Acad Dermatol 2023 Nov;89(5):885-892. [doi: 10.1016/j.jaad.2022.03.053] [Medline: 35390429]

Abbreviations

FST: Fitzpatrick skin type
HCP: health care professional
SOC: skin of color

Edited by S Subhadarshani; submitted 09.Dec.2025; peer-reviewed by N Hu, Q Pang; revised version received 18.Jun.2026; accepted 04.Jul.2026; published 15.Sep.2026.

Please cite as:
Jaguan D, Siddiqui Z, Choi JN
Analysis of Melanoma Educational Content and Skin of Color Representation on TikTok: Cross-Sectional Study
JMIR Dermatol 2026;9:e87064
URL: https://derma.jmir.org/2026/1/e87064
doi:10.2196/87064

© Daniella Jaguan, Zoya Siddiqui, Jennifer Nam Choi. Originally published in JMIR Dermatology (http://derma.jmir.org), 15.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Interest in Janus Kinase (JAK) Inhibitors for Alopecia Areata: A Google Trend Analysis

Jade Howard[1], MD; Nourine A H Kamili[2], PhD; Hala Idris[2], BS; Loren D Krueger[2], MD

[1]Transitional Year Residency Program and BayCare Health System (Winter Haven), College of Medicine, Florida State University, 1201 1st St S, Suite 100A, Winter Haven, FL, United States

[2]Department of Dermatology, School of Medicine, Emory University, Atlanta, GA, United States

Corresponding Author:
Jade Howard, MD
Transitional Year Residency Program and BayCare Health System (Winter Haven), College of Medicine, Florida State University, 1201 1st St S, Suite 100A, Winter Haven, FL, United States

Abstract

Public interest in Janus kinase (JAK) inhibitors for alopecia areata increased following media coverage and the 2022 US Food and Drug Administration (FDA) approval of baricitinib, highlighting the need for patient education and physician guidance on appropriate indications and treatment selection for hair loss disorders.

(JMIR Dermatol 2026;9:e75119) doi:10.2196/75119

KEYWORDS

alopecia areata; hair loss; Janus kinase inhibitors; Google Trends; alopecia areata treatment; hair loss treatment

Introduction

Janus kinase (JAK) inhibitors are a class of medications that work by targeting and inhibiting JAK enzymes, which play a significant and diverse role in the immune system. The JAK system has been implicated in a variety of immune pathways including inflammation and autoimmunity [1]. On June 13, 2022, the US Food and Drug Administration (FDA) approved the use of a JAK inhibitor, baricitinib, commercially known as Olumiant, for adults with severe alopecia areata, an autoimmune form of hair loss [2]. Additionally, this medication was also approved for the use of severe alopecia areata in children aged 12 years and over on June 26, 2023. The objective of this study was to assess the impact of the FDA approval of the JAK inhibitor, baricitinib (Olumiant), for severe alopecia areata on public interest in JAK inhibitors for alopecia areata.

Methods

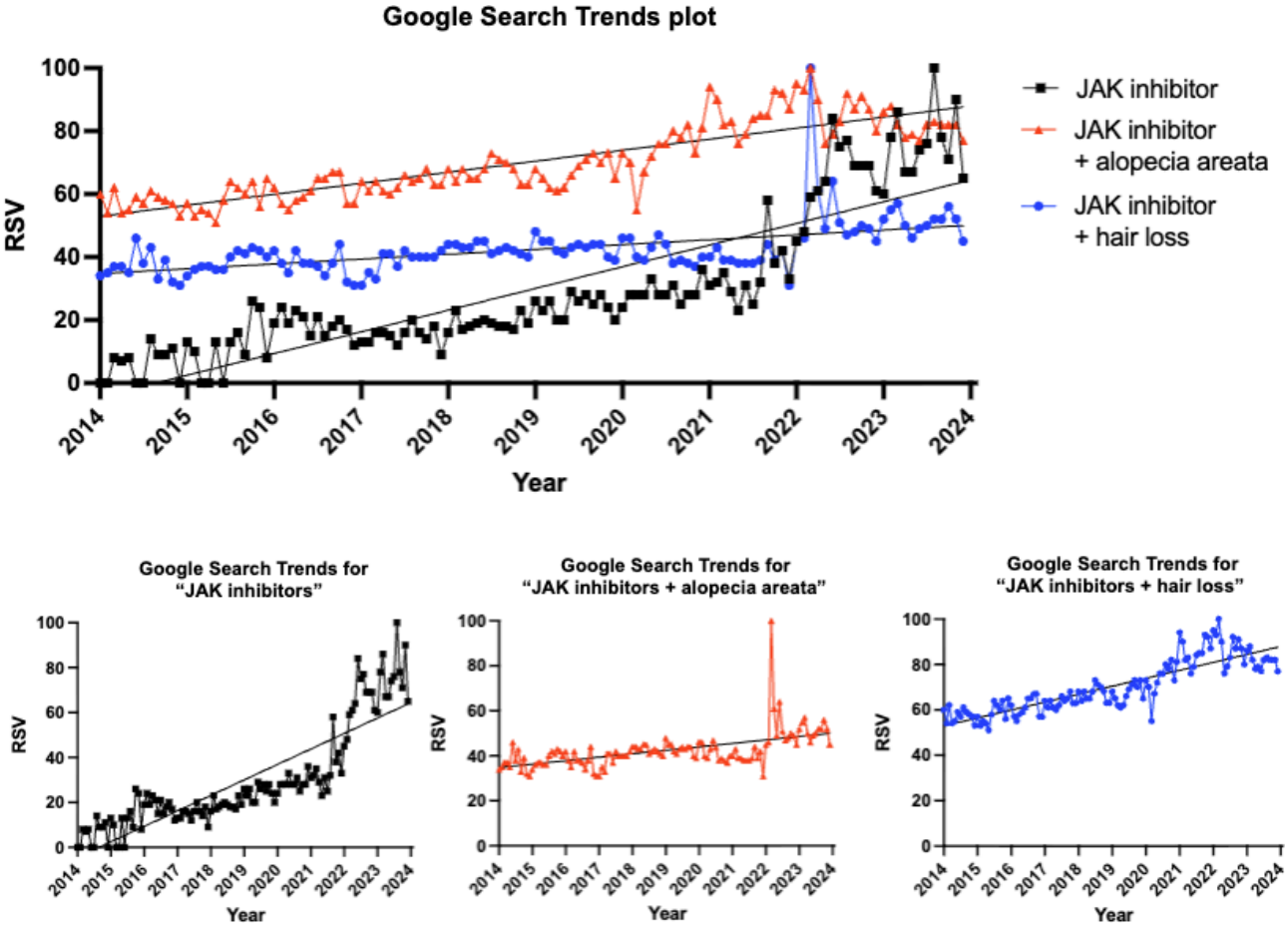
To evaluate public interest, we used Google Trends, a free analysis tool that provides insight into public interest for search terms. Google Trends calculates a relative search volume (RSV), ranging from 0 to 100. An RSV value of 0 indicates minimal interest in a keyword and a value of 100 represents the peak interest within a given period [3]. For our study, we searched the following keywords from January 2014 to January 2024: “JAK inhibitor,” “JAK inhibitor + hair loss,” and “JAK inhibitor + alopecia areata.” These keywords were chosen to encompass generic and specific terms related to the topic of JAK inhibitors and hair loss.

Results

A time series analysis was used to assess changes in RSV between January 2014 and December 2023. Linear regression demonstrated significant upward trends in RSV for all 3 search terms (Figure 1). Search interest for “JAK inhibitor” increased at a rate of 6.89 (95% CI 6.16 - 7.62; P<.001), RSV units per month with a strong model fit (R^2=0.748). Searches for “JAK inhibitor+ hair loss” also increased significantly, at a rate of 3.50 (95% CI 3.13 - 3.87; P<.001; R^2=0.749) RSV units per month. Similarly, “JAK inhibitor+ alopecia areata” demonstrated a statistically significant upward trend, though with a smaller magnitude of increase (1.54, 95% CI 1.12 - 1.96 RSV units per month; P<.001) and a more modest model fit (R^2=0.306).

XSL•FO
RenderX

Figure 1. Google Trends results for the search items "JAK inhibitor," "JAK inhibitor + alopecia areata," and "JAK inhibitor + hair loss" (data accessed April 30, 2024). Data are presented as relative search volume (RSV), where RSV of 100 represents peak search activity in a time period. The plot represents the time series analysis of Google Trends RSV per month between January 2014 and December 2023 for the terms "JAK inhibitor," "JAK inhibitor + alopecia areata," and "JAK inhibitor + hair loss." Search interest for "JAK inhibitor" increased at a rate of 6.89 (95% CI 6.16 - 7.62; P<.001; R^2=0.748) RSV units per month. Searches for "JAK inhibitor + hair loss" and "JAK inhibitor + alopecia areata" increased at a rate of 3.50 (95% CI 3.13 - 3.87; P<.001; R^2=0.749) RSV units per month and 1.54 (95% CI 1.12 - 1.96; P<.001; R^2=0.306) RSV units per month, respectively. JAK: Janus kinase.



Discussion

All 3 search terms ("JAK inhibitor," "JAK inhibitor + hair loss," "JAK inhibitor + alopecia areata") demonstrated significant increases over time (P<.001), indicating rising public interest from January 2014 to December 2023. Searches for "JAK inhibitor + hair loss" showed a strong and consistent upward trend (R^2≈0.75), while general searches for "JAK inhibitor" increased at the fastest rate, suggesting expanding overall awareness of the drug class. The increasing trend is not perfectly linear; rather, it reflects periods of spikes, drops, and plateaus. This variability is possibly influenced by social factors such as periods of media coverage, FDA approvals, etc. For example, there is an appreciable spike in all 3 search terms between 2020 - 2023, which may indicate increased public interest following increased media coverage about the FDA approval of JAK inhibitors for severe alopecia reported by news outlets as early as 2019 [4]. Baricitinib became the first FDA-approved JAK inhibitor for alopecia areata in 2022. This was followed by the approval of ritlecitinib in 2023 and deuruxolitinib in 2024 [5]. Additionally, social media applications like TikTok could be possible contributors to the increase in public interest, due to user-friendly content, the ability of content creators to freely express their journey with hair loss and JAK inhibitors, and users being able to engage in conversations with their peers anonymously. Although detailed TikTok data trends from 2014 - 2023 were not readily attainable, the TikTok Creator Search Insights tool, which provides trends for the past 6 months, indicates that more than 1 million searches were made on TikTok for JAK inhibitors and alopecia from September 2025 to February 2026. These searches yield a variety of videos, including content from patients describing their experience with JAK inhibitors. A more in-depth exploration of search trends and content on social media platforms would provide further valuable insight into the trends in public interest on this topic.

Given the increased public interest in JAK inhibitor treatment as indicated by Google Search Trends in this study, it is important that the public receives proper education regarding the implications of taking JAK inhibitors and the knowledge that JAK inhibitors are treatment for severe alopecia areata, which is an autoimmune condition, and not currently efficacious for other hair loss disorders. It is important for physicians to educate their patients on hair loss disorder treatments and increase patient awareness of certain hair loss treatments for the patient's form of hair loss.

XSL•FO
RenderX

Authors' Contributions

JH, NAHK, HI performed data analysis;
JH, NAHK, HI, LDK all contributed to manuscript preparation;
JH, NAHK, HI and LDK revised the manuscript, approved the final version, and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Ryguła I, Pikiewicz W, Kaminiów K. Novel Janus kinase inhibitors in the treatment of dermatologic conditions. Molecules 2023 Dec 13;28(24):8064. [doi: 10.3390/molecules28248064] [Medline: 38138551]
2. Freitas E, Guttman-Yassky E, Torres T. Baricitinib for the treatment of alopecia areata. Drugs (Abingdon Engl) 2023 Jun;83(9):761-770. [doi: 10.1007/s40265-023-01873-w] [Medline: 37195491]
3. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill 2019 May 29;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]
4. Abdelmalek M. New drug for alopecia shows promise: what you need to know. ABC News. 2019. URL: https://abcnews.com/GMA/Wellness/drug-alopecia-shows-promise/story?id=60128818&utm_source=chatgpt.com [accessed 2024-05-11]
5. Sanchez K, Englander H, Salloum L, et al. Evaluating current and emergent JAK inhibitors for alopecia areata: a narrative review. Dermatol Ther (Heidelb) 2025 Oct;15(10):2749-2764. [doi: 10.1007/s13555-025-01517-9] [Medline: 40794245]

Abbreviations

FDA: US Food and Drug Administration
JAK: Janus kinase
RSV: relative search volume

Edited by S Subhadarshani; submitted 01.Apr.2025; peer-reviewed by B Southwell; revised version received 02.Mar.2026; accepted 14.Mar.2026; published 24.Apr.2026.

Please cite as:
Howard J, Kamili NAH, Idris H, Krueger LD
Public Interest in Janus Kinase (JAK) Inhibitors for Alopecia Areata: A Google Trend Analysis
JMIR Dermatol 2026;9:e75119
URL: https://derma.jmir.org/2026/1/e75119
doi:10.2196/75119

© Jade Howard, Nourine A H Kamili, Hala Idris, Loren D Krueger. Originally published in JMIR Dermatology (http://derma.jmir.org), 24.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

What's Trending: Cross-Sectional Analysis of the Content and Quality of Popular Melanoma Videos on TikTok

Nazaneen Saleh[1], BA; Alexa Rahaman[2], BA; Emily Siegel[1], MD

[1]Rush Medical College, Rush University Medical Center, 1620 W Harrison St, Chicago, IL, United States
[2]School of Medicine and Health Sciences, George Washington University, Washington, DC, United States

Corresponding Author:
Nazaneen Saleh, BA
Rush Medical College, Rush University Medical Center, 1620 W Harrison St, Chicago, IL, United States

Abstract

In this cross-sectional analysis of popular melanoma-related TikTok videos, physician-created, educational, and longer videos were associated with higher information quality and audience engagement, although overall DISCERN scores remained low and treatment guidance was frequently absent, highlighting opportunities to improve comprehensive, physician-led melanoma education on social media.

(JMIR Dermatol 2026;9:e79250) doi:10.2196/79250

KEYWORDS

dermatology; melanoma; social media; TikTok; trends

Introduction

Background

As social media continues to integrate into daily life, platforms like TikTok, which now exceeds 1.8 billion users, create opportunities to efficiently share engaging health information [1]. While this can enhance health literacy, it also raises concerns about misinformation.

TikTok has become a popular platform for dermatology content, with both physicians and nonexperts sharing advice. Given the 42% rise in invasive melanoma cases in the United States from 73,870 cases in 2015 to 104,960 estimated cases in 2025, health literacy and early detection are increasingly important for improving survival [2-4]. The dermatology literature has gaps in analyzing the widespread utilization of TikTok videos for delivering educational health information. To the best of our knowledge, this is the first study examining melanoma-specific videos on TikTok.

Objective

This study aims to analyze the content and quality of popular melanoma-related videos on TikTok.

Methods

TikTok was queried using the hashtags #melanoma, #metastaticmelanoma, and #skincancer on two pre-existing accounts. In total, 32 videos were collected by two authors in September 2024 and independently evaluated by dermatology residents and attendings at Rush University (Chicago, United States) using the DISCERN tool, a validated measure for assessing consumer health information [5]. The DISCERN questionnaire is divided into 3 parts: reliability of the publication (questions 1-8), quality of information about treatment options (questions 9-15), and the overall score of the publication (question 16). Each item is rated on a scale from 1 (poor) to 5 (excellent). An adjusted DISCERN score was calculated by averaging only the reliability-related items (questions 1 - 8), since 28 of the 32 videos lacked treatment content. Only one video per TikTok creator was considered and every video was reviewed by both authors to ensure no repeated videos.

Results

The content creators of the 32 videos consisted of physicians (n=18, 56%), nonphysicians (n=11, 34%), private companies (n=2, 6%), and organizations (n=1, 3%) (Table 1). Among the 18 physician creators, the majority were dermatologists (13/18, 72%). The mean DISCERN score across all videos was 2.08, highlighting the overall need for improved quality in melanoma-related content on TikTok. When treatment-related questions were excluded, the adjusted DISCERN score rose to 2.68, indicating that guidance on treatment is frequently lacking.

XSL•FO
RenderX

Table . Summary of TikTok video characteristics, quality scores, and engagement metrics.

Category and subgroup	Number of videos, N=32 (%)	Mean DISCERN score	Mean adjusted DISCERN score	Mean number of likes	Mean number of saves	Mean number of shares
Content creator type						
Physician	18 (56)	2.3	2.98	17,499	2906	1525
Dermatology	13/18 (72)	2.14	2.87	17,753	2388	1428
Internal medicine	2/18 (11)	2.19	3.13	40,650	10,500	4351
Emergency medicine	1/18 (6)	3.69	4.38	1889	85	95
Oncology	1/18 (6)	2.44	2.5	661	167	83
Surgery	1/18 (6)	2.94	3.25	357	16	7
Nonphysician[a]	11 (34)	1.79	2.36	6515	1025	361
Private company[b]	2 (6)	2.19	2.44	8997	2729	1155
Nonprofit organization[c]	1 (3)	1.19	1.38	2518	994	165
Video length						
<1 minute	17 (53)	2.01	2.57	3973	768	313
>1 minute	15 (47)	2.15	2.81	22,642	3799	1906
Video type						
Anecdotal	7 (22)	1.72	2.13	5813	584	320
Educational	25 (78)	2.18	2.84	14,659	2638	1267

[a]Nonphysician: personal TikTok accounts run by any individuals who were not physicians.

[b]Private company: accounts promoting or selling products with a for-profit purpose.

[c]Nonprofit organization: organization affiliated with Oregon Health & Science University Dermatology (Start Seeing Melanoma).

Discussion

The analysis of content creators revealed that among the 18 physician creators, a significant majority (n=13, 72%) were dermatologists. This was expected as dermatology is the specialty with the most training in diagnosing and treating melanoma. However, the mean DISCERN score for videos created by dermatologists was 2.19, compared to 3.69 for emergency medicine physicians, indicating a general need for improved quality in melanoma-related content on TikTok regardless of physician training. Notably, when treatment-related questions were excluded, the adjusted DISCERN score increased from 2.08 across all videos to 2.68, suggesting that there is a frequent absence of guidance regarding treatment options in the content. Enhancing content with accurate, treatment-focused information may improve health literacy among viewers.

Physician-created videos outperformed those by nonphysicians in both quality (DISCERN: 2.3 vs 1.79, respectively; adjusted: 2.98 vs 2.36, respectively) and audience engagement metrics (mean number of likes: 17,499 vs 6515, respectively). This underscores the importance of professional involvement in online health education. Similarly, videos longer than one minute scored higher than shorter ones (DISCERN: 2.15 vs 2.01; adjusted: 2.81 vs 2.57), and educational videos outperformed anecdotal content (DISCERN: 2.18 vs 1.72; adjusted: 2.84 vs 2.13). Longer and educational videos also showed greater audience engagement. Together, these findings suggest that TikTok's short-form format does not inherently disadvantage content that prioritizes explanation and educational intent. Rather than implying viewer preference for longer videos, the observed associations indicate that informational clarity and depth can coexist with algorithm-driven visibility on the platform.

Limitations include the study's cross-sectional design, use of only three hashtags, platform-specific sample, and inability to verify creator credentials. Additionally, TikTok's proprietary search algorithm may affect the visibility and discoverability of videos among different users, introducing an element of unpredictability. Lastly, the DISCERN tool was originally developed to assess written health information and may not fully capture the unique elements of video-based content. Future research should explore tools tailored to assess video content and investigate strategies for optimizing physician-led education on social media platforms.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. TikTok says it has passed 1 billion users. The Verge. 2021. URL: https://www.theverge.com/2021/9/27/22696281/tiktok-1-billion-users [accessed 2024-10-05]
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA A Cancer J Clinicians 2015 Jan;65(1):5-29. [doi: 10.3322/caac.21254]
3. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45. [doi: 10.3322/caac.21871] [Medline: 39817679]
4. Jerant AF, Johnson JT, Sheridan CD, Caffrey TJ. Early detection and treatment of skin cancer. Am Fam Physician 2000 Jul 15;62(2):357-368. [Medline: 10929700]
5. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]

Edited by E Parker; submitted 17.Jun.2025; peer-reviewed by N Döring, S Bestavros; revised version received 28.Dec.2025; accepted 31.Dec.2025; published 25.Aug.2026.

Please cite as:
Saleh N, Rahaman A, Siegel E
What's Trending: Cross-Sectional Analysis of the Content and Quality of Popular Melanoma Videos on TikTok
JMIR Dermatol 2026;9:e79250
URL: https://derma.jmir.org/2026/1/e79250
doi:10.2196/79250

© Nazaneen Saleh, Alexa Rahaman, Emily Siegel. Originally published in JMIR Dermatology (http://derma.jmir.org), 25.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Asymptomatic Tattoo-Associated Cutaneous Nontuberculous Mycobacterial Infection: Case Report

Chris Alihosseini[1*], DO; Milana Stein[2*], DO; Ana Sofia Acosta Madiedo[1], MD; Roksana Hesari[1], DO; Matthew Kauffman[3], MS; Martin N Zaiac[4], MD

[1]Department of Dermatology, Larkin Community Hospital, Hialeah, FL, United States

[2]HCA Florida Westside Hospital, 8201 W Broward Blvd, Plantation, FL, United States

[3]Florida International University Herbert Wertheim College of Medicine, Miami, FL, United States

[4]Department of Dermatology, Herbert Wertheim College of Medicine, Miami, FL, United States

[*]these authors contributed equally

Corresponding Author:
Milana Stein, DO
HCA Florida Westside Hospital, 8201 W Broward Blvd, Plantation, FL, United States

Abstract

Tattoo-associated nontuberculous mycobacterial infections can present in a variety of ways, including papular, pustular, nodular, and abscess-like morphologies. We report the case of an immunocompetent 28-year-old male who presented with diffusely distributed monomorphic papules with collarettes of scale confined to recently tattooed skin. Histopathologic evaluation demonstrated an atypical lichenoid vacuolar interface dermatitis, and tissue culture ultimately confirmed infection with *Mycobacterium abscessus* complex following prolonged incubation. This case highlights the diagnostic challenges posed by atypical clinical and histopathologic features of tattoo-associated nontuberculous mycobacterial infections, emphasizing the importance of early biopsy, mycobacterial culture, and prolonged combination antimicrobial therapy for optimal patient outcomes.

(JMIR Dermatol 2026;9:e108587) doi:10.2196/108587

KEYWORDS

dermis; epidermis; microbiome; mycobacteria; tattoo-associated infection

Introduction

Tattoo-associated cutaneous reactions are increasingly encountered in dermatologic practice and are most commonly inflammatory or allergic in nature [1]. However, infectious complications must be considered in atypical cases. Nontuberculous mycobacteria (NTM) are environmental organisms capable of causing skin and soft tissue infections through direct inoculation, including during tattooing. Rapidly growing mycobacteria, such as *Mycobacterium abscessus (M abscessus)*, are being frequently implicated in tattoo-related cases. A 2026 systematic review identified 311 individuals across 15 countries with tattoo-associated NTM infection, with 29 of the 67 infection events being in the United States [2]. More specifically, in a 2015 Miami-Dade County outbreak, 38 of 226 interviewed tattoo clients were identified as outbreak-associated cases of *M abscessus* from contaminated graywash ink and tap water [3]. Exposure to gray tattoo ink was associated with an increased likelihood of developing a rash. These infections may be underrecognized due to their indolent course, variable morphology, and delayed culture growth. Awareness of tattoo-associated NTM infections is critical for timely diagnosis and appropriate management, as misdiagnosis can result in delayed appropriate therapy, unnecessary treatments, and prolonged morbidity [4]. We present a case of tattoo-associated cutaneous NTM infection in a young immunocompetent male, highlighting diagnostic challenges, atypical histopathologic findings, and important management considerations for dermatologists.

Case Report

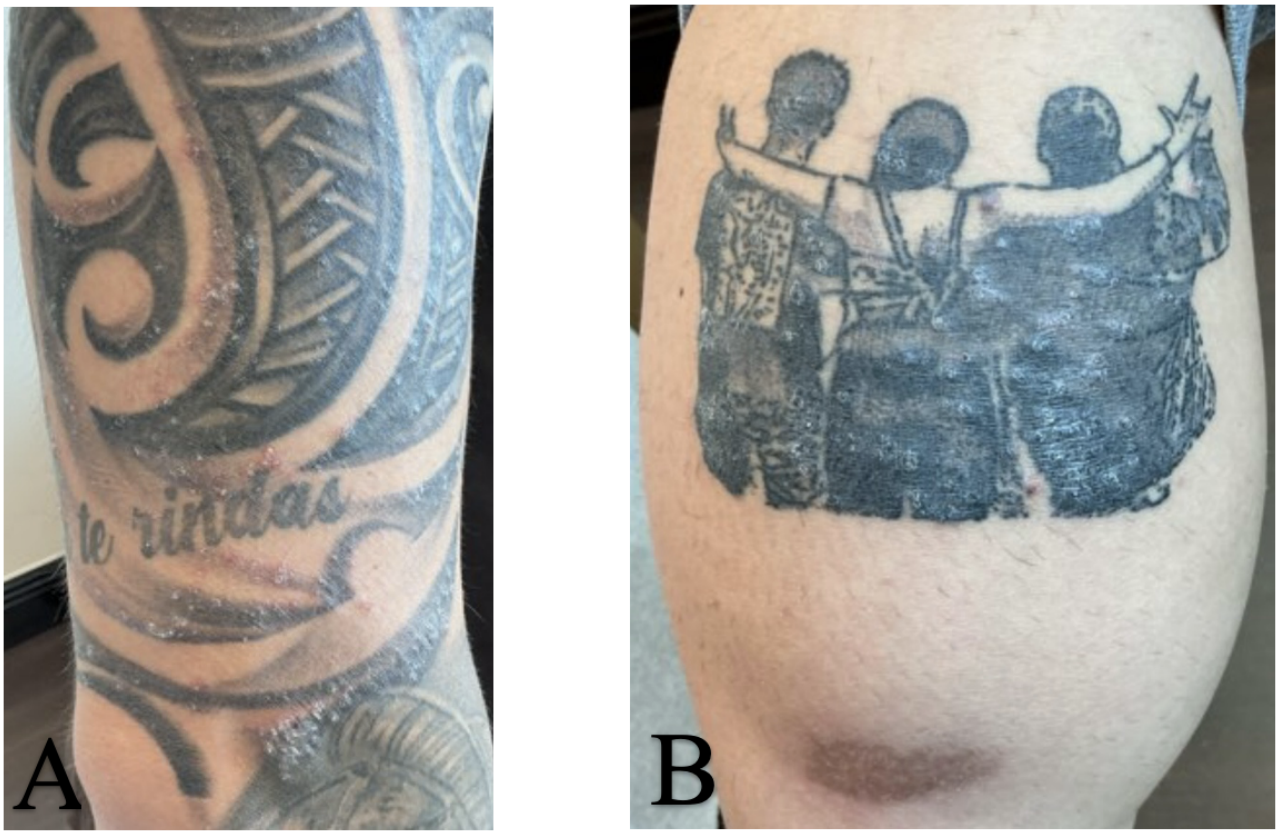
A 28-year-old male presented with an asymptomatic rash involving tattooed skin on the right upper arm and left calf 2 months after the initial tattoo session. The patient denied pain, pruritus, fever, chills, lymphadenopathy, drainage, or other systemic symptoms. He had no significant past medical history, was not immunocompromised, and reported no recent illness. The absence of constitutional or local symptoms, despite involvement of two anatomically distinct tattoo sites, made the presentation atypical for a clinically apparent infection and contributed to diagnostic complexity.

Physical examination demonstrated diffusely distributed monomorphic papules confined predominantly to areas of gray and blue tattooed skin, as well as some small papules with no erythema and with collarettes of white scale (Figure 1). The multifocal involvement of distinct tattoo sites raised suspicion for an inoculation-related etiology rather than a localized allergic

XSL•FO
RenderX

or inflammatory reaction. Given the chronicity, morphology, and confinement to tattoo pigment, an infectious process was included early in the differential diagnosis, despite the lack of classic infectious symptoms.

Figure 1. (A) Right upper arm showing diffusely distributed monomorphic papules confined predominantly to areas of gray and blue tattoo pigment and (B) left calf with evidence of small papules with and without surrounding erythema, some exhibiting collarettes of white scale consistent with multifocal tattoo-associated cutaneous involvement.



A punch biopsy (Figure 2) was obtained from the right upper arm for histopathologic evaluation and tissue culture, including an acid-fast bacilli culture. Histopathology revealed acute folliculitis with a lichenoid vacuolar lymphohistiocytic interface dermatitis associated with tattoo pigment (Figure 3). Although granulomatous inflammation is frequently described in cutaneous NTM infection, the presence of a lichenoid vacuolar interface pattern in this case was atypical and could have supported an inflammatory or lichenoid tattoo reaction in the absence of microbiologic testing.

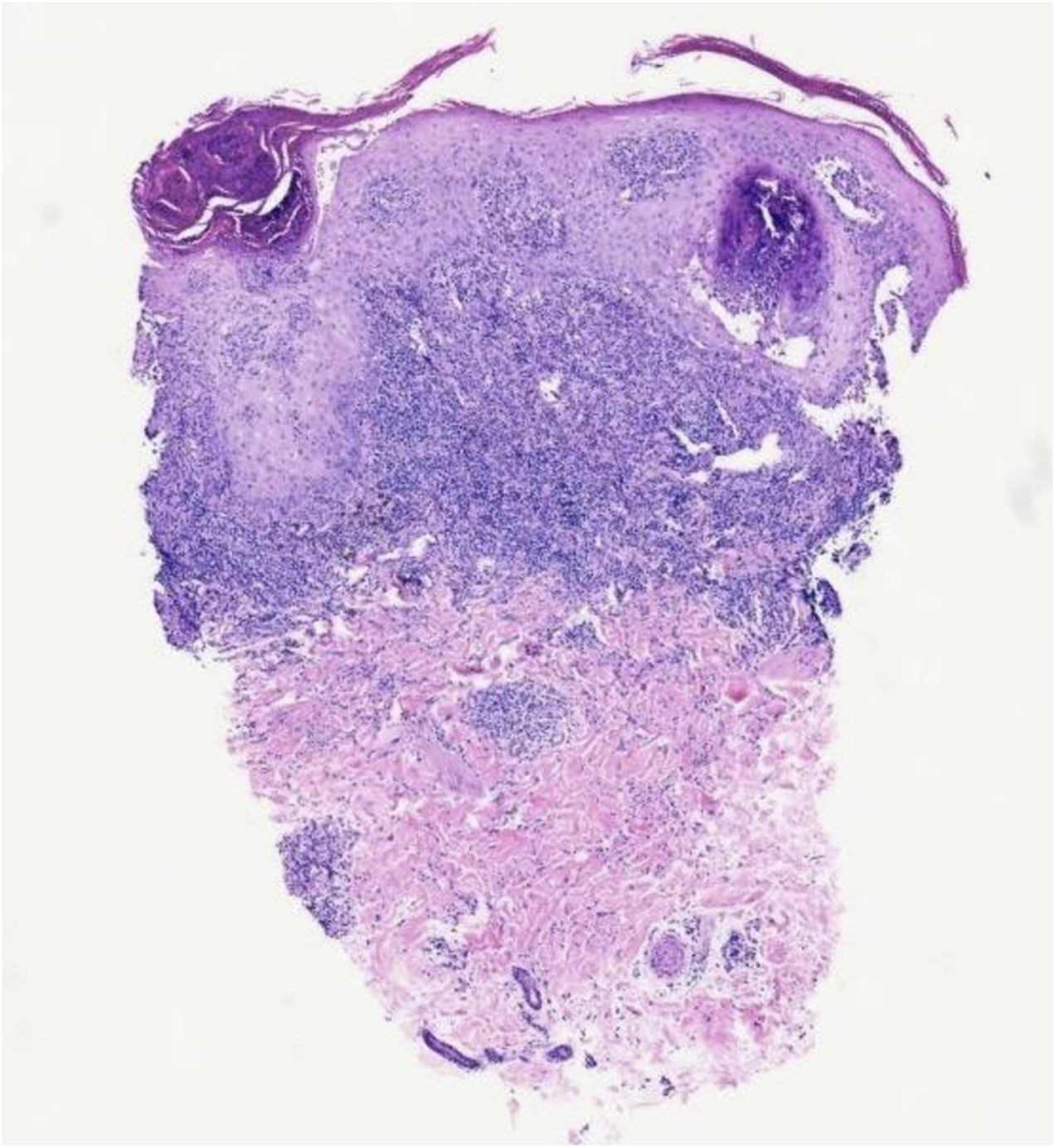
Initial empiric treatment consisted of twice-daily triamcinolone acetonide 0.1% ointment and doxycycline 100 mg for one week, reflecting standard management for inflammatory tattoo reactions while awaiting culture results. The acid-fast bacilli culture subsequently isolated a NTM species, confirming the diagnosis of tattoo-associated cutaneous NTM infection. The prolonged incubation period required for organism growth demonstrated a key diagnostic delay inherent to these infections.

Following microbiological confirmation of *M abscessus* complex, the treatment regimen was escalated. Prolonged dual antibiotic therapy was initiated with twice-daily doxycycline 100 mg and clarithromycin 500 mg for a planned duration of 3 months as per current recommendations on length of therapy for mild skin and soft tissue infections by NTM and dual-agent therapy to minimize the risk of antimicrobial resistance [5,6].

M abscessus complex is a group of NTM, characterized by fast-growing, nonchromogenic properties, and it is well known for its multidrug resistance [7]. For example, *M abscessus* complex accounted for 12 of 67 reported infection events in Miami in 2015, making it the second most frequently identified species after *Mycobacterium chelonae* [2].

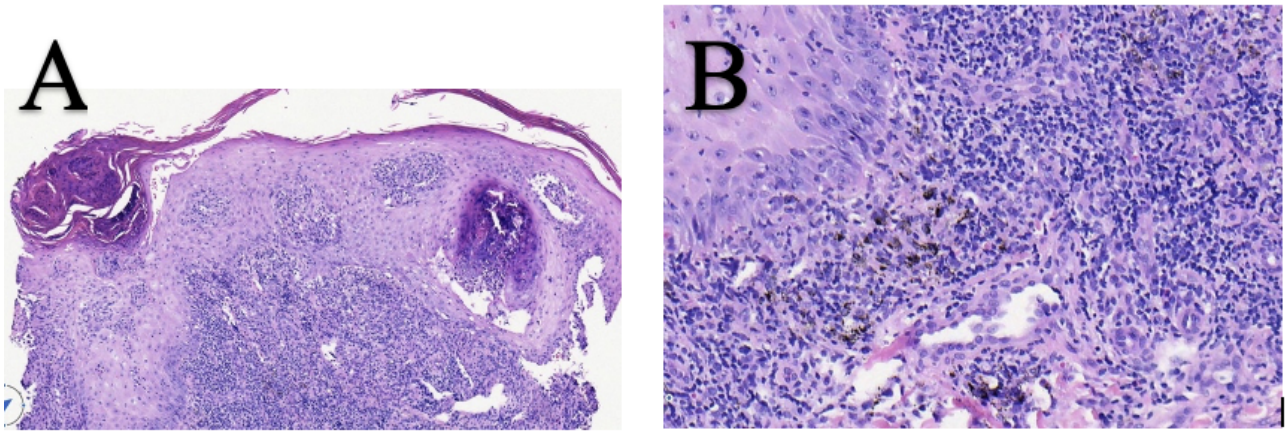
XSL•FO
RenderX

Figure 2. Punch biopsy from patient's right upper arm.



XSL•FO
RenderX

Figure 3. (A) Histopathology of punch biopsy demonstrating acute folliculitis with lichenoid vacuolar lymphohistiocytic interface dermatitis associated with tattoo pigment; hematoxylin and eosin stain at 4X magnification and (B) 20X magnification.



Ethical Considerations

Written informed consent for the publication of recognizable patient photographs and other identifiable material was obtained from the patient by the authors.

Discussion

NTM comprise all mycobacterial species other than *Mycobacterium tuberculosis* and *Mycobacterium leprae* and are widely distributed in the environment. Cutaneous NTM infections most commonly occur following trauma or direct inoculation, including tattooing, and rapidly growing mycobacteria are most frequently implicated in tattoo-associated infections [8]. Clinical manifestations vary widely, from papules and pustules to nodules or abscesses, and lesions may appear weeks to years after tattooing [8]. Distinctly, our patient presented with an entirely asymptomatic, predominantly nonerythematous eruption.

Diagnosis of cutaneous NTM infection remains challenging due to variable clinical presentations, frequent absence of systemic symptoms, and nonspecific histopathologic findings. In this case, the asymptomatic nature of the eruption and the presence of a lichenoid vacuolar interface dermatitis, an atypical histologic pattern for infection, further complicated diagnosis, which could have been mistaken for an inflammatory tattoo reaction. Skin biopsy with mycobacterial culture remains the diagnostic gold standard, although prolonged incubation may delay treatment [9]. Emerging diagnostic techniques, including polymerase chain reaction assays, DNA probes, and matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry can offer faster species identification once growth is achieved [9]. Nonetheless, antimicrobial susceptibility testing remains essential for therapy planning.

Management typically requires prolonged multidrug regimens, with treatment duration of 3 - 12 months, depending on disease severity [6]. This case illustrates the critical transition from empiric short-course therapy to guideline-supported prolonged combination therapy, emphasizing the importance of early infectious disease involvement. From a public health perspective, tattoo-associated NTM infections raise concerns regarding ink contamination and the need for sterile tattooing practices. Limitations of this case include its single-patient design, inability to establish the source of NTM contamination, and limited long-term follow-up.

Conclusion

This case describes an uncommon but increasingly seen presentation of tattoo-associated cutaneous NTM infection in an immunocompetent patient, characterized by asymptomatic multifocal lesions, atypical lichenoid vacuolar interface dermatitis, and delayed microbiologic diagnosis. Persistent, pigment-specific papules, particularly involving multiple tattoo sites or failing to respond to conventional therapy, should prompt consideration of NTM even in the absence of pain, erythema, or systemic symptoms. Early biopsy with mycobacterial culture and susceptibility testing can facilitate diagnosis and guide appropriate prolonged multidrug therapy. Recognizing these clinical clues and maintaining a low threshold for mycobacterial testing may help clinicians identify similar tattoo-associated infections earlier and avoid misdiagnosis as an inflammatory or allergic tattoo reaction.

Generative AI Use

The authors attest that there was no use of generative AI technology in the generation of text, figures, or other informational content of this manuscript.

Funding

The authors declare no financial support was received for this work.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: CA, MS
Supervision: MNZ
Writing - original draft: CA, MS (equal)
Writing - review & editing: ASAM, RH, MK

Conflicts of Interest

None declared.

References

1. Kazandjieva J, Tsankov N. Tattoos: dermatological complications. Clin Dermatol 2007;25(4):375-382. [doi: 10.1016/j.clindermatol.2007.05.012] [Medline: 17697920]
2. Daddy-Gaoh S, Foley SL, Kweon O, Kim SJ. Nontuberculous mycobacteria skin infections associated with tattoos. Front Public Health 2026;14:1854792. [doi: 10.3389/fpubh.2026.1854792] [Medline: 42487811]
3. Griffin I, Schmitz A, Oliver C, et al. Outbreak of tattoo-associated nontuberculous mycobacterial skin infections. Clin Infect Dis 2019 Aug 30;69(6):949-955. [doi: 10.1093/cid/ciy979] [Medline: 30452604]
4. Frew JW, Nguyen RTD. Tattoo-associated mycobacterial infections: an emerging public health issue. Med J Aust 2015 Sep 7;203(5):223-23e. [doi: 10.5694/mja15.00098] [Medline: 26852055]
5. Gonzalez-Santiago TM, Drage LA. Nontuberculous mycobacteria: skin and soft tissue infections. Dermatol Clin 2015 Jul;33(3):563-577. [doi: 10.1016/j.det.2015.03.017] [Medline: 26143432]
6. Wang XY, Jia QN, Li J. Treatment of non-tuberculosis mycobacteria skin infections. Front Pharmacol 2023;14:1242156. [doi: 10.3389/fphar.2023.1242156] [Medline: 37731736]
7. Nessar R, Cambau E, Reyrat JM, Murray A, Gicquel B. Mycobacterium abscessus: a new antibiotic nightmare. J Antimicrob Chemother 2012 Apr;67(4):810-818. [doi: 10.1093/jac/dkr578] [Medline: 22290346]
8. Chalarca-Cañas D, Caviedes-Cleves MA, Correa-Londoño LA, Ospina-Gómez JP, Velásquez-Lopera MM. Tattoos: risks and complications, clinical and histopathological approach. An Bras Dermatol 2024;99(4):491-502. [doi: 10.1016/j.abd.2023.07.004] [Medline: 38521707]
9. Pennington KM, Vu A, Challener D, et al. Approach to the diagnosis and treatment of non-tuberculous mycobacterial disease. J Clin Tuberc Other Mycobact Dis 2021 Aug;24:100244. [doi: 10.1016/j.jctube.2021.100244] [Medline: 34036184]

Abbreviations

***M abscessus*:** *Mycobacterium abscessus*
MALDI-TOF: matrix-assisted laser desorption ionization time-of-flight
NTM: nontuberculous mycobacteria

Edited by R Dellavalle; submitted 02.Aug.2026; peer-reviewed by X Nxele, Y Wang; revised version received 20.Sep.2026; accepted 22.Sep.2026; published 30.Oct.2026.

Please cite as:
Alihosseini C, Stein M, Madiedo ASA, Hesari R, Kauffman M, Zaiac MN
Asymptomatic Tattoo-Associated Cutaneous Nontuberculous Mycobacterial Infection: Case Report
JMIR Dermatol 2026;9:e108587
URL: https://derma.jmir.org/2026/1/e108587
doi:10.2196/108587

© Chris Alihosseini, Milana Stein, Ana Sofia Acosta Madiedo, Roksana Hesari, Matthew Kauffman, Martin N Zaiac. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Atypical Fibroxanthoma Presenting as a Rapidly Enlarging Tender Papule: Case Report

Elizabeth Tchernogorova[1], BA; Nicole Skalka[1], BSc; Mary Tidwell[1], BSc; Binh Minh Nguyen[2], BA; Julien Lanoue[3], MD; Robin Burger[4], MD; McKenna Abercrombie[3], DO

[1]Rocky Vista University, Ivins, UT, United States

[2]Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

[3]Advanced Dermatology and Cosmetic Surgery, Brandon, FL, United States

[4]Advanced Dermatology and Cosmetic Surgery, Delray Beach, FL, United States

Corresponding Author:
Binh Minh Nguyen, BA
Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

Abstract

Background: Atypical fibroxanthoma (AFX) is a rare dermal mesenchymal neoplasm that is most commonly identified on sun-exposed areas of the head and neck of older adults. Presentation on the dorsal hand is uncommon.

Objective: This report presents a rare case of AFX, its clinical presentation, histopathology, and treatment, and discusses the use of immunohistochemistry as a method in diagnosing AFX.

Methods: We evaluated a 67-year-old woman with a history of basal cell and squamous cell carcinomas who presented with a rapidly enlarging, tender papule on the dorsal aspect of her left hand. Prior topical 5-fluorouracil therapy was ineffective. Histopathologic examination demonstrated findings favoring AFX, and treatment was initiated. A literature review was performed to contextualize this case among previously reported instances of AFX.

Results: Clinical examination revealed a 1.2-cm, dome-shaped, erythematous papule. Histopathology demonstrated a dermal-based pleomorphic spindle cell neoplasm with brisk mitotic activity. Immunohistochemistry showed positivity for CD68 and CD10, with negative staining for SOX10, PRAME, pankeratin, p40, SMA, desmin, and CD34. The morphologic and immunohistochemical findings favored AFX. The tumor was completely excised via Mohs micrographic surgery (MMS) in a single stage with negative margins. No adjuvant therapy was required.

Conclusions: AFX remains a diagnosis of exclusion requiring comprehensive immunohistochemistry to distinguish it from morphologic mimics with aggressive biologic behavior. This case highlights that MMS is the preferred treatment for AFX, offering the lowest recurrence rates with maximal tissue conservation.

(JMIR Dermatol 2026;9:e84550) doi:10.2196/84550

KEYWORDS

atypical fibroxanthoma; mesenchymal neoplasm; sun-exposed skin; Mohs micrographic surgery; immunohistochemistry; skin cancer

History

A 67-year-old woman presented with a rapidly enlarging, tender papule on the dorsal aspect of her left hand. The lesion had evolved over several weeks, with no preceding trauma or known immunosuppression. The patient reported significant lifetime sun exposure and a history of non-melanoma skin cancer.

Ethical Considerations

Written informed consent was obtained from the patient pertaining to the use of their clinical photographs and medical information to be published online and with the understanding that this information may be publicly available and discoverable via search engines. Patient consent forms were provided to the journal.

Examination

Physical examination revealed a solitary, erythematous, dome-shaped papule measuring approximately 1.2 cm in diameter (Figure 1). The lesion was firm, nonulcerated, and tender to palpation. No regional lymphadenopathy was detected. The remainder of the cutaneous exam was unremarkable.

XSL•FO
RenderX

Figure 1. Papule on the dorsal aspect of patient's left hand.

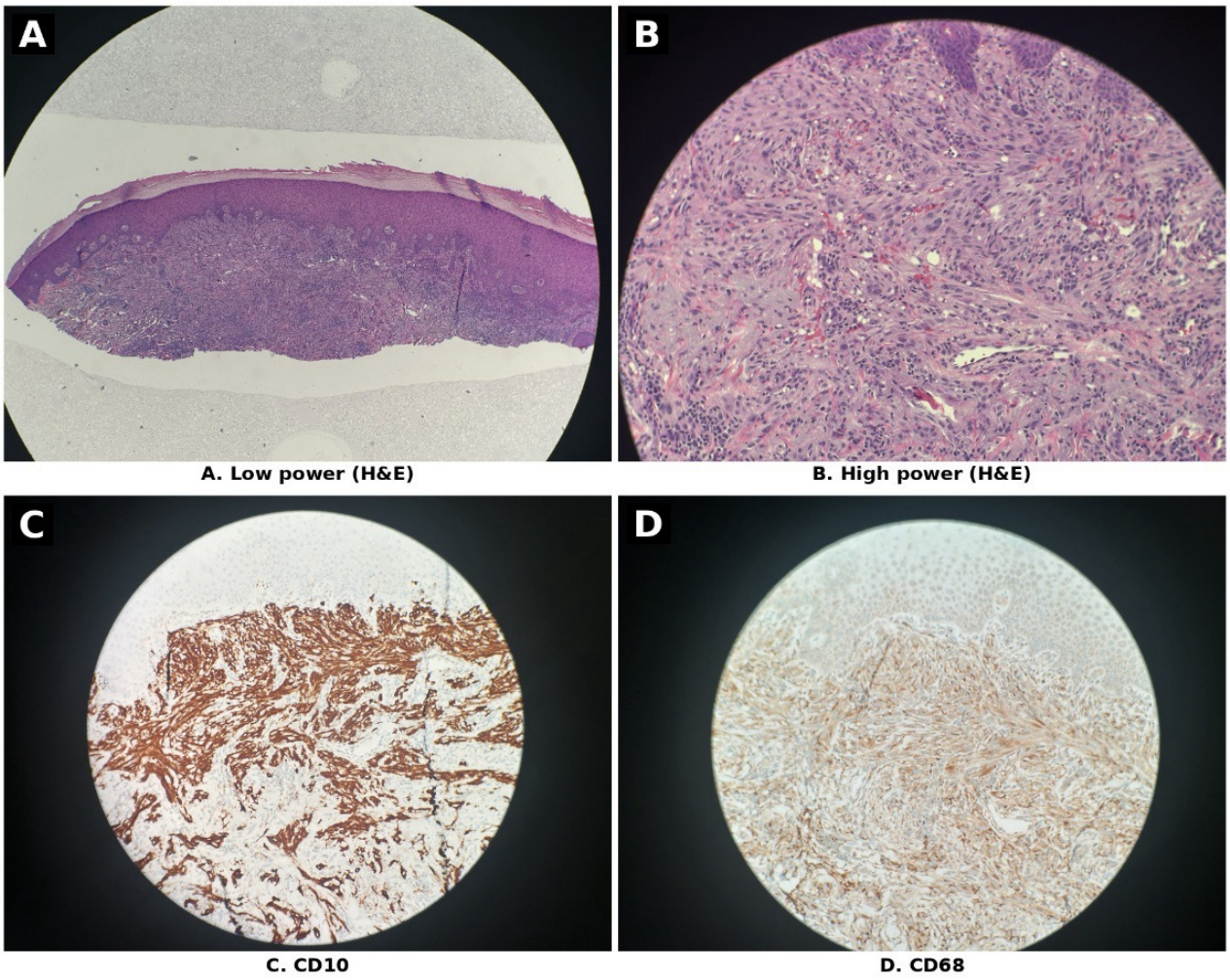


Histopathology

A shave biopsy demonstrated a dermal-based neoplasm composed of pleomorphic spindle and epithelioid cells with hyperchromatic nuclei and abundant cytoplasm (Figure 2, Figure 3). Histopathologic findings favored atypical fibroxanthoma (AFX). Given the superficial nature of the biopsy, however, the possibility that the lesion represented the superficial portion of a pleomorphic dermal sarcoma cannot be entirely excluded. Excision of any residual lesion was recommended for complete histologic evaluation. The lesion stained positive for CD68 and CD10 and negative for SOX10, PRAME, pankeratin, p40, SMA, desmin, and CD34. In the appropriate morphologic context, the immunohistochemical findings supported the diagnosis of atypical fibroxanthoma by helping to exclude epithelial, melanocytic, and smooth muscle differentiation.

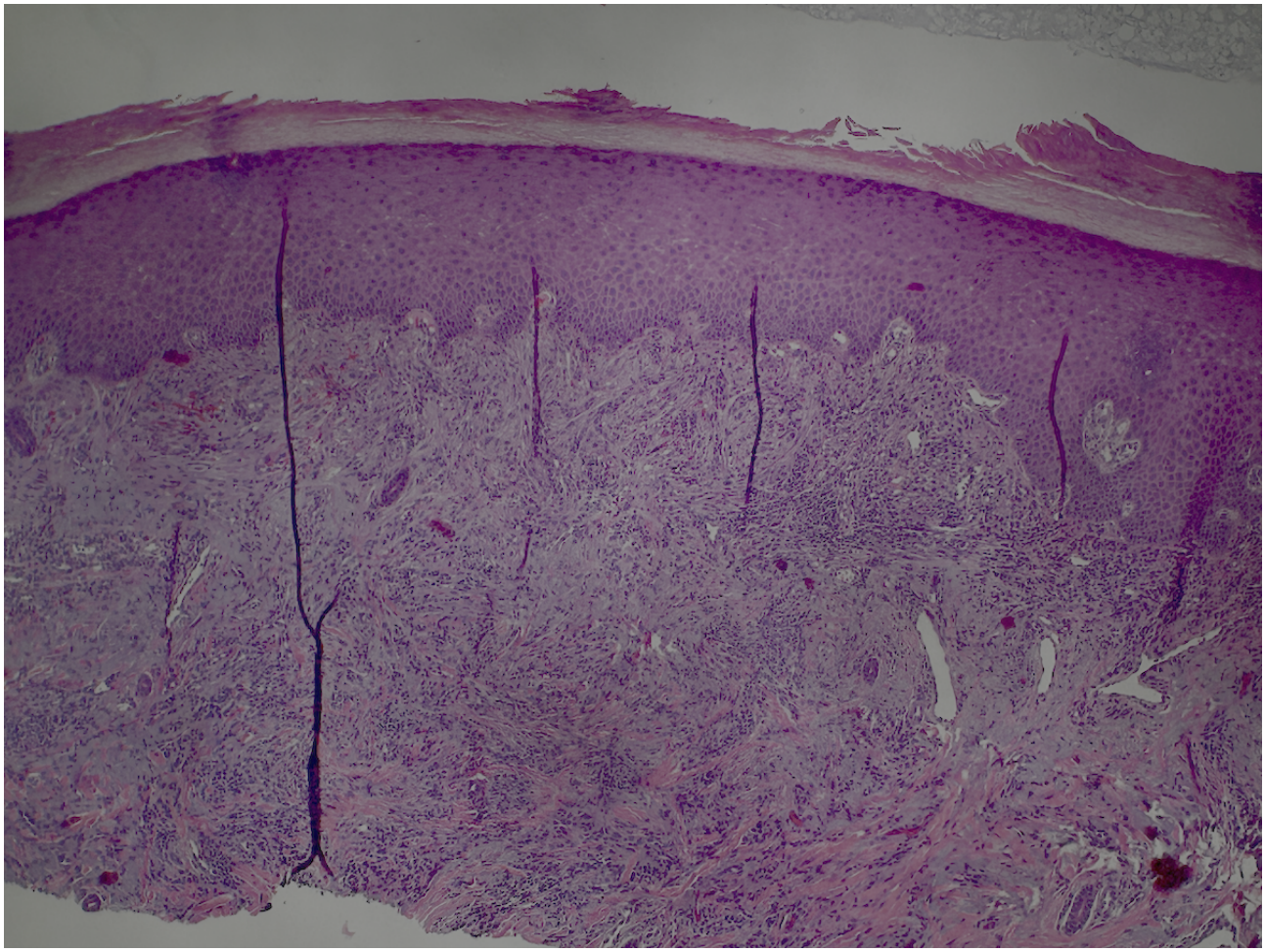
XSL•FO
RenderX

Figure 2. Histopathologic and immunohistochemical findings from the shave biopsy. (A) Low-power hematoxylin and eosin (H&E) stain demonstrates a dermal-based spindle cell proliferation (x4). (B) High-power H&E demonstrates a proliferation of atypical pleomorphic spindle cells (x20). (C) Lesional cells demonstrate strong immunoreactivity for CD10. (D) Lesional cells demonstrate immunoreactivity for CD68.



XSL•FO
RenderX

Figure 3. Histopathology of the lesion.



Course and Therapy

The patient was referred to Mohs micrographic surgery, a treatment associated with a high cure rate (Figures 4-6), where the tumor was cleared in a single stage with histologically negative margins. No adjuvant therapy was necessary. In the absence of residual tumor for further evaluation, and given the morphologic and immunohistochemical findings of the original biopsy, the lesion was regarded as atypical fibroxanthoma.

XSL•FO
RenderX

Figure 4. Pre-operative marking of the papule margins.



Figure 5. Papule site post operation showing negative margins.



XSL•FO
RenderX

Figure 6. Repair closure after completed Mohs micrographic surgery.



Discussion

AFX is a rare mesenchymal neoplasm that has been reported with increasing frequency over the past decade [1,2]. It is most commonly identified on sun-exposed areas of the head and neck of older adults, with a mean age of 75.8 years [1,3]. Approximately 89% of lesions arise in the head and neck; presentation on the dorsal hand is uncommon [4]. The diagnosis of AFX is often one of exclusion, as it shares overlapping morphological features with other neoplasms that develop as a result of chronic sun exposure [1]. Ultraviolet (UV) exposure is a known risk factor for the development of AFX due to its ability to induce mutations in p53, a tumor suppressor protein [5,6]. Histologically, AFX can resemble undifferentiated pleomorphic sarcoma [1,3,5,7,8]. These tumors are typically well-circumscribed but not encapsulated, showing an expansile, non-infiltrative growth pattern and are usually confined to the superficial dermis. In some cases, they may extend deeper into the reticular dermis. A diagnostically useful feature is the lack of a dermal “grenz zone,” a band of normal papillary dermis spared by the lesion [1]. The apparent rise in AFX diagnoses may in part be attributed to increased use of immunohistochemical panels, which aid in distinguishing AFX from histologic mimickers [2]. Commonly used markers include pan-cytokeratin, S100, desmin, CD34, CD10, procollagen-1, CD99, CD14, and CD68 [2,9]. Early biopsy is critical for accurate diagnosis and prompt surgical intervention.

When considering a differential diagnosis, the primary entities of concern include spindle cell squamous cell carcinoma, malignant melanoma, leiomyosarcoma, undifferentiated pleomorphic sarcoma, and dermatofibrosarcoma protuberans [5,6]. As noted, both clinical presentation and immunohistochemical staining are critical for distinguishing these tumors [2]. Standard treatment of AFX is complete surgical excision with margin control, most commonly achieved through Mohs micrographic surgery. Surgical treatment has shown lower local recurrence rates, as well as optimal cosmetic outcomes. A wide local excision with a margin of at least 1 - 2 cm can be used if MMS is not feasible. Margin control is critical when using a wide local excision, and incomplete excision can result in recurrence. Given AFX’s low metastatic potential, adjuvant therapy is generally not necessary. In general,

XSL•FO
RenderX

prognosis is excellent because of low metastatic potential and adequate excision margins [7].

In this report, we discuss a unique case of AFX presenting on the dorsal aspect of the hand. Typically, AFX occurs in sun-exposed areas of the head and neck, such as the cheeks, nose, forehead, and scalp [10]. Instances of AFX appearing outside the head and neck have rarely been documented. Despite this patient's atypical location, the patient achieved a successful outcome with single-stage MMS. In summary, this case underscores the importance of considering AFX in less common sites and emphasizes that prompt recognition and treatment with MMS are crucial to ensure the best possible outcome.

Acknowledgments

This work was presented as a poster at Rocky Vista University, Ivins, UT, as part of the Fall Research Day 2025.

During the preparation of this work the authors used OpenEvidence in order to find supporting articles and sources. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

ET: Contributed substantially to the acquisition and organization of clinicalcase information, coordination of the manuscript development process, and delegation of projectresponsibilities. Led the drafting of substantial portions of the manuscript and contributed extensively tocritical revision and editing for important intellectual content.
NS: Contributed substantially to manuscript development, including drafting significantportions of the manuscript, synthesis and interpretation of relevant information, and critical revision andediting for important intellectual content.
MT: Contributed substantially to manuscript development through drafting andrefinement of manuscript content, synthesis and interpretation of relevant information, and criticalrevision and editing for important intellectual content.
BMN: Contributed substantially to manuscript development through drafting andrefinement of manuscript content, synthesis and interpretation of relevant information, and criticalrevision and editing for important intellectual content.
JBL: Provided procedural and clinical expertise related to the surgicalmanagement of the patient. Contributed to the acquisition and interpretation of relevant clinical andprocedural information and critically reviewed and revised the manuscript for important intellectualcontent.
RB: Provided expertise in the histopathologic evaluation and interpretation of the case,including findings integral to the diagnosis of atypical fibroxanthoma. Contributed to the interpretationand presentation of the pathologic findings and critically reviewed and revised the manuscript forimportant intellectual content.
MA: Provided the clinical case that formed the basis of this report andcontributed substantially to the acquisition and interpretation of clinical information. Supervised theoverall development of the case report, provided clinical expertise and intellectual guidance throughoutthe project, and critically reviewed and revised the manuscript for important intellectual content.
All authors have reviewed and approved the final version of the manuscript and agree to be accountablefor their individual contributions and for the accuracy and integrity of the work.

Conflicts of Interest

None declared.

References

1. Agaimy A. The many faces of atypical fibroxanthoma. Semin Diagn Pathol 2023 Jul;40(4):306-312. [doi: 10.1053/j.semdp.2023.06.001] [Medline: 37438163]
2. Kuntz T, Siebdrath J, Hofmann SC, et al. Increase of atypical fibroxanthoma and pleomorphic dermal sarcoma: a retrospective analysis of four German skin cancer centers. J Dtsch Dermatol Ges 2022 Dec;20(12):1581-1588. [doi: 10.1111/ddg.14911] [Medline: 36442137]
3. Soleymani T, Aasi SZ, Novoa R, Hollmig ST. Atypical fibroxanthoma and pleomorphic dermal sarcoma: updates on classification and management. Dermatol Clin 2019 Jul;37(3):253-259. [doi: 10.1016/j.det.2019.02.001] [Medline: 31084719]
4. von Dannecker R, de Oliveira LML, Lapenda I, et al. Clinical outcomes and prognosis in atypical fibroxanthoma: a systematic review and meta‐analysis. Int J Dermatol 2025 Sep;64(9):1592-1598. [doi: 10.1111/ijd.17841] [Medline: 40350546]
5. Gru AA, Santa Cruz DJ. Atypical fibroxanthoma: a selective review. Semin Diagn Pathol 2013 Feb;30(1):4-12. [doi: 10.1053/j.semdp.2012.01.001] [Medline: 23327726]
6. Anders IM, Schimmelpfennig C, Wiedemann K, et al. Atypical fibroxanthoma and pleomorphic dermal sarcoma - gene expression analysis compared with undifferentiated cutaneous squamous cell carcinoma. J Dtsch Dermatol Ges 2023 May;21(5):482-491. [doi: 10.1111/ddg.15006] [Medline: 37035902]

XSL•FO
RenderX

7. Iorizzo LJ, Brown MD. Atypical fibroxanthoma: a review of the literature. Dermatol Surg 2011 Feb;37(2):146-157. [doi: 10.1111/j.1524-4725.2010.01843.x] [Medline: 21269345]
8. Hussein MRA. Atypical fibroxanthoma: new insights. Expert Rev Anticancer Ther 2014 Sep;14(9):1075-1088. [doi: 10.1586/14737140.2014.924401] [Medline: 24958232]
9. Helbig D, Ihle MA, Pütz K, et al. Oncogene and therapeutic target analyses in atypical fibroxanthomas and pleomorphic dermal sarcomas. Oncotarget 2016 Apr 19;7(16):21763-21774. [doi: 10.18632/oncotarget.7845] [Medline: 26943575]
10. Klein JC, Wilky BA, Ford HL. Molecular characterization of atypical fibroxanthoma and pleomorphic dermal sarcoma. Cancers (Basel) 2025 May 27;17(11):1785. [doi: 10.3390/cancers17111785] [Medline: 40507266]

Abbreviations

AFX: atypical fibroxanthoma
MMS: Mohs micrographic surgery

Edited by R Dellavalle; submitted 26.Sep.2025; peer-reviewed by C Tredici, R Phelps; revised version received 09.Feb.2026; accepted 10.Feb.2026; published 30.Oct.2026.

Please cite as:
Tchernogorova E, Skalka N, Tidwell M, Nguyen BM, Lanoue J, Burger R, Abercrombie M
Atypical Fibroxanthoma Presenting as a Rapidly Enlarging Tender Papule: Case Report
JMIR Dermatol 2026;9:e84550
URL: https://derma.jmir.org/2026/1/e84550
doi:10.2196/84550

© Elizabeth Tchernogorova, Nicole Skalka, Mary Tidwell, Binh Minh Nguyen, Julien Lanoue, Robin Burger, McKenna Abercrombie. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Chronic Facial Abscess Mimicking Cervicofacial Actinomyces From Dermal Filler Migration: Case Report

Monika Ziogaite[1], BS; Sarah Mannlein[2], BS; Nicole Bender[3], MD; Scott J Mahlberg[2], DO

[1]College of Osteopathic Medicine, Kansas City University, 1750 Independence Ave, Kansas City, MO, United States

[2]Colorado Center for Dermatology & Skin Surgery, Longmont, CO, United States

[3]Sagis Dermatopathology, Houston, TX, United States

Corresponding Author:
Monika Ziogaite, BS
College of Osteopathic Medicine, Kansas City University, 1750 Independence Ave, Kansas City, MO, United States

Abstract

Dermal fillers are commonly used for facial augmentation, but delayed complications such as granulomatous inflammation and filler migration can mimic chronic bacterial infections, such as cervicofacial actinomycosis, and lead to diagnostic misdirection. We present the case of a woman aged 56 years with a chronic, draining abscess on the right cheek that persisted for 3 years and was initially suspected to represent cervicofacial actinomycosis. Tissue cultures were negative, and histopathologic analysis following excisional biopsy revealed polymethyl methacrylate microspheres and hyaluronic acid surrounded by granulomatous inflammation and reactive lymphoid aggregates, consistent with a foreign body reaction to dermal filler. The patient experienced complete resolution after surgical excision. This case underscores the diagnostic challenges posed by delayed filler complications and highlights the importance of considering prior cosmetic procedures in patients with chronic facial abscesses.

(JMIR Dermatol 2026;9:e80278) doi:10.2196/80278

KEYWORDS

dermal filler; foreign body reaction; facial abscess; granulomatous inflammation; filler complications; cosmetic dermatology

Introduction

The use of dermal fillers for facial augmentation has increased significantly, with both temporary (hyaluronic acid) and semipermanent or permanent fillers (polymethyl methacrylate [PMMA], calcium hydroxylapatite, silicone) [1]. While most complications occur immediately or within weeks, delayed reactions, including granuloma formation and filler migration, can present months to years after injection [2].

Foreign body granulomas are a known complication of PMMA-based fillers (Bellafill/Artefill; Suneva Medical Inc) and result from a chronic inflammatory response to nondegradable microspheres [3]. These reactions may be triggered by delayed hypersensitivity, biofilm formation, or immune dysregulation and can resemble infectious or inflammatory processes. They can also closely mimic chronic infectious processes such as cervicofacial actinomycosis, characterized by draining sinuses and subcutaneous abscesses, often prompting an extensive infectious workup before the true etiology is recognized [4]. Diagnosis relies on histopathologic evaluation, which typically reveals multinucleated giant cells, lymphoid aggregates, and fibrosis surrounding filler particles [3,4].

This case highlights the importance of early recognition of iatrogenic causes in the differential diagnosis of chronic facial abscesses and underscores the long-term risks associated with semipermanent fillers, particularly PMMA-based products. Given the varied histopathologic presentations of different filler materials, distinguishing PMMA from other injectables is crucial for accurate diagnosis and management.

Ethical Considerations

The authors obtained written consent from the patient for their photographs and medical information to be published in print and online, with the understanding that this information may be publicly available. Patient consent forms were not provided to the journal but are retained by the authors.

Case Report

A woman aged 56 years presented with a chronic, nonhealing, draining abscess localized to the right cheek. Characterized by intermittent drainage, localized tenderness, and surrounding erythema, the nodule persisted for approximately 3 years, during which time the patient sought care from specialists on at least one occasion. The patient denied systemic symptoms such as fever, chills, dental caries, oral drainage, pain with salivation, or malaise. Past medical history was noncontributory, and the patient had no known history of immunosuppression, diabetes, or recurrent skin infections.

XSL•FO
RenderX

On physical examination, the deep nodule was ulcerated, with erythematous borders localized to the right inferior central malar cheek. The ulcer base exhibited crusting and purulent material, with 3 cm of surrounding induration (Figure 1). No regional lymphadenopathy was present.

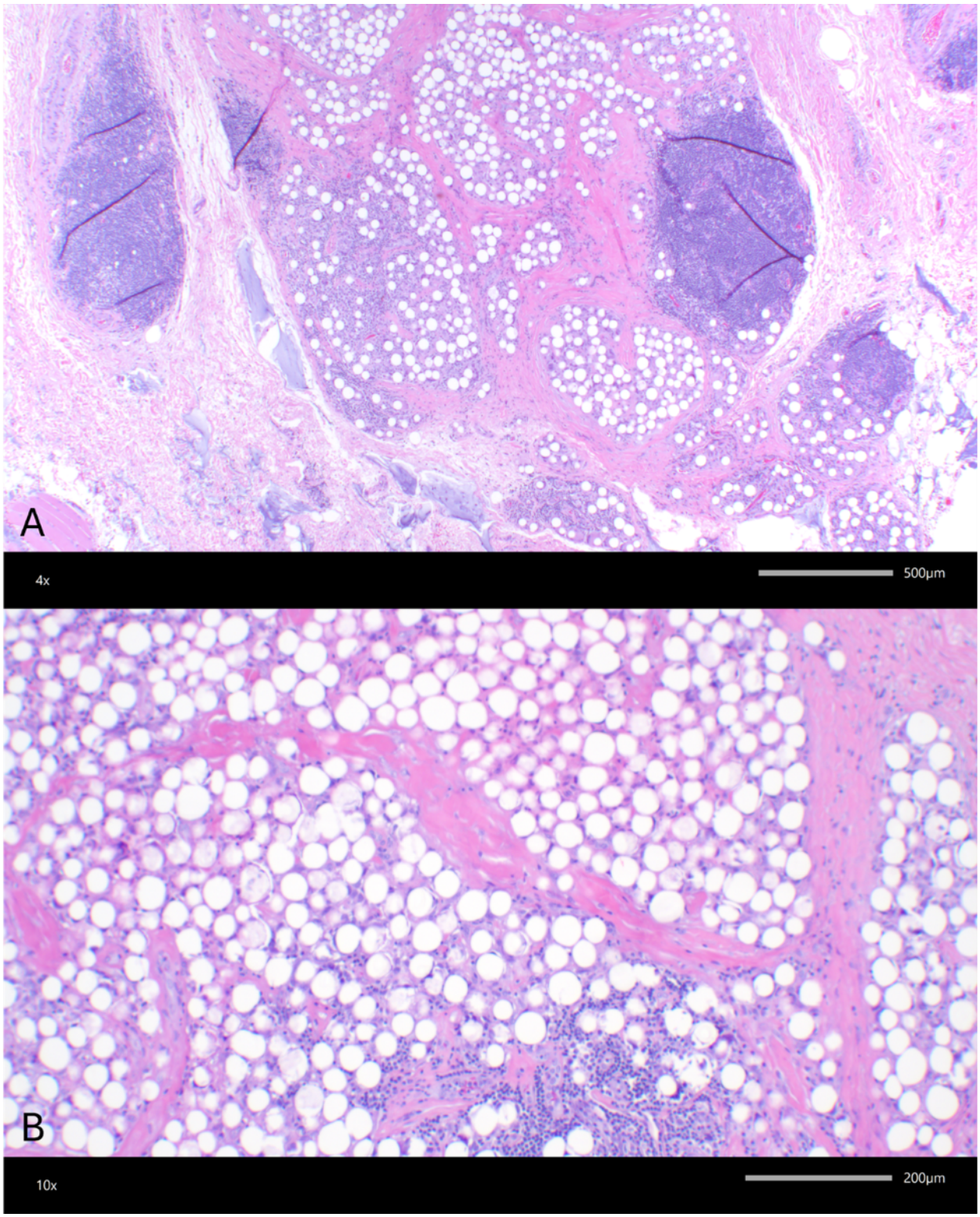
Figure 1. Nonhealing ulcer with surrounding erythema and induration on the right cheek.



Given the persistent nature of the abscess, a tissue culture was obtained via punch biopsy, which was negative for bacterial, fungal, and atypical mycobacterial growth. Due to the size and depth of the lesion, an excisional biopsy was performed to identify potential inflammatory or neoplastic pathology. Histopathologic analysis (Figure 2) revealed foreign body granulomas with abundant reactive lymphoid tissue, along with an accumulation of PMMA microspheres and hyaluronic acid, consistent with semipermanent dermal fillers. Surrounding granulomatous inflammation was also noted. Special stains, including Grocott methenamine silver, Fite, and Gram stains, were negative for fungi, mycobacteria, and bacteria, ruling out infectious etiologies.

XSL•FO
RenderX

Figure 2. Histologic features consistent with a foreign body reaction to polymethylmethacrylate (PMMA) filler. (A) Incisional biopsy with numerous rounded vacuolated spaces, 30 - 50 µm in size, consistent with PMMA spherules and surrounding granulomatous inflammation and reactive lymphoid aggregates. Adjacent homogeneous basophilic material consistent with hyaluronic acid is also present (hematoxylin and eosin stain; 4x). (B) Higher magnification of PMMA spherules (hematoxylin and eosin stain; 10×).



The patient was seen for suture removal at 1 week (Figure 3) with an uneventful postoperative course. At 4 weeks postoperatively, the surgical site was well healed, with no signs of infection or recurrence (Figure 3). Although initially unable to recall prior dermal filler use, review of the pathology report prompted the patient to remember having received Bellafill (Suneva Medical Inc) approximately 6 years prior and Restylane (Galderma Laboratories) approximately 10 years earlier, both for treatment of acne scarring on the cheeks.

XSL•FO
RenderX

Figure 3. (A) Postoperative appearance at 1 week showing early healing. (B) Postoperative appearance at 4 weeks with complete resolution and no recurrence.



Discussion

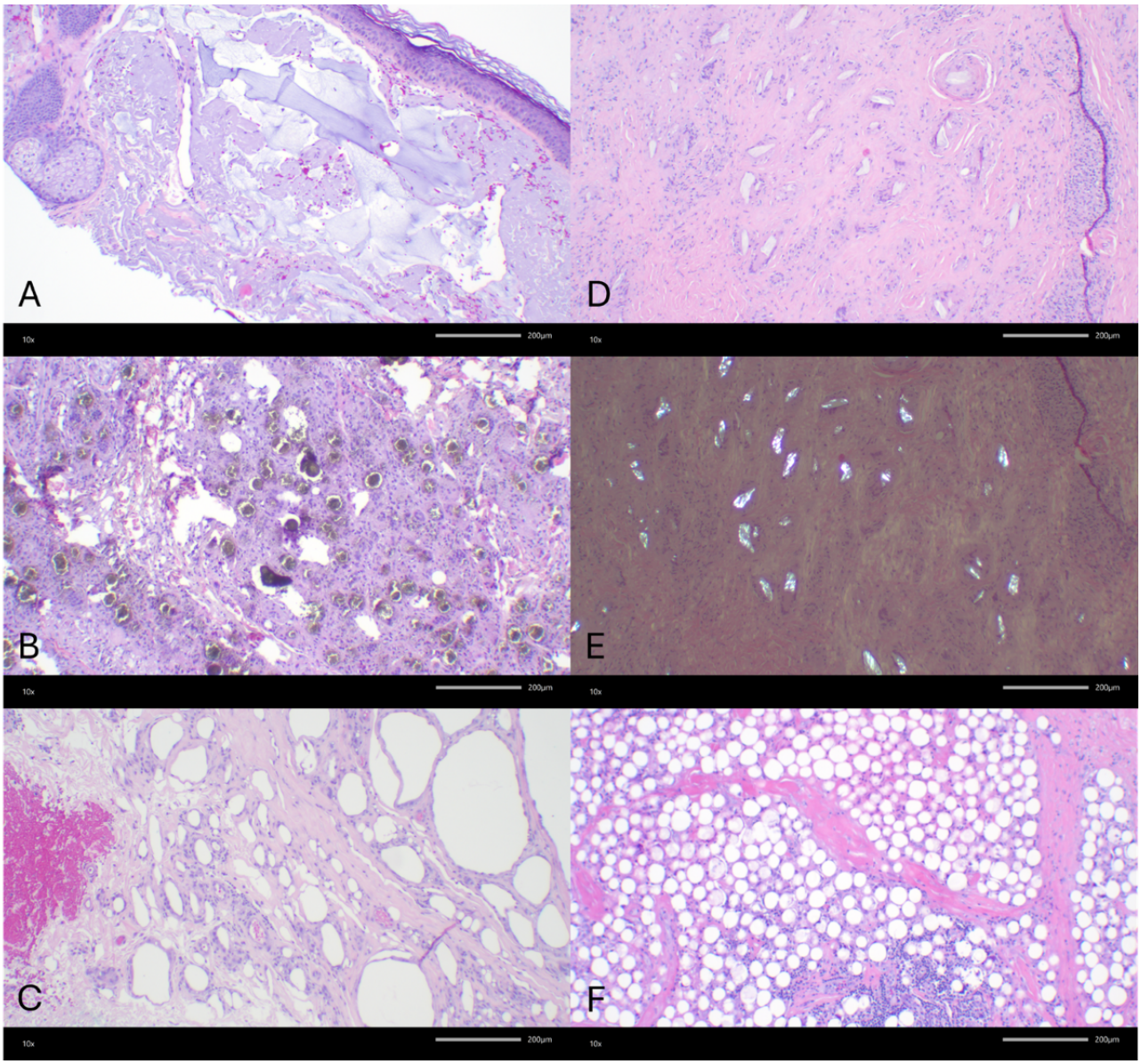
While dermal fillers are widely used for aesthetic enhancement, their delayed complications remain an evolving area of clinical concern. Although most adverse reactions occur shortly after injection, late-onset complications can develop months to years later, often leading to diagnostic uncertainty [5]. Unlike temporary hyaluronic acid–based fillers, PMMA and other nondegradable materials persist within tissues long-term, increasing the risk of prolonged inflammatory responses [6]. The delayed presentation of PMMA-related granulomas frequently results in misdiagnosis as infection or inflammatory dermatoses, delaying appropriate intervention [7].

While complications such as filler migration and granulomatous reactions are well documented, the development of a chronic filler reaction mimicking a cervicofacial actinomycetoma is rare. Actinomycetomas are chronic, subcutaneous infections caused by filamentous bacteria, characterized by abscesses, draining sinuses, and granule production [8]. The striking clinical resemblance between the foreign body granuloma in this case and a deep-seated actinomycotic infection underscores the diagnostic challenges posed by delayed filler reactions. This case highlights the need for broad infectious and histopathologic workups in atypical, chronic soft tissue infections to prevent unnecessary antibiotic treatment and delayed surgical intervention.

Histopathologic evaluation is essential for diagnosing PMMA-related granulomas, which are characterized by multinucleated giant cells, chronic lymphohistiocytic infiltrates, and fibrosis surrounding filler particles [9]. In this case, a negative infectious workup and biopsy findings of PMMA microspheres with reactive lymphoid tissue confirmed the diagnosis and guided treatment. Management remains challenging, as PMMA-based fillers lack a reversal agent comparable to that of hyaluronic acid fillers [10]. While intralesional corticosteroid injections may offer partial improvement, surgical excision is often required for definitive diagnosis and treatment, as in this case [4]. Given the histologic variability among filler types, distinguishing PMMA granulomas from reactions to calcium hydroxylapatite, poly-L-lactic acid, and silicone is critical for guiding optimal management strategies (Figure 4).

XSL•FO
RenderX

Figure 4. Histologic features consistent with foreign body reactions to soft tissue augmentation materials. (A) Pools of wispy homogeneous basophilic material consistent with hyaluronic acid (hematoxylin and eosin; 10×). (B) Gray-green granular nonrefractile microspheres consistent with calcium hydroxylapatite with surrounding granulomatous inflammation (hematoxylin and eosin; 10×). (C) Variably sized empty lipidlike vacuoles within histiocytes consistent with silicone granuloma (hematoxylin and eosin; 10×). (D) Oval, rhomboidal, and rice-shaped clear refractile and polarizable structures consistent with poly-L-lactic acid (PLLA) particles within histiocytes (hematoxylin and eosin; 10×). (E) Polarization of PLLA fragments (hematoxylin and eosin; 10×). (F) Fairly uniform 30 - 50-µm rounded vacuolated spaces consistent with polymethylmethacrylate spherules (hematoxylin and eosin; 10×).



While PMMA fillers are used less frequently than hyaluronic acid–based products, their potential for chronic inflammatory complications requires heightened clinical awareness and a detailed risk-benefit discussion prior to injection. Semipermanent fillers pose unique challenges due to their prolonged tissue retention and risk of delayed reactions. Clinicians should maintain a high index of suspicion for foreign body granulomas and probe for a history of prior filler use in cases of chronic, nonhealing facial abscesses, particularly when standard antimicrobial therapy fails or imaging reveals localized nodularity. Following excision, patients should be informed of and monitored for delayed recurrence as well as contralateral lesions, as these may occur months to years later. This case highlights the value of histopathologic microbiologic evaluation in diagnosing facial abscesses, the limitations of nonsurgical management for PMMA-induced granulomas, and the need for increased awareness of iatrogenic factors in chronic soft tissue reactions.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Jones DH. Semipermanent and permanent injectable fillers. Dermatol Clin 2009 Oct;27(4):433-444. [doi: 10.1016/j.det.2009.08.003] [Medline: 19850193]
2. Hong GW, Hu H, Chang K, et al. Review of the adverse effects associated with dermal filler treatments: part I nodules, granuloma, and migration. Diagnostics (Basel) 2024 Jul 30;14(15):1640. [doi: 10.3390/diagnostics14151640] [Medline: 39125515]
3. Sivam S, Mackay C, Humphrey C, Kriet JD. Giant PMMA foreign body granulomas with imaging. J Cutan Aesthet Surg 2023;16(3):256-258. [doi: 10.4103/JCAS.JCAS_194_20] [Medline: 38189071]
4. Lee JM, Kim YJ. Foreign body granulomas after the use of dermal fillers: pathophysiology, clinical appearance, histologic features, and treatment. Arch Plast Surg 2015 Mar;42(2):232-239. [doi: 10.5999/aps.2015.42.2.232] [Medline: 25798398]
5. Woodward J, Khan T, Martin J. Facial filler complications. Facial Plast Surg Clin North Am 2015 Nov;23(4):447-458. [doi: 10.1016/j.fsc.2015.07.006] [Medline: 26505541]
6. Funt D, Pavicic T. Dermal fillers in aesthetics: an overview of adverse events and treatment approaches. Clin Cosmet Investig Dermatol 2013 Dec 12;6:295-316. [doi: 10.2147/CCID.S50546] [Medline: 24363560]
7. Limongi RM, Tao J, Borba A, et al. Complications and management of polymethylmethacrylate (PMMA) injections to the midface. Aesthet Surg J 2016 Feb;36(2):132-135. [doi: 10.1093/asj/sjv195] [Medline: 26446059]
8. Welsh O, Vera-Cabrera L, Welsh E, Salinas MC. Actinomycetoma and advances in its treatment. Clin Dermatol 2012;30(4):372-381. [doi: 10.1016/j.clindermatol.2011.06.027] [Medline: 22682184]
9. Lemperle G, Gauthier-Hazan N, Wolters M, Eisemann-Klein M, Zimmermann U, Duffy DM. Foreign body granulomas after all injectable dermal fillers: part 1. Possible causes. Plast Reconstr Surg 2009 Jun;123(6):1842-1863. [doi: 10.1097/PRS.0b013e31818236d7] [Medline: 19483587]
10. Rao V, Chi S, Woodward J. Reversing facial fillers: interactions between hyaluronidase and commercially available hyaluronic-acid based fillers. J Drugs Dermatol 2014 Sep;13(9):1053-1056. [Medline: 25226005]

Abbreviations

PMMA: polymethyl methacrylate

Edited by J Solomon; submitted 07.Jul.2025; peer-reviewed by EA Sanchez-Romero, NC Balica; revised version received 16.Oct.2025; accepted 25.Dec.2025; published 03.Feb.2026.

Please cite as:
Ziogaite M, Mannlein S, Bender N, Mahlberg SJ
Chronic Facial Abscess Mimicking Cervicofacial Actinomyces From Dermal Filler Migration: Case Report
JMIR Dermatol 2026;9:e80278
URL: https://derma.jmir.org/2026/1/e80278
doi:10.2196/80278

© Monika Ziogaite, Sarah Mannlein, Nicole Bender, Scott J Mahlberg. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Anterior Cervical Hypertrichosis (Hairy Throat Syndrome): Pediatric Case Report and Brief Literature Review

Nancy Shehata[1], MD; Husna Irfan Thalib[2], MBBS; Heba Alahwal[1], MD

[1]Department of Dermatology, King Abdullah Medical Complex, Prince Nayef Street, Northern Abhor, Jeddah, Saudi Arabia

[2]General Medicine Practice Program, Batterjee Medical College, Jeddah, Saudi Arabia

Corresponding Author:
Nancy Shehata, MD
Department of Dermatology, King Abdullah Medical Complex, Prince Nayef Street, Northern Abhor, Jeddah, Saudi Arabia

Abstract

Anterior cervical hypertrichosis, also known as hairy throat syndrome, is a rare and typically benign condition characterized by a well-defined patch of terminal hair on the front of the neck. Although it is often an isolated finding, it may sometimes be associated with neurological or developmental abnormalities, which makes clinical awareness important. We report the case of a healthy 4-year-old girl who presented with congenital localized excessive hair growth over the anterior cervical region. The patch had remained unchanged since birth and was not associated with any skin changes or systemic symptoms. Her physical and neurological examinations were normal, and imaging studies ruled out underlying spinal or soft tissue anomalies. Although she had a resolved history of ptosis and was under follow-up for hypermetropia, no other abnormalities were identified. A sibling history of spina bifida prompted a more detailed evaluation, which returned normal results. This case highlights the importance of recognizing anterior cervical hypertrichosis as a rare but distinct clinical entity. Recognition of this benign entity is important to guide appropriate evaluation, avoid excessive investigations, and reduce caregiver anxiety.

(JMIR Dermatol 2026;9:e95391) doi:10.2196/95391

KEYWORDS

anterior cervical hypertrichosis; hairy throat syndrome; pediatric dermatology; congenital hypertrichosis; rare skin disorder

Introduction

Anterior cervical hypertrichosis (ACH), also known as “hairy throat syndrome,” is a rare condition characterized by the presence of a well-defined patch of excessive terminal hair on the anterior aspect of the neck, usually along the midline. It is most often congenital and discovered in early childhood, although some cases may present later [1,2]. The skin underneath the hair typically appears normal, without pigmentation changes, induration, or other cutaneous abnormalities. The condition is generally benign and asymptomatic; however, previous reports have described occasional associations with neurological, skeletal, and ophthalmological abnormalities including peripheral neuropathy, hallux valgus, optic atrophy, and spinal dysraphism [1,3,4]. Therefore, careful clinical assessment and selective investigations may be warranted in some patients depending on the clinical presentation and family history [1-5].

The pathogenesis of ACH is not completely understood but is hypothesized to arise from localized disturbances in hair follicle development during embryogenesis. There is no clear evidence of hereditary transmission in most cases, although rare familial occurrences have been described. Management is generally conservative, focusing on observation and reassurance. In cases where cosmetic concerns are significant, options such as laser hair removal may be considered [3,4].

It is essential to create increased awareness among dermatologists and pediatricians considering its rare occurrence and the potential for misinterpretation or unnecessary investigations. Early recognition and reassurance can prevent parental anxiety and help avoid excessive diagnostic procedures [5].

In this report, we present a case of ACH in a healthy child with no associated systemic abnormalities. We aim to highlight the typical clinical features, discuss the differential diagnoses, and the importance of clinical recognition and parental reassurance in managing this rare condition.

Case Report

A 4-year-old medically healthy female child presented to the outpatient clinic with a primary concern of localized excessive hair over the anterior aspect of the neck (Figure 1). The excessive hair growth was noted since birth and has remained unchanged in character or distribution. The excessive hair was localized and presented as a well-defined patch of terminal hair on the anterior neck. There were no associated skin lesions, swelling, pain, or other systemic symptoms. The child did not report any constitutional symptoms such as fever, weight loss,

XSL•FO
RenderX

or lethargy. A thorough systemic review was conducted and found to be unremarkable for any neurological complaints.

Figure 1. Clinical presentation of anterior cervical hypertrichosis (hairy throat syndrome) with localized excessive hair growth over the anterior neck, showing a well-defined distribution without underlying skin abnormalities.



Upon physical examination, the child appeared well and active. Examination of the neck showed the presence of a well-localized patch of thick terminal hair over the anterior cervical region, with no associated masses, discoloration, or tenderness. Neurological examination did not reveal any signs of central nervous system involvement such as abnormal reflexes, gait disturbance, or limb weakness.

The patient's past medical history is significant for ptosis, which has since been resolved without the need for surgical intervention. She also has a history of hypermetropia and is currently under regular follow-up with ophthalmology. Her developmental history is normal. There was no history of parental consanguinity. However, family history is notable for a sibling diagnosed with spina bifida, which prompted consideration of possible underlying spinal or neural tube abnormalities in this patient.

Based on the clinical findings and family history, a detailed evaluation plan was initiated. The patient was referred for further assessment by pediatric dermatology consultation. Parents were counseled and educated about the possible causes and implications of hypertrichosis, including rare associations with underlying spinal anomalies such as spinal dysraphism. Considering the reported associations of ACH with skeletal abnormalities and spinal dysraphism, along with the family history of spina bifida, selective investigations were performed. An x-ray of the foot was obtained to evaluate for skeletal abnormalities such as hallux valgus reported in previous cases, while a spinal x-ray was performed as an initial screening tool for occult spinal defects. Neck ultrasound was selected as a noninvasive modality to exclude underlying soft tissue abnormalities. As the patient had no neurological symptoms or abnormal neurological examination findings, advanced imaging such as magnetic resonance imaging was not deemed necessary. All investigations were within normal limits.

Ethical Considerations

Written informed consent for publication of the patient's clinical information and photographs was obtained from the patient's legal guardians. The study was conducted in accordance with institutional ethical standards and the principles of the Declaration of Helsinki. The authors obtained written consent from patients for their photographs and medical information to be published in print and online, and with the understanding that this information may be publicly available. Patient consent forms were not provided to the journal but are retained by the authors.

Discussion

ACH, also known as "hairy throat," is a rare form of localized hypertrichosis characterized by the presence of a localized patch of terminal hair on the front of the neck. Approximately 40 cases of ACH have been reported in the literature to date, as summarized in Table 1. While ACH is typically an isolated finding, it can be associated with systemic conditions such as neurological abnormalities (peripheral neuropathy, developmental delay, and mental retardation), ophthalmological disorders (optic atrophy, chorioretinal changes), hallux valgus,

XSL•FO
RenderX

and dorsal hypertrichosis. Therefore, it is highly recommended to take a detailed family history and conduct comprehensive clinical examinations and investigations (neurological and ophthalmological assessments, electromyography, and x-ray of the feet) for all patients with ACH to rule out any potential associated abnormalities.

XSL•FO
RenderX

Table . Literature review of reported cases of anterior cervical hypertrichosis: clinical features and associated findings.

Study	Age and gender	Description of ACH[a]	Neurological	Ophthalmological	Skeletal	Dermatological	Diagnosis	Treatment	Family history
Our case	4-year-old female	Localized excessive hair over anterior neck, present since birth, unchanged in character or distribution	No neurological complaints or findings	Ptosis (resolved), hypermetropia	X-ray of the foot normal (no skeletal anomalies)	Localized hypertrichosis over anterior cervical region	Clinical diagnosis	Referred to pediatric dermatology, parental counseling, investigations normal	Sibling with spina bifida
Chahoub et al [6] (2023)	12-year-old male	Localized patch of hair at the mid-neck region, fine, brown hairs, 3 cm in length	None reported	None reported	None reported	Congenital localized hypertrichosis	Dermoscopy and clinical examination	Laser hair removal (good response)	No family history mentioned
Sawatkar et al [7] (2023)	8-year-old female	Localized terminal coarse hair on anterior neck above the laryngeal prominence	None reported	None reported	None reported	ACH	Sporadic ACH, no consanguinity	None reported	No family history mentioned; sporadic ACH
Cutrone et al [8] (2022)	11-year-old girl and 3-year-old girl	Localized patch of terminal hair on the anterior neck, mostly placed on the hyoid region	None reported	None reported	None reported	None reported	Clinical examination	Laser hair removal	No family history mentioned
Kumar and Das [9] (2021)	8-year-old girl	Localized patch of hair over the laryngeal prominence	None reported	None reported	None reported	None reported	Clinical examination	Laser epilation offered but refused	No family history mentioned
Saini et al [10] (2021)	12-year-old girl	Patch of terminal hairs on the anterior neck above the laryngeal prominence	No	No	No	Yes (nevoid hypermelanosis)	Clinical examination	Laser hair removal	No family history mentioned
Blasco-Morente and Sánchez-Carpintero [5] (2017)	13-year-old female	Terminal hairs in the anterior midline region of the neck	None	None	None	Hypertrichosis	Isolated anterior cervical hypertrichosis	Laser epilation (partial improvement after 3 sessions)	None
Bostan et al [2] (2016)	15-year-old female	Localized patch of terminal hair on the anterior neck above the laryngeal prominence	None reported	None reported	None reported	None reported	Clinical examination, no special diagnostic approach mentioned	Laser hair removal	No family history mentioned
Megna et al (2015) [1][b]									

XSL•FO
RenderX

Study	Age and gender	Description of ACH[a]	Neurological	Ophthalmological	Skeletal	Dermatological	Diagnosis	Treatment	Family history
Megna et al (2015)	7-year-old Italian girl	Localized patch of terminal hair on the anterior neck	None reported	None reported	None reported	None reported	Clinical examination	Not specified	No history of hypertrichosis or any other skin or hair disease in her family
Meziane et al (2014)	4 patients aged 5 - 21 years	Patch of terminal hairs on anterior neck near cricoid cartilage	None reported	None reported	Yes (skeletal abnormalities), chronic juvenile idiopathic arthritis	Yes (hypertrichosis on other parts)	Clinical examination, x-rays, family history, histological examination	Laser hair removal, epilation	No family history mentioned
Reddy and Antaya (2010)	Two unrelated Hispanic females, aged 4 and 3 years	Solitary patch of excessive terminal hair growth in the midline of the neck	None reported	None reported	None reported	None reported	Clinical examination	Not specified	No family history mentioned
Echeverría et al (2010)	56-year-old woman and two daughters (27 and 35 years old, all female)	Congenital localized patch of terminal hair on anterior neck, hyperpigmented plaques with skin thickening on abdominal and lumbar areas	Dysesthesia (in daughter with Down syndrome), overlapping of fourth and third toes (daughter)	None reported	Overlapping toes (daughter with Down syndrome)	Morphea, congenital hypertrichosis	Familial anterior cervical hypertrichosis with systemic associations	Skin biopsy (morphea), liver function tests (chronic liver disease)	Present in 3 family members (maternal side)
Moreno-Giménez et al (2009)	27-year-old female (sporadic ACH case)	Congenital localized patch of terminal hair on anterior neck, no associated abnormalities	None reported	None reported	None reported	None reported	ACH	None reported	Sporadic, nonfamilial ACH in this case
Heitink et al (2007)	13-year-old female	Long, dark blond hair localized in anterior neck	None reported	None reported	None reported	None reported	Clinical examination, x-ray	Intense pulsed light treatment	One sister had spina bifida
Thienpont et al (2006)	No mention found	Anterior cervical hypertrichosis	Yes (mental retardation mildly dysmorphic facial)	Yes (hypermetropia)	Yes (skeletal abnormalities)	Yes (lumbosacral hypertrichosis)	Clinical examination, family history, imaging	Not specified	Family history of similar dysmorphic features and hypertrichosis in multiple affected relatives
Nanda et al (2006)	6 Arab children (aged 9, 3, 1, 12, 10, and 7 years, all female)	Localized hypertrichosis on anterior neck, present since birth or early childhood	None reported	None reported	None reported	Hypertrichosis	Isolated anterior cervical hypertrichosis	No treatment reported	Present in 5 cases (familial)

XSL•FO
RenderX

Study	Age and gender	Description of ACH[a]	Neurological	Ophthalmological	Skeletal	Dermatological	Diagnosis	Treatment	Family history
Corona-Rivera et al (2005)	Mexican boy, age not specified	Anterior cervical hypertrichosis	Yes (mental retardation, EEG[c] abnormalities, microcephaly)	None reported	Yes (hallux valgus)	Yes (hypertrichosis on back)	Clinical examination, EEG, brain MRI[d], x-rays	Laser hair removal	No family history mentioned
Hae-Woong Lee et al (2005)	28-year-old female	Increased fine hairs on the anterior cervical area	None reported	Mild myopia	None	Hypertrichosis	Familial anterior cervical hypertrichosis	Epilative laser therapy (refused)	Aunt and cousin also affected; no other abnormalities
Ardinger et al (1992)	13-year-old female and 38-year-old female (mother)	Patch of long hairs on anterior neck above the laryngeal prominence	None reported	None	None	ACH	None	None	None
Trattner et al (1991)	Case 1: 12-year-old male	Patch of long, fair, curly hair in the anterior cervical area	Impaired light touch and temperature perception, complete absence of pain perception in feet, reduced sensory action potentials, motor conduction abnormalities, spina bifida	Bilateral optic atrophy, central scotoma	Bilateral hallux valgus, osteolytic changes, fractures	Anterior cervical hypertrichosis	Osteomyelitis secondary to peripheral sensory neuropathy	Antibiotics (cefotaxime, piperacillin, amikacin, aztreonam, tobramycin, clindamycin), surgical drainage, immobilization	Multiple consanguineous marriages, affected mother and aunt
Trattner et al (1991)	Case 2: (mother) 30-year-old female	Long dark hair patch in the anterior cervical area, present since the first year of life	Decreased pain perception in both legs	Normal	Bilateral hallux valgus, mild kyphoscoliosis	Anterior cervical hypertrichosis	Peripheral sensory and motor neuropathy	No treatment mentioned	First cousin marriage, affected son and sister
Trattner et al (1991)	Case 3: (aunt) 32-year-old female	Localized patch of long, fair hair in the anterior cervical area, present since the first year of life	Subclinical peripheral sensory neuropathy in upper and lower extremities (discovered on EMG)	Normal	Bilateral hallux valgus	Anterior cervical hypertrichosis	Peripheral sensory neuropathy	No treatment mentioned	First cousin marriage, affected nephew and sister

[a]ACH: anterior cervical hypertrichosis.

[b]Studies in this category were identified in the review by Mengna et al [1].

[c]EEG: electroencephalogram.

[d]MRI: magnetic resonance imaging.

Trattner et al (as reviewed by Megna et al [1]) first reported ACH in three patients from Arab families with consanguineous marriages. All three patients exhibited peripheral neuropathy and bilateral hallux valgus, while one also had bilateral optic nerve atrophy and macular dysfunction. Tsukahara and Kajii [3] later documented ACH in seven members of a Japanese family across three generations with no other associated pathology found, except for one patient with Turner syndrome. ACH has been observed in individuals ranging from birth to early childhood [4], with both familial and sporadic cases reported. The inheritance patterns vary among hereditary cases, typically following an autosomal dominant pattern, though

XSL•FO
RenderX

autosomal recessive and X-linked dominant inheritance have also been noted [4]. Additionally, there are sporadic cases with no family history of the condition [2-5].

Our patient represents a sporadic presentation of ACH, as there was no family history of the condition or evidence of associated neurological or skeletal abnormalities on evaluation. Although the patient had a history of resolved ptosis and hypermetropia, no syndromic association was identified. Previous literature reviews have reported peripheral sensory and motor neuropathy as the most commonly associated abnormality, followed by hallux valgus and ophthalmological findings [4-10]. While ACH is primarily a benign cosmetic condition that may cause psychological distress, awareness of its occasionally reported systemic associations remains important. Appropriate clinical assessment, selective investigations guided by history and examination findings, and parental reassurance are essential components of management. Cosmetic treatment options including laser epilation, electrolysis, and intense pulsed light therapy may be considered in selected cases when aesthetic concerns arise [5].

Acknowledgments

All the views expressed in this paper are the authors' own views and not an official position of the institution.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Megna M, Balato N, Patruno C, Ayala F. Anterior cervical hypertrichosis: a case report and review of the literature. Pediatr Dermatol 2015;32(2):252-255. [doi: 10.1111/pde.12364] [Medline: 24916369]
2. Bostan S, Yaşar Ş, Serdar ZA, Gizlenti S. Anterior cervical hypertrichosis: a sporadic case. Turk Pediatri Ars 2016 Mar 1;51(1):49-51. [doi: 10.5152/TurkPediatriArs.2016.1315] [Medline: 27103865]
3. Tsukahara M, Kajii T. Hairy throat: a dominant trait affecting seven members of a family. Clin Dysmorphol 1992 Jul;1(3):165-167. [Medline: 1342865]
4. Nalluri R, Gilmour E, Brooke R. Anterior cervical hypertrichosis. Eur J Dermatol 2010;20(3):393-394. [doi: 10.1684/ejd.2010.0907] [Medline: 20400390]
5. Blasco-Morente G, Sánchez-Carpintero I. Hipertricosis cervical anterior aislada. Actas Dermosifiliogr 2017 Sep;108(7):672. [doi: 10.1016/j.ad.2016.02.027] [Medline: 27979307]
6. Chahoub H, Al Faker I, Marraha F, Rahmani N, Benyamna Y, Gallouj S. Anterior cervical hypertrichosis: a rare location. Our Dermatol Online 2023;14(1):100. [doi: 10.7241/ourd.20231.24]
7. Sawatkar G, Kowe PA. Anterior cervical hypertrichosis. Indian Pediatr 2023 Apr 15;60(4):328. [Medline: 37002853]
8. Cutrone M, Valerio E, Grimalt R. Clinical cases of anterior cervical hypertrichosis. Dermatol Open J 2022;7(1):6-8. [doi: 10.17140/DRMTOJ-7-148]
9. Kumar P, Das A. Anterior cervical hypertrichosis. Indian Dermatol Online J 2021;12(5):783. [doi: 10.4103/idoj.IDOJ_416_20] [Medline: 34667778]
10. Saini S, Kumar R, Yadav D. Anterior cervical hypertrichosis with linear and whorled nevoid hypermelanosis in a 12-year-old girl: a rare association. Indian J Paediatr Dermatol 2020;21(2):135-137. [doi: 10.4103/ijpd.IJPD_120_19]

Abbreviations

ACH: anterior cervical hypertrichosis

Edited by M Kirchberger; submitted 15.Mar.2026; peer-reviewed by A Roy, MK Drad; revised version received 29.May.2026; accepted 31.May.2026; published 14.Jul.2026.

Please cite as:
Shehata N, Thalib HI, Alahwal H
Anterior Cervical Hypertrichosis (Hairy Throat Syndrome): Pediatric Case Report and Brief Literature Review
JMIR Dermatol 2026;9:e95391
URL: https://derma.jmir.org/2026/1/e95391
doi:10.2196/95391

XSL•FO
RenderX

© Nancy Shehata, Husna Irfan Thalib, Heba Alahwal. Originally published in JMIR Dermatology (http://derma.jmir.org), 14.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Topical Recombinant Collagen Following Fractional Radio Frequency Microneedling for Nonsurgical Facelift: Case Report

Thien-Chong Marcus Wong[1*], MBBS, MMed, MRCS; Yizhi Ong[2*], MBChB, MSc

[1]Wong's Plastic Surgery Centre, Gleneagles Hospital, Singapore, Singapore

[2]Department for Continuing Education, University of Oxford, Rewley House, 1 Wellington Square, Oxford, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Yizhi Ong, MBChB, MSc
Department for Continuing Education, University of Oxford, Rewley House, 1 Wellington Square, Oxford, United Kingdom

Abstract

Background: Fractional radio frequency microneedling (FRM) is an effective nonsurgical facelift modality supported by growing clinical evidence. While generally well tolerated, postprocedural erythema, pain, swelling, and bruising typically last up to 7 days. Other serious complications include prolonged swelling, infection, persistent grid mark, burns, and hypo- or hyperpigmentation. Recombinant human collagen has demonstrated potent wound-healing and anti-inflammatory properties, which could be harnessed for post-FRM recovery.

Objective: To date, there is a lack of published studies evaluating the use of topical agents to improve postprocedural recovery following FRM. This case report describes the early experience of using topical recombinant collagen post FRM for nonsurgical facelift.

Methods: Two 40-year-old female patients with Fitzpatrick skin type IV underwent FRM for nonsurgical facelift. The procedure was performed with standard energy settings. Immediately post procedure, topical recombinant collagen was applied to the treated areas. Either clinical photography or cross-polarized imaging was used to objectively capture the changes in facial erythema and swelling.

Results: Both cases demonstrated a marked reduction of erythema and inflammation within 1 to 2 hours after application of topical recombinant collagen. These observations appeared earlier compared to the typical timeframe of 1 to 3 days. No adverse events were reported.

Conclusions: Topical recombinant collagen application post FRM was associated with an early reduction of skin erythema and inflammation.

(JMIR Dermatol 2026;9:e86892) doi:10.2196/86892

KEYWORDS

recombinant human collagen; postprocedure recovery; fractional radio frequency microneedling; wound healing; nonsurgical facelift

Introduction

In recent years, fractional radio frequency microneedling (FRM) has become a widely recognized and highly effective nonsurgical facelift treatment [1-3]. Studies have shown that it is a useful modality for skin tightening and dermal remodeling without the risks associated with invasive surgery [1-3]. For FRM, controlled thermal energy via microneedles is delivered into the dermis, inducing microinjury to stimulate neocollagenesis and elastin production [1-3]. Although its clinical efficacy and tolerability have been well documented, postprocedural erythema, edema, bruising, and discomfort are common and may persist for up to several days [4]. Other rare but serious complications that have been observed include prolonged swelling, persistent grid marks, burns, infection, and hypo- or hyperpigmentation [4].

Currently, no standardized postprocedural care protocol has been established to accelerate recovery and reduce downtime [4], underscoring the unmet clinical need for biologically active adjuncts that support tissue repair and reduce inflammation. Recombinant human collagen is an emerging innovative biomaterial that has shown promising potential for improving wound healing while minimizing inflammation [5-8], which could be harnessed for post-FRM recovery following energy-based facial rejuvenation procedures. To our knowledge, the use of topical recombinant collagen in the immediate post-FRM setting has not been described in the medical literature. Herein, we describe an innovative postprocedure

XSL•FO
RenderX

approach using topical recombinant collagen in two patients who underwent a nonsurgical facelift with FRM.

Case Presentation

Overview

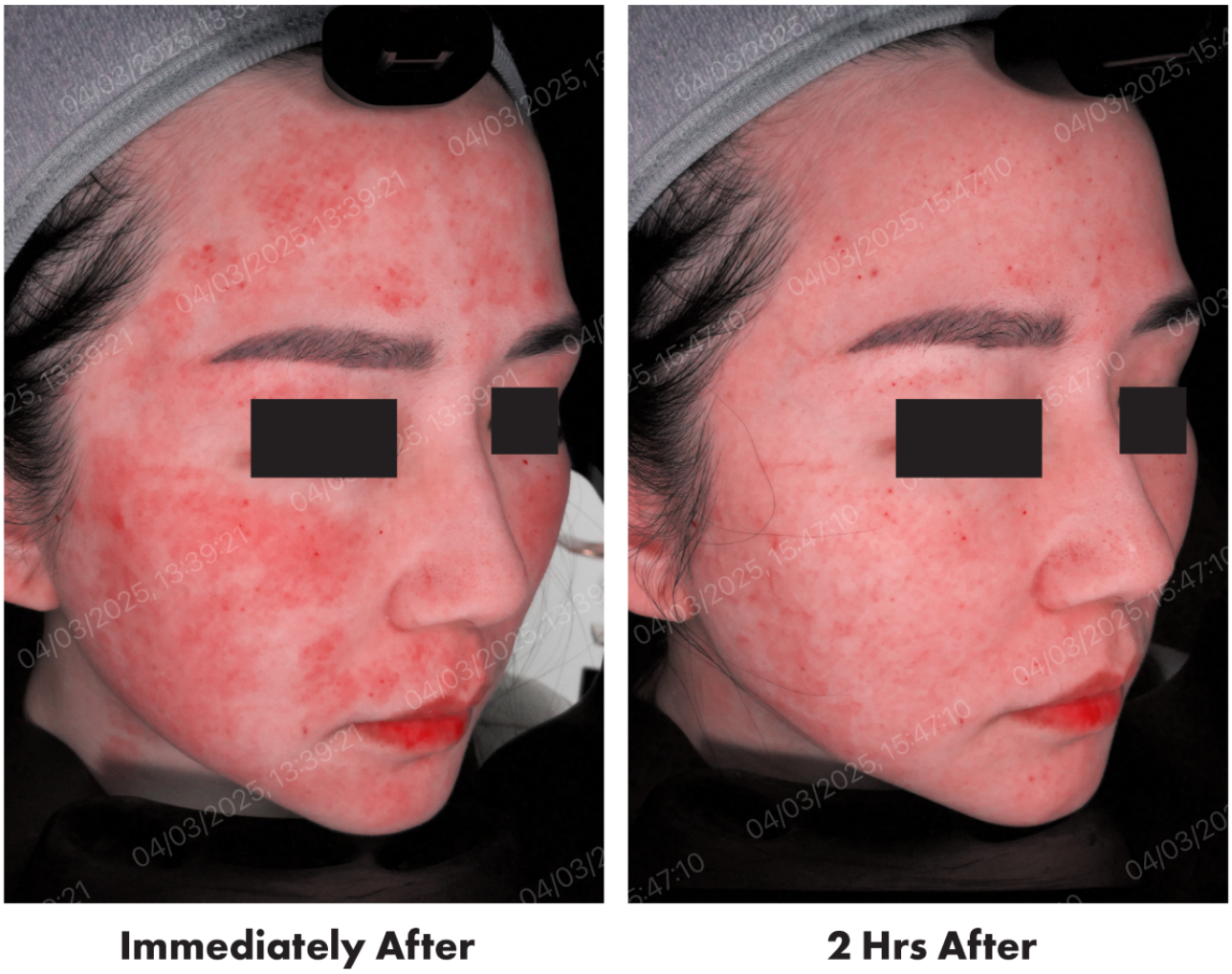
Two 40-year-old female patients with Fitzpatrick skin type IV presented with facial laxity and underwent FRM for nonsurgical facelifting. Both were nonsmokers, with no significant past medical history or drug allergies. The baseline assessment documented mild-to-moderate facial laxity with fine lines around various parts of the face. Standardized clinical photography under consistent lighting or cross-polarized imaging was performed before treatment, immediately after treatment, and 1 to 2 hours after treatment. The procedure was performed using a 24-pin tip in a fixed mode, with two passes and 35% to 50% overlap. They received a total of approximately 600 pulses each with standard energy settings of 20 to 25 mJ at a 2.0 mm depth for facial bony areas and 30 to 45 mJ at a 3.0 to 4.0 mm depth for facial soft tissue regions. Immediately post procedure, topical recombinant collagen was applied to both patients as part of their routine posttreatment care. The multimodal documentation showed early observable changes in erythema and inflammation within 1 to 2 hours after application of topical recombinant collagen. Case 1 demonstrated a clinically significant reduction in facial edema and erythema (Figure 1) that was observed on clinical photography within 1 hour post treatment. Additionally, the patient reported immediate soothing of skin discomfort and burning sensation after application of topical recombinant collagen. Case 2 showed a marked reduction of inflammation (Figure 2), visualized using a 3D skin analyzer under cross-polarized light within 2 hours post treatment. Subjectively, the patient reported rapid relief of skin redness and irritation after application of topical recombinant collagen. No adverse events or side effects were reported. The procedure and posttreatment care were well tolerated by both patients. During the follow-up review, both patients were satisfied with the treatment and rapid recovery results. The CARE Checklist has been completed for the drafting of this case report.

Figure 1. Clinical photography under standard lighting capturing apparent reduction in facial erythema and edema within 1-hour post fractional radio frequency microneedling.



XSL•FO
RenderX

Figure 2. Marked improvement of skin inflammation visualized by a 3D skin analyzer under cross-polarized light within 2 hours post fractional radio frequency microneedling.



Ethical Considerations

This case report did not require institutional ethics review according to the National Healthcare Group Domain Specific Review Board guidelines, as it describes observations from routine clinical care without systematic research intervention.

Written informed consent was obtained from the patients for publication of this case report and any accompanying images.

Discussion

This novel case report has shown that topical recombinant collagen applied immediately post FRM was associated with clinically apparent reduction in erythema, edema, and discomfort within 1 to 2 hours post treatment as compared to the expected 1 to 3 days [1-3]. The temporal relationship between the application of topical recombinant collagen and the earlier-than-expected postprocedural reduction in erythema and inflammation could represent a plausible effect superimposed on the normal recovery process. The reduction in skin erythema was apparent to the attending clinician, which was notably different from the typical recovery trajectory of FRM. The observed objective and subjective improvements suggest that topical recombinant collagen may be a feasible adjunct for postprocedural care following FRM. Collagen-based biomaterials have been shown to promote wound healing, stabilize the extracellular matrix, and reduce inflammation in both acute and chronic wounds [5-7]. While previous studies have largely focused on burns or surgical wounds [5-7], our findings have extended the potential application of recombinant collagen to aesthetic procedures involving controlled dermal injury from FRM. This early clinical experience corroborated with known in-vitro dermal repair and anti-inflammatory mechanisms of recombinant collagen, including the modulation of inflammatory cytokines such as inducible nitric oxide synthase, interleukin 1β (IL-1β), IL-10, and tumor necrosis factor α [9]. Postprocedural downtime is an important consideration. These preliminary observations warrant further investigation into strategies to support recovery following FRM.

Nevertheless, this report has several limitations. Only two patients were included with short-term follow-up. There was no control group for comparison. No standardized scoring systems for erythema, swelling, or patient-reported outcomes were used, as this was not part of routine clinical documentation. Randomized controlled studies will be required to confirm treatment safety and efficacy while optimizing treatment protocols before routine implementation.

XSL•FO
RenderX

In summary, topical recombinant collagen may represent a potential adjunct in postprocedural care following FRM, warranting further evaluation in prospective clinical trials.

Funding

This study was not supported by any sponsor or funder.

Data Availability

All data generated or analyzed during this study are included in this published article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest

None declared.

References

1. Tan MG, Jo CE, Chapas A, Khetarpal S, Dover JS. Radiofrequency microneedling: a comprehensive and critical review. Dermatol Surg 2021 Jun 1;47(6):755-761. [doi: 10.1097/DSS.0000000000002972] [Medline: 33577211]
2. Dayan E, Rovatti P, Aston S, Chia CT, Rohrich R, Theodorou S. Multimodal radiofrequency application for lower face and neck laxity. Plast Reconstr Surg Glob Open 2020 Aug 26;8(8):e2862. [doi: 10.1097/GOX.0000000000002862] [Medline: 32983756]
3. Flegontova E, Kreindel M, Vranis NM, Mulholland RS. Correction of age-related changes in the skin at the dermal and subdermal level using radiofrequency macroneedling therapy. J Cosmet Dermatol 2024 Jul;23(7):2401-2410. [doi: 10.1111/jocd.16361] [Medline: 38778550]
4. Shauly O, Marxen T, Menon A, Gould DJ, Miller LB, Losken A. Radiofrequency microneedling: technology, devices, and indications in the modern plastic surgery practice. Aesthet Surg J Open Forum 2023 Nov 6;5:ojad100. [doi: 10.1093/asjof/ojad100] [Medline: 38887534]
5. Mathew-Steiner SS, Roy S, Sen CK. Collagen in wound healing. Bioengineering (Basel) 2021 May 11;8(5):63. [doi: 10.3390/bioengineering8050063] [Medline: 34064689]
6. Chattopadhyay S, Raines RT. Review collagen-based biomaterials for wound healing. Biopolymers 2014 Aug;101(8):821-833. [doi: 10.1002/bip.22486] [Medline: 24633807]
7. Sundar G, Joseph J, C P, John A, Abraham A. Natural collagen bioscaffolds for skin tissue engineering strategies in burns: a critical review. Int J Polymeric Mater Polymeric Biomater 2020 Mar 24;70(9):593-604. [doi: 10.1080/00914037.2020.1740991]
8. Ben C, Liu X, Shen T, et al. A recombinant human collagen hydrogel for the treatment of partial-thickness burns: a prospective, self-controlled clinical study. Burns 2021 May;47(3):634-642. [doi: 10.1016/j.burns.2020.01.006] [Medline: 33402311]
9. Hu C, Li K, Li L, et al. Recombinant human collagen injection ameliorates skin photoaging via immunoregulation and collagen production. Collagen Leather 2025 Jul 2;7(22):1-18. [doi: 10.1186/s42825-025-00203-y]

Abbreviations

FRM: fractional radio frequency microneedling
IL-1β: interleukin-1β

Edited by M Kirchberger; submitted 31.Oct.2025; peer-reviewed by N Kumar, P Brieva; revised version received 15.Apr.2026; accepted 07.May.2026; published 29.Jun.2026.

Please cite as:
Wong TCM, Ong Y
Topical Recombinant Collagen Following Fractional Radio Frequency Microneedling for Nonsurgical Facelift: Case Report
JMIR Dermatol 2026;9:e86892
URL: https://derma.jmir.org/2026/1/e86892
doi:10.2196/86892

© Thien-Chong Marcus Wong, Yizhi Ong. Originally published in JMIR Dermatology (http://derma.jmir.org), 29.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Informed Consent Practices for Publication of Patient Images in Dermatology Journals

Toluwani Taiwo[1*], BA; Bianca Obiakor[1*], BA; Sarah McClung[2], MIS; Kanade Shinkai[3], MD, PhD

[1]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]University of California, San Francisco, San Francisco, CA, United States

[3]Department of Dermatology, University of California, San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Kanade Shinkai, MD, PhD
Department of Dermatology
University of California, San Francisco
1701 Divisadero St
San Francisco, CA, 94115
United States
Phone: 1 415 353 7800
Email: kanade.shinkai@ucsf.edu

Related Article:

Correction of: https://derma.jmir.org/2025/1/e60795

(JMIR Dermatol 2026;9:e94194) doi:10.2196/94194

In "Informed Consent Practices for Publication of Patient Images in Dermatology Journals" [1], the authors made one correction.

In Multimedia Appendix 1, we included a Dryad link to our raw dataset, which includes Clarivate data. We have recently learned that Clarivate's terms of use and Dryad's data reuse policy are incompatible, and we will not be able to deposit the raw data in Dryad. The raw dataset can be made available by contacting the corresponding author.

In Multimedia Appendix 1, the following text has been revised:

> *The raw dataset for this study is available for review via the Dryad platform and accessible via the following unique link:http://datadryad.org/share/k6Eiw-OfOGf0vJrxD6sOS8GC59H5SFuhtudjzHR1EH8.*

The text now reads:

> *The raw dataset for this study is available by contacting the corresponding author.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1
List of the top 50 dermatology journals ranked by the 2023 Clarivate Journal Citation Report and a link to the publicly available raw dataset used in the study.
[DOCX File , 17 KB - derma_v9i1e94194_app1.docx]

Reference

1. Taiwo T, Obiakor B, McClung S, Shinkai K. Informed Consent Practices for Publication of Patient Images in Dermatology Journals. JMIR Dermatol 2025 Apr 18;8:e60795 [FREE Full text] [doi: 10.2196/60795] [Medline: 40249653]

XSL•FO
RenderX

Edited by P Kharbanda; submitted 26.Feb.2026; this is a non–peer-reviewed article;accepted 27.Feb.2026; published 13.Mar.2026.

Please cite as:
Taiwo T, Obiakor B, McClung S, Shinkai K
Correction: Informed Consent Practices for Publication of Patient Images in Dermatology Journals
JMIR Dermatol 2026;9:e94194
URL: https://derma.jmir.org/2026/1/e94194
doi:10.2196/94194
PMID:

©Toluwani Taiwo, Bianca Obiakor, Sarah McClung, Kanade Shinkai. Originally published in JMIR Dermatology (http://derma.jmir.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Gender Wage Disparities in Dermatology

Danielle Duffle[1], MS, DO; Ramiro Rodriguez[2], MD; Benjamin Brooks[3], MBA, PhD; Robert Dellavalle[4], MSPH, MD, PhD

[1]Department of Dermatology, OhioHealth Riverside Methodist Hospital, Columbus, OH, United States

[2]Department of Dermatology, Anschutz Medical Campus, University of Colorado School of Medicine, Aurora, CO, United States

[3]Department of Clinical Education, Rocky Vista University College of Osteopathic Medicine, Irvins, UT, United States

[4]Department of Dermatology, University of Minnesota School of Medicine, 1 Veterans Drive, Minneapolis, MN, United States

Corresponding Author:
Robert Dellavalle, MSPH, MD, PhD
Department of Dermatology, University of Minnesota School of Medicine, 1 Veterans Drive, Minneapolis, MN, United States

Abstract

Differences in compensation between men and women exist, despite having similar experiences. Here, we present data from the Association of American Medical Colleges (AAMC) from 2022 of self-reported compensation rates of academic dermatologists. Women dermatologists comprised 36.95% (75/203) of full professors, 61.54% (152/247) of associate professors, and 63.47% (311/490) of assistant professors; among these, 52.17% (12/23) were service chiefs and 35.48 % (22/62) were department chairs. The data shows that, except for department chairs, women are reporting being paid less relative to men at all academic ranks. Division chiefs had the most significant (45.52%) wage difference (45.52%) between men and women at US $281,674. This is a $281,674 difference between men and women chiefs per year. Women earned US $91,738.73 less on average per year when considering all academic positions. Different career trajectories and societal pressures to take time off for family-related absences are two of many factors contributing to the observed differences. We conclude that engagement at a national, institutional, and personal level is needed to better understand the observed differences, and future research should focus on how to improve it.

(JMIR Dermatol 2026;9:e65003) doi:10.2196/65003

KEYWORDS

female physicians; women physicians; gender gap; dermatology; leadership in dermatology; pay disparities in medicine; nonlinear career trajectories; Claudia Goldin; Nobel Prize in Economic Sciences 2023; work-life balance in medicine; gender representation in dermatology; parental leave policies; women dermatologists; gender bias in medicine; career advancement for women physicians

"Female physicians do the impossible" [1]. Claudia Goldin, awarded the 2023 Nobel Prize in Economic Sciences for her research on gender disparities in labor markets, provides an economic framework that helps explain gender-based differences observed within dermatology [1]. Goldin provides insights into women's pay gap and leadership underrepresentation, which occurs despite more women entering dermatology (1013/1633, 62.03% in 2022 - 2023) [2]. While female representation has improved, we hypothesized that disparities in pay within leadership positions persist despite increased female representation [3,4]. Using data from the Association of American Medical Colleges (AAMC), we illustrate leadership and wage differences between men and women in dermatology and analyze these trends within Goldin's framework [3,5]. The purpose of this editorial is to help improve the understanding of the persistent differences between men and women in dermatology, despite similar academic rankings. Limitations of this analysis include its moment-in-time nature. Changes in salary trends could offer additional information on progress toward equity. Additionally, the AAMC did not include data on respondents' age or years of experience within leadership roles.

Percentages of women in dermatology by academic rank, leadership roles, and salary were contrasted with those of men using data from the AAMC Faculty Salary Report 2022 [3]. χ^2 analysis tested statistical significance ($P<.05$). Women comprise 36.95% (75/203) of full professors, 61.54% (152/247) of associate professors, 63.47% (311/490) of assistant professors, 52.17% (12/23) of chiefs, and 35.48 % (22/62) of department chairs (Figure 1). Division chiefs had a US $281,674.29 (45.52%) difference in mean income between female (–26.79%) and male dermatologists (+18.73%; Figure 2). Excluding Mohs training increased this gap to US $320,420.80.

XSL•FO
RenderX

Figure 1. Percentage of women and men in dermatology academic leadership roles in the Association of American Medical Colleges (AAMC) Faculty Salary Report 2022 [3] (57.73%). Department chair salaries had a US $1505.33 difference (Figure 2). There were significantly more men than women in chair and full professor positions, and significantly more women than men in associate and assistant professor positions in 2022. There was no significant difference between the number of women and men in chief roles in 2022.

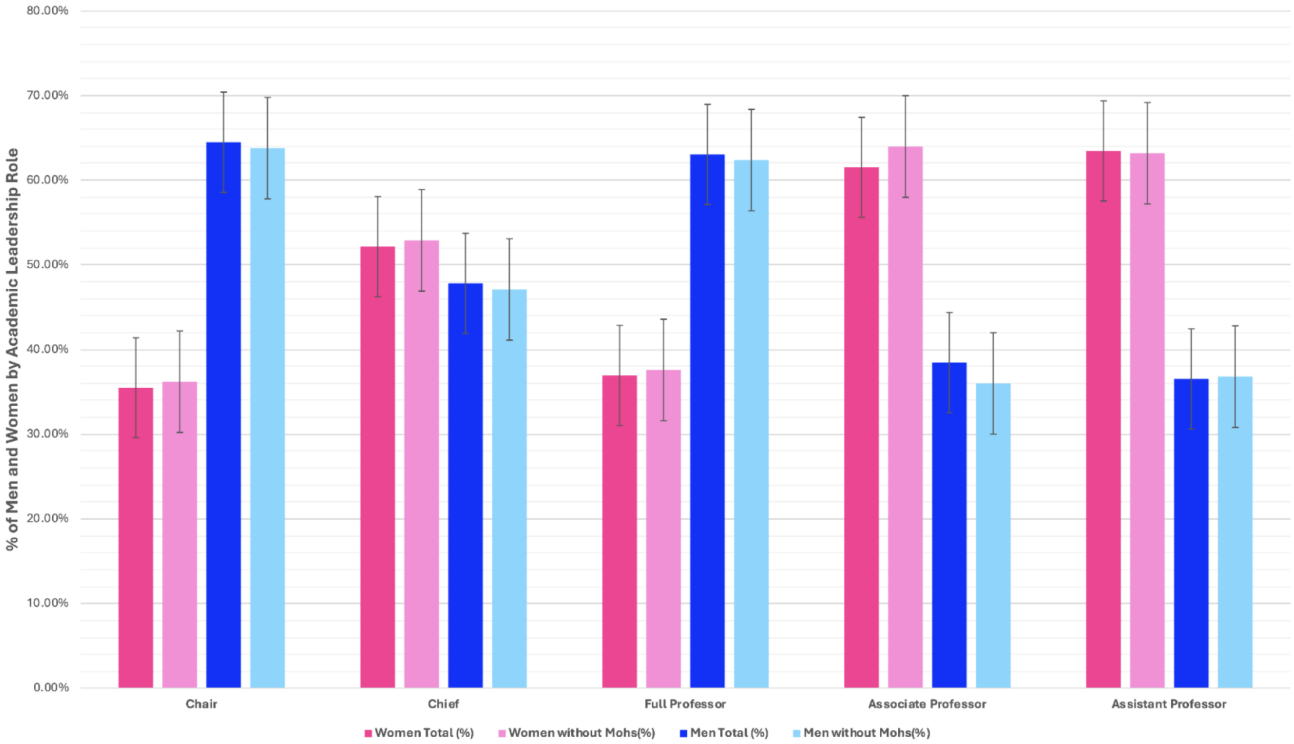
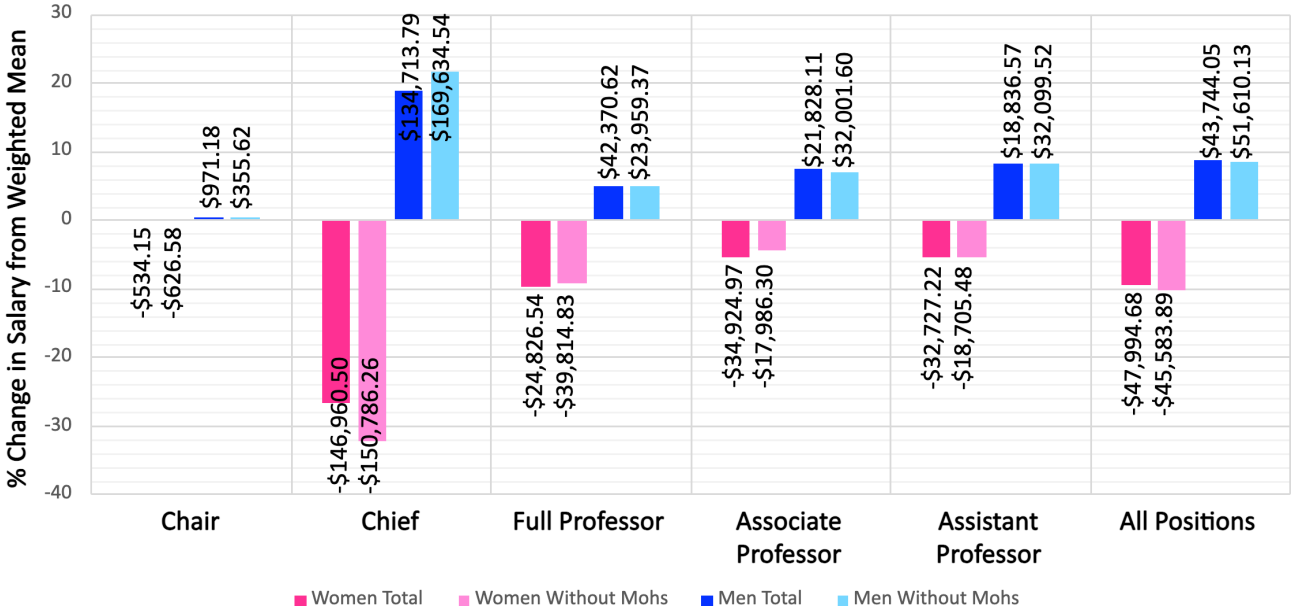


Figure 2. Percent change from weighted mean salary for women and men in dermatology academic leadership roles in 2022. Salary data from the Association of American Medical Colleges (AAMC) Faculty Salary Report 2022 [3]. A weighted mean salary was calculated for each position—compared between men and women—and the percentage change from the mean is displayed. Women reported salaries below the weighted mean, while men reported salaries above it in all dermatology academic leadership roles in 2022. For chair positions, the percent change for women (–0.15%) and men (0.08%) is too small to be visualized. All currency amounts are in US dollars.



These trends show increased female representation within academic ranks in dermatology relative to previous work [5,6] and increased female associate professors with a persistent and increased pay gap relative to prior AAMC surveys (2013 - 2018) [6]. To better understand these trends, Goldin [1,7] highlights different types of job/role trajectories: linear with long continuous hours and nonlinear with more flexibility. Goldin [7] describes "nonlinear" occupations as those offering flexibility but with reduced wage progression—a pattern consistent with dermatology's structure. Despite higher earnings enabling care services, female physicians still disproportionately take time off for family-related absences [1], leading to missed leadership opportunities, potentially explaining why female representation drops higher up dermatology's leadership ladder.

XSL•FO
RenderX

Given the time required for late-career promotions, we anticipate more women chairs and full professors in the coming decade. Despite differing representations among roles, women still earn less than men in all academic dermatology positions, which is consistent with the wage gap Goldin describes. This trend may reflect women's diverse roles but underscores a broader pattern of pay inequality within dermatology leadership.

Encouraging wage equity should occur at national, institutional, and personal levels. Nationally and specific to this survey, the AAMC could enhance the questions exploring what information was available during the contract negotiation and how family caretaking responsibilities impact career advancement. Individuals can engage with other national associations, such as the American Medical Women's Association, the Women's Dermatologic Society, and the American Academy of Dermatology, through their diversity/equity initiatives to explore mentorship opportunities, advocacy roles, and professional advancement. Institutionally, focus groups could evaluate how parental leave policies, flexible work (ie, teledermatology), mentorship programs, and leadership training impact promotion. Stay interviews could help identify department-specific areas of improvement, and assigning merit toward collaboration in publications could help enhance the academic development of women who are faculty members [8]. Focus groups could then be best informed on mitigating systemic bias against women, although Goldin [1] highlights that discrimination is not the primary driving force of the gender wage gap.

This editorial highlights persistent gender pay differences among academic dermatologists. We urge the AAMC to explore these differences through more nuanced survey questions in the faculty salary report. By integrating Goldin's insights, we aim to promote an equitable environment in dermatology, recognizing the complex interplay between gender, compensation, and career progression.

Acknowledgments

Generative AI was not used for this study or for writing the manuscript.

Funding

The authors declare no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

RD, DD, and BB participated in study conceptualization. DD and RR contributed to data curation; formal analysis; and writing, reviewing, and editing the original draft. RD and BB supervised and reviewed and edited the original draft.

Conflicts of Interest

RR is an Assistant Editor and RD is the Editor-in-Chief of *JMIR Dermatology* at the time of this publication. RR and RD had no involvement in the editorial review and processing of this manuscript. All other authors have no conflicts to declare.

References

1. Goldin C. Career and Family: Women's Century-Long Journey Toward Equity: Princeton University Press; 2021. [doi: 10.1515/9780691226736]
2. ACGME Data Resource Book 2022-2023: Accreditation Council for Graduate Medical Education; 2023. URL: https://www.acgme.org/about/publications-and-resources/graduate-medical-education-data-resource-book/ [accessed 2026-09-04]
3. AAMC Faculty Salary Report. Association of American Medical Colleges. 2022 Nov 15. URL: https://tinyurl.com/3v3mauj8 [accessed 2026-09-04]
4. Kamath P, Muthiah N, Venkatesh A, Pugliano-Mauro M. Gender and leadership: continued progress in academic dermatology. J Am Acad Dermatol 2023 Aug;89(2):394-396. [doi: 10.1016/j.jaad.2023.03.054] [Medline: 37080339]
5. Walter S, Murrell DF. Gender equity in academic dermatology: problems aplenty, yet paths ahead. J Eur Acad Dermatol Venereol 2024 Aug;38(8):1504-1513. [doi: 10.1111/jdv.20027] [Medline: 38606617]
6. Sachdeva M, Price KN, Hsiao JL, Shi VY. Gender and rank salary trends among academic dermatologists. Int J Womens Dermatol 2020 May 13;6(4):324-326. [doi: 10.1016/j.ijwd.2020.05.005] [Medline: 33015295]
7. Goldin C. A grand gender convergence: its last chapter. Am Econ Rev 2014 Apr 1;104(4):1091-1119. [doi: 10.1257/aer.104.4.1091]
8. Allan JM, Brooks AK, Crusto C, et al. Five strategies leaders in academic medicine can implement now to enhance gender equity. J Med Internet Res 2023 Jun 13;25:e47933. [doi: 10.2196/47933] [Medline: 37310782]

Abbreviations

AAMC: Association of American Medical Colleges

XSL•FO
RenderX

Edited by T Leung; submitted 07.Aug.2024; this is a non–peer-reviewed article;accepted 22.Mar.2026; published 30.Sep.2026.

Please cite as:
Duffle D, Rodriguez R, Brooks B, Dellavalle R
Understanding Gender Wage Disparities in Dermatology
JMIR Dermatol 2026;9:e65003
URL: https://derma.jmir.org/2026/1/e65003
doi:10.2196/65003

© Danielle Duffle, Ramiro Rodriguez, Benjamin Brooks, Robert Dellavalle. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Based Grading of Facial Aging Signs and Prediction of Visible Age From Standardized Facial Images: Algorithm Development and Internal Evaluation Study

Frederic Flament[1*], PhD; Panagiotis-Alexandros Bokaris[1*], PhD; Tao Li[1*], PhD; Aurelie Abric[2*], BSc; Sileye Ba[1*], PhD; Julien Despois[1*], MSc; Hussein Jouni[1*], MSc; Laudine Bertrand[1*], MSc; Paul Michel[1*], MSc; Thomas Dauxerre[1*], BSc; Diala Haykal[3*], MD; Hughes Cartier[4*], MD

[1]L'Oréal Research and Innovation, 9 rue Pierre Dreyfus, Clichy, France

[2]Eurosyn, Villebon-sur-Yvette, France

[3]Centre Médical Laser Palaiseau, private practice, Palaiseau, France

[4]Centre Médical Saint Jean, Arras, France

[*]all authors contributed equally

Corresponding Author:
Frederic Flament, PhD
L'Oréal Research and Innovation, 9 rue Pierre Dreyfus, Clichy, France

Abstract

Background: Facial appearance influences visible age (VA), defined here as the age perceived by others; this perception-based outcome differs from chronological age and biological age. Human evaluation of facial aging signs can be time-consuming and subject to interobserver variability. A standardized AI-based approach might offer a more reproducible assessment of visible facial aging in digital dermatology research.

Objective: This study aimed to develop and evaluate an AI-based process for automatically grading facial aging signs and predicting the VA of participants from standardized full-face images.

Methods: Evaluation of full-face photographs of 468 North American women aged 18 to 80 years recruited in Dallas and Phoenix, Texas. Fifteen trained experts, including dermatologists, graded 13 aging signs using Skin Aging Atlas scales. A separate, naïve group of 253 North Americans aged 20 to 60 years, recruited in New York City, estimated the VA of each participant. For each sign, facial areas were pulled out using facial landmarks and fed into ResNet-based regression models, which were trained to match the averaged expert scores. The automated facial sign grading models were tested on 90% of the image set, and the remaining 10% was used for evaluation. Results were assessed using the mean absolute error, the proportions of estimates within 0.5, 1.0, and 1.5 grading units of the expert reference, and Spearman correlations between chronological age, human-assessed VA, and predicted visible age (PVA).

Results: The mean absolute error was below 0.3 grading units for 11 of the 13 facial aging signs. The average proportion of automatic gradings within 1 grading unit of the expert reference was about 98% for participants of East Asian ancestry group, 96% for participants of African ancestry group, and 97% for participants of European ancestry group. Human-assessed VA correlated strongly with real age (r=0.94, 95% CI 0.928 - 0.950; P<.001), although participants were perceived as around 10 years older than their real age, on average. The AI-PVA also correlated strongly with the human-assessed VA (r=0.93, 95% CI 0.917 - 0.941; P<.001), with approximately 70% of estimates within 5 years.

Conclusions: The AI-based system showed low grading errors across most facial aging signs and also a fairly strong relationship between the PVA and human-assessed VA under standardized image acquisition conditions. Consequently, the system could provide a reproducible digital outcome for facial aging research and future longitudinal monitoring. However, external validation across separate populations, in men, and across image acquisition devices is required before clinical and consumer-facing implementation. As the predicted visible age is derived from quantified clinical facial signs, it reflects measurable features of visible facial aging, while its relationship with biological age and overall health remains to be established through future studies integrating biological biomarkers.

(JMIR Dermatol 2026;9:e93236) doi:10.2196/93236

KEYWORDS

AI; visible age; biological markers; longevity; image analysis; clinical grading; facial signs

XSL•FO
RenderX

Introduction

The lifespan of humans, otherwise called life expectancy, has shown a rather constant and significant progression in recent decades. According to the World Health Organization, global life expectancy increased from 66.8 years in 2000 to 73.1 years in 2019, and healthy life expectancy increased from 58.1 years to 63.5 years throughout the same time frame [1]. These figures vary, however, widely among adults with regard to genetic inheritance, exposure to life-threatening diseases, and geography. This large diversity introduces the notion of health span, that is, the years lived in good health, free from chronic disease and functional decline, including mental health [2-4].

This has reshaped the frontiers of the science of longevity, that is, the use of integrative strategies involving cosmetics, devices, digital services, coaching, oral supplements, and so on, far beyond a mere topical antiaging cosmetic solution. Facial skin, as a marker of the aging process, whether chronological and/or photoinduced, displays different aspects over decades and is an important element of health span, including well-being [5-8]. Previous studies have shown that these changes vary substantially across ethnic ancestries and genders [9-12]. Especially as a mirror of inner and invisible changes, facial skin could pave the way to exploring new holistic skin "clocks." Yet, in the absence of biological measurements, visible facial aging should not be deemed equivalent to biological aging.

In this framework, different notions of age need to be differentiated. Real age (RA) refers to chronological age, whereas visible age (VA) refers to more or less the age perceived by others from facial appearance. Predicted visible age (PVA), as used in the current study, refers to an algorithm-derived estimate of the same VA. Meanwhile, biological age represents physiological or molecular aging and typically depends on biological markers. Although these dimensions seem related, they are not quite interchangeable. Previous work has already shown that perceived age can be linked with age-related phenotypes, functional status, molecular markers, and survival, supporting its interest as a perception-based aging outcome [13,14]. Recently, deep learning methods have likewise suggested the possibility of predicting perceived facial age from facial images [15].

Within this evolving paradigm, cosmetic dermatology is thus emerging as more than a discipline of aesthetic refinement, and it has become a meaningful participant in promoting health span and contributing to broader integrative longevity strategies. VA (eg, age as perceived by others) may, therefore, represent a digital health and perception-based outcome. The extent to which VA mirrors biological age and could be compared with RA (chronological age, as presented on an ID card) is one facet, among others, of understanding individual aging and relationships with patients and consumers for all actors [16-27]. The present study does not investigate VA as a measure of biological age. In short, fitting the VA of a participant with some biological markers of the skin is an ambitious project that could ideally be performed in real-life conditions, such as points of sale (POS) for cosmetic products or clinics with doctors and practitioners. From a technical viewpoint, recent years have shown fruitful help from smartphones as image collectors, thanks to their efficient cameras and high resolution (≥5 megapixels). Smartphones equipped with adequate algorithms are increasingly used and can allow the prompt analysis of images from very large cohorts of participants in real life or at home [28-31]. Such image-based approaches are of special interest in digital dermatology, where more standardized and reproducible evaluation of visible facial features could complement conventional clinical or expert grading.

As for the biological aspect, the recent introduction of systems for fast assessments of biomarkers will hopefully yield valuable data related to the root causes of skin aging, viewed as longevity "clocks." Proteomic patterns, for example, can be extracted from adhesive patch sampling of the skin surface [32] and link invisible and visible aging manifestations. As these kinds of approaches are currently used to establish biological age, fitting their future results with a PVA remains instrumental in developing the most coherent ecosystem to accompany patients and consumers in engagement, compliance, and long-term monitoring. Hence, such integration represents a future perspective, as no biological markers were evaluated in the current study.

Although there is growing interest in AI-based facial assessment, approaches to combining standardized facial image capture, expert grading, and independent human assessment of VA are still needed. The question remains: do we really get more reliable and less variable results when using automatic grading? To increase diversity within the development set, the study included North American women spanning various ages and ethnic backgrounds, and covered a broad range of facial skin tones and visible aging indicators. Basically, the work aimed to design and test an AI-driven pipeline for automatically scoring facial aging signs and estimating VA using standardized full-face photographs. It also intended to compare this automatic grading with expert assessments and examine the relationships among RA, human-assessed VA, and AI-PVA across the ancestry groups represented in the study [33-35].

Methods

Study Design

The study followed a 5-fold structure: the acquisition of standardized full-face images from a cohort of North American women; the grading of visible facial signs by a panel of trained experts; the building of an AI-based automatic grading system; the evaluation of VA by an independent, naïve panel; and, finally, the calculation of the PVA.

Ethical Considerations

The study was approved by L'Agence nationale de sécurité du médicament et des produits de santé (ANSM, France) under reference number 2025-A02830-49. Oral and written consent, in accordance with the Declaration of Helsinki, was obtained from all participants related to the study objectives and procedures. The consent also included that images would be taken under anonymized conditions, blind-coded, and deleted after study completion. Participants also consented to the use of their images for publication.

XSL•FO
RenderX

Panel of Full-Face Photography

A total of 468 North American women (aged 18–80 y) were selected through a local agency and photographed in local facilities (Dallas, Texas, and Phoenix, Arizona, USA) using the Nexa device (Canfield Beauty) for standardization (position, lighting, etc). Figure 1 illustrates the device and 3 examples of the photographs provided, showing ethnic backgrounds and skin tones. Two inclusion criteria were used: (1) age between 18 and 80 years and (2) absence of any facial disorder (scars, acne, melasma, seborrheic dermatitis, telangiectasia, etc). The panel consisted of participants of different ages and ethnic ancestries, as presented in Table 1.

Figure 1. The Nexa device and 3 examples of full-face images from different skin tones.



Table . Distribution of the 468 North American women with clinical standardized pictures dataset according to ethnic ancestries and age groups.

Ethnic ancestry	18-30 y, n	31-40 y, n	41-50 y, n	51-60 y, n	61-70 y, n	71-80 y, n	Total, n
East Asian	16	16	14	26	20	8	100
African ancestry	28	25	23	15	28	6	125
European	52	56	48	29	42	16	243
Total	96	97	85	70	90	30	468

Panel of Trained Experts for Grading Facial Signs

A panel of 15 trained experts (including dermatologists), previously used on different occasions for comparable tasks, was asked to grade the severity of 13 facial signs (Table 2) according to reference scoring ranges obtainable from ethnic-specific skin atlases [36,37] under standardized conditions. Each expert grader was seated about 1 m away from a 24-inch high-resolution display with a resolution of 1920 × 1200 pixels. The screen was calibrated daily using a colorimeter. To ensure robustness of this process, several images were presented twice during the evaluation process without being identified as duplicates. The high number of experts allowed for the exclusion of graders in cases of erratic scores.

XSL•FO
RenderX

Table . Thirteen clinical facial signs and their respective photographic scale's ranges.

Facial signs	Definition	Scale range
Forehead wrinkles	Depth of the transverse wrinkles on the forehead	0 - 8
Glabellar wrinkles	Depth of vertical wrinkles between eyebrows	0 - 6
Crow's feet wrinkles	Depth of wrinkles at the area of the outer eye corner	0 - 6
Periorbital wrinkles (upper cheek area)	Depth of folds at malar area below Crow's feet, eye orbit excepted	0 - 9
Nasolabial fold	Depth of the fold present between the base of the nose and lips	0 - 7
Marionette lines	Depth of folds at the corner of lips	0 - 8
Pigmentary spots (density)	Number of spots per area unit on the cheek	0 - 7
Ptosis of lower part of the face	Sagging severity of the lower parts on each side of the chin	0 - 6
Diffuse redness	Diffuse redness and microvessels visible, especially on cheeks	0 - 4
Cheek skin pores	Size of visible pores on the cheek irrespective of their densities	0 - 5
Interocular wrinkles	Depth of horizontal folds between inner eye corners	0 - 7
Perioral lip wrinkles	Depth of wrinkles on superior labial area	0 - 6
Eye bags	Intensity of the ocular bulge	0 - 6

For each facial sign and photograph, the mean of the scores given by the experts was used as the reference score. Despite using integer grades on the photographic scales, the averaged expert scores provided continuous reference values and enabled the automatic system to account for subgrade variations.

Naïve Panel Attributing VA to Photographs

A total of 253 North American women of different ethnic backgrounds were recruited in New York City, USA, as a naïve panel by a local agency dedicated to part-time occupations. These participants were aged 20 to 60 years, with 48% (n=121) aged between 20 and 39 years and 52% (n=132) aged between 40 and 60 years, with good visual acuity as a unique inclusion criterion. They were informed about the objective of their task and signed an informed consent form, ensuring full anonymity. The whole panel, when successively viewing the 468 blind-coded photographs presented at random, was asked to answer a simple question: "How old is this woman?" Their answers then referred to VA and were linked to their RA.

AI-Based Automatic Grading System

Based on the expert's grading, an AI-based automatic grading system was developed to automatically compute the severity of the 13 facial signs. To assess facial clinical signs with high accuracy, the AI-based automatic grading system is framed as a regression task, where a dedicated model is trained to estimate a continuous severity score for each targeted clinical sign from a cropped facial image.

First, an off-the-shelf landmark detection algorithm is run to compute key points used as reference landmarks on the face, such as the tip of the nose, inner corners of the eyes, etc. Based on these landmarks, a cropping step uses handcrafted heuristics to extract the relevant zones of the face as defined by the clinical grading atlases.

The chosen backbone was ResNet, a streamlined convolutional neural network using residual connections to improve prediction accuracy [38,39]. The architecture consists of 5 convolutional residual blocks, progressively reducing the spatial dimensions to a $1 \times 1 \times 512$ feature map, followed by 3 fully connected layers with rectified linear unit (ReLU) activations that reduce the output dimensionality to a single scalar (the predicted clinical severity score). Each of the 13 facial signs was handled by a dedicated model instance, resulting in a holistic multistep diagnostic pipeline that takes a full-face image as input and yields a list of sign-specific predicted scores as output.

Functionally, each model takes as input a crop of the image that corresponds to the target clinical sign (eg, forehead wrinkles, periorbital wrinkles, and marionette lines) and predicts a scalar value within skin aging atlas grading scales (Table 2) [36,37]. The average grades provided by 15 experts were used as reference labels. Although the scales are defined with integer values (grade 0, grade 1, etc), the averaged expert annotations allowed the models to account for subgrade variations. The pooled labeled data from 15 expert annotators, trained to evaluate with high precision, allow us to account for subgrade accuracy by targeting the average prediction.

Model Training and Internal Evaluation

Each model was trained using the following configuration:

- Loss function: difference between the predicted score and the average of expert annotations
- Optimizer: Adam [40] with $\beta_1=0$, $\beta_2=0.99$
- Learning rate: 0.0005

XSL•FO
RenderX

- Batch size: 32
- Training duration: 300 epochs

To prevent overfitting and maximize generalizability, two complementary strategies were used. First, early stopping was monitored using an internal validation subset, carved out from the training data (10% of the 90% training subset), with training halted if validation loss did not improve over a defined patience window. Second, data augmentation was systematically applied during training, including the addition of chromatic noise, shifts in hue and lighting intensity, and small random rotations of input crops. These augmentations simulated real-world variability in capture conditions (eg, minor lighting inconsistencies and slight head movements) and reduced the risk of overfitting to the controlled photographic conditions of the training set, thereby improving robustness when deployed in live settings.

An automatic grading system was trained and evaluated on a standardized dataset that had been scored by experts and dermatologists. The dataset was split into training and test subsets, with 90% of the data used for training the model and 10% used to evaluate the capacity of the system to generalize to images not encountered during training [22,33]. Within the 90% training subset, a further internal validation split (10% of the training data) was used exclusively for early stopping and model selection purposes, while the remaining training data were used for model fitting. The split was stratified by ethnic ancestry and age group to ensure balanced representation across subgroups. Participant-level separation was strictly enforced to prevent any form of data leakage.

To ensure the resulting system is as robust and accurate as possible, the training procedure includes a data augmentation step, where input crops are randomly degraded or altered to teach the system to ignore such perturbations. Despite the use of controlled lighting and poses, data augmentation minimizes the risk that unforeseen circumstances degrade the performance of the system. The data augmentation includes the addition of chromatic noise, shifts in hue and lighting, as well as small rotations of the crops.

The crops are fed into the corresponding models to obtain a predicted grade of clinical severity, thereby linking with PVA. This results in a holistic algorithm that functions as a multistep diagnosis tool, taking a full-face image as input and yielding a list of predicted scores as the output.

PVA Derivation

Based on the validated automatic grading outputs, PVA was computed from the 10 facial signs found to be significantly correlated with VA as assessed by the naïve panel (Table 3; cheek skin pores (CSP), pigmentary spots (PS) density, and diffuse redness (DR) were excluded due to low correlation with VA). For each retained sign i, a second-order polynomial regression of the form $Y = A_i x_i^2 + B_i x_i + C_i$ was fitted using ordinary least squares on the training subset exclusively, where x_i was the automatically predicted facial sign grade, Y was the associated VA estimate, as assessed by the naïve panel, and A_i, B_i, and C_i were the regression coefficients. The PVA of each photographed participant was obtained by averaging the polynomial outputs across all 10 retained signs. Coefficients A_i, B_i, and C_i for each sign, as derived from the training set, are reported in Table 4. The generalization performance of the PVA model was then assessed on the held-out test subset by computing the Spearman correlation between the PVA and the VA assessed by the naïve panel (N=468), yielding r=0.93 (P<.001). This strict train/test separation ensures that the reported correlation reflects a genuine out-of-sample prediction and provides a robust safeguard against overfitting.

XSL•FO
RenderX

Table . Overall agreement statistics for the extended facial sign/imaging condition analyses.

Facial sign/imaging condition	MAE[a]	RMSE[b]	ICC[c] (95% CI)	Percentage within 1 grade (%)
Crow's feet wrinkles (n=173)	0.314	0.478	0.949 (0.930-0.960)	93
Eyebags (n=173)	0.184	0.365	0.906 (0.870-0.930)	96
Pigmentary spots (polarized; n=186)	0.879	1.080	0.659 (0.130-0.840)	66
Pigmentary spots (nonpolarized; n=186)	0.226	0.306	0.954 (0.940-0.970)	98
Cheek skin pores (n=186)	0.239	0.339	0.895 (0.860-0.920)	98
Marionette lines (n=186)	0.218	0.321	0.976 (0.970-0.980)	98
Forehead wrinkles (n=92)	0.278	0.359	0.937 (0.910-0.960)	98
Glabellar wrinkles (n=92)	0.240	0.468	0.912 (0.870-0.940)	98
Interocular wrinkles (n=92)	0.401	0.544	0.909 (0.860-0.940)	96
Nasolabial fold (n=186)	0.215	0.338	0.970 (0.960-0.980)	96
Ptosis (n=92)	0.303	0.466	0.886 (0.830-0.920)	93
Periorbital wrinkles (n=186)	0.201	0.283	0.967 (0.960-0.970)	99
Perioral lip wrinkles (n=92)	0.341	0.493	0.878 (0.820-0.920)	93
Diffuse redness (nonpolarized; n=92)	0.204	0.258	0.948 (0.920-0.970)	100
Diffuse redness (polarized; n=92)	0.177	0.218	0.963 (0.940-0.980)	100

[a]MAE: mean absolute error.

[b]RMSE: root mean squared error.

[c]ICC: intraclass correlation coefficient.

Table . Definition of coefficients for the equation for predicted visible age (PVA) based on facial clinical signs.

Facial signs/coefficients	A	B	C
Crow's feet wrinkles	−1.43	18.07	14.13
Eye bags	−4.70	24.26	34.38
Marionette lines	−1.07	12.13	34.22
Forehead wrinkles	−1.78	15.59	33.33
Glabellar wrinkles	−2.75	20.06	32.19
Interocular wrinkles	−1.60	16.03	28.73
Nasolabial fold	−1.55	14.90	34.99
Ptosis of the lower part of the face	−2.22	18.45	31.89
Periorbital wrinkles	−2.10	19.04	27.57
Perioral lip wrinkles	−3.27	22.82	30.71

Statistics

The level of agreement between the grades provided by the AI-based system and the experts was analyzed globally and stratified by ethnic ancestry based on metrics including MAE, root mean squared error (RMSE), intraclass correlation coefficient (ICC), Bland-Altman analysis, and calibration statistics. The number of paired observations varied by sign, reflecting the imaging protocol (not all signs were captured for every participant at every session). MAE was defined as the mean of the absolute differences between the automatically predicted grades and the averaged expert reference grades. For each facial sign, the proportions of individual predictions with an absolute error below 0.5, 1.0, and 1.5 grading units were also generated.

The correlations between PVA and VA, on the one hand, and PVA and RA, on the other hand (2 paired methods to be compared), were computed. Their significance was assessed

XSL•FO
RenderX

using Spearman coefficient correlation tests overall, considering all photographed volunteers (N=468).

Results

Assessment Panel

The standardized photographic dataset consisted of 468 North Americans aged between 18 and 80 years old, including 100 women of East Asian ancestry, 125 of African ancestry, and 243 of European ancestry, spread across the different age groups, as shown in Table 1.

The expert grading panel comprised 15 trained experts, including dermatologists with previous experience in facial sign grading. The naïve panel included 253 North American women from different ethnic profiles, aged between 20 and 60 years old, with approximately 50% aged between 20 and 39 years and 50% aged between 40 and 60 years.

Experts Panel Consensus and Variability

Table 5 illustrates the scoring dataset established by the expert panel, including MAEs calculated at different grade thresholds according to the Skin Aging Atlas classification. Most facial skin signs appear to be correctly assessed, with an acceptable MAE of around 0.5, except for periorbital wrinkles and the density of pigmentary spots (approximately 0.6).

XSL•FO
RenderX

Table . Values of the MAE in the scores of 13 facial signs recorded by 15 experts and dermatologists at different selected thresholds.

Facial sign	MAE[a]	Percent |error|<0.5	Percent |error|<1.0	Percent |error|<1.5
Panel A: East Asian ancestry				
Forehead wrinkles	0.44	66	93	99
Glabellar wrinkles	0.37	74	96	99
Crow's feet wrinkles	0.52	58	88	97
Periorbital wrinkles	0.59	54	82	94
Nasolabial fold	0.44	67	90	97
Marionette lines	0.52	59	88	96
Pigmentary spots (density)	0.62	51	80	93
Ptosis of the lower part of the face	0.49	57	91	99
Diffuse redness	0.37	72	97	100
Cheek skin pores	0.52	57	87	97
Interocular wrinkles	0.55	55	86	96
Perioral lip wrinkles	0.43	68	92	99
Eye bags	0.45	66	90	98
Panel B: African ancestry				
Forehead wrinkles	0.46	63	90	98
Glabellar wrinkles	0.39	73	95	99
Crow's feet wrinkles	0.53	57	87	96
Periorbital wrinkles	0.59	52	84	95
Nasolabial fold	0.49	65	88	96
Marionette lines	0.51	62	88	96
Pigmentary spots (density)	0.66	46	79	92
Ptosis of the lower part of the face	0.49	63	91	98
Diffuse redness	0.37	79	94	97
Cheek skin pores	0.59	53	82	94
Interocular wrinkles	0.51	57	88	97
Perioral lip wrinkles	0.47	62	91	99
Eye bags	0.43	69	92	97
Panel C: European ancestry				
Forehead wrinkles	0.48	63	89	97
Glabellar wrinkles	0.38	73	95	99
Crow's feet wrinkles	0.49	61	89	98
Periorbital wrinkles	0.61	53	82	93
Nasolabial fold	0.45	65	90	97
Marionette lines	0.53	57	87	96
Pigmentary spots (density)	0.55	55	85	95
Ptosis of the lower part of the face	0.55	54	87	97
Diffuse redness	0.41	67	94	99

XSL•FO
RenderX

Facial sign	MAE[a]	Percent |error|<0.5	Percent |error|<1.0	Percent |error|<1.5
Cheek skin pores	0.49	60	90	98
Interocular wrinkles	0.53	57	88	97
Perioral lip wrinkles	0.45	64	92	98
Eye bags	0.39	73	93	98

[a]MAE: mean absolute error.

Prediction Potential of the Automatic Scoring System

Table 6 summarizes the performance of the AI-based automatic grading system. On average across all 13 signs, considering the condition of MAE<1, which is the accepted error during a live evaluation by a clinician based on a standardized photographic scale, reliable accuracy is achieved for East Asian (~98%), African ancestry (~96%), and European (~97%) vs experts' clinical ground truth.

XSL•FO
RenderX

Table . Performance of the AI-based automatic grading system.

Facial sign	MAE[a]	Percent |error|<0.5	Percent |error|<1.0	Percent |error|<1.5
Panel A: East Asian ancestry				
Forehead wrinkles	0.24	88	100	100
Glabellar wrinkles	0.18	94	100	100
Crow's feet wrinkles	0.28	78	97	97
Periorbital wrinkles	0.21	90	100	100
Nasolabial fold	0.22	85	97	100
Marionette lines	0.23	87	95	100
Pigmentary spots (density)	0.25	87	95	100
Ptosis of the lower part of the face	0.27	88	94	100
Diffuse redness	0.15	100	100	100
Cheek skin pores	0.16	100	100	100
Interocular wrinkles	0.38	62	100	100
Perioral lip wrinkles	0.33	81	100	100
Eye bags	0.21	86	97	100
Panel B: African ancestry				
Forehead wrinkles	0.3	88	92	100
Glabellar wrinkles	0.4	77	92	96
Crow's feet wrinkles	0.21	94	98	100
Periorbital wrinkles	0.17	94	100	100
Nasolabial fold	0.22	88	94	100
Marionette lines	0.18	94	98	100
Pigmentary spots (density)	0.24	84	96	100
Ptosis of the lower part of the face	0.29	81	92	96
Diffuse redness	0.2	100	100	100
Cheek skin pores	0.29	90	96	96
Interocular wrinkles	0.44	77	92	96
Perioral lip wrinkles	0.24	85	96	100
Eye bags	0.16	93	96	100
Panel C: European ancestry				
Forehead wrinkles	0.28	87	100	100
Glabellar wrinkles	0.17	96	100	100
Crow's feet wrinkles	0.35	78	90	98
Periorbital wrinkles	0.22	86	99	100
Nasolabial fold	0.21	89	97	99
Marionette lines	0.24	91	99	100
Pigmentary spots (density)	0.21	92	100	100
Ptosis of the lower part of the face	0.31	81	96	100
Diffuse redness	0.18	98	100	100

XSL•FO
RenderX

Facial sign	MAE[a]	Percent |error|<0.5	Percent |error|<1.0	Percent |error|<1.5
Cheek skin pores	0.25	92	99	100
Interocular wrinkles	0.38	77	96	100
Perioral lip wrinkles	0.39	72	89	96
Eye bags	0.17	91	98	98

[a]MAE: mean absolute error.

Considering the average of all ethnicities, the MAE values were below 0.3 for 11 signs, except for interocular and perioral lip wrinkles, which showed higher values. Interestingly, the comparison of results between Tables 5 and 6 clearly indicates that the automatic grading system induces less variability and better reliability than that of the panel of trained experts, regardless of ethnicity or facial clinical sign. Indeed, the MAE values obtained suggest 2 or 3 times lower errors and variability for the automatic grading system compared with the experts' panel.

Additional Agreement Analyses

Additional analyses were conducted to further investigate how well the automatic grading system aligned with the expert assessments. Since pigmentary spots and vascular disorders were evaluated under both polarized and nonpolarized imaging conditions, these analyses included 15 facial aging sign/imaging condition pairings derived from the 13 facial signs assessed in the study.

Among these 15 sign/imaging condition combinations, the MAE was 0.292 grading units, and the RMSE was 0.419 grading units. The higher RMSE compared to MAE suggests the presence of some larger outlier prediction errors even though the average error remained low. Furthermore, more than 90% of the automatic predictions were within 1 grading unit of the expert reference for 14 of the 15 sign/imaging condition combinations. The main outlier was CPS under polarized conditions, where only 66% of the predictions were within 1 grading unit of the expert score.

The ICC values appeared to show good to excellent agreement for most of the evaluated signs. For CPS, especially under polarized lighting, the agreement was comparatively lower, with an ICC of approximately 0.66. Signs like upper lip wrinkles, ptosis, and cheek pores exhibited agreement within the good range, while most of the remaining sign/imaging condition combinations showed excellent agreement. The full ancestry-stratified statistics are provided in Multimedia Appendix 1.

Bland-Altman analyses showed mean biases close to zero for most signs, indicating little systematic overestimation or underestimation at the group level. Some proportional bias was observed at higher severity grades, whereas the limits of agreement were generally within approximately ±1 grading unit (Multimedia Appendix 2). Calibration plots indicate good linearity between the AI-automated grading and expert scores for most signs, with regression slopes close to 1.0 and intercepts close to 0. Similar to observations with higher MAE and lower ICC, deviations from the identity line were noted for CPS (polarized), whereas nonpolarized images demonstrated better-calibrated predictions for pigmentation-related signs (Multimedia Appendix 3).

PVA

Figure 2 presents correlation coefficients between the 13 facial clinical grades predicted by the automatic grading system and the VA assigned by the naive panel. It clearly shows that most facial signs are interrelated in the perception of age, with the exception of CSP, PS (density), and diffuse redness (DR). They will, therefore, not be retained in the subsequent calculation of PVA.

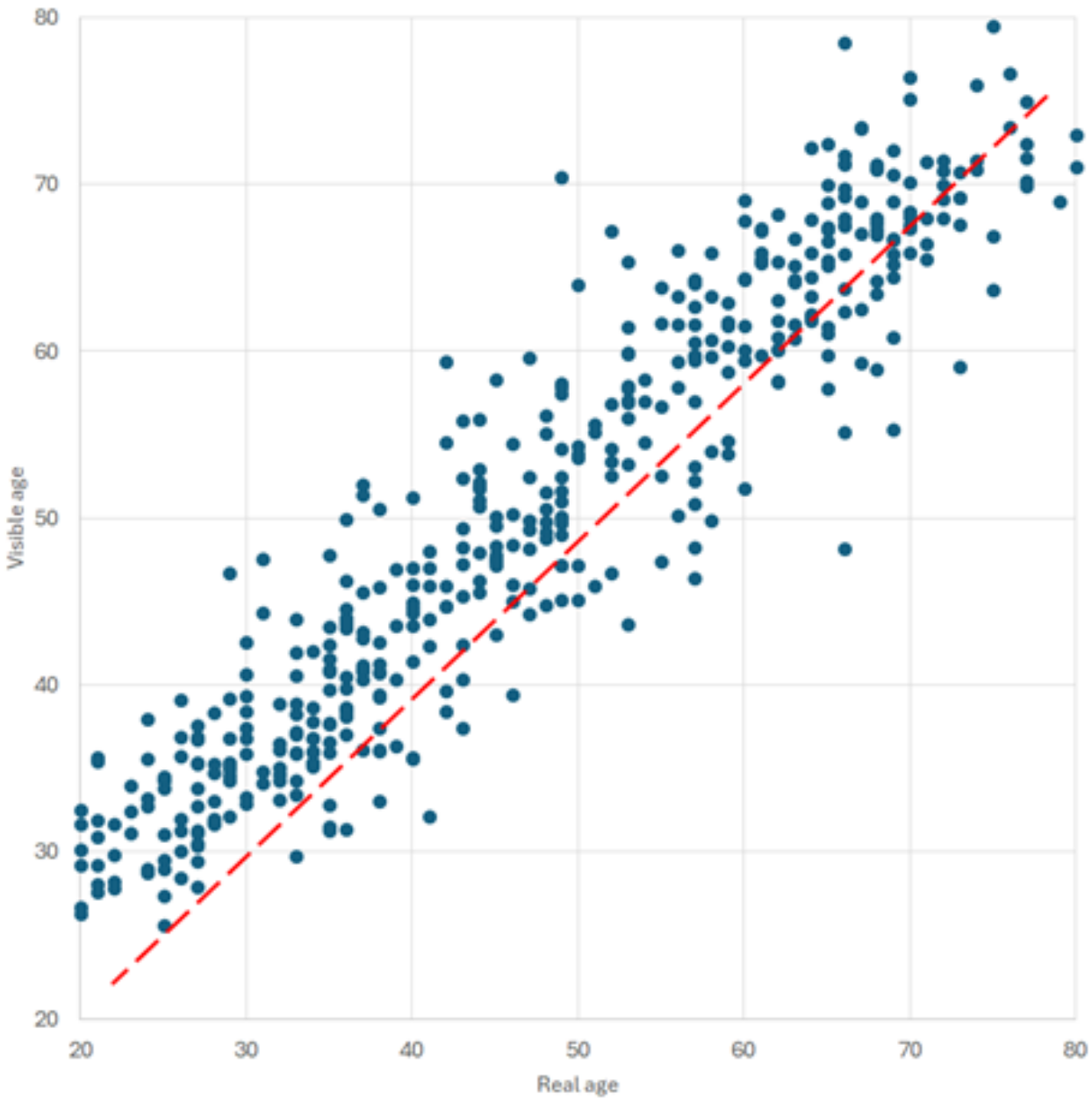
Figure 3 presents correlation coefficients between RA and the ages perceived by a panel of 253 US women. Although these appear strongly correlated (r=0.94; P<.001), the 468 photographed women were perceived to be older than they are by about 10 years in a large majority.

Figure 2. Correlation coefficients between all facial signs predicted by the AI-based automatic grading system and visible age (VA). The colored boxes show the levels of correlation with VA: green: ≥0.70; yellow: ≤0.5 and >0.7; orange: ≤0.2 and >0.5; red: ≥−0.1 and <0.2. CFW: crow's feet wrinkles; CSP: cheek skin pores; DR: diffuse redness; EB: eye bags; FW: forehead wrinkles; GW: glabellar wrinkles; IOW: interocular wrinkles; ML: marionette lines; NF: nasolabial fold; PLF: ptosis lower part face; PLW: perioral lip wrinkles; PS: pigmentary spots; PW: periorbital wrinkles.

	CFW	EB	PS	CSP	ML	FW	GW	IOW	NF	PLF	PW	PLW	DR	VA
CFW	1.00	0.47	0.25	0.01	0.63	0.48	0.68	0.71	0.60	0.70	0.71	0.70	0.19	0.75
EB	0.47	1.00	0.32	0.07	0.47	0.31	0.51	0.52	0.50	0.49	0.52	0.47	-0.04	0.66
PS	0.25	0.32	1.00	0.08	0.26	0.11	0.19	0.23	0.22	0.25	0.21	0.18	0.00	0.30
CSP	0.01	0.07	0.08	1.00	-0.02	0.06	0.06	0.02	0.07	-0.06	0.12	-0.10	0.00	0.06
ML	0.63	0.47	0.26	-0.02	1.00	0.40	0.56	0.59	0.58	0.79	0.59	0.67	0.22	0.70
FW	0.48	0.31	0.11	0.06	0.40	1.00	0.37	0.48	0.51	0.42	0.48	0.42	0.01	0.53
GW	0.68	0.51	0.19	0.06	0.56	0.37	1.00	0.66	0.61	0.61	0.60	0.60	0.08	0.69
IOW	0.71	0.52	0.23	0.02	0.59	0.48	0.66	1.00	0.65	0.62	0.61	0.62	0.08	0.76
NF	0.60	0.50	0.22	0.07	0.58	0.51	0.61	0.65	1.00	0.62	0.60	0.56	0.00	0.75
PLF	0.70	0.49	0.25	-0.06	0.79	0.42	0.61	0.62	0.62	1.00	0.66	0.71	0.20	0.74
PW	0.71	0.52	0.21	0.12	0.59	0.48	0.60	0.61	0.60	0.66	1.00	0.65	0.16	0.74
PLW	0.70	0.47	0.18	-0.10	0.67	0.42	0.60	0.62	0.56	0.71	0.65	1.00	0.23	0.73
DR	0.19	-0.04	0.00	0.00	0.22	0.01	0.08	0.08	0.00	0.20	0.16	0.23	1.00	0.08
VA	0.75	0.66	0.30	0.06	0.70	0.53	0.69	0.76	0.75	0.74	0.74	0.73	0.08	1.00

XSL•FO
RenderX

Figure 3. Correlation between the visible ages (assessed by the naïve panel) of the 468 US participants and their real ages (r=0.94).

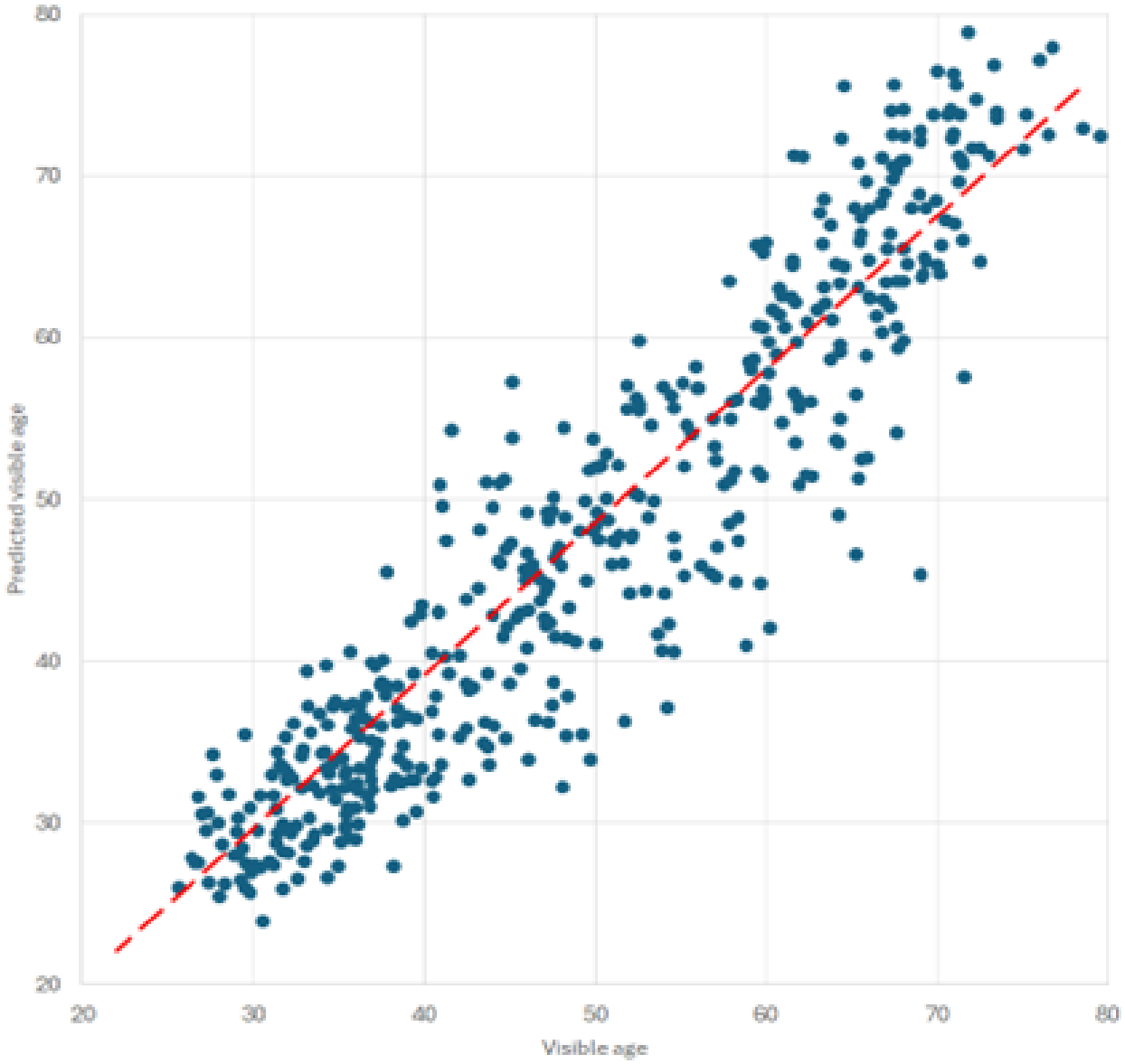


For each of the 10 retained signs, a second-order regression of the type $Y = Ax^2 + Bx + C$ (where x is the grade established by the automatic grading system) is adopted, where A, B, and C coefficients of each of the retained signs are indicated in Table 4. The sum of all values of all $A_i x_i^2 + B_i x_i + C$, when averaged, allows obtaining the PVA of each of the 468 US participants.

PVA was strongly correlated (r=0.93; P<.001) with the VA evaluated by the naïve panel (Figure 4). Nevertheless, this figure shows that despite the good correlation, gaps were found within the 40- and 60-year age ranges. However, about 70% of assessments were within 5 years of the VA assessed by the naïve panel.

XSL•FO
RenderX

Figure 4. Correlation between predicted visible ages of the 468 US participants and their visible ages, as assessed by the naïve panel (r=0.93).



Discussion

Principal Findings

The present work aimed at analyzing how an AI-based process could predict the VA of a given participant from a standardized full-face photograph in a robust manner, while reducing the maximum of biases arising from human assessment. The study paired standardized image capture, clinical grading by experts, an independent naïve panel for VA assessment, and automatic grading of facial aging signs. The NEXA photographic system used in this study provided standardized image acquisition and good image quality together with its ease of use (rapidity, reproducibility, etc), that is, the primary conditions to allow the most precise analysis by a naïve panel, an expert group, or an automatic AI-based process. It is nevertheless clear that attributing a precise VA to a person from a single facial aspect is by no means an easy task. In real life, other criteria are used, such as hair presence or style, body positioning (straight or bent), walking pace, etc, as well as reflecting cultural, geographical, or environmental specificities. Facial traits remain, however, a suitable model of skin aging, abundantly studied as they integrate intrinsic biological changes and external impacts (sun exposure and lifestyles), as extensively reviewed [41].

A relatively low prediction errors for most facial signs were obtained when the results provided by the automatic system were compared with those from the expert panel. These findings were further confirmed by additional analyses, with good to excellent agreement obtained for most evaluated sign/imaging condition combinations. In the Bland-Altman analyses, the mean biases approached zero for most signs, and calibration overall was satisfactory. Still, performance was not homogeneous across all facial signs. Pigmentary spots exhibited lower agreement when assessed under polarized lighting. These findings suggest that the performance of the automatic grading system partially depends on the facial sign and imaging condition evaluated.

XSL•FO
RenderX

VA and Human Perception

With regard to the assessments by the naïve panel, comparing VA and RA, although highly correlated (r=0.94; P<.001), leads to blurred results. Why this panel "viewed" almost all participants as much older (by 10 y or more at younger ages) than they really were is an open question [9,12]. In contrast to the blind/neutral feature of the automatic system, the perception of the panel is processed at a glance that rapidly integrates all facial aspects, which might attribute more "weight" to some facial regions or signs (upper or lower face, ptosis, etc) than others; this is a possibility that could be checked through an eye-tracking methodology (not explored here). To what extent such focus on specific facial regions obeys cultural or socioeconomic influences is another open question. In contrast, the automatic system provides, in the end, a realistic predictive estimate of the VAs of the 468 women studied, although Figure 3 shows larger gaps in the 40 and 60 years range, noting that these follow a normal Gaussian distribution. It is likely that such a result first illustrates the period when the process of facial aging leads to a larger variability among women with regard to hormonal changes (perimenopausal and postmenopausal periods). The latter indeed induce profound and widely varying effects among the women concerned in both physiological (skin included) and psychological status [42-45].

Second, the spontaneous attribution of a VA to a person likely differs in complexity according to age; that is, it is probably more precisely established at age ranges where markers of aging are either not very pronounced (18–30 y) or already settled (>60 y). It is noteworthy that 3 signs related to homogeneity (PS, CSP, and DR) showed low correlations with VA and were not included in the PVA calculation. When different ancestries are considered together, facial signs such as wrinkles and sagging may contribute more consistently to the overall perception of VA.

Performance Across Ancestry Groups

The choice of performing the study in the United States enabled the inclusion of women from different ancestry groups and age ranges. The automatic grading system showed generally high agreement with expert assessments across the ancestry groups represented in the study. The additional analyses, however, showed variation depending on both facial sign and ancestry group. Hence, the overall performance estimates should not be read as exactly the same for every subgroup. In a diverse US female population, similar work had also demonstrated the feasibility of automated grading of facial signs using selfies, which depended on facial features, especially those associated with pigmentation [46].

These findings are important in the context of potential algorithmic bias. Although the inclusion of women of East Asian ancestry, African ancestry, and European ancestry increased diversity within the development dataset, it does not establish fairness or generalizability to all populations. In particular, performance across age groups was not separately characterized to the same extent as ancestry-related performance. Further evaluation on larger and more independent datasets will, therefore, be needed to assess whether performance stays comparable across the different demographic groups.

Potential Relevance to Digital Dermatology and Health Care Setting

These results demonstrate that the application of standardized image analysis provides a reproducible digital outcome to assess visible facial aging. This approach might be useful for complementing traditional clinical grading by providing standardized quantitative assessment of facial aging signs and VA in the fields of dermatology and cosmetic dermatology.

Further potential applications could be found in the longitudinal monitoring of visible facial changes or in standardized assessments of appearance-related outcome measures in research and clinical practice. From this perspective, emerging concepts such as intelligent mirrors can offer a framework for the repeated, standardized, and noninvasive acquisition of skin phenotypes over time, supporting a transition from single time point assessments toward longitudinal skin monitoring [47]. These tools might additionally aid practitioners', patients', and consumers' communication about visible change.

The present study has not evaluated diagnostic accuracy for diseases, nor has it explored the influence on clinical decisions, treatment outcomes, patient management, or health system outcomes. Thus, at this stage, the system cannot be regarded as being a diagnostic or prognostic tool. Furthermore, since the evaluation used standardized photographs captured with one imaging device, applicability for use at POS, in clinics, on smartphones, or in consumer-facing situations under less standardized conditions requires further demonstration.

The present approach may potentially be integrated with other measurement modalities (eg, cutaneous biomarkers) to explore relationships between visible facial aging and biological processes. However, such work represents a future research perspective.

Ethical Considerations

While the application of AI-based facial grading and PVA in dermatology offers great promise, several ethical considerations must be addressed. First, more longitudinal studies are needed to validate the long-term imbrication between biological age and VA. While some treatments have been shown to influence biomarkers in the short term, their long-term impact on skin health and longevity remains unclear. Additionally, the use of biological age or VA as metrics for treatment decisions raises ethical questions, such as whether patients should be informed of their biological age if it differs significantly from their chronological age, and if so, how this information should be communicated alongside VA and how it might influence a patient's perception. Another challenge lies in the accessibility of these advanced, expensive, and specialized tools in POS or clinics, limiting their availability to a broader patient population.

Addressing these barriers, for example, with nomad solutions on smartphones, is essential to ensure equitable access to these innovations, preventing disparities between patients who can afford high-end aesthetic or dermo-cosmetic treatments and those who cannot. Additionally, differences in life expectancy between men and women, influenced by hormonal and environmental factors, add complexity to longevity research. These variations underscore the importance of developing

XSL•FO
RenderX

sex-specific approaches for personalized and effective treatments [48].

Limitations and Future Perspectives

Several limitations of the study should be noted. The data are restricted to women in North America who were photographed with a standardized imaging device. Although women across various ages and ethnicities were included, this diversity does not, by itself, establish generalizability to men, populations outside North America, other image-capture devices, or less-standardized real-life conditions. Additionally, there was no validation of results on an independently collected dataset.

Further investigations will be necessary, including the use of independent datasets containing more varied human populations, as well as alternative image acquisition devices under variable imaging conditions, before broader clinical or consumer-facing applications can be considered.

Overall performance was satisfactory across all ancestry groups included in the study, with variations depending on the facial signs and subgroup. To further assess the robustness of, and possible biases associated with, each group, additional investigations would be necessary, as the current analyses are based on limited datasets that lack balance. The naïve panel consisted solely of North American women aged 20 and 60 years, and thus the VAs might partially represent their specificities and perceptions.

It is worth noting that the algorithm trained to predict VA is based on subjective human evaluations. As such, PVA reflects an assessment of visible changes related to chronological age rather than objectively measurable signs of biological aging. No biological, proteomic, or physiological markers were assessed in the present study, which further implies that while the idea of VA being a potential "skin clock" is appealing, its relationship with biological age, health span, or overall health should currently be considered exploratory. This AI-based process could, in future studies, be associated concomitantly with biological measurements extracted from the skin surface. This new tool will support engagement, compliance, and long-term monitoring. It is nevertheless clear that this system calls for some extensions, such as its application to men—for whom technical challenges regarding beards, for example, could be addressed in a new AI-based model—or a deeper study of countries' specificities to tackle cultural shifts through larger cohorts of the major ethnic backgrounds.

In addition, building a global ecosystem to accompany patients and consumers through integrative longevity strategies is at hand, as we compare the MAE obtained here for the automatic grading system across all ethnicities (~97%) with that obtained with a similar system developed with Skin Aging Atlas references for smartphone selfies (~98%) [38,49]. This aspect will allow us to go further into North American women and controlled imaging. Another instrumental introduction will be leveraging technologies such as GANs for photorealistic and accurate personalized simulations [17,22,50] of future visible manifestations of longevity treatments of aging root causes—or the absence of them.

Conclusions

The present study demonstrated that an AI-based process can automate the evaluation of some features of visible facial aging with relatively low prediction errors and provide a predicted VA that has a strong correlation with human-assessed VAs under standardized image acquisition conditions. The automatic grading system demonstrated good agreement with expert assessments across diverse ancestral backgrounds represented in the study, although performance varied by facial sign and subgroup.

These findings support the potential use of AI-based facial assessment as a reproducible digital outcome for facial aging research and longitudinal monitoring. Nevertheless, external validation in independent, more diverse, and less standardized conditions is required before any wider clinical or consumer-facing rollout. As predicted, VA is derived from quantified facial clinical signs; it reflects measurable features of visible facial aging, while its relationship with biological age, overall health, or its potential as a validated "skin clock" remains to be established through future studies integrating biological biomarkers.

Acknowledgments

The authors want to thank Mr Didier Saint-Leger, Mr Soheil Seyfaie, Mr Aurelien Utzeri, Mr Paul Tartrat, Mr Cedric Larcher, Mr Chengda Ye, Mrs Celeste Grossgold, Mrs Olivia De Drezy, Mr Rawad Abdayem, Mrs Madiiha Mandary, Mr Thierry Lageat, Mr Frederic Woodland, Mr Adrien Chretien, Mrs Komal Patel, Mrs Margarita Yatskayer, Mrs Vania Lacascade, and Mr Guive Balooch not only for their great help in completion of this paper and global study but also for their strong support and enthusiasm.

AI was used as part of the study methodology for facial assessment and analysis, as described in the manuscript. No generative AI tools were used to generate scientific data, results, or conclusions.

Funding

All costs of the present work were entirely covered by L'Oréal Research & Innovation.

Data Availability

The data that support the findings are available from the corresponding author upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

FF, PAB, and TL were involved in the design, conduct, interpretation, preparation, and decision to submit the manuscript for publication. They had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. AA was involved in the conduct, collection, management, and analysis of the quiz study. TL, PM, JD, and SB designed algorithms and improved them through the training and validation phases. HJ and LB conducted data analysis for assessments of automatic grading system performance.
TD was involved in the design and interpretation. DH and HC made substantial contributions to revising the manuscript.
All authors reviewed and approved the manuscript before submission and have agreed to be accountable for their contributions.

Conflicts of Interest

FF, PAB, TL, SB, JD, HJ, LB, PM, and TD are employees of the L'Oréal group. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Full agreement statistics stratified by ancestry group.
[DOCX File, 34 KB - derma_v9i1e93236_app1.docx]

Multimedia Appendix 2
Bland-Altman plots.
[DOCX File, 437 KB - derma_v9i1e93236_app2.docx]

Multimedia Appendix 3
Calibration plots.
[DOCX File, 524 KB - derma_v9i1e93236_app3.docx]

References

1. World Health Statistics 2024: monitoring health for the SDGs, Sustainable Development Goals. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/74b12494-f213-4b5b-9533-18442147e1fb/content [accessed 2026-10-03]
2. Masfiah S, Kurnialandi A, Meij JJ, Maier AB. Definitions of healthspan: a systematic review. Ageing Res Rev 2025 Sep;111:102806. [doi: 10.1016/j.arr.2025.102806] [Medline: 40523529]
3. Garmany A, Yamada S, Terzic A. Longevity leap: mind the healthspan gap. NPJ Regen Med 2021 Sep 23;6(1):57. [doi: 10.1038/s41536-021-00169-5] [Medline: 34556664]
4. Haykal D, Flament F, Othman K, Siow R. Lifespan, healthspan, and the expanding role of cosmetic dermatology in longevity science. J Cosmet Dermatol 2025 Dec;24(12):e70610. [doi: 10.1111/jocd.70610] [Medline: 41392941]
5. Haykal D, Flament F, Mora P, Balooch G, Cartier H. Unlocking longevity in aesthetic dermatology: epigenetics, aging, and personalized care. Int J Dermatol 2025 Dec;64(12):2204-2214. [doi: 10.1111/ijd.17725] [Medline: 40064617]
6. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. Cell 2023 Jan 19;186(2):243-278. [doi: 10.1016/j.cell.2022.11.001] [Medline: 36599349]
7. Jacques E, Herzog C, Ying K, et al. Invigorating discovery and clinical translation of aging biomarkers. Nat Aging 2025 Apr;5(4):539-543. [doi: 10.1038/s43587-025-00838-w] [Medline: 40164770]
8. Biomarkers of Aging Consortium, Herzog CMS, Goeminne LJE, Poganik JR, et al. Challenges and recommendations for the translation of biomarkers of aging. Nat Aging 2024 Oct;4(10):1372-1383. [doi: 10.1038/s43587-024-00683-3] [Medline: 39285015]
9. Matts PJ, Fink B, Grammer K, Burquest M. Color homogeneity and visual perception of age, health, and attractiveness of female facial skin. J Am Acad Dermatol 2007 Dec;57(6):977-984. [doi: 10.1016/j.jaad.2007.07.040] [Medline: 17719127]
10. Porcheron A, Mauger E, Soppelsa F, et al. Facial contrast is a cross-cultural cue for perceiving age. Front Psychol 2017;8:1208. [doi: 10.3389/fpsyg.2017.01208] [Medline: 28790941]
11. Russell R, Porcheron A, Sweda JR, Jones AL, Mauger E, Morizot F. Facial contrast is a cue for perceiving health from the face. J Exp Psychol Hum Percept Perform 2016 Sep;42(9):1354-1362. [doi: 10.1037/xhp0000219] [Medline: 27100405]
12. Flament F, Abric A, Adam AS. Evaluating the respective weights of some facial signs on perceived ages in differently aged women of five ethnic origins. J Cosmet Dermatol 2021 Mar;20(3):842-853. [doi: 10.1111/jocd.13612] [Medline: 32649786]
13. Christensen K, Thinggaard M, McGue M, et al. Perceived age as clinically useful biomarker of ageing: cohort study. BMJ 2009 Dec 10;339:b5262. [doi: 10.1136/bmj.b5262] [Medline: 20008378]
14. Mekić S, Pardo LM, Gunn DA, et al. Younger facial looks are associated with a lower likelihood of several age-related morbidities in the middle-aged to elderly. Br J Dermatol 2023 Feb 22;188(3):390-395. [doi: 10.1093/bjd/ljac100] [Medline: 36763776]

XSL•FO
RenderX

15. Turner C, Pardo LM, Gunn DA, et al. Deep learning predicted perceived age is a reliable approach for analysis of facial ageing: a proof of principle study. J Eur Acad Dermatol Venereol 2024 Dec;38(12):2295-2302. [doi: 10.1111/jdv.20365] [Medline: 39360788]
16. Dobos G, Lichterfeld A, Blume-Peytavi U, Kottner J. Evaluation of skin ageing: a systematic review of clinical scales. Br J Dermatol 2015;172(5):1249-1261. [doi: 10.1111/bjd.13509] [Medline: 25363020]
17. Flament F, Bokaris PA, Despois J, et al. Ethnic differences in women's perception of simulated facial aging over a 15-year horizon: a GAN-based model approach. Cosmetics 2025;12(4):154. [doi: 10.3390/cosmetics12040154]
18. Raynaud E, Bertrand L, Flament F, et al. A combined approach to predict tobacco-induced facial aging using dermatologist knowledge elicitation and generative models. Dermatol Ther (Heidelb) 2026 Jan;16(1):629-636. [doi: 10.1007/s13555-025-01570-4] [Medline: 41134452]
19. Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial intelligence in dermatology: a primer. J Invest Dermatol 2020 Aug;140(8):1504-1512. [doi: 10.1016/j.jid.2020.02.026] [Medline: 32229141]
20. du Crest D, Madhumita M, Enbiale W, et al. Skin and digital—the 2024 narrative. Mayo Clin Proc Digit Health 2024 Sep;2(3):322-330. [doi: 10.1016/j.mcpdig.2024.05.008] [Medline: 40206126]
21. Haykal D, Garibyan L, Flament F, Cartier H. Hybrid cosmetic dermatology: AI generated horizon. Skin Res Technol 2024 May;30(5):e13721. [doi: 10.1111/srt.13721] [Medline: 38696225]
22. Despois J, Flament F, Perrot M. AgingMapGAN (AMGAN): high-resolution controllable face aging with spatially aware conditional GANS. Presented at: European Conference on Computer Vision; Aug 23-28, 2020. [doi: 10.1007/978-3-030-67070-2_37]
23. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine learning in dermatology: current applications, opportunities, and limitations. Dermatol Ther (Heidelb) 2020 Jun;10(3):365-386. [doi: 10.1007/s13555-020-00372-0] [Medline: 32253623]
24. Jain A, Way D, Gupta V, et al. Development and assessment of an artificial intelligence-based tool for skin condition diagnosis by primary care physicians and nurse practitioners in teledermatology practices. JAMA Netw Open 2021 Apr 1;4(4):e217249. [doi: 10.1001/jamanetworkopen.2021.7249] [Medline: 33909055]
25. Elder A, Ring C, Heitmiller K, Gabriel Z, Saedi N. The role of artificial intelligence in cosmetic dermatology—current, upcoming, and future trends. J Cosmet Dermatol 2021 Jan;20(1):48-52. [doi: 10.1111/jocd.13797] [Medline: 33151612]
26. Goh CL. Will digitization, big data and artificial intelligence- and deep learning-based algorithm govern the practice of medicine? J Eur Acad Dermatol Venereol 2022 Jul;36(7):947. [doi: 10.1111/jdv.18223] [Medline: 35712908]
27. Liu Y, Jain A, Eng C, et al. A deep learning system for differential diagnosis of skin diseases. Nat Med 2020 Jun;26(6):900-908. [doi: 10.1038/s41591-020-0842-3] [Medline: 32424212]
28. Kim S, Cho D, Kim J, et al. Smartphone-based multispectral imaging: system development and potential for mobile skin diagnosis. Biomed Opt Express 2016 Dec 1;7(12):5294-5307. [doi: 10.1364/BOE.7.005294] [Medline: 28018743]
29. Seité S, Khammari A, Benzaquen M, Moyal D, Dréno B. Development and accuracy of an artificial intelligence algorithm for acne grading from smartphone photographs. Exp Dermatol 2019 Nov;28(11):1252-1257. [doi: 10.1111/exd.14022] [Medline: 31446631]
30. Sangers TE, Kittler H, Blum A, et al. Position statement of the EADV artificial intelligence (AI) task force on AI‐assisted smartphone apps and web‐based services for skin disease. J Eur Acad Dermatol Venereol 2024 Jan;38(1):22-30. [doi: 10.1111/jdv.19521] [Medline: 37766502]
31. Flament F, Jacquet L, Ye C, et al. Artificial intelligence analysis of over half a million European and Chinese women reveals striking differences in the facial skin ageing process. J Eur Acad Dermatol Venereol 2022 Jul;36(7):1136-1142. [doi: 10.1111/jdv.18073] [Medline: 35279898]
32. Foucher A, Nouveau S, Piffaut V, Marque S, Aguilar L, Cavusoglu N. Clinical vs. chronological skin age: exploring determinants and stratum corneum protein markers of differential skin ageing in 351 healthy women. Sci Rep 2024 Oct 9;14(1):23643. [doi: 10.1038/s41598-024-65083-4] [Medline: 39384829]
33. Klare BF, Burge MJ, Klontz JC, Vorder Bruegge RW, Jain AK. Face recognition performance: role of demographic information. IEEE Trans Inf Forensics Secur 2012;7(6):1789-1801. [doi: 10.1109/TIFS.2012.2214212]
34. Drozdowski P, Rathgeb C, Dantcheva A, Damer N, Busch C. Demographic bias in biometrics: a survey on an emerging challenge. IEEE Trans Technol Soc 2020;1(2):89-103. [doi: 10.1109/TTS.2020.2992344]
35. Buolamwini J, Gebru T. Gender shades: intersectional accuracy disparities in commercial gender classification. Presented at: 1st Conference on Fairness, Accountability and Transparency (FAccT); Feb 23-24, 2018 URL: https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf [accessed 2026-10-03]
36. Bazin R, Flament F. Skin Aging Atlas Volume 2: Asian Type: Editions Med'Com; 2010.
37. Flament F, Bazin R, Qiu H. Skin Aging Atlas Volume 5: Photo-Aging Face & Body: Éditions Med'Com; 2017.
38. Jiang R, Kezele I, Levinshtein A, et al. A new procedure, free from human assessment that automatically grades some facial skin structural signs. Comparison with assessments by experts, using referential atlases of skin ageing. Int J Cosmet Sci 2019 Feb;41(1):67-78. [doi: 10.1111/ics.12512] [Medline: 30664236]
39. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016. [doi: 10.1109/CVPR.2016.90]

XSL•FO
RenderX

40. Kingma DP, Ba J. Adam: a method for stochastic optimization. Presented at: 3rd International Conference on Learning Representations (ICLR 2015); May 7-9, 2015 URL: https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1611214 [accessed 2026-10-03]
41. Swift A, Liew S, Weinkle S, Garcia JK, Silberberg MB. The facial aging process from the “inside out”. Aesthet Surg J 2021 Sep 14;41(10):1107-1119. [doi: 10.1093/asj/sjaa339] [Medline: 33325497]
42. McNeil MA, Merriam SB. Menopause. Ann Intern Med 2021 Jul;174(7):ITC97-ITC112. [doi: 10.7326/AITC202107200] [Medline: 34251902]
43. Talaulikar V. Menopause transition: physiology and symptoms. Best Pract Res Clin Obstet Gynaecol 2022 May;81:3-7. [doi: 10.1016/j.bpobgyn.2022.03.003] [Medline: 35382992]
44. US Preventive Services Task Force, Mangione CM, Barry MJ, Nicholson WK, et al. Hormone therapy for the primary prevention of chronic conditions in postmenopausal persons: US Preventive Services Task Force recommendation statement. JAMA 2022;328(17):1740-1746. [doi: 10.1001/jama.2022.18625]
45. Soares CN. Depression and menopause: an update on current knowledge and clinical management for this critical window. Med Clin North Am 2019 Jul;103(4):651-667. [doi: 10.1016/j.mcna.2019.03.001] [Medline: 31078198]
46. Flament F, Jiang R, Houghton J, et al. Accuracy and clinical relevance of an automated, algorithm-based analysis of facial signs from selfie images of women in the United States of various ages, ancestries and phototypes: a cross-sectional observational study. J Eur Acad Dermatol Venereol 2023 Jan;37(1):176-183. [doi: 10.1111/jdv.18541] [Medline: 35986708]
47. Haykal D, Flament F. Intelligent mirrors for longitudinal skin phenotyping. J Invest Dermatol 2026 Jul;146(7):1733-1735. [doi: 10.1016/j.jid.2026.02.016] [Medline: 41931069]
48. Engelbrecht HR, Merrill SM, Gladish N, et al. Sex differences in epigenetic age in Mediterranean high longevity regions. Front Aging 2022;3:1007098. [doi: 10.3389/fragi.2022.1007098] [Medline: 36506464]
49. Flament F, Lee YW, Lee DH, et al. The continuous development of a complete and objective automatic grading system of facial signs from selfie pictures: Asian validation study and application to women of three ethnic origins, differently aged. Skin Res Technol 2021 Mar;27(2):183-190. [doi: 10.1111/srt.12922] [Medline: 32686236]
50. Kips R, Jiang R, Ba S, et al. Real‐time virtual‐try‐on from a single example image through deep inverse graphics and learned differentiable renderers. Comput Graph Forum 2022 May;41(2):29-40. [doi: 10.1111/cgf.14456]

Abbreviations

CSP: cheek skin pores
DR: diffuse redness
ICC: intraclass correlation coefficient
MAE: mean absolute error
POS: points of sale
PVA: predicted visible age
RA: real age
ReLU: rectified linear unit
RMSE: root mean squared error
VA: visible age

Edited by R Dellavalle; submitted 12.Feb.2026; peer-reviewed by L González, TAK Manne, V Palama; revised version received 21.Sep.2026; accepted 23.Sep.2026; published 28.Oct.2026.

Please cite as:
Flament F, Bokaris PA, Li T, Abric A, Ba S, Despois J, Jouni H, Bertrand L, Michel P, Dauxerre T, Haykal D, Cartier H
AI-Based Grading of Facial Aging Signs and Prediction of Visible Age From Standardized Facial Images: Algorithm Development and Internal Evaluation Study
JMIR Dermatol 2026;9:e93236
URL: https://derma.jmir.org/2026/1/e93236
doi:10.2196/93236

© Frederic Flament, Panagiotis-Alexandros Bokaris, Tao Li, Aurelie Abric, Sileye Ba, Julien Despois, Hussein Jouni, Laudine Bertrand, Paul Michel, Thomas Dauxerre, Diala Haykal, Hughes Cartier. Originally published in JMIR Dermatology (http://derma.jmir.org), 28.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Local Deployment of Open-Weight Language Models in Dermatology: Viewpoint on Privacy, Equity, and Practical Implementation

William J Nahm[1], MD, MBA; Emily S Yin[2], MD, MHS; Emily C Milam[2], MD; Jason G Weed[2], MD

[1]Dr. Phillip Frost Department of Dermatology and Cutaneous Surgery, University of Miami Miller School of Medicine, 1295 N.W. 14th Street, Miami, FL, United States

[2]The Ronald O. Perelman Department of Dermatology, New York University Grossman School of Medicine, New York, NY, United States

Corresponding Author:
William J Nahm, MD, MBA
Dr. Phillip Frost Department of Dermatology and Cutaneous Surgery, University of Miami Miller School of Medicine, 1295 N.W. 14th Street, Miami, FL, United States

Abstract

Generative AI, particularly large language models (LLMs), is reshaping clinical workflows in dermatology. However, cloud-based commercial models pose persistent challenges to Health Insurance Portability and Accountability Act (HIPAA) compliance, especially in dermatology, where protected health information (PHI) extends beyond text to clinical photographs, dermoscopic images, and total-body photography that may capture identifiable anatomical features and document conditions carrying social stigma. Locally hosted, open-weight LLMs that are run within the institution's own infrastructure offer dermatology practices a pathway to leverage AI capabilities while retaining full control of their data. This viewpoint synthesizes evidence on when locally hosted, open-weight LLMs should be preferred for dermatologic workflows, when cloud deployment may remain preferable, and how multimodal AI fits into a coherent local deployment strategy. We advance 4 arguments. First, model compression techniques (knowledge distillation, structured pruning, and low-bit quantization) together with mixture-of-experts architectures have lowered hardware thresholds enough that 7- to 33-billion-parameter models now run on consumer-grade workstations with modest neural processing units or graphics processing units. Second, dermatology is fundamentally a visual specialty, and a credible local deployment strategy must integrate LLMs with vision models, including convolutional neural networks, vision transformers, vision-language models, and dermatology-specific foundation models such as PanDerm and medical multimodal models such as MedGemma. Third, locally hosted, open-weight models confer specific advantages for dermatology, including complete institutional control of clinical images, freedom from vendor model deprecation that disrupts validated workflows, and the ability to audit and fine-tune models to address well-documented performance gaps in skin of color. Fourth, local deployment is not a panacea; cloud models remain preferable for some tasks, and local deployment introduces governance challenges (heterogeneity across practices, model drift, and quantization-induced accuracy loss) that require structured mitigation through validated reporting frameworks such as CONSORT-AI (Consolidated Standards of Reporting Trials), SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials), DECIDE-AI (Developmental and Exploratory Clinical Investigations of Decision support systems driven by AI), and TRIPOD+AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis), as well as retrieval-augmented generation and federated learning approaches. We situate these arguments within the international regulatory landscape, including the European Union's General Data Protection Regulation, the European Union AI Act, and Germany's *Digitale Gesundheitsanwendungen* (DiGA) framework, in addition to HIPAA. We provide quantitative cost examples showing that current consumer hardware capable of running 14- to 33-billion-parameter models can be acquired for roughly the price of 1 to 2 years of enterprise cloud-AI subscriptions. We close by mapping a practical implementation pathway and identifying near-term research priorities. Locally hosted, open-weight LLMs that are deployed thoughtfully and within governance frameworks offer dermatology practices a credible route to harness generative AI while preserving regulatory compliance, equity across skin types, and the dermatologist-patient relationship.

(JMIR Dermatol 2026;9:e94764) doi:10.2196/94764

KEYWORDS

large language models; model compression; artificial intelligence; AI; dermatology; Health Insurance Portability and Accountability Act; HIPAA; teledermatology; local deployment; open-weight models; clinical documentation; data privacy

XSL•FO
RenderX

Introduction

Generative AI, particularly large language models (LLMs), has considerable potential to enhance health care delivery, from generating patient summaries to assisting with medical coding and optimizing practice workflows [1]. In dermatology, AI applications are expanding rapidly across clinical image analysis, teledermatology, dermatopathology, patient education, and clinical decision support [2,3]. The dermatology field has witnessed exponential growth in deep learning research since 2016, culminating in the US Food and Drug Administration's landmark 2024 approval of DermaSensor, the first AI skin cancer detection device authorized for primary care use [4].

Despite this momentum, the cloud-based deployment of leading commercial models, including OpenAI's GPT-5.2, Anthropic's Claude Opus 4.6, xAI's Grok, and Google's Gemini 3 Pro, raises substantive concerns about data security and Health Insurance Portability and Accountability Act (HIPAA) compliance [5]. HIPAA mandates that protected health information (PHI), including personally identifiable information, remain confidential and within the control of health care entities or verified business associates [6]. While cloud-based commercial LLMs can technically achieve HIPAA compliance through Business Associate Agreements (BAAs) with vendors, such as those available through enterprise tiers of OpenAI, Microsoft Azure, and Amazon Web Services, this approach introduces dependency on external vendors' data governance policies, limited institutional autonomy over AI infrastructure, and vulnerability to pricing and policy changes [5]. Frequent update cycles of proprietary models compound these concerns; when closed models are updated to optimize performance for general audiences, those changes can disrupt validated clinical workflows, and model versions are routinely deprecated within months, requiring health care organizations to revalidate clinical tools built upon them [5].

These concerns are particularly acute in dermatology, where they extend beyond textual records to include clinical photographs, dermoscopic images, and total-body photography that may capture identifiable anatomical features [7]. Dermatologic images of conditions such as sexually transmitted infections, autoimmune blistering diseases, or psychodermatologic manifestations carry heightened sensitivity due to social stigma, making the privacy of these data a paramount clinical and ethical concern [8]. The risks of poor data stewardship in AI-driven dermatology were starkly illustrated by a 2021 case in which a United Kingdom–based teledermatology platform disclosed patient photographs without obtaining appropriate authorization [9]. Locally hosted, open-weight LLMs, deployed within the institution's own infrastructure rather than transmitted to vendor cloud end points, present a compelling alternative for dermatology practices seeking to integrate generative AI while maintaining the highest standards of data protection. We propose that locally hosted, open-weight LLMs, when integrated with complementary vision models and embedded in adequate governance, may be the preferred deployment model for many dermatologic workflows. We articulate where this preference holds, where cloud models remain better suited, and what implementation steps clinical practices can realistically take.

The remainder of this viewpoint is organized as follows. We first describe our approach to the literature. We then consider open-weight model characteristics and the hardware thresholds at which locally hosted deployment becomes practical, followed by a discussion of multimodal AI relevant to dermatology. Subsequent sections address model compression and its dermatologic implications, practical local deployment infrastructure, and scenarios in which cloud-based deployment may remain preferable. We then examine regulatory frameworks beyond the HIPAA, review performance evidence and benchmarking, and discuss bias, equity, and the specific role of open-weight models in dermatology. We close with governance, validation, and reporting standards, reliability and data quality considerations, risks of overreliance and mitigation strategies, and future directions.

Methodological Approach to the Literature Review

We searched the PubMed, Embase, and arXiv databases between January 2020 and January 2026 using the terms "large language model," "open-weight," "local deployment," "on-premises," "vision-language model," "foundation model," "federated learning," "dermatology," "skin of color," "HIPAA," and "GDPR," alone and in combination. Peer-reviewed primary research and consensus reports were prioritized. We also drew on official model documentation, hardware vendor technical disclosures, and reporting framework publications (CONSORT-AI [Consolidated Standards of Reporting Trials], SPIRIT-AI [Standard Protocol Items: Recommendations for Interventional Trials], DECIDE-AI [Developmental and Exploratory Clinical Investigations of Decision Support Systems driven by AI], and TRIPOD+AI [Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis]) to characterize evolving infrastructure and governance. Where non–peer-reviewed technical disclosures were used, they were cross-checked against peer-reviewed analyses; web pages were retained only where no peer-reviewed equivalent existed.

Open-Weight Language Model Architectures and Hardware Requirements

Local deployment of AI solutions predominantly relies on open-weight LLMs, which provide trained model weights for inference while typically withholding full training methodology details [10]. The term "open-weight" is preferred here over "open-source," as most popular models, including the Llama and Mistral families, do not meet the full criteria of open-source software, which would require disclosure of training code, data, and architecture [10]. A model being closed does not inherently preclude local deployment; closed models are typically unavailable for on-premises use because vendors do not distribute their weights for local inference. Notably, the landscape is shifting: in 2025, OpenAI released its first

XSL•FO
RenderX

open-weight models for self-hosting, signaling that even developers of traditionally closed models now recognize the value of local deployment.

The advancement of LLMs has been driven by the continuous expansion of training datasets and parameter counts. For context, OpenAI's GPT-3 comprised 175 billion parameters; competing open-weight LLMs now exceed 600 billion parameters (eg, DeepSeek V3.2). The latest generation of open-weight models increasingly use mixture-of-experts (MoE) architectures that activate only a subset of parameters per query. Alibaba's Qwen 3 contains 235 billion total parameters but activates only 22 billion per token, reducing inference costs substantially compared with dense models of equivalent capability. Meta's Llama 4 introduced natively multimodal MoE models with 17 billion active parameters across up to 128 experts, fitting on a single graphics processing unit (GPU) with 4-bit integer (INT4) quantization, and Mistral AI's Mistral Large 3, released in December 2025 under the Apache 2.0 license, offers 41 billion active parameters within a 675-billion-parameter total MoE architecture. These developments are important precisely because they reduce the hardware threshold for clinically relevant inference.

Practical hardware thresholds, however, are not adequately captured by parameter counts alone. Table 1 summarizes representative open-weight model classes against memory footprint after INT4 quantization (a process to make an AI model smaller and faster by storing some of its numbers using INT4 instead of larger formats) and current hardware capable of running them at clinically usable speeds. A representative dermatology workstation, an Apple Mac Studio with M5 Max (64 GB unified memory, approximately US $3500 in early 2026) or a desktop with a single 24 GB consumer GPU (eg, NVIDIA RTX 5090, approximately US $2000), can currently run dense 14- to 33-billion-parameter models or MoE models with comparable active-parameter counts at interactive latencies. By contrast, enterprise cloud-AI subscriptions with HIPAA-eligible end points typically run US $30 to US $60 per user per month for clinical-grade tiers, yielding 3-year costs of US $1000 to US $2000 per clinician. For a 4-clinician dermatology practice, 2 years of cloud-AI subscriptions therefore approximate the one-time cost of a workstation capable of supporting the entire practice. Although a formal cost-effectiveness analysis is beyond the scope of this viewpoint, this comparison illustrates that the affordability argument is no longer hypothetical.

Table . Representative open-weight model classes and indicative hardware requirements after 4-bit integer (INT4) quantization[a].

Model class (examples)	Active parameters (approximately)	Memory after INT4 (approximately)	Representative hardware	Suitable dermatology tasks
Llama 3 8B and Qwen 3 8B [11,12]	7 - 8 billion	5 GB	Laptop with 16 GB RAM and modern NPU[b] [13]	Note structuring, basic patient education, and deidentification [3,14,15]
Llama 4 Scout, Qwen 3 14B and Mistral Small 3 [12,16,17]	14 - 17 billion	10 - 12 GB	Workstation with single 24 GB GPU[c] or M5 Pro/Max with 32 - 64 GB [18,19]	Drafting clinical letters, teledermatology triage, and RAG[d] over local references [20-22]
Qwen 3 32B and DeepSeek V3.2 (MoE[e]) [12,23]	22 - 37 billion (active)	20 - 30 GB	Workstation with 48 GB GPU or M5 Max/Ultra with 64 - 128 GB [18,19]	Differential diagnosis drafting (with HITL[f]) and dermatopathology summaries [24,25]
Llama 4 Maverick and Mistral Large 3 (MoE) [16,26]	>41 billion (active)	40 - 60 GB	Server with 1 - 2 enterprise GPUs (eg, H100 or H200) or Blackwell Ultra [27]	Group-practice or department-wide deployment and multitenant inference [5,28,29]

[a]Memory estimates assume INT4 quantization and standard context windows; vision and embedding models add to the budget. Hardware configurations are from early 2026.

[b]NPU: neural processing unit.

[c]GPU: graphics processing unit.

[d]RAG: retrieval-augmented generation.

[e]MoE: mixture-of-experts.

[f]HITL: human-in-the-loop.

Multimodal AI Systems for Dermatologic Applications

Dermatology is fundamentally a visual specialty, and the AI tools most relevant to dermatologic care span LLMs, image-only convolutional neural networks (CNNs), vision transformers, vision-language models (VLMs), and emerging dermatology-specific foundation models. Mehta et al [24] trace a clear evolution in dermatopathology from first-generation CNNs to hybrid CNN-transformer systems and large-scale foundation models such as Virchow and PanDerm, with several systems (eg, DermAI and PathAssist Derm) now in real-world deployment.

Yan et al [25] recently described PanDerm, a multimodal vision foundation model pretrained on more than 2 million skin disease images from 11 institutions across 4 imaging modalities. Across

XSL•FO
RenderX

28 benchmarks spanning skin cancer screening, differential diagnosis, lesion segmentation, and longitudinal monitoring, PanDerm achieved state-of-the-art performance, often outperforming existing models with as little as 10% of labeled data, and improved nondermatologist health care professionals' differential diagnosis by 16.5% across 128 skin conditions on clinical photographs [25]. Although PanDerm is at the research stage, its architecture, model weights, and broader trajectory are directly relevant to local deployment, as a dermatology-specific foundation model that can be downloaded and fine-tuned on institution-controlled data is precisely the type of tool local deployment is designed to support. Medical multimodal LLMs such as Google's MedGemma offer a complementary class of models combining vision and language capabilities under permissive licenses suitable for local deployment.

Two implications follow. First, hardware planning for local deployment must account not only for LLM memory footprints but also for vision-model inference budgets; a workstation provisioned to run a 14-billion-parameter LLM may need additional GPU memory to host a CNN ensemble or a VLM concurrently. Second, the regulatory and governance considerations developed below for LLMs largely apply, with appropriate adaptation, to local vision-model deployment as well. We therefore use "locally hosted models" inclusively in subsequent sections, while continuing to emphasize LLM-specific considerations where they differ.

Model Compression for Dermatologic Applications

Model compression techniques have emerged as a solution to these computational and financial barriers by transforming resource-intensive general-purpose models into compact variants deployable on constrained hardware [30,31]. These techniques enable the distillation of large models into smaller variants optimized for specific clinical use cases, trading broad generalization for targeted performance in defined tasks [30]. Key approaches include knowledge distillation, pruning, and quantization [30,31]. Knowledge distillation transfers knowledge from a large, complex "teacher" model to a smaller, more efficient "student" counterpart, reducing computational requirements while preserving task-relevant performance [31]. Pruning eliminates nonessential parameters within neural networks, reducing storage requirements and inference costs [30]. Quantization compresses models by decreasing the bit precision of model weights, with INT4 quantization capable of reducing a 140 GB model to approximately 35 GB [30]. Hardware-accelerated quantization formats, such as NVIDIA 4-bit floating-point (NVFP4), introduced with the Blackwell Ultra GPU architecture in 2025, further optimize low-precision inference by approaching 8-bit floating-point accuracy at a substantially reduced memory footprint.

Quantization carries known costs that warrant explicit acknowledgment. Aggressive low-bit quantization can degrade performance, particularly on tasks requiring nuanced reasoning, such as differential diagnosis support for complex inflammatory dermatoses or interpretation of subtle dermoscopic features. The multidimensional safety evaluation of LLM compression by Xu et al [30] demonstrates that perplexity alone does not capture compression-induced safety degradations relevant to medical use, including increased rates of biased or harmful outputs. Newer quantization techniques (eg, Generative Pre-trained Transformer Quantization, Activation-aware Weight Quantization, and NVFP4) have narrowed but have not closed this gap. Practices should therefore (1) prefer the highest-precision variant their hardware supports for clinical tasks, (2) validate the specific quantized variant they deploy on dermatology-relevant prompts before clinical use, and (3) revalidate when changing quantization formats or hardware. For text-based clinical documentation tasks, such as generating dermatologic consultation notes or summarizing pathology reports, knowledge distillation typically preserves high accuracy while significantly reducing parameter counts, and accuracy losses from INT4 quantization are usually small. For diagnostic reasoning tasks, the performance gap between compressed and full-scale models may widen, and organizations should not assume that compression yields equivalent performance across all use cases without empirical validation [28,30].

Infrastructure Requirements for On-Premises Model Deployment

A growing ecosystem of desktop applications, including LM Studio and Ollama, facilitates local deployment of small- and large-parameter open-weight LLMs on personal computers [32]. These applications provide intuitive graphical user interfaces for implementing and evaluating various LLMs on local hardware, eliminating dependence on external cloud computing. Neural processing units (NPUs) are specialized microprocessors designed to accelerate AI workloads, enabling efficient local inference without incurring cloud service costs [33]. Models in the Mistral, Llama, and Qwen families with 7 to 33 billion parameters now run efficiently on personal or small-office computers with 16 to 64 GB of memory, including Apple's M5-series Mac models with embedded neural accelerators. The Apple M5 chip embeds neural accelerators directly within each GPU core, delivering meaningful improvements in time-to-first-token for dense 14-billion-parameter models on a laptop. The NVIDIA Jetson platform delivers cost-effective AI processing in compact, energy-efficient form factors for edge deployments.

At the data center and enterprise scale, NVIDIA's Blackwell Ultra GB300 GPU offers 288 GB of HBM3e memory and substantial NVFP4 throughput, enabling hosting of 300-billion-plus parameter models without memory offloading. AMD's Instinct MI350 series provides a competitive alternative for institutions seeking vendor diversity. Latency depends on 3 factors: model size, hardware match, and network conditions. For small-parameter models running on adequately provisioned local hardware, local latency is competitive with, and often better than, cloud alternatives because no network round-trip is required. For very large dense models on consumer hardware, cloud inference on specialized accelerators is typically faster. We therefore recommend that practices select a model size that matches their hardware, as a properly matched local deployment

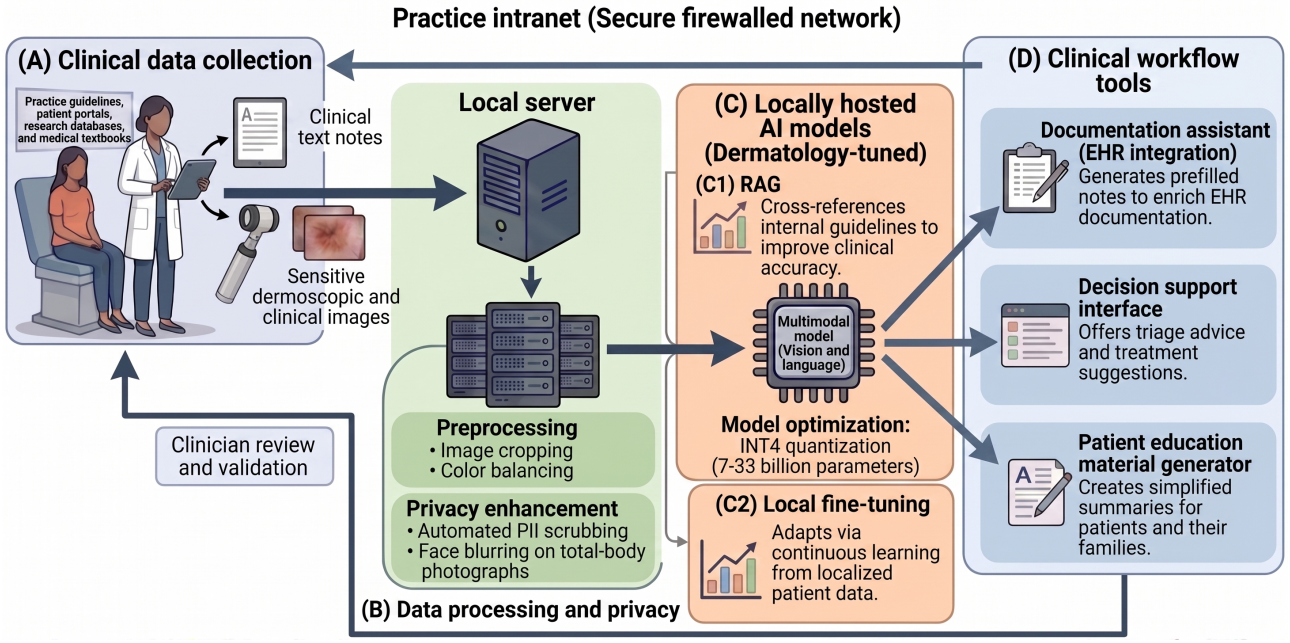
XSL•FO
RenderX

offers both privacy and competitive latency, while a mismatched one offers neither.

Dermatology-Specific Use Cases and Clinical Workflows

Local deployment creates powerful synergies with complementary technologies in dermatologic settings. A dermatology office can use specialized speech-to-text models to transcribe physicians' dictation during skin examinations, including detailed descriptions of lesion morphology, distribution patterns, and dermoscopic features. These transcriptions can then be processed by local LLMs that transform raw conversational text into structured clinical notes conforming to dermatologic documentation standards. By deploying these systems entirely on local hardware, practices eliminate cloud-based processing of sensitive clinical descriptions and images, avoid service interruptions from third-party cloud downtime, and address patient concerns about external access to descriptions of their skin conditions. Source recordings could be retained and encrypted in local storage for a defined verification window. Where downstream queries of the documentation are needed, retrieval-augmented generation (RAG) over the institution's own structured note repository (and, where appropriate, curated dermatologic references) should be preferred over reliance on the LLM's parametric memory, as RAG materially reduces hallucination rates and provides citable provenance [34]. Figure 1 illustrates this end-to-end local workflow, from data capture through documentation, retrieval, and human review.

Figure 1. Operationalizing locally hosted, dermatology-tuned AI models (LLM plus vision) within a practice's secure, firewalled intranet. (A) Clinical data collection: patient encounters and reference sources (practice guidelines, patient portals, research databases, and medical textbooks) generate clinical text notes and sensitive dermoscopic and clinical images that remain on local infrastructure. (B) Data processing and privacy: a local server preprocesses data (image cropping and color balancing) and applies privacy safeguards (automated personally identifiable information [PII] scrubbing and face blurring on total-body photographs). (C) Locally hosted AI models: a single multimodal (vision and language) model, quantized to 4-bit integer (INT4) at 7 to 33 billion parameters, draws on (C1) retrieval-augmented generation (RAG) against internal guidelines and (C2) local fine-tuning on localized patient data. (D) Clinical workflow tools: outputs power a documentation assistant for electronic health record (EHR) integration, a decision support interface offering triage and treatment suggestions, and a patient education material generator. A feedback loop returns model outputs to the clinician in panel A for human-in-the-loop review and validation before they are finalized or acted upon.



Beyond clinical documentation, locally deployed LLMs offer dermatology-specific applications across multiple domains. In teledermatology, local models can assist with triaging store-and-forward consultations by generating preliminary assessments from submitted images and clinical histories, reducing dermatologist review time while keeping all patient data within the institution [3,20]. For dermatopathology, small-parameter models can assist in structuring pathology reports, cross-referencing diagnostic criteria, and generating educational summaries for referring clinicians [24]. In the management of chronic inflammatory conditions such as psoriasis and atopic dermatitis, local AI can support treatment documentation, track biologic therapy responses, and generate patient-facing educational materials tailored to individual treatment plans [21]. Empirical evidence specifically supports LLM use for dermatologic diagnostic assistance, although with important caveats. In a comparative study using 100 histopathology-confirmed dermoscopic images from the International Skin Imaging Collaboration archive, Liu et al [35] reported that Claude 3 Opus and ChatGPT (GPT-4) achieved primary diagnostic accuracies of 56% and 48%, respectively, with the correct diagnosis appearing among the top 3 differentials in 76% and 78% of cases. Such performance is not adequate for autonomous diagnosis, but it indicates that current-generation LLMs (including locally deployable open-weight successors) can plausibly support shortlist generation, second-opinion prompts, and patient education when used with human-in-the-loop validation.

XSL•FO
RenderX

Comparative Analysis of Locally Hosted and Cloud-Based Deployment Models

Local deployment is not a universal answer. We outline 4 scenarios in which cloud-based deployment, with appropriate BAAs, may be the better choice for a dermatology practice. First, when a workflow requires the highest available reasoning performance, frontier proprietary models running in vendor data centers continue to outperform locally deployable open-weight alternatives on the most demanding diagnostic reasoning tasks. Second, when a practice cannot reliably maintain on-premises infrastructure (no IT staff, frequent power interruptions, or aging hardware), cloud reliability may exceed what local deployment can offer. Third, when a workflow requires elastic burst capacity (eg, processing thousands of teledermatology cases during a public health surge), cloud resources scale more easily than fixed local hardware. Fourth, model drift, the gradual divergence of a deployed model's performance from validation due to changes in patient population, documentation styles, or disease prevalence, is not solved by local deployment; if anything, isolated local deployments without centralized monitoring can experience drift unobserved for longer. Practices should weigh these considerations and may reasonably adopt hybrid architectures that route routine documentation locally and challenging diagnostic queries to a HIPAA-compliant cloud end point with explicit consent. Table 2 summarizes when each deployment model is preferable.

Table . When to prefer locally hosted versus cloud-based large language model deployment in dermatology.

Scenarios	Prefer locally hosted	Consider cloud-based (with BAA[a])
Routine documentation and encounter summarization	Yes; sensitive content remains on premises [6]	Acceptable if BAA is in place and audited [6]
Stigmatized conditions (eg, sexual health and psychodermatology)	Strongly preferred [8]	Avoid unless clearly justified
Frontier diagnostic reasoning queries	Possible with ≥32 billion parameters and MoE[b] if hardware permits	Often preferable; cloud frontier models still lead
Burst capacity (mass triage and public health surge)	Limited by fixed local hardware	Cloud elasticity advantageous
Practice without on-site IT support	Risk of unmaintained models and drift	Vendor-managed cloud may be more reliable
Cross-border (GDPR[c] and European Union AI Act) workflows	Helps with data localization obligations [34]	Requires careful data processing agreements [34]

[a]BAA: Business Associate Agreement.

[b]MoE: mixture-of-experts.

[c]GDPR: General Data Protection Regulation.

International Regulatory Frameworks and Cross-Jurisdictional Considerations

Although HIPAA shapes US dermatology practice, the international regulatory landscape is broader and increasingly relevant for cross-border teledermatology, multisite research, and academic-industry collaboration. The European Union's General Data Protection Regulation (GDPR) imposes principles of data minimization, purpose limitation, and subject rights (including the right of access and erasure) that frequently align well with local deployment architectures. Data that never leave the institution is more readily reconciled with these obligations than data shared with third-country processors [34]. The European Union AI Act (Regulation EU 2024/1689) imposes additional obligations on high-risk AI systems used in health care, including risk management, data governance, transparency, and human oversight requirements, and requires conformity assessment under the Medical Device Regulation when the AI system constitutes a medical device. Germany's *Digitale Gesundheitsanwendungen* (DiGA) framework, which provides reimbursement for prescription-grade digital health applications, illustrates how local deployment architectures can intersect with national reimbursement pathways. For example, DiGA-listed applications must satisfy stringent data protection requirements that local processing helps to meet.

For dermatology practices that operate across borders or partner with European researchers, 3 points are practically salient. First, local deployment is helpful but not sufficient; GDPR and the European Union AI Act impose obligations (documentation, conformity assessment, and postmarket monitoring) regardless of where computation occurs. Second, regulatory frameworks increasingly converge on transparency, auditability, and human oversight, all of which are easier to achieve with open-weight models. Third, locally hosted models can also raise their own questions about user input parameters; data the user types into a locally hosted system may be retained for fine-tuning, audited internally, or inadvertently exposed through weak access controls. Local deployment shifts the locus of risk from external transmission to internal governance, and practices must design access controls, audit trails, and prompt-logging policies to match [36].

Performance Evidence and Benchmarking

Recent research demonstrates that locally deployed models can achieve strong performance in health care–relevant tasks. The

XSL•FO
RenderX

LLM-Anonymizer study showed that local Llama 3 models achieved accuracies and sensitivities exceeding 98% for medical document deidentification, with both the 8-billion- and 70-billion-parameter variants performing comparably [14]. These findings suggest that smaller models can match larger counterparts for well-defined tasks, enabling deployment on standard modern hardware without substantial infrastructure investment.

However, comprehensive performance benchmarking across dermatologic use cases remains limited. Existing data focus primarily on document deidentification, and performance metrics for tasks such as diagnostic assistance, clinical note generation, and dermatopathology report summarization are lacking. The field urgently needs comparative data quantifying precision, recall, and clinical efficiency gains for locally deployed models versus commercial alternatives in dermatology-specific scenarios. Drawing on adjacent specialties, Luo et al [22] demonstrated that a locally hosted 13-billion-parameter model with retrieval augmentation achieved performance comparable to larger models in ophthalmology, noting that models smaller than 13 billion parameters may lack proficiency for complex medical reasoning but remain viable for many clinical and research applications. Analysis of LLM use patterns in health care shows that studies using open-weight models frequently focus on mental health applications, conversational agents, and documentation support, suggesting these models currently offer their strongest value proposition for customizable, privacy-sensitive workflows rather than high-complexity diagnostic reasoning [10]. Independent benchmarking platforms such as the Chatbot Arena and Open LLM Leaderboard show that leading open-weight models now approach the performance of proprietary models on many reasoning tasks, a trajectory that bodes well for clinical applications.

Algorithmic Bias, Health Equity, and the Role of Open-Weight Models in Dermatologic Practice

The biases embedded within training data demand critical examination, particularly in dermatology, where these biases have well-documented and quantitatively measurable clinical consequences. Daneshjou et al [37] evaluated dermatology AI on the Diverse Dermatology Images set and reported substantial accuracy drops for darker skin tones, with state-of-the-art deep learning models classifying skin lesions less accurately on Fitzpatrick types V and VI than on lighter skin types. A subsequent scoping review of skin cancer image datasets found pervasive underrepresentation of darker skin tones and other minoritized groups across publicly available datasets, with limited or absent metadata on Fitzpatrick type, ethnicity, or geographic origin in many datasets [38]. A more recent dataset review by Wen et al [39] confirmed that dataset shifts arising from demographic variations and inconsistent labeling continue to compromise clinical utility and propagate discriminatory biases when algorithms are implemented in mismatched populations. Algorithmic bias resulting from feature selection and model weighting decisions can amplify these training data biases, intensifying health care disparities when AI systems are deployed in dermatologic practice [40].

These factors strengthen the argument for open-weight models in dermatology in 3 concrete ways that go beyond general claims of "transparency." First, accessibility of weights enables institutions to fine-tune models on locally curated, demographically representative datasets, addressing population-specific performance gaps that vendors of closed models may not prioritize. Second, weight access permits empirical bias auditing through controlled probing, including stratified accuracy testing across Fitzpatrick types, which is impossible with closed models accessed only through APIs. Third, federated learning approaches over open-weight models enable multi-institutional fine-tuning to broaden representation without centralizing sensitive imaging data. Haggenmüller et al [41] demonstrated this approach concretely for melanoma versus nevus classification across 6 German university hospitals, with federated training achieving performance comparable to centralized training while preserving institutional data sovereignty. The dermatology community has both the opportunity and the obligation to use these capabilities deliberately so that AI tools serve patients across all skin types equitably [10].

Governance Frameworks, Validation Protocols, and Reporting Standards for Clinical AI

While local deployment presents compelling opportunities, health care organizations must carefully evaluate inherent trade-offs. Compressed models may exhibit degradation in downstream task performance and amplification of safety concerns related to information accuracy, bias, and the generation of harmful content [30]. These evaluation challenges compound medicolegal uncertainties, as physicians, including dermatologists, lack clear guidance on incorporating AI technologies without increasing liability exposure [42,43]. Established AI reporting and evaluation frameworks should anchor local deployments. CONSORT-AI and SPIRIT-AI provide standards for AI-related randomized trials and trial protocols, TRIPOD+AI standardizes reporting of AI prediction model studies, and DECIDE-AI provides reporting guidance for early clinical evaluation of AI-driven decision support systems [44,45]. Although none of these frameworks were written specifically for locally deployed open-weight LLMs, their items (data provenance, model versioning, performance reporting across subgroups, and human-AI interaction documentation) translate directly to local deployment and should be incorporated into dermatology practice protocols and academic publications.

The possibility of numerous dermatology practices independently deploying different small models creates governance challenges that benefit from concrete, feasible mechanisms. We propose a tiered governance approach. At the practice level, every locally deployed model should have a model card recording the weights identifier, quantization, hardware, deployment date, intended use, validation dataset,

XSL•FO
RenderX

and contraindications, plus an audit log capturing prompts, outputs, clinician overrides, and adverse events. At the network level, regional dermatology associations or academic consortia could host opt-in registries that aggregate model cards and adverse event signals; participation requires only minimal infrastructure (a shared form, a data-use agreement, and quarterly reporting) and could leverage federated approaches to share statistical signals without sharing PHI. At the national level, frameworks such as the Coalition for Health AI's blueprint for trustworthy AI implementation, originally designed for centrally deployed systems, can be adapted to provide guidance and benchmarking standards for distributed local deployments [46]. We acknowledge that smaller practices may lack data scientists; this is precisely why federated learning models with explicit, prescriptive data-processing guidelines are more realistic than expecting each practice to design its own monitoring program [41].

Reliability and Data Quality

Most LLMs lack health care–specific design or training. Unlike established clinical resources such as UpToDate, which use peer-reviewed content and expert curation, LLMs typically train on vast, unvetted web data without distinguishing reliable from questionable sources [47]. High-quality training data are optimal, yet the field has not established standardized thresholds for data reliability, as requirements vary substantially based on training data complexity and the intended clinical use case [48].

LLMs frequently lack transparent citation trails, and generated references may fail to support content or prove entirely fictional, leaving dermatologists unable to verify AI-generated clinical information [42]. The probabilistic nature of text generation introduces further reliability concerns; identical queries can produce varying responses across attempts, with outputs influenced by query phrasing and context [42]. Such variability can generate interaction biases affecting both clinical performance and objectivity [40]. The "black box" nature of machine learning algorithms compounds trust deficits, with global research indicating that 63% of participants express unwillingness or ambivalence toward trusting AI in health care [49,50]. Implementing human-in-the-loop frameworks where dermatologists validate AI outputs is crucial for mitigating these limitations [15]. Explainable AI technologies offer promising solutions; notably, Chanda et al [51] demonstrated through eye-tracking that dermatologist-like explainable AI enhanced melanoma diagnostic accuracy, suggesting that advances in interpretability can directly benefit dermatologic practice.

Automation Bias and Mitigation Strategies

Clinician overreliance on LLMs poses documented risks. In a multicenter randomized clinical vignette study, clinicians experienced an 11.3% decrease in diagnostic accuracy when shown systematically biased AI model predictions, demonstrating how uncritical acceptance of AI outputs can compromise clinical judgment [52]. In dermatology specifically, nonspecialists may increasingly use AI tools for diagnosis despite lacking expertise in comprehensive management, a concern particularly relevant given the Food and Drug Administration's recent approval of AI devices for primary care skin cancer screening [53,54]. Such overreliance threatens to diminish diagnostic competence and compromise essential pattern-recognition skills central to dermatologic practice [55].

Addressing these risks requires multifaceted strategies. Greater investment in human-computer interaction research within dermatology is needed to understand how clinicians engage with AI recommendations and where decision-making is most vulnerable to automation bias [52]. Mandatory human-in-the-loop validation protocols should be integrated into all clinical AI deployments, and locally hosted LLMs are well-suited to enforce these workflows in practice (eg, by requiring clinician acknowledgment before downstream actions). Additionally, research examining the impact of AI tools on patient trust and willingness to disclose information, particularly regarding sensitive skin conditions, is essential for guiding technical design. Patient perceptions matter not only ethically but also operationally. If patients perceive AI tools as protecting their privacy, they may be more forthcoming about stigmatized conditions, ultimately supporting clinical accuracy.

Future Directions and Conclusions

Computational requirements for locally hosted models are expected to decrease as hardware advances and model compression techniques mature, particularly with continued progress in NPU capabilities and the open-weight ecosystem. Concurrently, the open-weight ecosystem continues to mature, with organizations including Meta, Alibaba, Mistral AI, DeepSeek, and OpenAI contributing high-performance models under permissive licenses, ensuring that locally deployable alternatives keep pace with proprietary offerings.

For dermatology specifically, 4 near-term priorities follow from this analysis. First, dermatology-specific benchmarking suites are needed that test locally deployable models on documentation, triage, and visual reasoning tasks, with performance stratified by Fitzpatrick type and condition rarity. Second, structured cost-effectiveness analyses should compare local and cloud deployments under realistic dermatology practice assumptions, including IT staffing and incident response costs. Third, federated learning collaborations across academic dermatology centers should be expanded, building on emerging precedents in melanoma diagnostics, to develop fine-tuned open-weight models that better represent diverse skin types without centralizing PHI [41]. Fourth, dermatology professional societies should issue practical guidance on model cards, audit logs, and adverse event reporting that smaller practices can realistically adopt.

Funding

No funding was received to assist with the preparation of this article.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: WJN (lead),
ESY (supporting), ECM (supporting), JGW (supporting)
Data curation: not applicable
Formal analysis: not applicable
Funding acquisition: not applicable
Investigation: WJN
Methodology: WJN
Project administration: WJN
Resources: not applicable
Software: not applicable
Supervision: ESY, ECM, JGW
Validation: not applicable
Visualization: WJN
Writing – original draft: WJN
Writing – review & editing: WJN (lead), ESY (supporting), ECM (supporting), JGW (supporting)

Conflicts of Interest

None declared.

References

1. Whig P, Dutta PK, Yathiraju N, Bhattacharya P, Sharma S. Transforming healthcare operations with generative AI: optimizing resource allocation, workflow management, and predictive analytics. In: Bhattacharya P, Liu H, Dutta PK, Rodrigues JJ, Sethi G, editors. Revolutionizing Healthcare 5.0: The Power of Generative AI: Springer; 2024:289-308. [doi: 10.1007/978-3-031-75771-6_19]
2. Omiye JA, Gui H, Daneshjou R, Cai ZR, Muralidharan V. Principles, applications, and future of artificial intelligence in dermatology. Front Med (Lausanne) 2023;10:1278232. [doi: 10.3389/fmed.2023.1278232] [Medline: 37901399]
3. Nahm WJ, Sohail N, Burshtein J, Goldust M, Tsoukas M. Artificial intelligence in dermatology: a comprehensive review of approved applications, clinical implementation, and future directions. Int J Dermatol 2025 Sep;64(9):1568-1583. [doi: 10.1111/ijd.17847] [Medline: 40387622]
4. Venkatesh KP, Kadakia KT, Gilbert S. Learnings from the first AI-enabled skin cancer device for primary care authorized by FDA. NPJ Digit Med 2024 Jun 15;7(1):156. [doi: 10.1038/s41746-024-01161-1] [Medline: 38879640]
5. Dennstädt F, Hastings J, Putora PM, Schmerder M, Cihoric N. Implementing large language models in healthcare while balancing control, collaboration, costs and security. NPJ Digit Med 2025 Mar 6;8(1):143. [doi: 10.1038/s41746-025-01476-7] [Medline: 40050366]
6. Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. JAMA 2023 Jul 25;330(4):309-310. [doi: 10.1001/jama.2023.9458] [Medline: 37410450]
7. Yadav N, Pandey S, Gupta A, Dudani P, Gupta S, Rangarajan K. Data privacy in healthcare: in the era of artificial intelligence. Indian Dermatol Online J 2023;14(6):788-792. [doi: 10.4103/idoj.idoj_543_23] [Medline: 38099022]
8. Gordon ER, Trager MH, Kontos D, et al. Ethical considerations for artificial intelligence in dermatology: a scoping review. Br J Dermatol 2024 May 17;190(6):789-797. [doi: 10.1093/bjd/ljae040] [Medline: 38330217]
9. Zbrzezny AM, Krzywicki T. Artificial intelligence in dermatology: a review of methods, clinical applications, and perspectives. Appl Sci 2025;15(14):7856. [doi: 10.3390/app15147856]
10. Xu J, Ding Y, Bu Y. Position: open and closed large language models in healthcare. arXiv. Preprint posted online on Jan 17, 2025. [doi: 10.48550/arXiv.2501.09906]
11. Introducing Meta Llama 3: the most capable openly available LLM to date. Meta. 2024. URL: https://ai.meta.com/blog/meta-llama-3/ [accessed 2026-08-10]
12. Qwen3: think deeper, act faster. Qwen Studio. 2025. URL: https://qwen.ai/blog?id=qwen3 [accessed 2026-08-10]
13. Apple unleashes M5, the next big leap in AI performance for Apple silicon. Apple. 2025. URL: https://www.apple.com/newsroom/2025/10/apple-unleashes-m5-the-next-big-leap-in-ai-performance-for-apple-silicon/ [accessed 2026-08-10]
14. Wiest IC, Leßmann ME, Wolf F, et al. Deidentifying medical documents with local, privacy-preserving large language models: the LLM-anonymizer. NEJM AI 2025 Mar 27;2(4). [doi: 10.1056/AIdbp2400537]
15. Bakken S. AI in health: keeping the human in the loop. J Am Med Inform Assoc 2023 Jun 20;30(7):1225-1226. [doi: 10.1093/jamia/ocad091] [Medline: 37337923]
16. The Llama 4 herd: the beginning of a new era of natively multimodal AI innovation. Meta AI. 2025. URL: https://ai.meta.com/blog/llama-4-multimodal-intelligence/ [accessed 2026-08-10]
17. Mistral Small 3. Mistral AI. 2025. URL: https://mistral.ai/news/mistral-small-3/ [accessed 2026-08-10]

XSL•FO
RenderX

18. Apple introduces MacBook Pro with all-new M5 Pro and M5 Max, delivering breakthrough pro performance and next-level on-device AI. Apple Inc. 2026. URL: https://www.apple.com/newsroom/2026/03/apple-introduces-macbook-pro-with-all-new-m5-pro-and-m5-max/ [accessed 2026-08-10]
19. GeForce RTX 5090 graphics cards. NVIDIA. URL: https://www.nvidia.com/en-us/geforce/graphics-cards/50-series/rtx-5090/ [accessed 2026-08-10]
20. Shapiro J, Lyakhovitsky A. Revolutionizing teledermatology: exploring the integration of artificial intelligence, including Generative Pre-trained Transformer chatbots for artificial intelligence-driven anamnesis, diagnosis, and treatment plans. Clin Dermatol 2024;42(5):492-497. [doi: 10.1016/j.clindermatol.2024.06.020] [Medline: 38942153]
21. Jairath N, Pahalyants V, Shah R, Weed J, Carucci JA, Criscito MC. Artificial intelligence in dermatology: a systematic review of its applications in melanoma and keratinocyte carcinoma diagnosis. Dermatol Surg 2024 Sep 1;50(9):791-798. [doi: 10.1097/DSS.0000000000004223] [Medline: 38722750]
22. Luo MJ, Pang J, Bi S, et al. Development and evaluation of a retrieval-augmented large language model framework for ophthalmology. JAMA Ophthalmol 2024 Sep 1;142(9):798-805. [doi: 10.1001/jamaophthalmol.2024.2513] [Medline: 39023885]
23. DeepSeek-V3.2 release. DeepSeek API Docs. 2025. URL: https://api-docs.deepseek.com/news/news251201/ [accessed 2026-08-10]
24. Mehta A, Motavaf M, Raza D, et al. Deep learning image processing models in dermatopathology. Diagnostics (Basel) 2025 Oct 4;15(19):2517. [doi: 10.3390/diagnostics15192517] [Medline: 41095736]
25. Yan S, Yu Z, Primiero C, et al. A multimodal vision foundation model for clinical dermatology. Nat Med 2025 Aug;31(8):2691-2702. [doi: 10.1038/s41591-025-03747-y] [Medline: 40481209]
26. Introducing Mistral 3. Mistral AI. 2025. URL: https://mistral.ai/news/mistral-3/ [accessed 2026-08-10]
27. NVIDIA H100 GPU. NVIDIA. URL: https://www.nvidia.com/en-us/data-center/h100/ [accessed 2026-08-10]
28. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. JAMA 2023 Sep 5;330(9):866-869. [doi: 10.1001/jama.2023.14217] [Medline: 37548965]
29. Umeton R, Kwok A, Maurya R, et al. GPT-4 in a cancer center — institute-wide deployment challenges and lessons learned. NEJM AI 2024 Mar 28;1(4). [doi: 10.1056/AIcs2300191]
30. Xu Z, Gupta A, Li T, Bentham O, Srikumar V. Beyond perplexity: multi-dimensional safety evaluation of LLM compression. In: Findings of the Association for Computational Linguistics: EMNLP 2024: Association for Computational Linguistics; 2024. [doi: 10.18653/v1/2024.findings-emnlp.901]
31. Yang C, Zhu Y, Lu W, et al. Survey on knowledge distillation for large language models: methods, evaluation, and application. ACM Trans Intell Syst Technol 2025;16(6):1-27. [doi: 10.1145/3699518]
32. Local LLM deployment: privacy-first AI complete guide. Digital Applied. 2025. URL: https://www.digitalapplied.com/blog/local-llm-deployment-privacy-guide-2025 [accessed 2026-08-11]
33. Whig P, Dutta PK, Yathiraju N, Bhattacharya P, Batra I. Enhancing healthcare operations: harnessing generative AI for efficiency and innovation. In: Bhattacharya P, Liu H, Dutta PK, Rodrigues JJ, Sethi G, editors. Revolutionizing Healthcare 5.0: The Power of Generative AI: Springer; 2025:273-287. [doi: 10.1007/978-3-031-75771-6_18]
34. Knott M, Krebs M, Kerscher A. Large language models in healthcare quality management: a European perspective on process automation and compliance. Front Digit Health 2026;8:1761641. [doi: 10.3389/fdgth.2026.1761641] [Medline: 41970523]
35. Liu X, Duan C, Kim MK, et al. Claude 3 Opus and ChatGPT with GPT-4 in dermoscopic image analysis for melanoma diagnosis: comparative performance analysis. JMIR Med Inform 2024 Aug 6;12:e59273. [doi: 10.2196/59273] [Medline: 39106482]
36. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Med Ethics 2021 Sep 15;22(1):122. [doi: 10.1186/s12910-021-00687-3] [Medline: 34525993]
37. Daneshjou R, Vodrahalli K, Novoa RA, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv 2022 Aug 12;8(32):eabq6147. [doi: 10.1126/sciadv.abq6147] [Medline: 35960806]
38. Wen D, Khan SM, Ji Xu A, et al. Characteristics of publicly available skin cancer image datasets: a systematic review. Lancet Digit Health 2022 Jan;4(1):e64-e74. [doi: 10.1016/S2589-7500(21)00252-1] [Medline: 34772649]
39. Wen D, Soltan A, Trucco E, Matin RN. From data to diagnosis: skin cancer image datasets for artificial intelligence. Clin Exp Dermatol 2024 Jun 25;49(7):675-685. [doi: 10.1093/ced/llae112] [Medline: 38549552]
40. Hanna MG, Pantanowitz L, Jackson B, et al. Ethical and bias considerations in artificial intelligence/machine learning. Mod Pathol 2025 Mar;38(3):100686. [doi: 10.1016/j.modpat.2024.100686] [Medline: 39694331]
41. Haggenmüller S, Schmitt M, Krieghoff-Henning E, et al. Federated learning for decentralized artificial intelligence in melanoma diagnostics. JAMA Dermatol 2024 Mar 1;160(3):303-311. [doi: 10.1001/jamadermatol.2023.5550] [Medline: 38324293]
42. Mello MM, Guha N. ChatGPT and physicians' malpractice risk. JAMA Health Forum 2023 May 5;4(5):e231938. [doi: 10.1001/jamahealthforum.2023.1938] [Medline: 37200013]

XSL•FO
RenderX

43. Nelson CA, Pachauri S, Balk R, et al. Dermatologists' perspectives on artificial intelligence and augmented intelligence - a cross-sectional survey. JAMA Dermatol 2021 Jul 1;157(7):871-874. [doi: 10.1001/jamadermatol.2021.1685] [Medline: 34037674]
44. Shiferaw KB, Roloff M, Balaur I, Welter D, Waltemath D, Zeleke AA. Guidelines and standard frameworks for artificial intelligence in medicine: a systematic review. JAMIA Open 2025;8(1):ooae155. [doi: 10.1093/jamiaopen/ooae155] [Medline: 39759773]
45. Frutuoso J. Building a safe and transparent workflow for large language model (LLM)-assisted clinical trials and prediction models: a technical report. Cureus 2025 Sep;17(9):e92571. [doi: 10.7759/cureus.92571] [Medline: 41111869]
46. Blueprint for trustworthy AI implementation guidance and assurance for healthcare. : Coalition for Health AI; 2023 URL: https://assets.ctfassets.net/7s4afyr9pmov/4AXIWGIlcrjWDaW2ueTaRS/f98e5cb2528187635895cce6ba5ec309/Blueprint_for_Trustworthy_AI.pdf [accessed 2026-08-11]
47. Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. EBioMedicine 2023 Apr;90:104512. [doi: 10.1016/j.ebiom.2023.104512] [Medline: 36924620]
48. Al Hleewa SO, Al Mubarak M. Success factors of using artificial intelligence. In: Al Mubarak M, Hamdan A, editors. Technological Sustainability and Business Competitive Advantage: Springer; 2023:169-184. [doi: 10.1007/978-3-031-35525-7_11]
49. Shevtsova D, Ahmed A, Boot IW, et al. Trust in and acceptance of artificial intelligence applications in medicine: mixed methods study. JMIR Hum Factors 2024 Jan 17;11:e47031. [doi: 10.2196/47031] [Medline: 38231544]
50. Gillespie N, Lockey S, Ward T, Macdade A, Hassed G. Trust, attitudes and use of artificial intelligence: a global study 2025. : The University of Melbourne and KPMG; 2025 URL: https://kpmg.com/in/en/insights/2025/05/trust-attitudes-and-use-of-ai.html [accessed 2026-08-11]
51. Chanda T, Haggenmueller S, Bucher TC, et al. Dermatologist-like explainable AI enhances melanoma diagnosis accuracy: eye-tracking study. Nat Commun 2025 May 21;16(1):4739. [doi: 10.1038/s41467-025-59532-5] [Medline: 40399272]
52. Jabbour S, Fouhey D, Shepard S, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. JAMA 2023 Dec 19;330(23):2275-2284. [doi: 10.1001/jama.2023.22295] [Medline: 38112814]
53. Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. JAMA 2023 Dec 19;330(23):2255-2257. [doi: 10.1001/jama.2023.22557] [Medline: 38112824]
54. Daneshjou R, Smith MP, Sun MD, Rotemberg V, Zou J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. JAMA Dermatol 2021 Nov 1;157(11):1362-1369. [doi: 10.1001/jamadermatol.2021.3129] [Medline: 34550305]
55. Quinn TP, Senadeera M, Jacobs S, Coghlan S, Le V. Trust and medical AI: the challenges we face and the expertise needed to overcome them. J Am Med Inform Assoc 2021 Mar 18;28(4):890-894. [doi: 10.1093/jamia/ocaa268] [Medline: 33340404]

Abbreviations

BAA: Business Associate Agreement
CNN: convolutional neural network
CONSORT: Consolidated Standards of Reporting Trials
DECIDE: Developmental and Exploratory Clinical Investigations of Decision Support Systems
DiGA: *Digitale Gesundheitsanwendungen*
GDPR: General Data Protection Regulation
GPU: graphics processing unit
HIPAA: Health Insurance Portability and Accountability Act
INT4: 4-bit integer
LLM: large language model
MoE: mixture-of-experts
NPU: neural processing unit
NVFP4: NVIDIA 4-bit floating-point
PHI: protected health information
RAG: retrieval-augmented generation
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
VLM: vision-language model

XSL•FO
RenderX

Edited by S Yasamin Parvar; submitted 05.Mar.2026; peer-reviewed by A Yilmaz, B Eapen, I Aryendu, O Okeda; revised version received 27.Jul.2026; accepted 01.Aug.2026; published 21.Aug.2026.

Please cite as:
Nahm WJ, Yin ES, Milam EC, Weed JG
Local Deployment of Open-Weight Language Models in Dermatology: Viewpoint on Privacy, Equity, and Practical Implementation
JMIR Dermatol 2026;9:e94764
URL: https://derma.jmir.org/2026/1/e94764
doi:10.2196/94764

© William J Nahm, Emily S Yin, Emily C Milam, Jason G Weed. Originally published in JMIR Dermatology (http://derma.jmir.org), 21.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Agentic AI in Dermatology: A Call to Action

Brian Chu[1], MD; Heather Shen[2], MS; Ivy Lee[3], MD; Jules B Lipoff[4], MD

[1]Department of Dermatology, Brown University, 593 Eddy St, APC 10th Floor, Providence, RI, United States

[2]Praxis Labs, Providence, RI, United States

[3]Comprehensive Dermatology Center of Pasadena, Pasadena, CA, United States

[4]Department of Dermatology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, United States

Corresponding Author:
Brian Chu, MD
Department of Dermatology, Brown University, 593 Eddy St, APC 10th Floor, Providence, RI, United States

Abstract

Artificial intelligence (AI) tools are shifting from passive, user-initiated tools to proactive *agentic AI* systems that are capable of autonomous, multi-step actions. These agents can independently gather information, execute sequential tasks, and collaborate with humans or other agents without requiring constant prompting from humans. Early adopters in health care have demonstrated early feasibility across multiple specialties and clinical settings. Dermatology is well-positioned to benefit given its high patient volumes, administrative burdens, and clinicopathological workflows. To guide responsible adoption of agentic AI, we propose a risk-stratification framework based on clinical risk and task reversibility. Barriers to widespread adoption of agentic AI include limitations in model reliability, interoperability across health records, and unresolved questions around liability, privacy, and regulation. Dermatologists must proactively engage via professional organizations and industry partnerships to ensure that agentic AI is developed safely, equitably, and in alignment with our values.

(JMIR Dermatol 2026;9:e94909) doi:10.2196/94909

KEYWORDS

agentic AI; AI agents; AI in dermatology; artificial intelligence; technology in dermatology

From AI Tools to Agentic Teammates

Medicine is at an inflection point as artificial intelligence (AI) evolves from passive tools to proactive teammates, a concept known as *agentic AI* (throughout this article, the terms “agentic AI,” “AI agents,” and “agentic teammates” are used interchangeably) [1]. While current AI tools are powerful, they are passive: a user has to initiate every interaction, provide all context, and interpret all outputs. In comparison, autonomous AI agents act without prompting, perform sequential steps independently, gather information from multiple sources, collaborate with humans or other agents, and self-improve through feedback loops [1]. Essentially, AI agents can behave as teammates with assigned roles and responsibilities. Of note, agentic AI is distinct from traditional rule-based automation, which executes fixed logic on structured inputs. In contrast, AI agents use large language models (LLMs) to reason across unstructured data and handle novel situations, which go beyond what deterministic automation can achieve.

Large health care systems have already begun leveraging agentic AI to reduce administrative burden and improve quality of care. For example, Duke Health intends to deploy AI agents in cardiology for care coordination that autonomously handle scheduling requests, optimize resource allocation, and connect patients to clinical trials [2]. At Ochsner Health, AI agents continuously review patient panels of primary care providers, reaching out to patients for screenings, lab monitoring, and postdischarge care coordination [3]. Oxford University Hospitals is piloting AI agents in oncology that summarize charts, determine tumor staging, and draft guideline-compliant plans for multidisciplinary tumor boards [4]. These early implementations demonstrate the feasibility of agentic AI across multiple specialties and clinical settings.

Opportunities and Risks for Dermatology

Dermatology is particularly well-positioned to benefit from these emerging AI capabilities given its high patient volumes, substantial administrative burdens, and clinicopathological diagnostic approach [5]. In a future dermatology practice, AI agents could tackle high patient volumes through intelligent scheduling (eg, spacing out procedures and complex visits), previsit charting, scribing and coding, and automated patient follow-up. Voice agents can reach even more patients through phone calls, including older adults and patients with limited English proficiency. AI agents can identify and recruit patients for clinical trials, particularly to address the critical need for diverse representation. Medication access (eg, prior authorization, appeals, patient assistance programs, and laboratory monitoring) for specialty drugs like biologics is also well-suited for an agentic workflow that eases the burden of manual labor on support staff. The unique clinicopathological workflow in dermatology also lends itself to multistep agentic

XSL•FO
RenderX

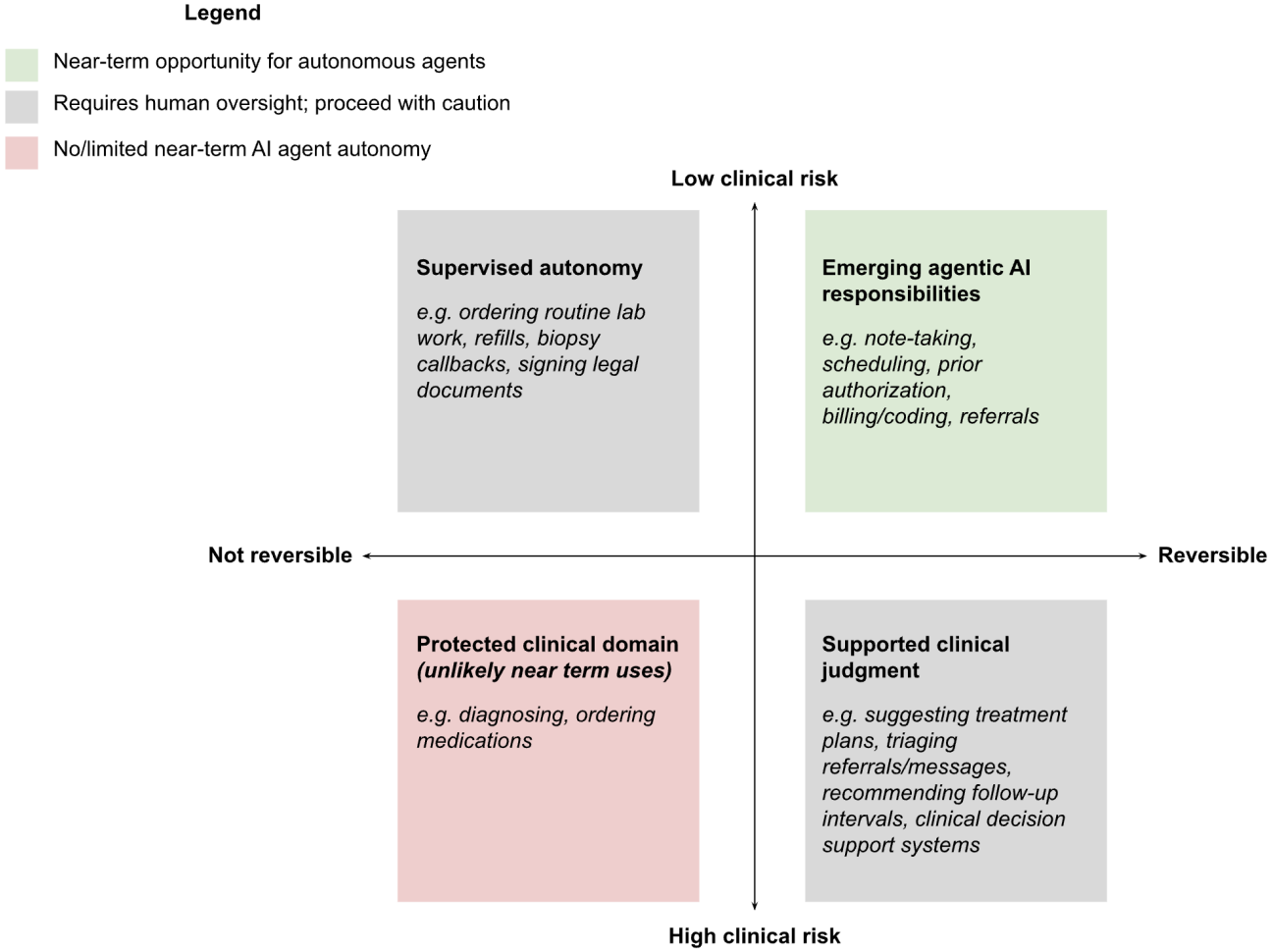
coordination: an agentic system could receive a pathology result, verify the correct patient and anatomic site against prior documentation and images, alert the dermatologist, and coordinate definitive treatment (eg, reaching out to a Mohs surgeon).

Despite the promising applications, there are limited published case studies of agentic AI in the dermatology literature. Nonetheless, there are numerous companies that have been launched in the past few years that are targeting every aspect of the dermatology workflow with an agentic approach. As health care systems and other specialties begin establishing the workflows, norms, and business models for agentic AI, dermatologists must actively educate themselves on this technology and engage to shape its real-world implementation.

Deploying agentic AI requires balancing risk; when considering pilot implementations, dermatologists should take a graduated approach for delegating tasks to autonomous agents. To assess readiness for agentic AI in dermatology practices, we propose the following framework (Figure 1): promising agentic applications ready for pilots focus on reversible, low-risk tasks. Tasks that are irreversible or carry higher clinical risk warrant stronger oversight by dermatologists or staff and therefore require additional caution before fully adopting agentic teammates. High-stakes, irreversible decisions with major clinical risk are not suitable for delegation to AI agents in near-term, real-world settings.

Significant barriers also impede the widespread deployment of AI agents in dermatology. First, current AI models are not ready for the reliability and complexity necessary in real-world clinical settings. A recent study demonstrated that the best-performing LLMs successfully completed 70% of representative clinical tasks, such as medication and test ordering, referrals, documentation, and patient communication. Models struggled most on tasks that required modifying records rather than simply retrieving information, highlighting the need for stepwise adoption [6]. Second, interoperability must be improved so that agents can gather data from siloed health records. Even with full access, agents need the right permissions to actually take action on behalf of dermatologists. Third, fundamental questions need to be addressed about liability, privacy, and regulation, as historically, regulatory agencies judge an AI-enabled medical device for a specific task, rather than a broad range of interconnected tasks [1,7].

Figure 1. Autonomous agentic AI is most ready for tasks that are reversible and of low clinical risk. When tasks are irreversible or of high clinical risk, stronger oversight with a human dermatologist is warranted. Tasks that are both irreversible and of high clinical risk (eg, diagnosing, ordering medications) are unlikely to be ready for near-term adoption and should have limited AI agent autonomy. The placement of example tasks within this framework is intended as a general guide, and exact positioning varies depending on institutional workflows, local staffing models, or electronic health record (EHR) capabilities.



XSL•FO
RenderX

Why Dermatologists Must Lead Now

Though this technology may seem distant for the practicing dermatologist, decisions about agentic AI are being made now. National and international dermatology organizations should include agentic AI on their agendas for AI use, seek partnerships with companies, and write guidelines for deploying AI agents. Unlike previous AI guidance focused on diagnostic performance, agentic AI guidelines must address multistep reasoning, minimum performance standards, mechanisms requiring human oversight, and protocols for monitoring ongoing safety. Concrete near-term actions include establishing a dedicated agentic AI task force within professional societies such as the American Academy of Dermatology, developing standardized reporting criteria for pilot programs in dermatology practices, and creating a voluntary registry of deployments to facilitate shared learning.

Research should also specifically examine agent performance across diverse patient populations to mitigate bias [7]. Although there are well-documented performance disparities of *diagnostic* AI tools on darker skin tones, it is unclear whether agentic AI systems that focus on nondiagnostic tasks will face the same challenges. However, agentic AI systems that rely on flawed components risk amplifying these disparities at scale, making them invisible in automated workflows. Safeguards should include ongoing postdeployment monitoring across demographic subgroups.

Agentic AI reflects an emerging paradigm where intelligent, autonomous agents act on behalf of physicians and patients, enhancing efficiency, continuity, and patient experience. There is no substitute for the human physician–patient relationship, and agentic AI offers dermatologists an opportunity to strengthen this relationship, not replace it. Dermatologists must not only stay informed of advances in agentic AI; we must also actively shape decisions about agentic AI to ensure appropriate, ethical, and optimized deployment to reflect our values and patients' needs.

Acknowledgments

All ideas, perspectives, and content are entirely original and written by humans. Generative AI tools were used as a collaborator to discuss strengths and weaknesses of the article, perform literature review, and suggest copyedits. Models used included OpenAI ChatGPT 5.1 and Claude Sonnet 4.5.

Conflicts of Interest

None declared.

References

1. Zou J, Topol EJ. The rise of agentic AI teammates in medicine. Lancet 2025 Feb 8;405(10477):457. [doi: 10.1016/S0140-6736(25)00202-8] [Medline: 39922663]
2. Talhelm M. DukeHealth. Duke Health, Trase Systems partner to develop AI health care tools. 2025 Sep. URL: https://corporate.dukehealth.org/news/duke-health-trase-systems-partner-develop-ai-health-care-tools [accessed 2026-05-20]
3. Madden B. AI-enabled primary care is no longer optional. Hospitalogy. 2025 Dec. URL: https://hospitalogy.com/articles/2025-12-04/ai-enabled-primary-care-is-no-longer-optional/ [accessed 2025-12-02]
4. Oxford-built multi-agent assistant for cancer care to be piloted in collaboration with Microsoft.: University of Oxford; 2025 Dec. URL: https://www.ox.ac.uk/news/2025-12-16-oxford-built-multi-agent-assistant-cancer-care-be-piloted-collaboration-microsoft [accessed 2025-12-16]
5. Omiye JA, Gui H, Daneshjou R, Cai ZR, Muralidharan V. Principles, applications, and future of artificial intelligence in dermatology. Front Med 2023;10:1278232. [doi: 10.3389/fmed.2023.1278232]
6. Jiang Y, Black KC, Geng G, et al. MedAgentBench: a virtual EHR environment to benchmark medical LLM agents. NEJM AI 2025 Aug 28;2(9). [doi: 10.1056/AIdbp2500144]
7. Arza A, Lebhar J, Lipoff JB. Applications of artificial intelligence in dermatology: ethical considerations. Dermatol Clin 2025 Oct;43(4):529-540. [doi: 10.1016/j.det.2025.05.003] [Medline: 41093475]

Abbreviations

AI: artificial intelligence
LLM: large language model

XSL•FO
RenderX

Edited by M Kirchberger; submitted 08.Mar.2026; peer-reviewed by M Hasselwander, S Drew; revised version received 19.Apr.2026; accepted 19.Apr.2026; published 05.Jun.2026.

Please cite as:
Chu B, Shen H, Lee I, Lipoff JB
Agentic AI in Dermatology: A Call to Action
JMIR Dermatol 2026;9:e94909
URL: https://derma.jmir.org/2026/1/e94909
doi:10.2196/94909

© Brian Chu, Heather Shen, Ivy Lee, Jules B Lipoff. Originally published in JMIR Dermatology (http://derma.jmir.org), 5.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Beta-Alanine and Aquagenic Pruritus: Proposed Neuroimmune Mechanism

Natalie Piserchio[1], BS; Bailey Baratta[2], MBA, DO; Benjamin Brooks[3], PhD; Brandon Muse[4], DO; Katelin Ball[5], DO

[1]Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

[2]Internal Medicine Residency HEB/Denton, Texas Health Resources, Bedford, TX, United States

[3]Department of Biomedical Sciences, Rocky Vista University, Ivins, UT, United States

[4]Department of Obstetrics and Gynecology, Orlando Health, Orlando, FL, United States

[5]Department of Neurology, University of New Mexico, Albuquerque, NM, United States

Corresponding Author:
Natalie Piserchio, BS
Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

Abstract

Aquagenic pruritus (AP) is a rare itch disorder with limited effective treatments, and emerging clinical observations suggest that oral β-alanine may reduce symptoms. The purpose of this viewpoint is to propose a biologically plausible mechanism through which β-alanine may alleviate primary AP. We reviewed published case reports and patient-reported survey data describing β-alanine use in AP and integrated these clinical observations with experimental data on MAS-related G protein–coupled receptor D (MrgprD)–expressing sensory neurons and their role in mast-cell regulation. Published case reports describe marked improvement in water-induced pruritus following prophylactic oral β-alanine administration, and a recent survey of patients with idiopathic AP reported substantial symptom relief among β-alanine users. Preclinical data indicate that MrgprD-neuronal glutamate release suppresses mast cell hyperresponsiveness, suggesting a potential pathway for the observed antipruritic effect. Additional mechanisms, including β-alanine metabolism to carnosine and its potential mast cell–stabilizing effects, may also contribute. β-alanine may act through modulation of a nonhistaminergic neuroimmune circuit and represents a promising therapeutic candidate for further investigation in AP.

(JMIR Dermatol 2026;9:e90737) doi:10.2196/90737

KEYWORDS

aquagenic pruritus; beta-alanine; primary aquagenic pruritus; MrgprD; nonhistaminergic itch; mast-cell hyperresponsiveness; neuroimmune signaling; glutamatergic signaling; pruriceptive pathways; neurocutaneous mechanisms; itch modulation; MAS-related G protein–coupled receptor D

Aquagenic pruritus (AP) is a rare skin disorder characterized by intense itching, tingling, or burning after contact with water, which can significantly impact an individual's quality of life [1]. Although AP is classically associated with polycythemia vera, idiopathic (primary) AP likely represents the most common form in the general population [2]. Therapeutic approaches to alleviate this condition remain challenging and often ineffective. Existing treatments primarily focus on symptom relief through topical emollients, phototherapy, antihistamines, and avoidance of water exposure, but they provide only partial relief and do not address the underlying mechanisms triggering the pruritus [1].

The pathogenesis of idiopathic AP is incompletely understood. Several mechanisms have been proposed, including abnormal mast cell activation, dysregulated cutaneous nerve signaling, and altered neuroimmune communication within the skin [3]. A small biopsy study found increased acetylcholinesterase activity in eccrine-associated nerve fibers after water exposure in patients with AP and polycythemia vera, suggesting increased acetylcholine release and possible involvement of eccrine-associated pathways in AP pathogenesis [4]. AP has also been described in familial cases, suggesting a potential genetic predisposition [1].

Recent case reports have shed light on the role of β-alanine, a nonessential amino acid, in modulating the sensation of itch. β-alanine is a known agonist at the pruriceptive receptor MAS-related G protein–coupled receptor D (MrgprD) and induces a weak nonhistaminergic itch in humans and mice [5]. A survey-based study investigating β-alanine supplementation among 75 patients with idiopathic AP reported that 70.7% of participants who used β-alanine described substantial symptom relief, with an average relief score of 8.84 out of 10 (95% CI 8.52 - 9.16) [6].

In one case report, a 33-year-old adult male with primary AP found that prophylactic oral β-alanine taken 5 to 15 minutes before water exposure markedly reduced pruritus with sustained improvement at the 20-week follow-up [5]. In a second case, a 16-year-old male with AP refractory to traditional treatments

XSL•FO
RenderX

also found relief from oral β-alanine taken prior to water exposure [7].

In these reports, β-alanine was administered orally in powder form [5,7]. Survey data suggest that patients most commonly use doses averaging approximately 1.59 g per day during acute exacerbations. In this survey, 66% of users preferred powder formulations and 73.6% reported taking β-alanine on an as-needed basis [6].

The mechanism responsible for this improvement is unknown, and because MrgprD activation is typically pruritogenic, the antipruritic effect is unlikely to result from direct receptor agonism. Other mechanisms may also contribute to the observed therapeutic benefit. For example, β-alanine is a precursor to carnosine, a dipeptide with antioxidant and intracellular pH-buffering properties that may help stabilize mast cell responses. These effects could complement or act independently of MrgprD-mediated neural pathways [6].

Although the mechanisms underlying primary AP are not fully established, the delay between water exposure and itch onset reported in many patients may suggest a nonhistaminergic process involving abnormal neuroimmune interactions between sensory afferents and epidermal or immune cells such as keratinocytes, basophils, or mast cells [5]. A mouse study from 2021 demonstrated that MrgprD-expressing cutaneous sensory neurons help maintain skin homeostasis by releasing glutamate, which dampens mast cell hyperresponsiveness [3]. One possibility is that β-alanine may transiently enhance MrgprD-neuronal activity enough to increase local glutamate release, thereby strengthening this homeostatic inhibitory circuit. In primary AP, where mast cell hyperresponsiveness and nonhistaminergic pathways are suspected, enhanced MrgprD-mediated glutamatergic signaling could suppress mast cell activation and reduce downstream pruriceptive input. Recent work from the same research group further supports this model by demonstrating that glutamate can act directly on both mouse and human mast cells to suppress their activation, reinforcing the potential role of glutamatergic signaling in cutaneous immune homeostasis [8].

Available patient-reported data suggest that β-alanine is generally well tolerated, with transient paresthesia being the most frequently reported side effect. This is consistent with the known safety profile of β-alanine in sports nutrition studies and is typically mild and self-limited. Importantly, 90% of surveyed users reported sustained therapeutic benefit without loss of efficacy over time, suggesting minimal tachyphylaxis [6,9].

Although speculative, this mechanism aligns with both the delayed nature of AP symptoms reported by some individuals and the observed clinical benefit of prophylactic β-alanine administration. Given the debilitating nature of AP and the scarcity of effective therapies, the emerging evidence supporting β-alanine as a fast-acting, well-tolerated, inexpensive, and accessible intervention is noteworthy.

We encourage further controlled studies to characterize optimal dosing, duration of effect, long-term safety, and the underlying neurocutaneous mechanisms. If β-alanine consistently alleviates symptoms in primary AP, this observation may also support the concept that non–FcεR-mediated mast cell activation and neuroimmune dysregulation play important roles in disease pathogenesis.

β-alanine may represent a promising therapeutic avenue for a condition that currently lacks reliable treatment options. These observations highlight the need to further explore β-alanine as a targeted modulator of nonhistaminergic itch pathways in primary AP.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Siegel FP, Tauscher J, Petrides PE. Aquagenic pruritus in polycythemia vera: characteristics and influence on quality of life in 441 patients. Am J Hematol 2013 Aug;88(8):665-669. [doi: 10.1002/ajh.23474] [Medline: 23657863]
2. Steinman HK, Greaves MW. Aquagenic pruritus. J Am Acad Dermatol 1985 Jul;13(1):91-96. [doi: 10.1016/s0190-9622(85)70149-1] [Medline: 2411768]
3. Zhang S, Edwards TN, Chaudhri VK, et al. Nonpeptidergic neurons suppress mast cells via glutamate to maintain skin homeostasis. Cell 2021 Apr 15;184(8):2151-2166. [doi: 10.1016/j.cell.2021.03.002] [Medline: 33765440]
4. Bircher AJ, Meier-Ruge W. Aquagenic pruritus: water-induced activation of acetylcholinesterase. Arch Dermatol 1988;124(1):84-89. [doi: 10.1001/archderm.1988.01670010048020] [Medline: 3337547]
5. Alinaghi F, Elberling J. Primary aquagenic pruritus successfully treated by β-alanine: a case report. Case Rep Dermatol 2024;17(1):252-254. [doi: 10.1159/000545842] [Medline: 40625687]
6. Concistrè A. Beta-alanine as a potential treatment for aquagenic pruritus: an online social media-based survey study. Acta Dermatovenerol Croat 2024 Dec;32(4):123-134. [Medline: 40657651]
7. Degala J, Degala A. The use of beta-alanine for the management of aquagenic pruritus. Ann Allergy Asthma Immunol 2025 Nov;135(5):S156-S157. [doi: 10.1016/j.anai.2025.08.469]

XSL•FO
RenderX

8. Zhang YR, Keshari S, Kurihara K, et al. Agonism of the glutamate receptor GluK2 suppresses dermal mast cell activation and cutaneous inflammation. Sci Transl Med 2024 Dec 11;16(777):eadq9133. [doi: 10.1126/scitranslmed.adq9133] [Medline: 39661706]
9. Saunders B, Elliott-Sale K, Artioli GG, et al. β-alanine supplementation to improve exercise capacity and performance: a systematic review and meta-analysis. Br J Sports Med 2017 Apr;51(8):658-669. [doi: 10.1136/bjsports-2016-096396] [Medline: 27797728]

Abbreviations

AP: aquagenic pruritus
MrgprD: MAS-related G protein–coupled receptor D

Edited by R Dellavalle; submitted 02.Jan.2026; peer-reviewed by A Concistre, D Kaplan; revised version received 10.Mar.2026; accepted 10.Mar.2026; published 25.Mar.2026.

Please cite as:
Piserchio N, Baratta B, Brooks B, Muse B, Ball K
Beta-Alanine and Aquagenic Pruritus: Proposed Neuroimmune Mechanism
JMIR Dermatol 2026;9:e90737
URL: https://derma.jmir.org/2026/1/e90737
doi:10.2196/90737

© Natalie Piserchio, Bailey Baratta, Benjamin Brooks, Brandon Muse, Katelin Ball. Originally published in JMIR Dermatology (http://derma.jmir.org), 25.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Current Management and Unmet Needs in Alopecia Areata: Viewpoint From a Malaysian Perspective

Azura Mohd Affandi[1], AdvMDerm; Adawiyah Jamil[2], AdvMDerm; Noor Zalmy Azizan[3], AdvMDerm; Felix Boon-Bin Yap[4], AdvMDerm; Kin Fon Leong[5]; Agnes Yoke Hui Heng[6], MBBS; Siti Nuraihan[7], MPaed; Wooi Chiang Tan[8], AdvMDerm; Latha Selvarajah[9], AdvMDerm; Teeba Raja[10], AdvMDerm; Evelyn Wen Yee Yap[11], AdvMDerm; Peter Wee Beng Ch'ng[12], AdvMDerm; Tze Yuen Teoh[3], AdvMDerm; Mei Ying Tey[13], RPh

[1]Department of Dermatology, Ministry of Health, Hospital Kuala Lumpur, Kuala Lumpur, Malaysia
[10]Department of Dermatology, Ministry of Health, Hospital Selayang, Selangor, Malaysia
[11]Department of Dermatology, Mahkota Medical Centre, Melaka, Malaysia
[12]Department of Dermatology, Gleneagles Hospital, Kuala Lumpur, Malaysia
[13]Medical Affairs, Pfizer, Level 10 & 11, Wisma Averis, Tower 2, Avenue 5, Bangsar South, No 8, Jalan Kerinchi, Kuala Lumpur, Malaysia
[2]Faculty of Medicine, University Kebangsaan Malaysia, Kuala Lumpur, Malaysia
[3]Central Dermatology Specialist Clinic, Kuala Lumpur, Malaysia
[4]Sunway Medical Centre, Selangor, Malaysia
[5]Dermatology Unit, Hospital Tunku Azizah, Kuala Lumpur, Malaysia
[6]Agnes Heng Dermatology, Ipoh, Malaysia
[7]Pediatric Dermatology Unit, Pediatric Department, Ministry of Health, Hospital Selayang, Selangor, Malaysia
[8]Department of Dermatology, Ministry of Health, Hospital Pulau Pinang, Pulau Pinang, Malaysia
[9]Department of Dermatology, Ministry of Health, Hospital Sultan Ismail, Johor Bahru, Malaysia

Corresponding Author:
Mei Ying Tey, RPh
Medical Affairs, Pfizer, Level 10 & 11, Wisma Averis, Tower 2, Avenue 5, Bangsar South, No 8, Jalan Kerinchi, Kuala Lumpur, Malaysia

Abstract

Alopecia areata (AA) is a chronic, immune-mediated disease characterized by nonscarring hair loss and a substantial clinical and psychosocial burden, yet it is still widely perceived as a purely cosmetic concern. In this viewpoint, we draw on the available Malaysian studies, selected international evidence, and our collective clinical experience to propose that AA warrants greater clinical, research, and policy attention in Malaysia. Malaysia-specific evidence remains limited and is derived largely from retrospective, single-center studies; nonetheless, the available data point to variable disease presentation, inconsistent use of standardized severity assessment, reliance on conventional therapies, substantial loss to follow-up, and a meaningful psychosocial impact among affected patients. Current local management appears to rely mainly on conventional therapies, while access to advanced therapies such as Janus kinase inhibitors remains constrained by affordability and reimbursement challenges. We therefore call for 5 priorities: strengthening Malaysia-specific epidemiological and real-world evidence, including through a national AA registry; adopting standardized disease-assessment tools in clinical studies and routine practice; integrating psychosocial support into AA care pathways; developing locally relevant clinical guidance; and improving public, clinician, and policymaker awareness that AA is a medical condition rather than a cosmetic concern. Together, these steps would support earlier diagnosis, more equitable access to effective treatment, and holistic, patient-centered care for adults and children living with AA in Malaysia.

(JMIR Dermatol 2026;9:e89583) doi:10.2196/89583

KEYWORDS

alopecia; alopecia areata; Malaysia; reimbursement; Janus kinase inhibitors; JAK inhibitors

Introduction

Alopecia areata (AA) is a chronic autoimmune condition marked by nonscarring hair loss, primarily affecting the scalp, face, and body [1]. The disorder arises when the immune system mistakenly targets active hair follicles, resulting in sharply defined patches of hair loss without permanently damaging the follicular structure. Although many patients may experience spontaneous hair regrowth, AA often follows a relapsing course requiring continuous medical management [2]. AA affects approximately 1 in every 1000 individuals and carries a lifetime risk of around 2% [2]. The condition may occur across all age

XSL•FO
RenderX

groups; however, it is somewhat more frequent in children, with AA rates of 1.92% compared with 1.47% in adults [3].

The condition is frequently associated with comorbidities, including vitiligo, lupus, psoriasis, atopic dermatitis, and thyroid disorders [2]. Additionally, other comorbidities such as metabolic syndrome, iron deficiency anemia, vitamin D deficiency, psychiatric diseases, audiologic, and ophthalmic abnormalities are more prevalent in patients with AA, who also have a higher risk of developing other autoimmune conditions [4].

The average age of onset of AA is approximately 30 years, and patients with severe AA tend to have a longer disease duration [5]. Male , nail involvement, and disease duration longer than 1 year are identified as significant risk factors for severe AA [5]. Although AA is not life-threatening, its psychological impact is significant, and it can lead to anxiety disorder, depression, and stress, often necessitating consultation with mental health professionals [6,7]. Anxiety is significantly increased in patients with AA, who are reported to have 2.5 times higher odds of experiencing anxiety compared with control groups [3]. Children and adolescents with AA are equally affected [8], with much greater susceptibility to bullying [9], and recent research indicates that 51% of pediatric patients meet the criteria for an anxiety disorder and 62.9% show at least one psychiatric condition [10].

Malaysia, an upper-middle-income nation with a population of approximately 34 million, is strategically situated in Southeast Asia as a regional hub for information and communication technology and medical tourism [11]. In Malaysia, patients with AA may initially present to primary care settings, including government health clinics (Klinik Kesihatan), private general practitioners, and aesthetic physicians. While some patients with mild, limited hair loss may be managed initially with topical corticosteroids, mild disease may also warrant dermatology referral for diagnostic confirmation, assessment of disease activity, and specialist management. Referral to dermatology is particularly important in cases of diagnostic uncertainty, extensive or rapidly progressive disease, inadequate response to initial treatment, or when systemic treatment is being considered. The public health care system follows a structured referral pathway from primary care to dermatology departments in major hospitals, whereas patients willing to pay out of pocket may access private dermatologists directly. Despite this structure, there is a notable gap in epidemiological data, comprehensive assessments of disease burden, and locally tailored clinical guidelines for AA in Malaysia. Given the significant impact of AA on quality of life and overall health, this Viewpoint aims to consolidate current knowledge on AA epidemiology, clinical burden, and management practices in Malaysia and to set out the key unmet needs within the Malaysian context. Through this Viewpoint, we aim to inform health care professionals, researchers, and policymakers, and to guide future research and health care initiatives that enhance care for patients with AA.

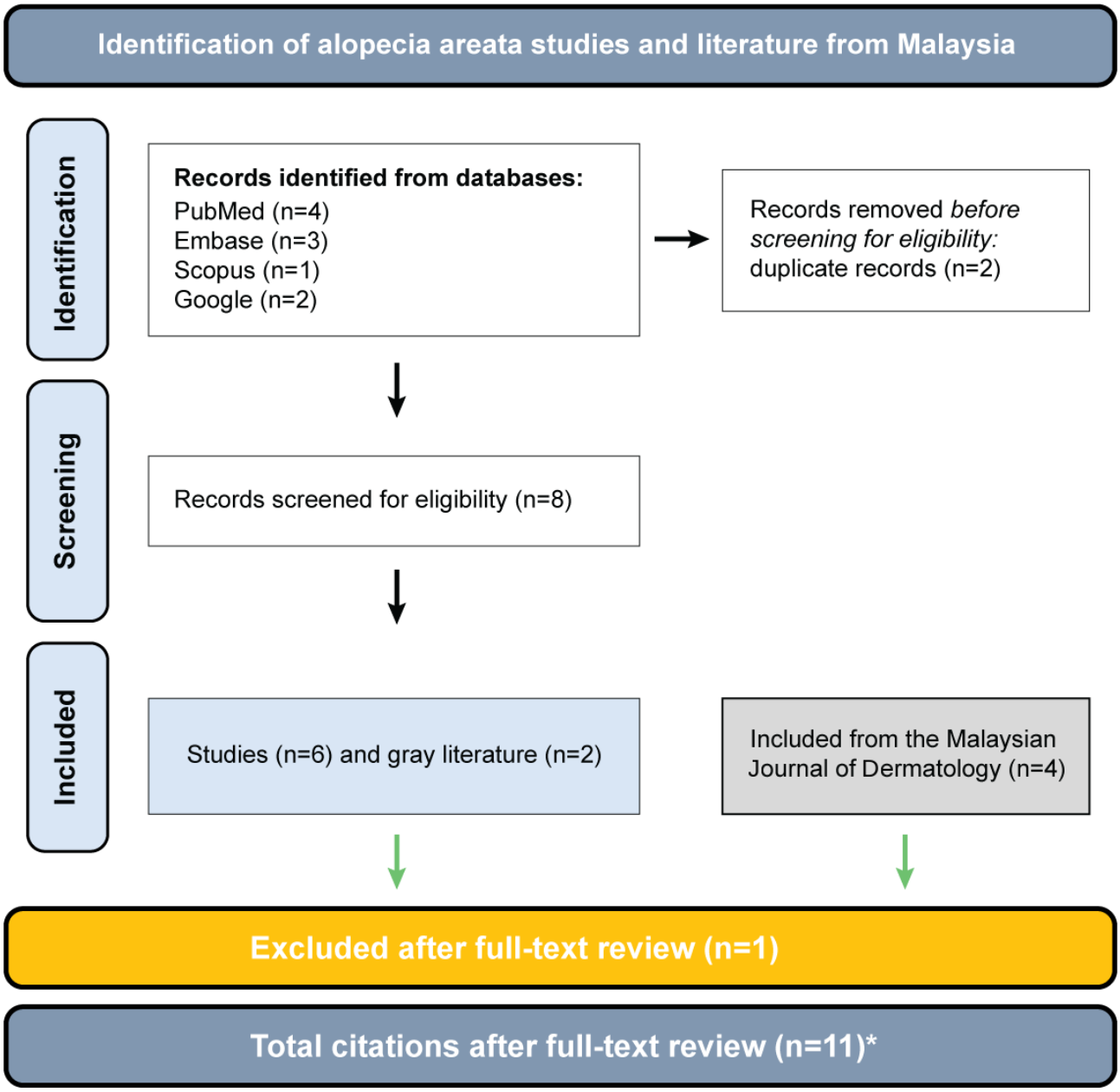
Opportunity to Strengthen Malaysia-Specific Evidence Base for AA

AA has been briefly studied in Malaysia across multiple patient cohorts, providing insight into the disease's demographic and ethnic distribution [12-22] (Figure 1; Multimedia Appendix 1). A study conducted at a tertiary dermatology center in Kuala Lumpur analyzed 200 patients with AA [15]. The findings from this cohort revealed a median age of onset of 25 years, with a male-to-female ratio of 1.38:1. The median duration from the onset of symptoms to presentation was 3 months. The frequency of AA in a dermatology outpatient clinic at Hospital Kuala Lumpur in Kuala Lumpur, Malaysia, was 1.32% in 58,252 patients seen over 7 years from 2008 to 2014 [19]. Another study from the Universiti Malaya Medical Centre reviewed the medical records of 154 patients with nonscarring alopecia (28.6% with AA) from 2000 to 2009 [16,22]. This study reported a mean age at presentation of 24.5 (range 4 - 62) years, with 30.1% of patients aged <18 years and 20.5% aged <16 years [16,22]. The male-to-female ratio in this study was nearly equal, at 1:1.05 [16].

The ethnic distribution across the studies was variable but relatively balanced, with most studies showing a higher frequency in Malays, followed by Indians. Heah et al [19] found that Malays represented 69.2% of AA cases, followed by Indians 28% and Chinese 10%, while Yin and Chew [22] reported the highest proportion of patients as Malays (34.1%), followed by Indians (31.8%), and Chinese (22.7%). However, Tang et al [15] reported that Indians had the highest frequency of AA, accounting for 35.5% of cases. Meanwhile, the study by Lee and Lee [16] found that Malays and Indians each represented 32.7% of cases. However, ethnic variations reported in these retrospective Malaysian AA cohorts may also reflect the catchment populations of individual hospitals rather than actual national distributions. Moreover, these frequencies were not reported against the background of Malaysia's overall ethnic composition, and as such, the degree to which certain ethnic groups are overrepresented is not discernible.

XSL•FO
RenderX

Figure 1. Literature search for the identification of alopecia areata studies from Malaysia. The search used the keywords "alopecia areata" and "Malaysia," and databases were searched through March 2025. *One citation did not include patients from Malaysia and was excluded after full-text review [23].



AA exhibits a range of distinct clinical patterns that can vary in severity and progression [3]. Recognizing these different presentations is essential for accurate diagnosis, prognosis, and management, as they may reflect underlying differences in disease activity or response to treatment. Tang et al [15] have best characterized clinical presentations at Kuala Lumpur Hospital in Malaysia: 98.5% (197/200) of patients presented with patchy scalp involvement—often extending to eyebrows, eyelashes, or facial hair—while ophiasis was seen in 3 cases. More extensive disease was uncommon, with alopecia totalis in 0.5% and universalis in 1.0% of patients [15].

On the basis of the available Malaysian studies, existing local evidence is derived mostly from single-center cohorts rather than nationwide data sources. Current evidence warrants further characterization of national prevalence, incidence, severity distribution, treatment patterns, long-term outcomes, and patient-reported burden. Future priorities may include establishing national AA registries to capture robust epidemiological data and the clinical spectrum of disease, to better define the national disease burden, and to support clinical care, treatment decisions, and guideline development.

Standardize Disease-Assessment Tools in Clinical Studies and Routine Clinical Use

Various tools have been developed to quantify AA severity, and accurate assessment is essential for monitoring progression and guiding treatment [24]. The Severity of Alopecia Tool (SALT), first proposed in 1999, is widely used in clinical trials to quantify hair loss [24,25]. Other objective assessments, such as body surface area (BSA), are also critical for determining eligibility for systemic treatments [26].

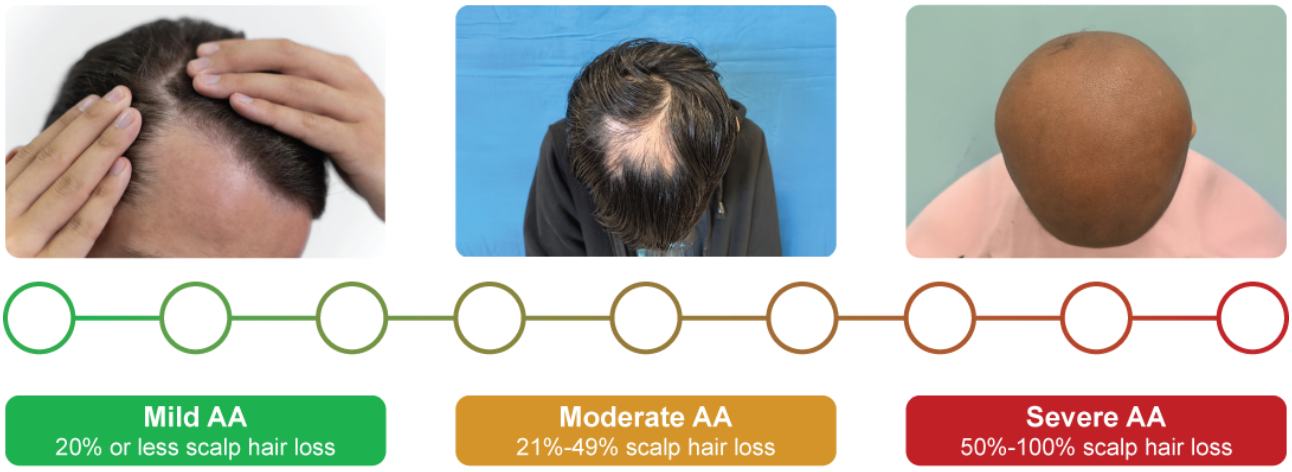
XSL•FO
RenderX

While some Malaysian studies have incorporated disease-assessment measures, these have varied across studies, limiting standardization and comparability of disease severity and treatment outcomes. Albela et al [14] categorized the extent of hair loss and regrowth based on assessment criteria by Olsen et al [27], using the SALT score, and defined severe AA as involvement of more than 50% of the total scalp area (S3-S5). Ang and Tang [20] adopted the clinical response measurement by McDonald Hull and Norris grading system. Grades 1 and 2 were considered no response, while grades 3 and 4 were considered a good response. In Kaur [18], disease activity was measured using the hair pull test. The absence of a consistent severity scoring system may hinder accurate disease monitoring, treatment response evaluation, and the development of localized clinical guidelines.

Given the potentially high burden of AA, accurate diagnosis and assessment of disease severity are important to guide effective intervention [28]. Several formal measures of disease severity are used in AA clinical trials, including clinician-reported outcomes such as Severity of Alopecia Tool (SALT), Eyebrow Assessment Scale (EBA), Eyelash Assessment Scale (ELA), Alopecia Areata Investigator Global Assessment (AA-IGA), and patient-reported outcomes such as Alopecia Areata Patient Priority Outcome (AAPPO) and Patient's Global Impression of Change (PGI-C) [27,29-31].

Additionally, a recently developed Alopecia Areata Severity Scale offers a practical, clinic-friendly framework for grading AA. It uses the extent of scalp hair loss as the primary basis for a severity rating combined with 4 additional secondary clinical features, which, when present, permit an increase in the severity level (Figure 2) [32]. In local clinical practice, measures such as BSA involvement, the number of affected patches, serial photography, and overall clinical presentation are used to support assessment of disease severity and monitoring of treatment response. Standardized disease-assessment tools are recommended to guide care, monitor response, facilitate clinical decision-making, and support access or reimbursement discussions.

Figure 2. The Alopecia Areata Severity Scale for disease assessment in routine clinical practice. Scalp hair loss is graded as mild, moderate, or severe; 4 secondary features each permit an increase of 1 severity level [32]. Photographs courtesy of Dr Kin Fon Leong (images for moderate and severe AA) and Freepik (image for mild AA). AA: alopecia areata.



*If AA is mild or moderate, increase the severity rating by one level if any of the following are present: negative impact on psychosocial functioning; noticeable involvement of eyebrows or eyelashes; inadequate response after at least 6 months of treatment; or a diffuse (multifocal) positive hair pull test indicating rapidly progressive AA.

Recognition of Clinical Burden and Psychosocial Impact of AA

AA presents a notable clinical burden in Malaysia, both in terms of treatment and its impact on patients' quality of life [18]. Khoo et al [21] studied the psychosocial impact of alopecia in 240 patients with nonscarring alopecia attending a tertiary outpatient dermatology clinic from mid-2021 to mid-2022 in Penang, Malaysia. The median age was 38.5 (range 18 - 70) years, the majority were of Malay ethnicity (42.5%), and 18.8% of patients were diagnosed with AA [21]. Most patients expressed significant concern about their hair loss, with younger individuals and unmarried respondents reporting higher levels of anxiety, depression, and impaired quality of life and female patients expressing greater concern about hair loss than male patients [21]. These findings emphasize that the psychosocial burden of alopecia is profound and requires targeted interventions in clinical practice [21].

Albela et al [14] recognized that AA can substantially affect health-related quality of life, particularly mental health. Patients with AA are also at increased risk of psychiatric comorbidities, including anxiety, depression, social phobia, and personality disorders, compared with the general population.

These findings highlight the complex emotional toll of AA and the importance of addressing mental health as a core component of patient care. Integrating psychological support into clinical management is essential to improve outcomes and provide more holistic care for patients.

Sex has been proposed as a factor in the psychosocial experience of AA, although the available evidence does not indicate a consistent difference in burden between male and female. In the local study by Khoo et al [21], female patients expressed

XSL•FO
RenderX

greater concern about their hair loss than male patients, but there was no significant difference between male and female patients in Dermatology Life Quality Index or Hospital Anxiety and Depression Scale scores. A plausible explanation for the present observation is that men's attitudes toward their hair and appearance are now comparable to those of women. Male patients should therefore be regarded as equally concerned and equally vulnerable, and their emotional burden should not be assumed to be lesser. In the cohort reported by Tang et al [15], 41% of patients did not return after the initial consultation, but this loss to follow-up was not stratified by gender, nor were the underlying reasons, such as inadequate treatment response, cost, or accessibility, systematically captured. The management of AA may also benefit significantly from a multidisciplinary, team-based approach that addresses both the physical and psychological effects of AA. Beyond dermatological treatment, individuals with AA often experience significant emotional and psychosocial challenges, including anxiety, depression, and reduced self-esteem [6,8]. Integrating mental health professionals, such as counselors or psychologists, into the care team may help provide holistic support [6]. Collaborative care models that bring together dermatologists, mental health specialists, and primary care providers could enhance patient outcomes. As such, there is a critical need to strengthen the integration of mental health services within existing AA management frameworks, promoting psychological intervention and ongoing support as standard components of care.

Current Treatment Landscape and Implications for Targeted Therapies

Managing AA remains challenging due to unpredictable disease progression, variable treatment response, and high recurrence rates, especially with systemic steroids [33,34]. Additionally, long-term use of immunosuppressants such as methotrexate and azathioprine raises safety concerns [35,36].

The primary treatment modalities for AA in Malaysia range from corticosteroids (topical, oral or intralesionals), immunosuppressive therapy to phototherapy in selected hospitals, depending on age, subtype, and severity [20]. Yin and Chew [22] reported the use of topical steroids in more than half of 44 patients with AA in Malaysia (53.3%), followed by intralesional steroids (26.7%), topical minoxidil (17.8%), oral steroids (11.1%), oral finasteride (2.2%) and azathioprine (used concurrently with oral steroids) in 2.2% of patients. According to the retrospective 5-year audit at Hospital Kuala Lumpur between 2011 and 2015, AA accounted for 0.7% of cases treated with phototherapy [13]. Tang et al [15] found that topical corticosteroids were the most commonly used treatment, although other therapies, such as topical agents and phototherapy, were also used. However, the study also noted that 41% of patients did not return for follow-up care after the initial consultation, possibly suggesting challenges with treatment adherence. Ang et al [20] reported a case series of 11 patients with AA treated with topical immunotherapy using squaric acid dibutyl ester (SADBE). The median age of the patients was 13 (IQR 9-33) years; the majority were female (73%) and of Malay ethnicity (55%). The response rates to SADBE application were better in less severe forms of AA (80%) than in severe forms of AA (66%), and 28% of patients did not respond to SADBE [20]. A retrospective case series by Albela et al [14] reviewed 13 pediatric patients (6 boys and 7 girls, aged 4-16 years) with severe AA treated with oral methotrexate administered once weekly, with a mean dose of 0.4 mg/kg, at a center from January 2019 to December 2020. The authors observed that 5 of 12 assessable patients experienced more than 50% regrowth, while the remaining 7 had treatment failure.

Current local management appears to rely heavily on corticosteroids and conventional therapies, with treatment approaches often shaped by availability, cost, clinician experience, and disease severity. Globally, intralesional steroids remain the most widely used treatment for AA across all severity levels, followed by topical steroids, oral steroids, topical minoxidil, and Janus kinase (JAK) inhibitors [37,38], although intralesional steroid injections are often avoided in younger children due to discomfort and pain [39]. However, recent trends indicate a decline in steroid and minoxidil use as JAK inhibitors are gaining preference among dermatologists [40]. European expert consensus prioritize JAK inhibitors as a first-line systemic treatment for nonacute AA, before considering alternative systemic agents such as cyclosporine or methotrexate [41].

The introduction of JAK inhibitors has transformed the treatment landscape, offering a tolerable and effective option for patients with severe AA [42,43]. JAK inhibitors have shown promising results in improving treatment outcomes with early intervention in severe AA [43]. While JAK inhibitors have changed the global treatment landscape for moderate to severe AA, affordability, reimbursement, and equitable access remain major barriers in Malaysia. Treatment access discussions should be positioned around medical need, disease burden, validated severity assessment, and patient-centered outcomes. Attention should also be given to the ongoing challenges surrounding the reimbursement and accessibility of novel treatment options, particularly advanced therapies such as JAK family kinase inhibitors. These treatments, while showing promise in clinical outcomes, remain financially out of reach for many patients due to limited insurance coverage.

The absence of clearly defined national clinical practice guidelines highlights the need for more robust management strategies to support dermatologists, including pediatric dermatologists, and patients in shared treatment decision-making. National clinical guidelines could provide evidence-based recommendations for diagnosis, severity assessment, referral, treatment sequencing, monitoring, psychosocial care, and access discussions.

In the absence of country- or region-specific AA disease management guidelines, international guidelines and consensus statements can guide disease assessment, appropriate therapy selection, and treatment duration [41,44-46].

XSL•FO
RenderX

Awareness of AA as a Chronic Immune-Mediated Disease

AA is not merely a cosmetic condition; it imposes a significant psychosocial burden, often resulting in social isolation, impaired productivity, and reduced quality of life [6,8,47,48]. Additionally, AA is linked to systemic conditions such as thyroid dysfunction and atopic dermatitis[4]. The public, medical professionals, and decision-makers must recognize that an active, immune-driven process leads to ongoing hair loss and distress in AA, and patients desperately seek treatments that halt disease progression and restore their original appearance [49,50]. Reframing AA as a chronic inflammatory disease rather than just a cosmetic concern could help drive policy changes and improve treatment accessibility [48].

Increased awareness of and advocacy about AA among the general population are needed through public education, which may aid in combating misconceptions about AA [18]. Hair fulfills both functional (eg, loss of eyelashes predisposes to eye irritation) and aesthetic roles and is crucial to self-image, self-expression, identity, and social perception, and its loss has been linked to increased anxiety, depression, and a diminished quality of life, particularly among younger individuals [21,51]. As a result, many patients turn to nonprescription and alternative treatments, a trend reflected by the fast-growing market for such products and supported by annual expenditures of US $30.2 billion on alternative medicine [21]. Khoo et al [21] noted that a significantly higher proportion of Malay patients opted for nonprescription and alternative therapies compared with other ethnic groups, suggesting hesitancy to try prescription medications. A common reason for avoiding consultation was the perception that hair loss is a natural part of aging, cited by 35.7% of respondents. Additionally, almost half of the patients (46.5%) reported dissatisfaction with their consultations, primarily due to the lack of specific treatment recommendations and unanswered questions [21]. Recognizing these attitudes and behaviors is essential for physicians and policymakers to improve consultation strategies and tailor interventions that address alopecia's psychological and physical dimensions. Awareness of AA among patients, hair care professionals, and health care providers and prompt referral to a dermatologist are also important [52,53]. The rapid rise of hair salons and centers offering unapproved therapies risks delaying proper medical management and undermining effective care [54]. Public education campaigns can help reframe AA as an autoimmune disease with physical and psychological impacts rather than a purely cosmetic issue [55].

In the Malaysian setting, the reluctance to seek medical consultation reported by Khoo et al [21] appeared to reflect a knowledge gap, with hair loss commonly regarded as a natural part of aging and uncertainty as to whether it constitutes a medical condition.

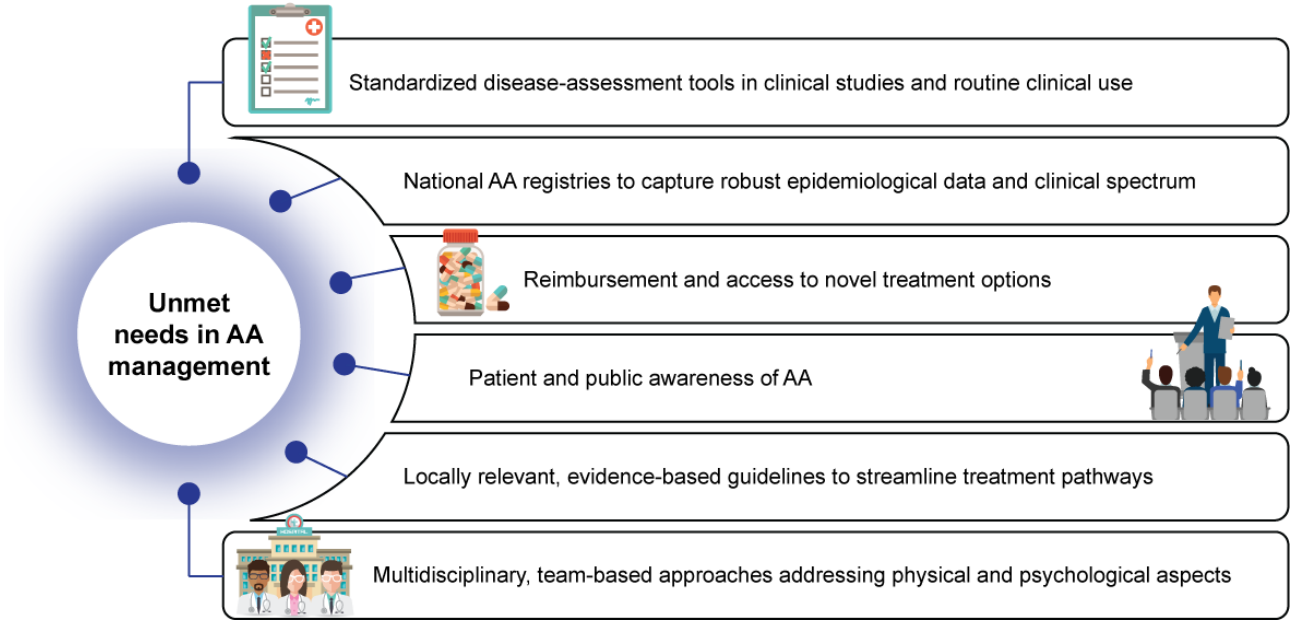
The findings highlight the need for health care initiatives that improve public education and raise awareness and understanding of AA, as low awareness and stigmatization may contribute to psychological distress, delayed care-seeking, and uncertainty among patients about whether AA is a medical condition requiring clinical management. Public and clinician awareness should emphasize that medical treatment options are available and that timely referral may support appropriate assessment and care.

Conclusions

AA is a chronic, immune-mediated disease with substantial clinical and psychosocial impact, yet Malaysia-specific evidence remains limited and is largely derived from retrospective, single-center studies. Available local data suggest variable disease presentation, the need for standardized severity-assessment tools, reliance on conventional therapies, substantial loss to follow-up, and a meaningful psychosocial burden among affected patients. These findings highlight the need to strengthen local epidemiological and real-world evidence, adopt standardized disease-assessment tools in clinical studies and routine practice, and integrate psychosocial support into AA care pathways (Figure 3). Improving public and clinician awareness is also important to reframe AA as a medical condition rather than a cosmetic concern and to encourage timely referral and appropriate management. Looking ahead, locally relevant clinical guidance to support evidence-based management, multidisciplinary care models for patient-centered AA care, and equitable access to effective treatment options, including advanced therapies where appropriate, will be essential to improve outcomes for both adult and pediatric patients with AA in Malaysia.

XSL•FO
RenderX

Figure 3. Unmet needs in the management of alopecia areata (AA) in Malaysia. Priority actions to address gaps in AA epidemiological data, severity assessment, treatment access, public awareness, and multidisciplinary care in Malaysia.



Acknowledgments

The authors sincerely appreciate the scientific support and medical writing assistance provided by Dr Sajita Setia and the editorial support provided by Dansoa Tabi-Amponsah from Transform Medical Communications, New Zealand.

The authors would like to thank the Director General of Health Malaysia for his permission to publish this article.

The views and opinions expressed in this article are those of the authors and do not represent or reflect in any way the official policy or position of their current or previous employers or collaborators. A meeting convened by Pfizer to address the unmet needs in the management of alopecia areata in Malaysia and attended by the authors served as a springboard for subsequent manuscript development by all authors.

Funding

The development of this manuscript and the open access publication fee were supported by Pfizer. Medical writing support was provided by Transform Medical Communications and funded by Pfizer. The authors did not receive any honoraria, personal fees, or direct financial compensation for the preparation of this manuscript.

Data Availability

All data generated for this manuscript are included in this published article and its supplementary material. Other datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors made substantial contributions to the conception and design of the study, acquisition, analysis, and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; gave final approval of the version to be published; and agreed to be accountable for all aspects of the work.

Conflicts of Interest

MYT is an employee of Pfizer, a manufacturer of ritlecitinib, which is indicated for the treatment of alopecia areata. The authors report no other conflicts of interest in this work. No author received an honorarium for the preparation of the article.

Multimedia Appendix 1
Summary of the identified alopecia areata studies and literature from Malaysia.
[DOCX File, 28 KB - derma_v9i1e89583_app1.docx]

References

XSL•FO
RenderX

1. Jolkovsky EL, Goldberg DJ. Alopecia areata treatment: past, present, and future. Dermatol Rev 2024 Dec;5(6):e70008. [doi: 10.1002/der2.70008]
2. Lepe K, Syed HA, Zito PM. Alopecia areata. In: StatPearls: StatPearls Publishing; 2025.
3. Sibbald C. Alopecia areata: an updated review for 2023. J Cutan Med Surg 2023;27(3):241-259. [doi: 10.1177/12034754231168839] [Medline: 37340563]
4. Lee S, Lee H, Lee CH, Lee WS. Comorbidities in alopecia areata: a systematic review and meta-analysis. J Am Acad Dermatol 2019 Feb;80(2):466-477.e16. [doi: 10.1016/j.jaad.2018.07.013] [Medline: 30031145]
5. You HR, Kim SJ. Factors associated with severity of alopecia areata. Ann Dermatol 2017 Oct;29(5):565-570. [doi: 10.5021/ad.2017.29.5.565] [Medline: 28966512]
6. Mulinari-Brenner F. Psychosomatic aspects of alopecia areata. Clin Dermatol 2018;36(6):709-713. [doi: 10.1016/j.clindermatol.2018.08.011] [Medline: 30446192]
7. García-Hernández MJ, Ruiz-Doblado S, Rodriguez-Pichardo A, Camacho F. Alopecia areata, stress and psychiatric disorders: a review. J Dermatol 1999 Oct;26(10):625-632. [doi: 10.1111/j.1346-8138.1999.tb02063.x] [Medline: 10554427]
8. Bilgiç Ö, Bilgiç A, Bahalı K, Bahali AG, Gürkan A, Yılmaz S. Psychiatric symptomatology and health-related quality of life in children and adolescents with alopecia areata. J Eur Acad Dermatol Venereol 2014 Nov;28(11):1463-1468. [doi: 10.1111/jdv.12315] [Medline: 24237476]
9. Christensen T, Yang JS, Castelo-Soccio L. Bullying and quality of life in pediatric alopecia areata. Skin Appendage Disord 2017 Aug;3(3):115-118. [doi: 10.1159/000466704] [Medline: 28879186]
10. Altunisik N, Ucuz I, Turkmen D. Psychiatric basics of alopecia areata in pediatric patients: evaluation of emotion dysregulation, somatization, depression, and anxiety levels. J Cosmet Dermatol 2022 Feb;21(2):770-775. [doi: 10.1111/jocd.14122] [Medline: 33797195]
11. Julian GS, Shau WY, Chou HW, Setia S. Bridging real-world data gaps: connecting dots across 10 Asian countries. JMIR Med Inform 2024 Aug 15;12:e58548. [doi: 10.2196/58548] [Medline: 39026427]
12. Isa AR, Noor M. Non-ionizing radiation exposure causing ill-health and alopecia areata. Med J Malaysia 1991 Sep;46(3):235-238. [Medline: 1839918]
13. Vaani VV, Tang MM, Tan LL, Asmah J. The utilization of phototherapy in the department of dermatology, Hospital Kuala Lumpur: a 5-year audit. Med J Malaysia 2018 Jun;73(3):125-130. [Medline: 29962494]
14. Albela H, Begum S, Wee AL, Ponnuthurai N, Leong KF. Efficacy and tolerability of methotrexate in the treatment of severe paediatric alopecia areata. Skin Appendage Disord 2022 May;8(3):206-210. [doi: 10.1159/000521238] [Medline: 35707294]
15. Tang MM, Gill P, Thevarajah S. Clinical pattern, management and outcome of alopecia areata of a multi ethnic population in Kuala Lumpur Hospital, Malaysia. Exp Dermatol 2010;19(6):600 [FREE Full text]
16. Lee YY, Lee CK. A 10-year retrospective review of non-scarring alopecia in a tertiary hospital, Malaysia. Exp Dermatol 2010;19(6 Suppl):600. [doi: 10.1111/j.1600-0625.2010.01097.x]
17. Azlan A, Khaliddin NB. A paediatric case of severe allergic conjunctivitis with underlying alopecia areata. Malays J Med Res 2023;07(1):19-21. [doi: 10.31674/mjmr.2023.v07i01.004]
18. Kaur I. Alopecia areata. Clin Dermatol J 2022;7(2). [doi: 10.23880/CDOAJ-16000274]
19. Heah SK, Mohd Noor N, Johar A. Prevalence of skin diseases in dermatology outpatient clinic, Hospital Kuala Lumpur. Malays J Dermatol 2017;38:19-24 [FREE Full text]
20. Ang TS, Tang JJ. Efficacy of topical immunotherapy with squaric acid dibutylester in the treatment of alopecia areata – a case series of 11 patients. Malays J Dermatol 2018;40:65-68 [FREE Full text]
21. Khoo VL, Loo CH, Khor YH, Mohd Ali N, Tan WC. Attitudes, health seeking behaviors, expectations and psychosocial impact of patients with non-scarring alopecia: results of tertiary outpatient skin specialist clinics in Penang. Malays J Dermatol 2023;50:29-39 [FREE Full text]
22. Yin YL, Chew KL. A 10-year retrospective review of non-scarring alopecia in a tertiary hospital in Malaysia. Malays J Dermatol 2012;29(2):12-15 [FREE Full text]
23. Luk NM, Chiu LS, Lee KC, et al. Efficacy and safety of diphenylcyclopropenone among Chinese patients with steroid resistant and extensive alopecia areata. J Eur Acad Dermatol Venereol 2013 Mar;27(3):e400-e405. [doi: 10.1111/jdv.12009] [Medline: 23057682]
24. Olsen EA, Roberts J, Sperling L, et al. Objective outcome measures: collecting meaningful data on alopecia areata. J Am Acad Dermatol 2018 Sep;79(3):470-478.e3. [doi: 10.1016/j.jaad.2017.10.048] [Medline: 29128463]
25. Olsen E, Hordinsky M, McDonald-Hull S, et al. Alopecia areata investigational assessment guidelines. National Alopecia Areata Foundation. J Am Acad Dermatol 1999 Feb;40(2 Pt 1):242-246. [doi: 10.1016/S0190-9622(99)70195-7] [Medline: 10025752]
26. Wambier CG, King BA. Rule of thumb: a simple tool to estimate 1% scalp surface area. J Am Acad Dermatol 2019 Aug;81(2):630-631. [doi: 10.1016/j.jaad.2019.01.022] [Medline: 30682394]
27. Olsen EA, Hordinsky MK, Price VH, et al. Alopecia areata investigational assessment guidelines–part II. National Alopecia Areata Foundation. J Am Acad Dermatol 2004 Sep;51(3):440-447. [doi: 10.1016/j.jaad.2003.09.032] [Medline: 15337988]
28. Apalla Z, Harfmann K. Clinical assessment of alopecia areata severity and validating the patient experience: a vodcast. Am J Clin Dermatol 2025 Nov;26(6):1055-1059. [doi: 10.1007/s40257-025-00973-5] [Medline: 40810760]

XSL•FO
RenderX

29. Tziotzios C, Sinclair R, Lesiak A, et al. Long-term safety and efficacy of ritlecitinib in adults and adolescents with alopecia areata and at least 25% scalp hair loss: results from the ALLEGRO-LT phase 3, open-label study. J Eur Acad Dermatol Venereol 2025 Jun;39(6):1152-1162. [doi: 10.1111/jdv.20526] [Medline: 39846397]
30. Wyrwich KW, Kitchen H, Knight S, et al. The Alopecia Areata Investigator Global Assessment scale: a measure for evaluating clinically meaningful success in clinical trials. Br J Dermatol 2020 Oct;183(4):702-709. [doi: 10.1111/bjd.18883] [Medline: 31970750]
31. Winnette R, Martin S, Harris N, Deal LS. Development of the alopecia areata patient priority outcomes instrument: a qualitative study. Dermatol Ther (Heidelb) 2021 Apr;11(2):599-613. [doi: 10.1007/s13555-021-00508-w] [Medline: 33751494]
32. King BA, Mesinkovska NA, Craiglow B, et al. Development of the alopecia areata scale for clinical use: results of an academic-industry collaborative effort. J Am Acad Dermatol 2022 Feb;86(2):359-364. [doi: 10.1016/j.jaad.2021.08.043] [Medline: 34474079]
33. Rudnicka L, Trzeciak M, Alpsoy E, et al. Disease burden, clinical management and unmet treatment need of patients with moderate to severe alopecia areata; consensus statements, insights, and practices from CERTAAE (Central/Eastern EU, Russia, Türkiye AA experts) Delphi panel. Front Med (Lausanne) 2024;11:1353354. [doi: 10.3389/fmed.2024.1353354] [Medline: 38741770]
34. Alhanshali L, Buontempo MG, Lo Sicco KI, Shapiro J. Alopecia areata: burden of disease, approach to treatment, and current unmet needs. Clin Cosmet Investig Dermatol 2023;16:803-820. [doi: 10.2147/CCID.S376096] [Medline: 37025396]
35. Ansari F, Singh S. Azathioprine in combination with methotrexate may not be a good therapeutic alternative in severe and recalcitrant forms of alopecia areata. J Eur Acad Dermatol Venereol 2020 Dec;34(12):e814-e815. [doi: 10.1111/jdv.16675] [Medline: 32458573]
36. Nohria A, Zhang YH, Desai D, et al. Real-world experience quantifying access to JAK inhibitor care for alopecia areata patients: a patient-centered survey study. Int J Womens Dermatol 2024;10(2):e145. [doi: 10.1097/JW9.0000000000000145] [Medline: 38606145]
37. Lee H, Huang KP, Mostaghimi A, Choudhry NK. Treatment patterns for alopecia areata in the US. JAMA Dermatol 2023 Nov 1;159(11):1253-1257. [doi: 10.1001/jamadermatol.2023.3109] [Medline: 37728940]
38. Anderson P, Piercy J, Austin J, et al. Alopecia areata treatment patterns and satisfaction: results of a real-world cross-sectional survey in Europe. Dermatol Ther (Heidelb) 2024 Dec;14(12):3243-3258. [doi: 10.1007/s13555-024-01280-3] [Medline: 39414726]
39. Manchanda Y, Ramamoorthy R. Revisiting pediatric alopecia areata: newer insights. Indian J Paediatr Dermatol 2021;22(4):301-305. [doi: 10.4103/ijpd.ijpd_109_21]
40. Chen LC, Senna MM, Aguh C. Emergence of Janus kinase inhibitors led to increase in proportion of severe alopecia areata patients receiving treatment: a retrospective cohort study. J Am Acad Dermatol 2024 Sep;91(3):579-581. [doi: 10.1016/j.jaad.2024.05.056] [Medline: 38810935]
41. Rudnicka L, Arenbergerova M, Grimalt R, et al. European expert consensus statement on the systemic treatment of alopecia areata. J Eur Acad Dermatol Venereol 2024 Apr;38(4):687-694. [doi: 10.1111/jdv.19768] [Medline: 38169088]
42. Ismail FF, Sinclair R. JAK inhibition in the treatment of alopecia areata - a promising new dawn? Expert Rev Clin Pharmacol 2020 Jan;13(1):43-51. [doi: 10.1080/17512433.2020.1702878] [Medline: 31865802]
43. Chim I, Ghiya R, Sinclair RD, Eisman S. Novel investigational drugs for alopecia areata and future perspectives. Expert Opin Investig Drugs 2024 May;33(5):441-449. [doi: 10.1080/13543784.2024.2348062] [Medline: 38682280]
44. Harries MJ, Ascott A, Asfour L, et al. British Association of Dermatologists living guideline for managing people with alopecia areata 2024. Br J Dermatol 2025 Jan 24;192(2):190-205. [doi: 10.1093/bjd/ljae385] [Medline: 39432739]
45. Patel HA, Revankar RR, Koblinski JE, et al. A clinician's guide to pediatric and adolescent alopecia areata treatments. JAAD Rev 2024 Dec;2:57-66. [doi: 10.1016/j.jdrv.2024.08.013]
46. Fatani MI, Alkhalifah A, Alruwaili AF, et al. Diagnosis and management of alopecia areata: a Saudi expert consensus statement (2023). Dermatol Ther (Heidelb) 2023 Oct;13(10):2129-2151. [doi: 10.1007/s13555-023-00991-3] [Medline: 37558830]
47. Toussi A, Barton VR, Le ST, Agbai ON, Kiuru M. Psychosocial and psychiatric comorbidities and health-related quality of life in alopecia areata: a systematic review. J Am Acad Dermatol 2021 Jul;85(1):162-175. [doi: 10.1016/j.jaad.2020.06.047] [Medline: 32561373]
48. Rodgers AR. Why finding a treatment for alopecia areata is important: a multifaceted perspective. J Investig Dermatol Symp Proc 2018 Jan;19(1):S51-S53. [doi: 10.1016/j.jisp.2017.10.008] [Medline: 29273107]
49. Tervonen T, Whichello C, Law E, et al. Treatment preferences of adults and adolescents with alopecia areata: a discrete choice experiment. J Dermatol 2024 Feb;51(2):243-252. [doi: 10.1111/1346-8138.17056] [Medline: 38087841]
50. Hoffman L, Fabi S. Look better, feel better, live better? The impact of minimally invasive aesthetic procedures on satisfaction with appearance and psychosocial wellbeing. J Clin Aesthet Dermatol 2022 May;15(5):47-58. [Medline: 35642226]
51. Nguyen B, Hu JK, Tosti A. Eyebrow and eyelash alopecia: a clinical review. Am J Clin Dermatol 2023 Jan;24(1):55-67. [doi: 10.1007/s40257-022-00729-5] [Medline: 36183302]

XSL•FO
RenderX

52. Akintilo L, Sally R, Shaw K, Ugonabo N, Shapiro J, Lo Sicco K. Hair care professionals as community health partners in detection and management of alopecia: a pilot survey study. J Drugs Dermatol 2021 Aug 1;20(8):908-910. [doi: 10.36849/JDD.5643] [Medline: 34397193]
53. Mandelin JM, Rehn L, Kosunen M, Väkeväinen M, Rannanpää S. The diverse experiences and challenges of individuals with alopecia areata: survey results from Finland. J Patient Exp 2025;12:23743735251346633. [doi: 10.1177/23743735251346633] [Medline: 40470311]
54. Trüeb RM, Schweizerische Arbeitsgruppe für Trichologie. Interrelations between physician, hairdresser and mass media in management of hair loss [Article in German]. Hautarzt 2000 Oct;51(10):729-732. [doi: 10.1007/s001050051205] [Medline: 11153357]
55. Mostaghimi A, Napatalung L, Sikirica V, et al. Patient perspectives of the social, emotional and functional impact of alopecia areata: a systematic literature review. Dermatol Ther (Heidelb) 2021 Jun;11(3):867-883. [doi: 10.1007/s13555-021-00512-0] [Medline: 33770385]

Abbreviations

AA: alopecia areata
AA-IGA: Alopecia Areata Investigator Global Assessment
AAPPO: Alopecia Areata Patient Priority Outcome
BSA: body surface area
EBA: Eyebrow Assessment Scale
ELA: Eyelash Assessment Scale
JAK: Janus kinase
PGI-C: Patient's Global Impression of Change
SADBE: squaric acid dibutyl ester
SALT: Severity of Alopecia Tool

Edited by R Dellavalle; submitted 15.Dec.2025; peer-reviewed by M Pritchard; revised version received 25.Aug.2026; accepted 27.Aug.2026; published 21.Sep.2026.

Please cite as:
Mohd Affandi A, Jamil A, Azizan NZ, Yap FBB, Leong KF, Heng AYH, Nuraihan S, Tan WC, Selvarajah L, Raja T, Yap EWY, Ch'ng PWB, Teoh TY, Tey MY
Current Management and Unmet Needs in Alopecia Areata: Viewpoint From a Malaysian Perspective
JMIR Dermatol 2026;9:e89583
URL: https://derma.jmir.org/2026/1/e89583
doi:10.2196/89583

© Azura Mohd Affandi, Adawiyah Jamil, Noor Zalmy Azizan, Felix Boon-Bin Yap, Kin Fon Leong, Agnes Yoke Hui Heng, Siti Nuraihan, Wooi Chiang Tan, Latha Selvarajah, Teeba Raja, Evelyn Wen Yee Yap, Peter Wee Beng Ch'ng, Tze Yuen Teoh, Mei Ying Tey. Originally published in JMIR Dermatology (http://derma.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI and Digital Tools in Dermatology: Addressing Access and Misinformation

Dominique du Crest[1*], MBE; Monisha Madhumita[2*], MD; Wendemagegn Enbiale[3*], MD, MPH, PhD; Jose Antonio Ruiz Postigo[4*], MD, PhD; Josep Malvehy[5*], MD; Shannon Wongvibulsin[6*], MD, PhD; Somesh Gupta[7*], MD; Harald Kittler[8*], MD; Charbel Skayem[9], MD; Anjali Mahto[10*], MD; Adewole S Adamson[11*], MD, MPP; Jules B Lipoff[12*], MD; Art Papier[13*], MD; Hugues Cartier[14*], MD; Sébastien Garson[15], MD; Esther Freeman[16*], MD, PhD

[1]SkinAid SAS, 50 Rue Danton, Levallois Perret, France

[10]Self London, London, United Kingdom

[11]University of Texas at Austin Dell Medical School, Austin, TX, United States

[12]Department of Dermatology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, United States

[13]Department of Dermatology, University of Rochester College of Medicine, Rochester, NY, United States

[14]Saint-Jean Clinic of Dermatology, Arras, France

[15]SELARL Dr GARSON, Senlis, France

[16]Massachusetts General Hospital, Department of Dermatology, Boston, MA, United States

[2]Department of Dermatology, Saveetha Medical College, Chennai, India

[3]College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia

[4]World Health Organization, Geneva, Switzerland

[5]Centro de Investigación Biomédica en Red de Enfermedades Raras Instituto de Salud Carlos III, Barcelona, Spain

[6]Division of Dermatology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

[7]All India Institute of Medical Sciences, New Delhi, New Delhi, India

[8]Department of Dermatology, Medical University of Vienna, Vienna, Austria

[9]Hôpital Ambroise-Paré, Boulogne Billancourt, France

[*]these authors contributed equally

Corresponding Author:
Dominique du Crest, MBE
SkinAid SAS, 50 Rue Danton, Levallois Perret, France

Abstract

Digital dermatology, which is defined as the use of digital technologies that leverage individual- and population-level skin data to improve the diagnosis, treatment, and prevention of skin diseases, has emerged as a critical frontier for bridging persistent gaps in dermatologic care. This transformation holds particular promise for addressing long-standing inequities linked to geography, income, and skin type. According to the Global Burden of Disease 2023 study, skin and subcutaneous diseases remain among the most prevalent global health conditions, contributing substantially to disability-adjusted life years. Digital tools (including teledermatology, artificial intelligence [AI], and large language models) offer new ways to extend diagnosis, education, and patient empowerment to historically underserved populations. However, these same innovations risk amplifying disparities if they are not designed and deployed intentionally. Algorithmic bias, uneven digital access, and the absence of culturally responsive models can undermine progress. In this conceptual and narrative review, we draw on expert dialogues and illustrative literature, including multistakeholder exchanges at the *Skin and Digital Summit (2023-2025*) and related global forums, to examine how digital dermatology can promote equitable skin health. We focus on 3 interlinked priorities: expanding access through scalable digital platforms, ensuring AI fairness via comprehensive and diverse datasets, and countering dermatological misinformation. Central to the latter is a bot concept described here as a dynamic cycle that analyzes scientific literature; ranks evidence; translates complex research into clear language; and delivers trustworthy, personalized guidance to both consumers and clinicians. By embedding expert oversight and evidence prioritization, such tools can ensure that accurate, actionable information reaches users at the speed and scale of the internet. Drawing on case studies (including lessons from the World Health Organization's AI skin health app) and insights from the Skin and Digital Summit, we highlight both the transformative potential and the ethical complexities of these digital solutions. To navigate this evolving landscape, we propose the concept of radical dermatology**,** which confronts the reality that big tech is reshaping skin health whether we like it or not and insists that dermatologists and stakeholders lead the transformation through bold collaboration and unwavering clinical relevance.

XSL•FO
RenderX

(JMIR Dermatol 2026;9:e79044) doi:10.2196/79044

KEYWORDS

digital dermatology; AI in health care; global health; teledermatology; artificial intelligence; AI; large language model; LLM; health tech; skin; dermatology; digital health; misinformation; ethical AI; radical dermatology; social media

Introduction

Globally, skin and subcutaneous diseases are among the most common health conditions. In 2019 alone, there were an estimated 4.86 billion new cases globally, with fungal (34%) and bacterial (23%) infections making up the majority of cases [1]. Skin conditions accounted for nearly 43 million disability-adjusted life years, a burden that is both physically and psychosocially disabling [1].

Despite this high burden, dermatologic care remains inaccessible for many. An estimated 70% of people living with chronic skin diseases, including psoriasis and rosacea, do not receive professional medical care, largely due to limited dermatology workforce capacity, cost barriers, geographic distance, and sociocultural factors in care delivery [2]. These barriers affect both low-resource settings and underserved communities in low-, middle-, and high-income countries [3].

At the same time, rapid smartphone penetration, the maturation of generative artificial intelligence (AI) tools, and the World Health Organization's (WHO) prioritization of skin health have created an unprecedented window for digital solutions. Digital dermatology refers to the application of digital technologies and skin-related data (including telemedicine, mobile health [mHealth], AI, and digital imaging) to improve the diagnosis, treatment, and prevention of skin diseases. While it offers opportunities to bridge persistent gaps in access to care, challenges such as bias in algorithms, limited infrastructure, and variable digital literacy must be addressed through deliberate equity-focused design and governance. Unlike traditional models that rely on in-person specialist visits concentrated in urban centers, digital dermatology can decentralize expertise, shorten wait times, enable earlier diagnosis, extend specialist reach, and support patient self-management across geographies [4]. However, equity is not guaranteed. If bias goes unaddressed, if infrastructure remains weak, or if digital literacy lags, digital dermatology may deepen the very divides it seeks to close. In the following sections, we highlight strategies to counter these risks and to align innovation with equity.

This conceptual and narrative review synthesizes expert insights, stakeholder dialogue, and illustrative case examples to explore equity-centered strategies in digital dermatology. Input was drawn from international experts through 4 panels and 3 roundtables at the Skin and Digital Summit held virtually (IMCAS, December 2024) and in Paris (IMCAS, January 2025) [5]. Data were captured through structured notes and partial transcripts and then analyzed using inductive thematic coding by 2 reviewers, with disagreements resolved by consensus and findings organized through framework synthesis [6,7].

As discussions occurred in public professional forums, institutional review board approval was not required. Any direct speaker quotations were deidentified or quoted with permission, and related files were stored securely with restricted access, in compliance with data protection best practices.

To complement these expert perspectives, we conducted a targeted literature review (PubMed, Scopus, and Web of Science; English-language studies), including studies on equity, implementation, and ethics in digital dermatology and related fields, thereby reducing selection bias.

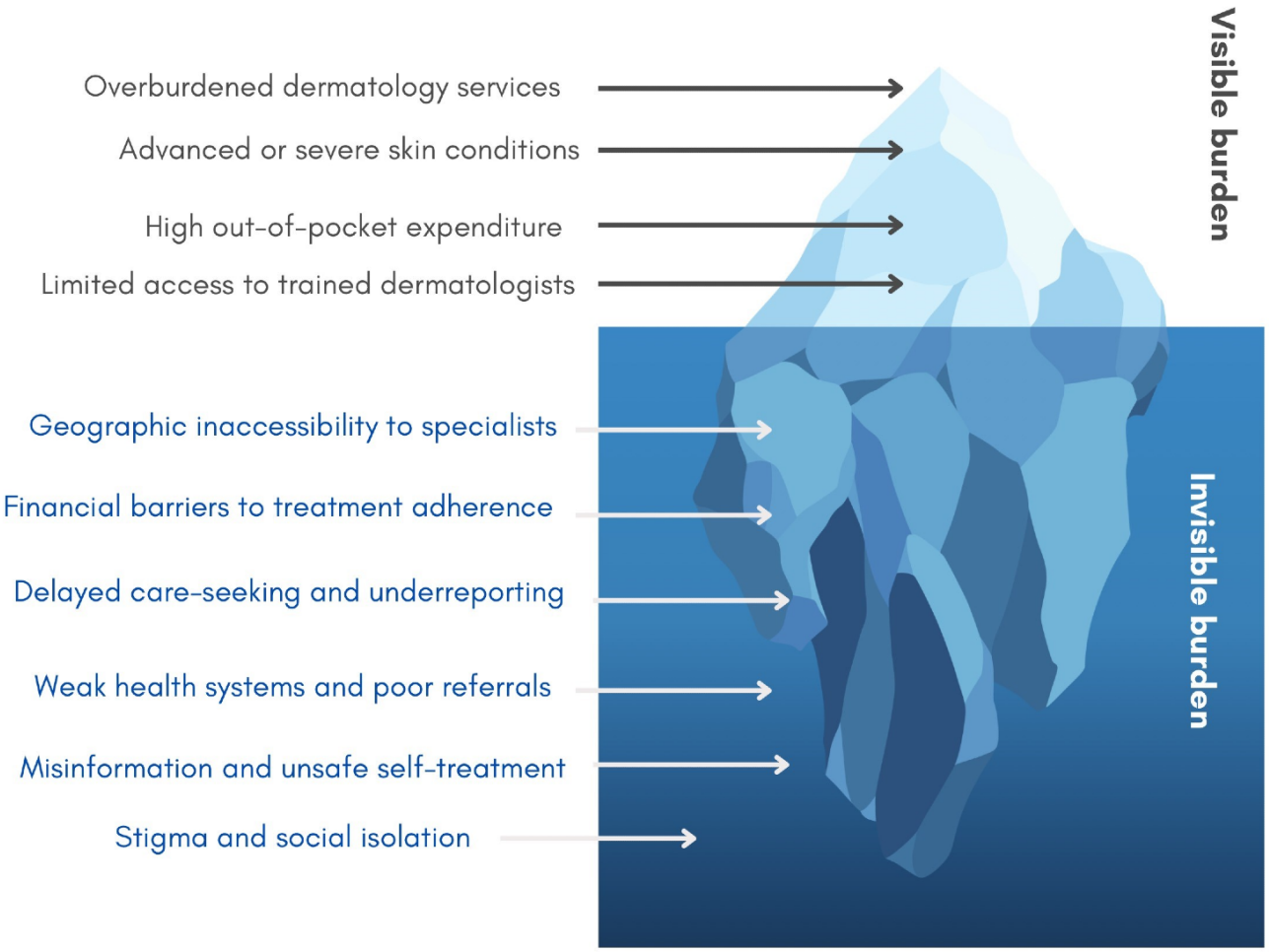
We discuss the WHO-neglected tropical disease mobile app as an example of ongoing implementation efforts and insights from global experts with references, where applicable. These examples were selected to highlight emerging real-world dynamics that are not yet captured in peer-reviewed literature. A narrative approach was used to thematically analyze insights across the digital dermatology continuum involving access to dermatologic care through digital innovation, datasets in dermatology, and combating skin-related misinformation.

Expanding Access to Dermatologic Care Through Digital Innovation

Despite the global burden of skin disease, dermatologic care remains underprioritized. Structural barriers, workforce shortages, and high out-of-pocket costs continue to limit access, especially in underserved and rural areas (Figure 1). These challenges are particularly acute for skin of color and populations in low- and middle-income countries [8,9].

XSL•FO
RenderX

Figure 1. The iceberg phenomenon of dermatologic care disparities.



Digital health innovations, including teledermatology, AI-assisted diagnostics, and mHealth apps, are increasingly recognized as promising tools to improve dermatology service delivery and care for all [10]. Given its reliance on visual diagnosis, dermatology makes it more amenable to remote care compared with many other specialties. High-quality clinical images and contextual data can enable remote triage, diagnosis, and follow-up, offering an opportunity to decentralize expertise [11,12].

One area is the proliferation of consumer-facing dermatology apps that offer mole monitoring, acne advice, and other self-assessment tools for timely triage and basic education, particularly in contexts where specialist access is limited. A 2024 review identified more than 900 such apps, including 41 with AI capabilities, and found their performance to be variable [13]. A 2021 analysis also showed sensitivity for melanoma to be highly variable across top apps, often with little dermatologist oversight or regulatory validation [14].

Alongside commercial tools, institutional efforts are advancing [15]. In 2024, the WHO piloted an AI-assisted mobile app across 5 Kenyan counties to help frontline workers identify 13 neglected tropical diseases and 24 common skin conditions. Forty Ministry of Health workers captured 605 patient images, which were reviewed by dermatologists. Although formal peer-reviewed accuracy data are pending, preliminary findings indicate more than 80% diagnostic agreement and strong usability for triage purposes rather than as a diagnostic replacement. Importantly, the app includes offline functionality and multilingual support and design features critical for resource-limited contexts [16,17]. However, several infrastructural and regulatory barriers persist as challenges, including unstable networks, device costs, limited digital literacy, and fragmented data governance frameworks that constrain uptake [18]. In some regions, legal restrictions on data sharing further complicate deployment. Addressing these challenges will require coordinated public-private partnerships and context-specific policy support.

Additionally, we must acknowledge that although digital dermatology tools can improve diagnosis and education, their impact is limited if patients cannot access the medications they are prescribed. In underserved areas, medication shortages, high costs, and weak supply chains often prevent timely treatment. Closing gaps in dermatologic care requires strategic, multisector partnerships to ensure that digital solutions are not just short-term pilot projects, but long-lasting, locally led interventions supported by active partnerships with big tech and big pharma [19]. To strengthen implementation, policy levers such as teledermatology reimbursement and data protection standards must be embedded within digital health initiatives. Such initiatives must work in tandem with public health systems, with continued monitoring of program success (such as shorter time to treatment after digital triage), through demand forecasting, community health worker training, and pharmacy

XSL•FO
RenderX

partnerships for sustained equitable access to both diagnosis and treatment [20].

Datasets: The Key to Unlocking AI Potential in Dermatology

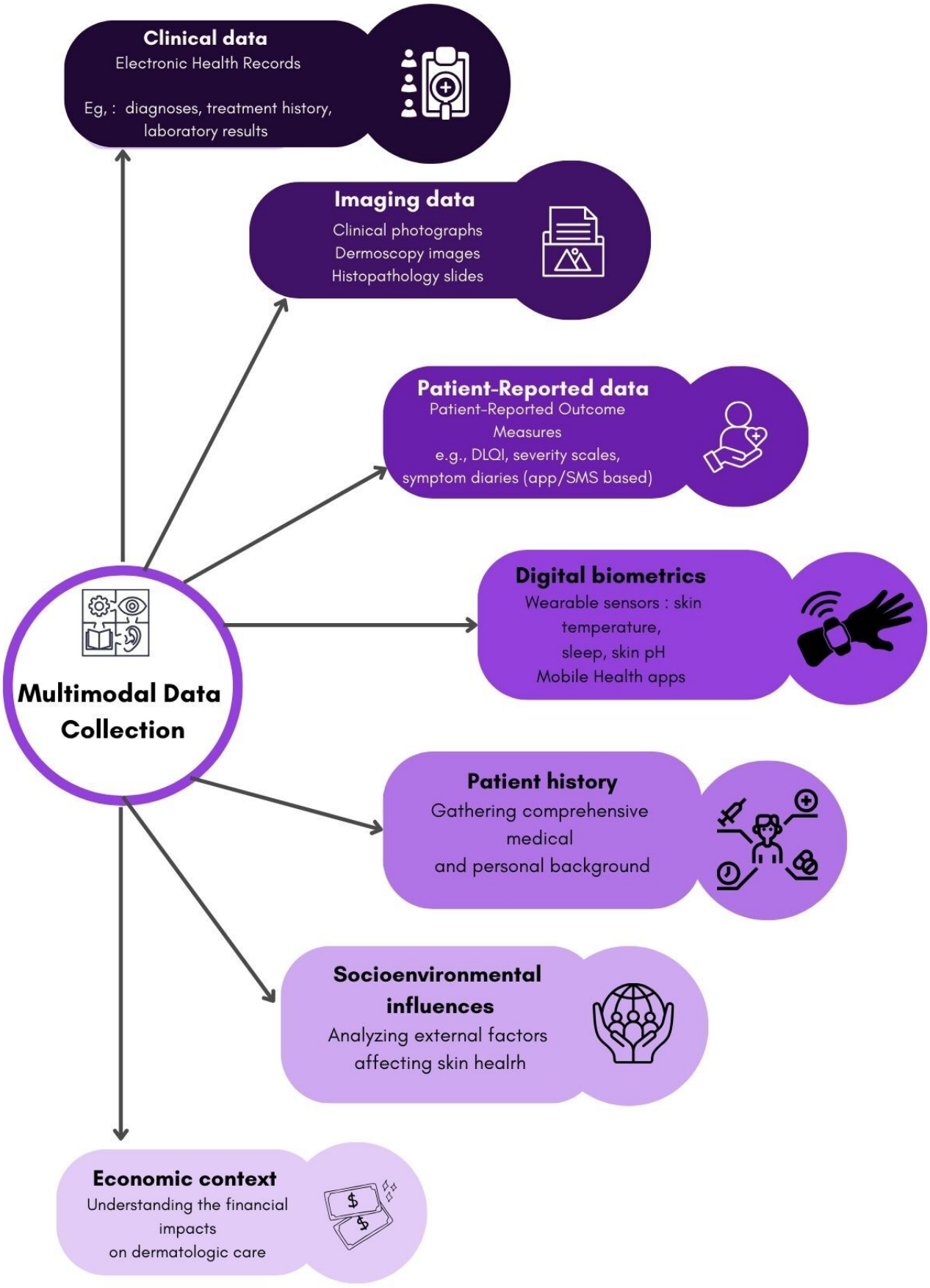
Complementing access-focused interventions, the success of AI in digital dermatology hinges on one critical factor: data. Dermatologic AI systems predominantly rely on visual data, including clinical photographs and dermatoscopy images, for training and validation [21]. However, the lack of large, diverse, and representative datasets poses a major limitation [22]. Most existing datasets overrepresent light-skinned patients from high-income settings and focus heavily on conditions such as melanoma [23]. Even newer collections, such as the Diverse Dermatology Images dataset, while addressing skin tone gaps, are limited in diagnostic breadth and geographic diversity [24,25]. To strengthen fairness and reproducibility, skin tone annotation should evolve beyond the traditional Fitzpatrick phototypes toward newer scales such as the Monk Skin Tone scale, using standardized labeling and consistent documentation protocols across datasets [26].

Without datasets that reflect all segments of the population and systems that ensure access to recommended care, AI solutions may offer precise diagnoses that are ultimately inaccessible to those who need them most. Relying on single-modality datasets, such as static images alone, restricts the ability of AI systems to capture the real-world complexity of dermatological care [27]. Dermatologic decision-making is not merely visual; it involves psychometric, environmental, and cultural factors. For instance, 2 patients with the same lesion may receive different treatment recommendations depending on their medical history, stress levels, socioeconomic context, or cultural expectations.

To address these issues, a multimodal data approach is essential (Figure 2) [28]. This includes integrating clinical photographs with patient histories, self-reported outcomes, environmental exposures, and behavioral data. For example, in conditions such as psoriasis or vitiligo, combining lesion photographs with validated severity indices (eg, Dermatology Life Quality Index and Pruritus Numerical Rating Scale), environmental parameters (UV index and humidity), comorbidities, and adherence telemetry can yield more personalized, context-aware interventions. These inputs can be captured via mHealth apps and wearables [29,30]. AI models trained on such holistic datasets, encompassing lesion images, lifestyle habits, stress levels, and treatment adherence, can deliver more accurate, individualized care. Importantly, these insights must be coupled with robust privacy safeguards, including on-device data processing and granular patient consent. Furthermore, to translate these innovations into improved outcomes in underserved areas, they must be linked to reliable access to medications through community health workers, local pharmacies, or public-sector distribution programs.

XSL•FO
RenderX

Figure 2. Expanding dermatology with multimodal data.



Enabling this shift requires interoperable and transparent data infrastructure with an emphasis on terminology mapping and handling of metadata. Platforms such as OpenMHealth, Apple HealthKit, and Fast Healthcare Interoperability Resources–compliant electronic health records can help structure and standardize multimodal data [31]. Future

XSL•FO
RenderX

dermatology-specific tools must enable seamless integration of patient-reported outcomes and subjective (eg, pruritus scores) and objective (eg, lesion photos) metrics while ensuring data privacy. Ethical development and deployment of such AI tools must follow established frameworks such as FATE (fairness, accountability, transparency, and ethics) [32] and the OECD AI Principles [33]. These frameworks call for human oversight, robust transparency about model limitations, and equitable access to benefits. Poorly curated or biased data can not only reduce diagnostic accuracy but also worsen outcomes for already marginalized groups.

While appealing, the belief in autonomous, data-driven AI-based decision-making has significant limitations; data alone are not the answer. Although increasing data volume often enhances predictive accuracy, the quality and representativeness of the datasets remain crucial [34]. Poorly curated data can reinforce existing biases and, in some cases, worsen clinical outcomes [35]. Human clinicians inherently recognize that accurate interpretation of data requires understanding its context, particularly in dermatology, where the psychosocial impact of visible skin conditions demands sensitivity beyond numerical analysis alone [36]. Effective dermatological decision-making thus integrates objective data with patient-specific considerations, including emotional, social, and cultural factors. Identical skin lesions may necessitate different diagnostic or therapeutic approaches depending on individual patient circumstances and preferences, underscoring the necessity of collaboration between data-driven insights, clinical expertise, and patient perspectives.

Combating Skin-Related Misinformation

Maximizing the impact of digital dermatology also requires confronting the parallel crisis of misinformation [37]. Fueled by algorithmic amplification and low digital health literacy, skin-related myths such as “sunscreen causes cancer” or “natural remedies are always safer” spread faster than verified medical guidance [38]. Platforms that reward engagement often privilege sensational content over accuracy, and dermatology is no exception. This dynamic is compounded by the low visibility of authoritative voices: only 4% of dermatology influencers on Instagram are board-certified dermatologists, and on TikTok, more than one-third of dermatology-related videos are created by nonprofessionals [39], frequently promoting unverified treatments such as raw potatoes for acne. On Reddit [40] and other forums, patients increasingly crowdsource diagnoses—sometimes for sensitive conditions such as sexually transmitted infections—reflecting growing distrust in formal health care and a shift toward peer-based digital health advice [41]. Studies of parenting blogs have shown that posts containing sunscreen misinformation consistently receive more engagement than scientifically accurate content [42].

Efforts to counter this phenomenon must go beyond reactive myth-busting and instead adopt structured, measurable, and trust-centered approaches and strategies grounded in successful digital literacy campaigns [43]. Public health initiatives in adjacent domains offer useful models. For example, the WHO’s “Pause Before You Share” campaign effectively encouraged users to reflect before forwarding unverified information during the COVID-19 pandemic [44]. The United Nations Children’s Fund and Gavi’s #VaccinesWork initiative combined localized messaging with influencer partnerships to rebuild vaccine confidence [45]. In sexual health, the AI-powered chatbot “Roo,” developed by Planned Parenthood, has demonstrated how conversational interfaces can deliver reliable, age-appropriate, and stigma-free health information at scale [46]. These campaigns share a common thread: they prioritize trust, accessibility, and proactive engagement, principles replicable in dermatology.

Drawing inspiration from these models, we propose an AI-powered dermatology chatbot (still in its conceptual phase) designed to address misinformation while empowering users with accurate, evidence-based content. The chatbot would operate on a retrieval-augmented generation architecture, fine-tuned on dermatology-specific content. It would draw from validated sources such as Cochrane systematic reviews, the American Academy of Dermatology (AAD), and the European Academy of Dermatology and Venereology (EADV) guidelines, patient association websites, and public health repositories from the WHO and CDC (Centers for Disease Control and Prevention). To ensure scientific rigor, the chatbot would internally rank information using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) framework. It would also include medical disclaimers, clinician escalation paths, hallucination checks, and refusal policies for diagnostic or treatment queries. Bias testing**,** data minimization**,** and content updates aligned with AAD and EADV guidelines would maintain fairness, privacy, and accuracy. Its performance should be tracked through reach, engagement, accuracy, harmful-advice rate, and user trust metrics. Dermatologists should lead content creation and partnerships to amplify credible, evidence-based information [47].

It is crucial to reiterate that dermatologists need to proactively develop accurate digital content to combat the spread of misinformation [48]. This can involve structured social media initiatives such as verified content series, myth-busting reels, or question-and-answer sessions on platforms where misinformation is most widespread. Collaborating with trusted influencers, patient advocacy groups, or public health campaigns can further expand reach and ensure that evidence-based information is both accessible and engaging [49].

Beyond Content Curation: Developing a Technological Solution to Address Skin Misinformation

The rapid, unchecked spread of skin-related misinformation is emerging as a significant public health concern—one that no dermatologist, no matter how dedicated, can tackle alone [48-51]. In this digital era, where falsehoods travel faster than facts, expert voices must be amplified by bold technological innovation. What is urgently needed is not just another fact sheet, but an intelligent, real-time solution: an AI-powered dermatology conversational agent (Figure 3) capable of delivering science-backed answers at the speed of the internet

XSL•FO
RenderX

[52]. This chatbot could serve as a critical tool, helping to dispel myths, provide accurate information, and guide individuals toward appropriate care, with empathy [53].

Figure 3. Cycle of artificial intelligence (AI)–powered dermatological chatbot. LLM: large language model; ML: machine learning; NLP: natural language processing.

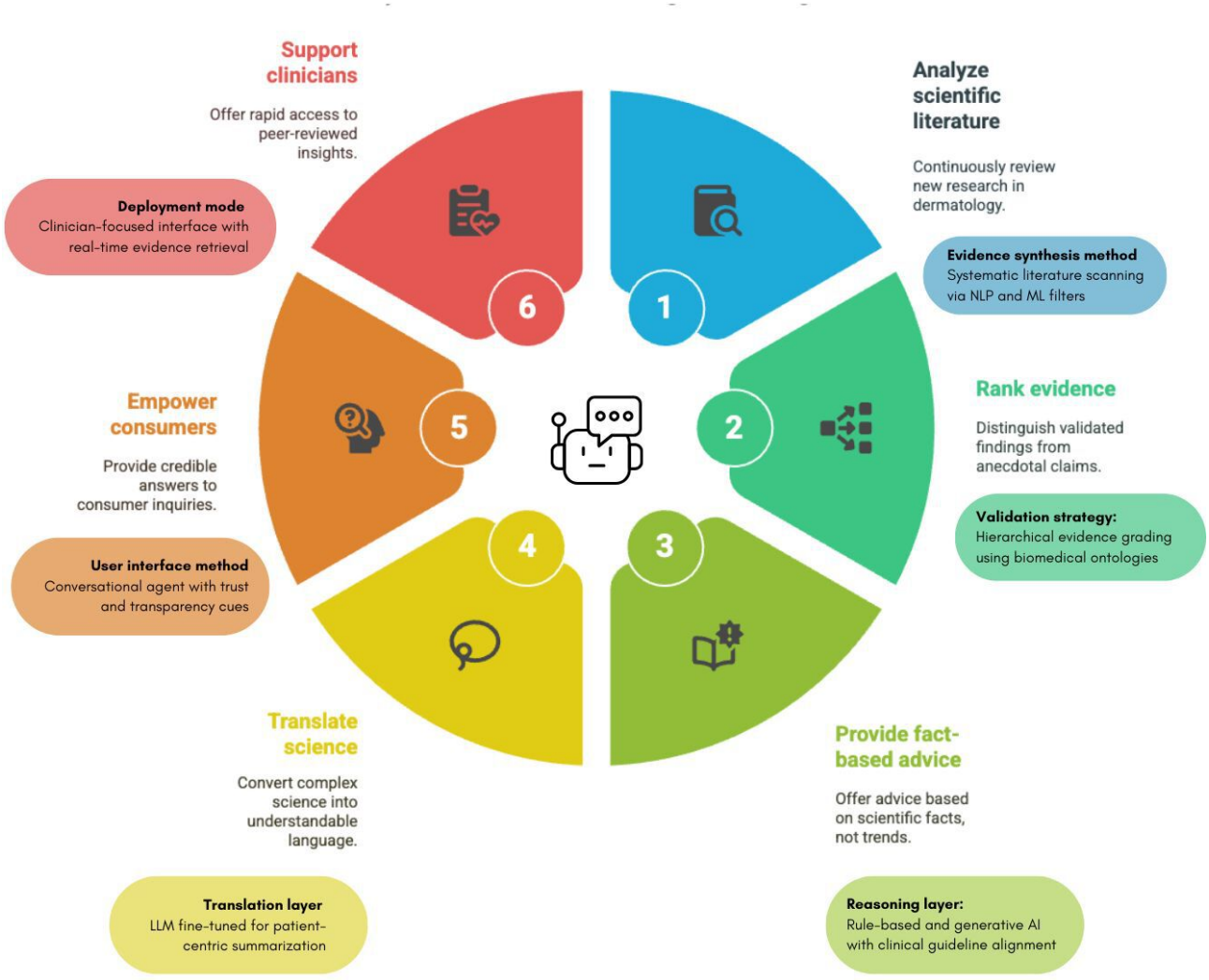


Figure 3 outlines the engine behind such a tool—a dynamic cycle that analyzes scientific literature, ranks evidence, translates complex research into clear language, and delivers trustworthy, personalized guidance to both consumers and clinicians. This is not just an upgrade; it is a paradigm shift, and it demands a multidisciplinary alliance of dermatologists, AI developers, skin scientists, tech experts, and health communicators. Only together can we build the digital frontline dermatology desperately needs.

The concept of "radical dermatology" underscores this transformation. Radical dermatology refers to a future-oriented framework that calls on dermatologists and skin health stakeholders to actively lead, rather than passively follow, the digital transformation of their field. It recognizes the growing influence of big tech in reshaping skin health and emphasizes that dermatologists must drive this change through bold collaboration, clinical relevance, and one that is ethically governed through measures such as equity dashboards and data diversity thresholds.

While conceptual in origin, radical dermatology lays the groundwork for concrete steps such as pilot studies, comparative evaluations of AI- and clinician-led care, and assessments of cost-effectiveness and equity outcomes. However, limitations of this paper include its presentation of a conceptual framework and a hypothetical chatbot model, with insights derived exclusively from expert dialogue rather than empirical clinical trials. To advance this research, future work should aim at pilot studies across diverse populations, comparative analyses of AI-supported versus clinician-led care, and comprehensive cost-effectiveness research. Dermatologists must play a central role in leading this transformative shift, guiding not only the application of digital tools but also their fundamental design, rigorous testing, and governance. This crucial role requires dermatologists to do more than just use digital tools; they must actively shape the technological, ethical, and societal landscape of digital dermatology.

Conclusions

Leveraging individual- and population-level skin data ethically through the convergence of computer vision, mHealth, and generative AI presents a pivotal opportunity for dermatology.

XSL•FO
RenderX

Conversational tools such as dialogic AI or chatbots, when well designed, can interpret validated evidence, provide real-time responses, and deliver guidance in ways that are accessible, accurate, and empathetic [52,53]. This is particularly relevant in dermatology, where care is visual, is time sensitive, and depends heavily on patient education and reassurance.

The promise of digital dermatology will only be achieved if these tools are developed for everyone, with transparency, accuracy, and trust as core principles and with an emphasis on ongoing safety auditing. People without access to dermatology care or reliable information stand to gain the most, provided their needs are addressed from the outset. Deployment should therefore prioritize underserved users through multilingual interfaces, offline functionality, and low-bandwidth design as integral components of equity-centered implementation. Dermatologists must remain central to this process, guiding design, overseeing implementation, and setting ethical standards to ensure that digital solutions are clinically sound and culturally appropriate.

The future of dermatology will be shaped not by algorithms alone but through collaboration among clinicians, technologist researchers, policymakers, and patients to ensure technology meaningfully advances skin health for all.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Szeto MD, Alhanshali L, Rundle CW, et al. Dermatologic data from the global burden of disease study 2019 and the patientslikeme online support community: comparative analysis. JMIR Dermatol 2024 Dec 11;7:e50449. [doi: 10.2196/50449] [Medline: 39661989]
2. Wehausen B, Hill DE, Feldman SR. Most people with psoriasis or rosacea are not being treated: a large population study. Dermatol Online J 2016 Jul 15;22(7):13030/qt4nc3p4q2. [Medline: 27617716]
3. Yakupu A, Aimaier R, Yuan B, et al. The burden of skin and subcutaneous diseases: findings from the global burden of disease study 2019. Front Public Health 2023;11:1145513. [doi: 10.3389/fpubh.2023.1145513] [Medline: 37139398]
4. Skin diseases as a global public health priority. World Health Organization. 2025 May 27. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA78/A78_R15-en.pdf [accessed 2025-05-30]
5. Skin and Digital Summit. URL: https://www.skindigitalsummit.com/ [accessed 2026-02-13]
6. International Masters Course on Aging Science (IMCAS). URL: https://www.imcas.com/en [accessed 2925-02-20]
7. du Crest D, Madhumita M, Enbiale W, et al. Skin and digital-the 2024 narrative. Mayo Clin Proc Digit Health 2024 Sep;2(3):322-330. [doi: 10.1016/j.mcpdig.2024.05.008] [Medline: 40206126]
8. Adepoju O, Dang P, Nguyen H, Mertz J. Equity in digital health: assessing access and utilization of remote patient monitoring, medical apps, and wearables in underserved communities. Inquiry 2024;61:469580241271137. [doi: 10.1177/00469580241271137] [Medline: 39323052]
9. Gwillim EC, Azzawi S, Aigen AR. Underserved populations and health equity in dermatology: digital medicine and the role of artificial intelligence. Clin Dermatol 2024;42(5):498-506. [doi: 10.1016/j.clindermatol.2024.06.021] [Medline: 38944248]
10. Schaekermann M, Spitz T, Pyles M, et al. Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. EClinicalMedicine 2024 Apr;70:102479. [doi: 10.1016/j.eclinm.2024.102479] [Medline: 38685924]
11. Ahuja S, Briggs SM, Collier SM. Teledermatology in rural, underserved, and isolated environments: a review. Curr Dermatol Rep 2022;11(4):328-335. [doi: 10.1007/s13671-022-00377-2] [Medline: 36310767]
12. Chang AY, Kiprono SK, Maurer TA. Providing dermatological care in resource-limited settings: barriers and potential solutions. Br J Dermatol 2017 Jul;177(1):247-248. [doi: 10.1111/bjd.15372] [Medline: 28731250]
13. Wongvibulsin S, Yan MJ, Pahalyants V, Murphy W, Daneshjou R, Rotemberg V. Current state of dermatology mobile applications with artificial intelligence features. JAMA Dermatol 2024 Jun 1;160(6):646-650. [doi: 10.1001/jamadermatol.2024.1342] [Medline: 38452263]
14. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. Br J Dermatol 2022 Apr;186(4):744-746. [doi: 10.1111/bjd.20903] [Medline: 34811727]
15. Alami H, Rivard L, Lehoux P, et al. Artificial intelligence in health care: laying the foundation for responsible, sustainable, and inclusive innovation in low- and middle-income countries. Global Health 2020 Jun 24;16(1):52. [doi: 10.1186/s12992-020-00584-1] [Medline: 32580741]

XSL•FO
RenderX

16. The WHO Skin NTDs app shows encouraging results in Kenya study. World Health Organization. 2024. URL: https://www.who.int/news/item/04-12-2024-the-who-skin-ntds-app-shows-encouraging-results-in-kenya-study [accessed 2025-05-25]
17. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med 2022 Aug 18;5(1):119. [doi: 10.1038/s41746-022-00663-0] [Medline: 35982146]
18. Sylla B, Ismaila O, Diallo G. 25 years of digital health toward universal health coverage in low- and middle-income countries: rapid systematic review. J Med Internet Res 2025 May 29;27:e59042. [doi: 10.2196/59042] [Medline: 40440696]
19. Duniphin DD. Limited access to dermatology specialty care: barriers and teledermatology. Dermatol Pract Concept 2023 Jan 1;13(1):e2023031. [doi: 10.5826/dpc.1301a31] [Medline: 36892370]
20. Orton M, Agarwal S, Muhoza P, Vasudevan L, Vu A. Strengthening delivery of health services using digital devices. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S61-S71. [doi: 10.9745/GHSP-D-18-00229] [Medline: 30305340]
21. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. Bioengineering (Basel) 2023 Dec 18;10(12):1435. [doi: 10.3390/bioengineering10121435] [Medline: 38136026]
22. Daneshjou R, Vodrahalli K, Novoa RA, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv 2022 Aug 12;8(32):eabq6147. [doi: 10.1126/sciadv.abq6147] [Medline: 35960806]
23. Narla A, Kuprel B, Sarin K, Novoa R, Ko J. Automated classification of skin lesions: from pixels to practice. J Invest Dermatol 2018 Oct;138(10):2108-2110. [doi: 10.1016/j.jid.2018.06.175] [Medline: 30244720]
24. Groh M, Harris C, Daneshjou R, Badri O, Koochek A. Towards transparency in dermatology image datasets with skin tone annotations by experts, crowds, and an algorithm. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-26. [doi: 10.1145/3555634]
25. Jain T. Evaluating machine learning-based skin cancer diagnosis. arXiv. Preprint posted online on Sep 4, 2024. [doi: 10.48550/arxiv.2409.03794]
26. Arora A, Alderman JE, Palmer J, et al. The value of standards for health datasets in artificial intelligence-based applications. Nat Med 2023 Nov;29(11):2929-2938. [doi: 10.1038/s41591-023-02608-w] [Medline: 37884627]
27. Mishra S, Chaudhury S, Imaizumi H, Yamasaki T. Assessing robustness of deep learning methods in dermatological workflow. arXiv. Preprint posted online on Jan 15, 2020. [doi: 10.48550/arXiv.2001.05878]
28. Krones F, Marikkar U, Parsons G, Szmul A, Mahdi A. Review of multimodal machine learning approaches in healthcare. Information Fusion 2025 Feb;114:102690. [doi: 10.1016/j.inffus.2024.102690]
29. Hong J, Mosca M, Hadeler E, Hakimi M, Bhutani T, Liao W. The future of personalized medicine in psoriasis. Dermatological Reviews 2021 Oct;2(5):282-288. [doi: 10.1002/der2.87]
30. Ahuja R, Narayanan B, Gupta S. Precision medicine and personalized approach in vitiligo. Dermatological Reviews 2022 Oct;3(5):308-312. [doi: 10.1002/der2.159]
31. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14. [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
32. Singhal A, Neveditsin N, Tanveer H, Mago V. Toward fairness, accountability, transparency, and ethics in ai for social media and health care: scoping review. JMIR Med Inform 2024 Apr 3;12:e50048. [doi: 10.2196/50048] [Medline: 38568737]
33. Yeung K. Recommendation of the Council on Artificial Intelligence (OECD). Int leg mater 2020 Feb;59(1):27-34. [doi: 10.1017/ilm.2020.5]
34. Okolo CT. Optimizing human-centered AI for healthcare in the Global South. Patterns (N Y) 2022 Feb 11;3(2):100421. [doi: 10.1016/j.patter.2021.100421] [Medline: 35199066]
35. Green BL, Murphy A, Robinson E. Accelerating health disparities research with artificial intelligence. Front Digit Health 2024;6:1330160. [doi: 10.3389/fdgth.2024.1330160] [Medline: 38322109]
36. Breen N, Berrigan D, Jackson JS, et al. Translational health disparities research in a data-rich world. Health Equity 2019;3(1):588-600. [doi: 10.1089/heq.2019.0042] [Medline: 31720554]
37. Nelson EE, Black TA, Rousseau MA, Rashid RM. Combating misinformation in dermatology. Dermatol Pract Concept 2023 Oct 1;13(4):e2023234. [doi: 10.5826/dpc.1304a234] [Medline: 37992336]
38. Ecker U, Roozenbeek J, van der Linden S, et al. Misinformation poses a bigger threat to democracy than you might think. Nature New Biol 2024 Jun 6;630(8015):29-32. [doi: 10.1038/d41586-024-01587-3] [Medline: 38840026]
39. Ranpariya V, Chu B, Fathy R, Lipoff JB. Dermatology without dermatologists? Analyzing Instagram influencers with dermatology-related hashtags. J Am Acad Dermatol 2020 Dec;83(6):1840-1842. [doi: 10.1016/j.jaad.2020.05.039] [Medline: 32416205]
40. Chu B, Fathy R, Nobles AL, Lipoff JB. Patient crowdsourcing of dermatologic consults on a Reddit social media community. J Am Acad Dermatol 2021 Jul;85(1):226-227. [doi: 10.1016/j.jaad.2020.08.009] [Medline: 32777316]
41. Nobles AL, Leas EC, Althouse BM, et al. Requests for diagnoses of sexually transmitted diseases on a social media platform. JAMA 2019 Nov 5;322(17):1712-1713. [doi: 10.1001/jama.2019.14390] [Medline: 31688875]
42. O'Connor C, Rafferty S, Murphy M. A qualitative review of misinformation and conspiracy theories in skin cancer. Clin Exp Dermatol 2022 Oct;47(10):1848-1852. [doi: 10.1111/ced.15249] [Medline: 35514125]
43. Sood A, Sangari A, Stoff BK. Navigating internet-based misinformation with patients in the clinic. J Am Acad Dermatol 2023 Apr;88(4):961-962. [doi: 10.1016/j.jaad.2022.03.038] [Medline: 35346757]

XSL•FO
RenderX

44. Rodrigo P, Arakpogun EO, Vu MC, Olan F, Djafarova E. Can you be mindful? The effectiveness of mindfulness-driven interventions in enhancing the digital resilience to fake news on COVID-19. Inf Syst Front 2022 Mar 2:1-21. [doi: 10.1007/s10796-022-10258-5] [Medline: 35250364]
45. Germani F, Biller-Andorno N. How to counter the anti-vaccine rhetoric: filling information voids and building resilience. Hum Vaccin Immunother 2022 Nov 30;18(6):2095825. [doi: 10.1080/21645515.2022.2095825] [Medline: 35802046]
46. Mills R, Mangone ER, Lesh N, Jayal G, Mohan D, Baraitser P. Chatbots that deliver contraceptive support: systematic review. J Med Internet Res 2024 Feb 27;26:e46758. [doi: 10.2196/46758] [Medline: 38412028]
47. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490. [doi: 10.1136/bmj.328.7454.1490] [Medline: 15205295]
48. Szeto MD, Mamo A, Afrin A, Militello M, Barber C. Social media in dermatology and an overview of popular social media platforms. Curr Dermatol Rep 2021;10(4):97-104. [doi: 10.1007/s13671-021-00343-4] [Medline: 34692234]
49. El Mikati IK, Hoteit R, Harb T, et al. Defining misinformation and related terms in health-related literature: scoping review. J Med Internet Res 2023 Aug 9;25:e45731. [doi: 10.2196/45731] [Medline: 37556184]
50. Mehta-Ambalal SR, Nisarta M. Dermatology 2.0- how the internet is changing us, our patients and our practice. Indian Dermatol Online J 2021;12(4):593-596. [doi: 10.4103/idoj.IDOJ_788_20] [Medline: 34430468]
51. O'Connor C, O'Grady C, Murphy M. Spotting fake news: a qualitative review of misinformation and conspiracy theories in acne vulgaris. Clin Exp Dermatol 2022 Sep;47(9):1707-1711. [doi: 10.1111/ced.15222] [Medline: 35434841]
52. Mariani MM, Hashemi N, Wirtz J. Artificial intelligence empowered conversational agents: a systematic literature review and research agenda. J Bus Res 2023 Jun;161:113838. [doi: 10.1016/j.jbusres.2023.113838]
53. Chen D, Chauhan K, Parsa R, et al. Author correction: patient perceptions of empathy in physician and artificial intelligence chatbot responses to patient questions about cancer. NPJ Digit Med 2025 Jun 21;8(1):382. [doi: 10.1038/s41746-025-01793-x] [Medline: 40544179]

Abbreviations

AAD: American Academy of Dermatology
AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
EADV: European Academy of Dermatology and Venereology
FATE: fairness, accountability, transparency, and ethics
GRADE: Grading of Recommendations, Assessment, Development and Evaluations
mHealth: mobile health
WHO : World Health Organization

Edited by J Solomon, R Dellavalle; submitted 26.Jun.2025; peer-reviewed by H Peter Soyer, S Khosravi; revised version received 15.Dec.2025; accepted 16.Jan.2026; published 20.Feb.2026.

Please cite as:
du Crest D, Madhumita M, Enbiale W, Ruiz Postigo JA, Malvehy J, Wongvibulsin S, Gupta S, Kittler H, Skayem C, Mahto A, Adamson AS, Lipoff JB, Papier A, Cartier H, Garson S, Freeman E
AI and Digital Tools in Dermatology: Addressing Access and Misinformation
JMIR Dermatol 2026;9:e79044
URL: https://derma.jmir.org/2026/1/e79044
doi:10.2196/79044

© Dominique du Crest, Monisha Madhumita, Wendemagegn Enbiale, Jose Antonio Ruiz Postigo, Josep Malvehy, Shannon Wongvibulsin, Somesh Gupta, Harald Kittler, Charbel Skayem, Anjali Mahto, Adewole Adamson, Jules B Lipoff, Art Papier, Hugues Cartier, Sébastien Garson, Esther Freeman. Originally published in JMIR Dermatology (http://derma.jmir.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Online Discussions About Alopecia Areata in 5 Arabian Gulf Countries: Social Media Listening Study

Ashraf Reda[1*], Dr med, MBBSMD, MBBS; Ghadear AL Alawi[2*], Dr med, MBBSMD, MBBS; Ameen Al Awadhi[3*], Dr med, MBBSMD, MBBS; Amina Al Obaidli[4*], Dr med, MBBSMD, MBBS; Fadwa Ebeid[5*], BPharm; Haytham Mohamed Ahmed[5*], BPharm

[1]Department of Dermatology, Mediclinic Welcare Hospital, Dubai, United Arab Emirates

[2]Department of Dermatology, Adan hospital, Hadiya, Kuwait

[3]Department of Dermatology, Salmaniya Medical Complex, Manama, Bahrain

[4]Department of Dermatology, Rumaila hospital, Doha, Qatar

[5]Department of Medical Affairs, Pfizer Inc Ltd, Building 6, Dubai Media City, Dubai, United Arab Emirates

[*]all authors contributed equally

Corresponding Author:
Haytham Mohamed Ahmed, BPharm
Department of Medical Affairs, Pfizer Inc Ltd, Building 6, Dubai Media City, Dubai, United Arab Emirates

Abstract

Background: Alopecia areata (AA) is an autoimmune disease that can affect the quality of life of affected individuals. Publicly available online conversations and search behavior provide valuable insights into public awareness, understanding, and information related to AA. An infodemiology approach, therefore, offers an opportunity to characterize these perspectives across Gulf countries. This study evaluated the understanding and awareness of AA in Bahrain, Kuwait, Oman, Qatar, and the United Arab Emirates based on AA-related mentions and internet searches.

Objective: Despite its prevalence, AA is often misunderstood, particularly in the Middle East. By examining the data from social media platforms, online forums, and internet searches, this study provided insight into the regional AA perceptions and awareness in these 5 Gulf countries.

Methods: Online discussions in English and Arabic mentioning AA and related terms were analyzed from the 5 countries between January 2022 and April 2023.

Results: The analysis of more than 8500 mentions of AA over 16 months revealed significant online engagement, especially from the United Arab Emirates. Educational content dominated the discussions, with a notable increase following the Oscars incident in April 2022. Gulf news platforms and X (formerly Twitter) were the primary discussion platforms, with health care professionals mainly using English and patients using both English and Arabic. Common topics included potential contributors to AA, emotional distress, and the AA patient journey.

Conclusions: The study identified gaps in awareness and understanding of AA among patients, health care professionals, and medical facilities in the Gulf region through online discussions and search patterns. These insights can facilitate the development of targeted educational initiatives, support strategies, improve awareness, reduce stigma, and enhance support networks for those affected by AA.

(JMIR Dermatol 2026;9:e78262) doi:10.2196/78262

KEYWORDS

alopecia areata; social media; patient concerns; search patterns; online discussions; search trends

Introduction

Alopecia areata (AA) is an autoimmune disease that results in nonscarring hair loss, commonly manifesting as well-defined patches on the scalp or other body areas [1]. It can occur at any age and affects both men and women, leading to significant psychosocial distress due to its visible effect on the physical appearance. AA affects approximately 2% of the general population at some stage in their lives, as evidenced by numerous epidemiological studies conducted across Europe, North America, and Asia [1]. In the Middle East, Saudi Arabia reported prevalence rates between 2.3% and 13.8%, while in Africa, prevalence rates were lower, such as 0.1% in Nigeria and 0.2% in South Africa [2]. These variations can be attributed to differences in genetic, environmental, and health care factors.

XSL•FO
RenderX

A study conducted in 2021 showed that Facebook, Twitter, Instagram, TikTok, YouTube, Snapchat, and other social media platforms help users share dermatology-related content, seek medical advice, share photos and videos, and build social networks by interacting with other users, thereby benefiting greatly from the platforms' visual nature [3]. Another study showed that a significant number of participants with dermatological conditions (82.4%) used the internet for information, and 65.4% specifically used social media [4]. AA is a frequently searched term on Instagram, indicating high engagement among those affected [5]. By analyzing online discussions and search trends related to AA, researchers and health care professionals (HCPs) can gain insights into regional perceptions, common beliefs, challenges, and levels of awareness. A 2019 study reported that social media can help develop instruction for dermatology trainees, facilitate new research methods, and foster patient-centered perspectives and advocacy groups [6]. Using online data analytics, key themes and patterns in AA-related interactions can be identified, aiding the development of public health campaigns and educational initiatives, enhancing awareness, reducing stigma, and improving support for those affected by AA.

Infodemiology refers to the study of the distribution and determinants of information shared through electronic media, particularly the internet, with the aim of informing public health strategies and policy decisions [7]. Social media analysis provides large-scale, real-time data that allows researchers to explore human behavior, track emerging trends, and understand public perceptions [8]. Unlike traditional surveys or focus groups, which are limited by sample size and response bias [9], using data from social media to capture the real-world patient journey, alongside traditional epidemiological studies, sheds light on unmet needs, improves patient care pathways, develops tailored treatment pathways, and helps future researchers address the complexities of AA management [10]. Online discussions often illuminate dimensions that may be underrepresented in clinical datasets, including experiences of stigma, emotional distress, identity challenges, and social burden. By capturing these experiences in real time, social media analysis complements conventional research methods and contributes to a more holistic framework for addressing the complexities of AA management. The study was grounded in infodemiology, and the findings informed the study design and research questions, justifying a focus on real-world online discourse and platform-based differences in AA communication. The framework further guided the interpretation of results by contextualizing observed patterns within established evidence on patient information-seeking behavior and the variability of online health content.

Methods

Study Design

The AA social analytics study was commissioned by Pfizer in April 2023 and conducted by Rila Global Consulting. The company scanned more than 100 million websites, analyzing up to 16 months of data from January 2022 to April 2023. This research focused on the United Arab Emirates, Kuwait, Bahrain, Oman, and Qatar and included content in both English and Arabic.

Data Sources and Collection

Rila Global Consulting crafted specific queries to retrieve mentions of AA-related terms, abbreviations, and hashtags associated with AA. They conducted searches for related products and identified key influencers by monitoring online engagement metrics. Authors with significant impact or visibility were selected as top influencers. The system scanned publicly available data on social media sites such as X (formerly Twitter), blogs, online forums, news outlets, and health information websites. Content originating from the United Arab Emirates, Kuwait, Bahrain, Oman, and Qatar was included in the analysis. Both English- and Arabic-language content was captured. All data were obtained exclusively from publicly accessible sources. No private accounts, password-protected content, or personally identifiable information were accessed or retained.

Search Strategy and Query Development

A structured query library was developed to retrieve AA-related content with high specificity. The query set included medical terminology related to alopecia areata, common abbreviations such as AA, lay expressions, and colloquial terms related to hair loss, hashtags associated with AA discussions, and product-related keywords associated with management, supplements, cosmetic solutions, and more. Queries were refined to minimize irrelevant results, particularly where "AA" appeared as an unrelated abbreviation; therefore, the dataset reflected genuine AA-related discourse.

Data Cleaning and Filtering

After data extraction, the dataset underwent a multistage cleaning process. This included the removal of duplicate posts and syndicated content to prevent repeated material. Language filtering was applied to retain only English- and Arabic-language content. Geolocation filtering was conducted using available metadata and platform-specific location indicators to ensure relevance to the targeted Gulf Cooperation Council countries. Bot-generated, spam, and automated content were excluded using algorithmic detection methods. The dataset was restricted to publicly accessible posts to maintain ethical compliance. No personal or identifying information was collected, stored, or analyzed at any stage of the study.

Content Classification and Analysis

Content processing and classification were performed using vendor-implemented automation, including language identification, keyword and query matching, and rule-based parsing rather than trainable machine learning models. No supervised or unsupervised statistical natural language processing models (topic modeling, embeddings, or large language models) were trained or evaluated as part of this study. Posts were categorized across multiple dimensions, including topic (symptoms, diagnosis, emotional impact, and treatment discussions), sentiment (positive, negative, or neutral), audience type (patients, caregivers, HCPs, influencers, or the public), and platform characteristics (X, forums, or news websites). Automated outputs were supplemented by expert review to support contextual interpretation, particularly for

XSL•FO
RenderX

Arabic-language posts and culturally specific expressions. Manual review was conducted to contextualize ambiguous posts, particularly in Arabic, but was not designed as a formal annotation exercise with a predefined sampling frame or dual independent coding. Accordingly, interrater agreement metrics were not calculated, and the analysis should be interpreted as descriptive rather than as a validated classification model.

Analytic Approach and Tooling

The classification approach was explicitly described as rule-based, using vendor-implemented parsing, keyword and query matching, and language identification. No supervised or unsupervised machine learning model training was conducted in this study. The term “hybrid approach” was defined as a combination of automated, rule-based processing and manual contextual review, particularly for Arabic-language content.

Bilingual Data Processing

English- and Arabic-language content was processed within a unified analytic workflow. Language identification was applied during data cleaning, and posts were classified using a shared rule-based framework across both languages without requiring full translation into a single language. Topic categories and labels were aligned through consistent keyword mapping and conceptual equivalence of medical and colloquial terms related to alopecia and hair loss. Expert manual review was conducted to support contextual interpretation, particularly for Arabic-language content. This review addressed dialectal variation, idiomatic expressions, and ambiguities that could not be fully resolved through automated rule-based processing. Although efforts were made to ensure consistency across languages, differences in linguistic structure and expression may have led to variability in classification.

Influencer Identification

Influential authors were identified using a combination of quantitative and qualitative indicators. These included the frequency of AA-related posts, engagement metrics (likes, comments, shares, and reposts), follower count and overall visibility, cross-platform presence, and the relevance and consistency of AA-related content. Authors demonstrating high reach, sustained engagement, or impact on online discussions were classified as key influencers.

Quantitative Metrics

Standardized social analytics metrics were applied to quantify the scale and dynamics of AA-related conversations. These metrics included mentions, impressions, reach, engagement, traffic, and follower-based indicators, as defined in Table 1. The metrics were used to compare patterns across countries and digital platforms and to identify trends over the study period.

Table . Key terms and metrics on alopecia areata study in Gulf countries.

Term	Definition
Mention	Mentioning a brand, a product, or an account in an online post
Author	The nickname, username, or full name of the individual or organization who posted a mention
Tweets (post)	A post on X (formerly Twitter)
Retweets (repost)	Resharing another account’s post to all followers by clicking on the repost (retweet) button
Following	The number of users the author follows
Followers	The number of users following the author
Impressions	The combined sum of the followers of the original author and the followers of any user who reshared the post
Reach	An estimate of the number of individuals who may have seen a piece of content based on engagement (eg, likes, comments, retweets, and replies) and traffic (eg, average site visitors and monthly visitors), allowing comparison across data sources and platforms
Engagement	The total number of times a user interacts with a tweet or post, as measured by a range of metrics that include likes, comments, replies, shares or retweets, click-throughs, and saves
Traffic	The average number of visitors to a website, mobile app, or profile

Descriptive Nature of Analysis

The analytical approach used in this study was both descriptive and exploratory. Data were gathered from publicly available online discussions via query-based retrieval rather than structured sampling methods. As a result, the findings reflect trends in digital discourse and are not meant to represent population-level estimates or support inferential statistical conclusions.

Ethical Considerations

This study analyzed publicly available data from online platforms and social media sources in English and Arabic. No private, restricted, or password-protected content was accessed. No personal identifiers were collected, stored, or reported. All data were analyzed in aggregate form, and individual users were not identifiable. The study adhered to established ethical principles for internet-mediated research and complied with applicable data protection standards. As the research involved

XSL•FO
RenderX

the analysis of publicly available content and did not include human participants or intervention, formal ethics committee approval was not required.

Results

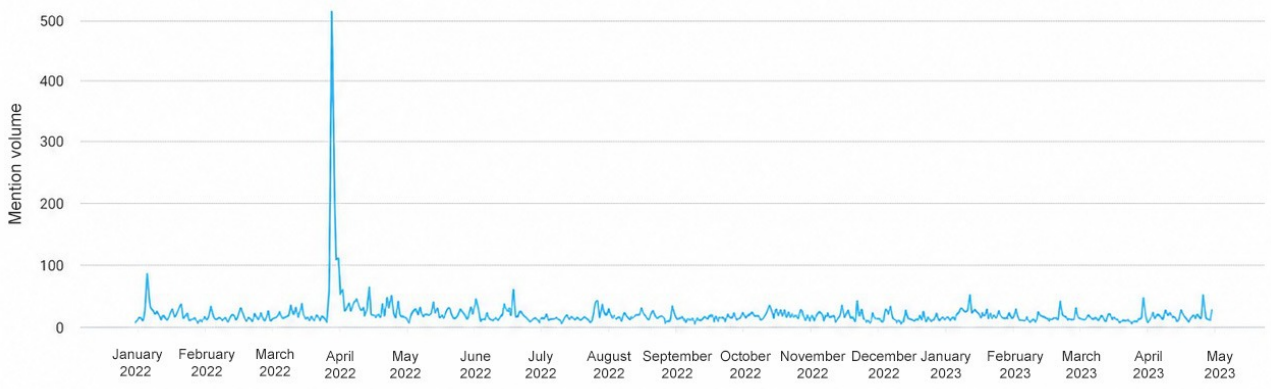
Overview

The analysis of online discussions about AA revealed several prevalent themes, misconceptions, and regional differences. Key topics included AA discussions in the Gulf, contributors to hair loss, treatment options, experiences with AA, emotional impacts, and coping strategies. Search trend analysis demonstrated differences in query volume and topic patterns across Gulf countries. English was predominantly used by HCPs and medical institutions, whereas patients used a combination of Arabic and English. Interchangeable use of terms such as "baldness," "hair loss," and "alopecia" was observed across discussions.

Impact of the Oscars Incident on AA Discourse in the Gulf Region

The volume of discussions on AA in the Gulf region remained relatively stable month over month, with a marked spike occurring during the Oscars in 2022, as shown in Figure 1. During the same period, alopecia was referenced in widely circulated media coverage [11]. However, this observation reflected a temporal association and should be interpreted as hypothesis generating rather than indicative of a causal relationship.

Figure 1. Monthly volume of alopecia areata mentions in the Gulf region, showing a marked spike in April 2022.



Content Source Analysis: AA Discussions in the Gulf (2022-2023)

Between 2022 and 2023, discussions on AA in the Gulf were primarily concentrated on news platforms and X (formerly Twitter). News sources accounted for 4505 (53%) of total mentions (N=8500), followed by X at 3655 (43%), and Tumblr at 340 (4%).

Gulf Region Country Analysis: AA Discussions

Discussions on AA in the Gulf region predominantly originated from the United Arab Emirates, which accounted for 5695 (67%) of total mentions (N=8500). Bahrain and Kuwait each accounted for 850 (10%) of the share of voice, followed by Qatar 680 (8%) and Oman 425 (5%).

Social Media Analysis: Dominance of Educational Content in AA Discussions

Social media discussions on AA were predominantly educational or informational in nature, accounting for 4692 (55.2%) of analyzed content (N=8500). Personal stories comprised 2015 (23.7%), followed by HCP interactions (n=1071, 12.6%) and references to medical journals or research articles (n=722, 8.5%), as shown in Table 2.

Table . Content types shared online about alopecia areata (N=8500).

Category	Value, n (%)
Educational or informational	4692 (55.2)
Personal stories	2015 (23.7)
Health care professional interactions	1071 (12.6)
Medical journals and research articles	722 (8.5)

Causes and Risk Factors

Posts included references to stress, dietary practices, hijab wearing, and environmental exposures as potential contributors to hair loss or alopecia. Users shared personal experiences and coping strategies. These observations reflect online discourse and are not intended to imply causal relationships.

Social media platforms are increasingly used by individuals experiencing hair loss as sources of medical information and patient education. However, much of the content available on these platforms promotes hair care products and treatments that lack scientific evidence [12]. HCPs use social media to educate patients about alopecia, clarify misconceptions about hair oils, and emphasize evidence-based treatments such as minoxidil,

XSL•FO
RenderX

finasteride, and spironolactone pills. They also provide insights into the effects of the COVID-19 pandemic on hair health and discuss preventive measures against baldness, cautioning against hairstyles such as the "ponytail." Furthermore, they share information about new treatments such as microscopic cell technology and emphasize the importance of consulting with physicians for tailored treatment plans, including options such as hair transplantation, Regenera injections, platelet-rich plasma (PRP) therapy, and mesotherapy.

Patients with alopecia typically seek advice from dermatologists and cosmetologists. Cosmetic surgeons were mentioned mainly in promotional contexts. Most educational content about alopecia from HCPs and medical facilities is in English, which may restrict access for Arabic-speaking individuals. Some facilities address this by offering multilingual resources.

Patients' journeys in seeking solutions for alopecia vary in duration Figure 2, with many starting with medical treatments but eventually exploring natural products or alternative methods if conventional treatments prove ineffective [13]. Some patients may initially try medical treatments such as injections and creams provided by hospitals or dermatologists but may turn to alternative remedies based on recommendations from relatives or social circles if these treatments fail. There was a disparity in the use of clinical language when discussing alopecia compared to nonclinical language, indicating potential gaps in understanding and awareness of alopecia as a medical condition in the Gulf region and suggesting a need for improved education and awareness efforts.

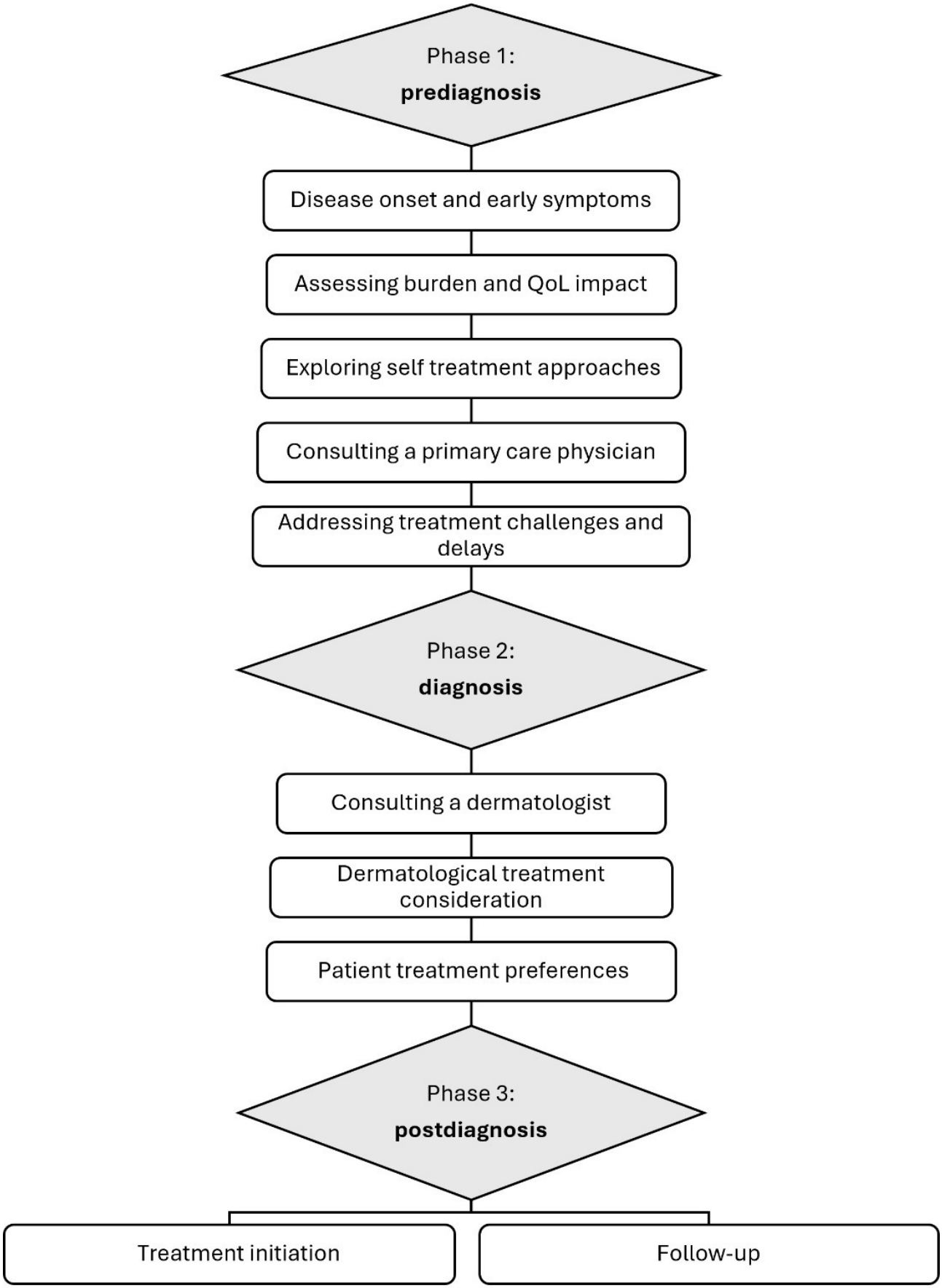
Common treatments included topical corticosteroids, minoxidil, and intralesional corticosteroids, which help reduce inflammation, promote hair regrowth, and manage hair loss [14,15]. In the Gulf region, clinics frequently advertised hair transplant services using search terms such as "hair loss." In contrast, despite the availability of PRP therapy, patient discussion regarding this modality appeared limited. Homeopathy was also promoted as a treatment option, alongside various products such as dietary supplements for hair loss, shampoos, and conditioners. Natural remedies and foods believed to aid hair regrowth, such as green tea, aloe vera, and garlic, are frequently discussed. Dermarollers were promoted for hair regrowth, with stress, pollution, and harsh chemicals identified as contributing factors to hair loss.

News about the approval of JAK (janus kinase) inhibitors for the treatment of alopecia was widely shared across the Gulf region, with foreign HCPs, local authors, and media outlets disseminating the information [16].

Authors in the Gulf region also shared news about a South Korean politician promising free baldness treatments in exchange for votes. The ruling party's presidential candidate, Lee Jae-myung, proposed covering baldness treatment under public health insurance, which generated mixed reactions among voters [17]. This news sparked discussions among Gulf authors, highlighting the distinction between hair loss and alopecia. Prominent organizations, including dermatology societies across the Gulf, actively engaged in discussions on this topic, contributing to awareness and understanding in the region.

XSL•FO
RenderX

Figure 2. Alopecia areata patient journey during the prediagnosis, diagnosis, and postdiagnosis phases. QoL: quality of life.



XSL•FO
RenderX

Discussion

Principal Results

Over a span of 16 months, from January 2022 to April 2023, we thoroughly analyzed approximately 8500 mentions of AA, with a significant 67% (5695) of these coming from the United Arab Emirates. Most content was educational (n=4692, 55.2%). The incident at the Oscars in April 2022 led to a regional spike in alopecia-related discussions. Users discussed AA across the region on platforms such as Gulf News (n=4505, 53%) and X (n=3655, 43%). English was predominantly used by HCPs, while patients used a combination of English and Arabic. Authors interchangeably used “baldness,” “hair loss,” and “alopecia” to enhance audience engagement. Fungal infections of the scalp, stress, diet, and environmental factors were discussed as leading contributors to alopecia. Many patients expressed emotional distress associated with alopecia. Natural remedies and self-treatment were frequently explored by patients, while beauty and medical centers advertised nonpharmacological treatments along with topical minoxidil.

Causes and Risk Factors

In the Gulf region, discussions on hair loss and alopecia highlighted several factors: scalp fungal infections were mentioned in 1% of all conversations, while stress, dietary choices, hijab wearing, and environmental factors were recognized as possible contributors to hair loss.

Scalp Hair Fungal Infection

According to reports from news outlets, the German Association of Dermatologists has identified head fungus as one of the skin conditions likely leading to baldness or alopecia. Fungal infections are known to trigger hair loss, often accompanied by itching. Tinea capitis, a condition affecting both children and adults, manifests as scalp eruptions and hair loss [18]. Discussions on social media concerning both hair and head fungus and their effects on hair loss and alopecia constituted 1% of all posts analyzed.

Stress

Users engaged in discussions regarding the correlation between stress and hair loss, highlighting the importance of managing stress to reduce or reverse the effects of alopecia. Chronic stress leads to hair loss, although the exact mechanism is unknown [19]. Personal stories were shared by users on how stressful life events, such as school exams, divorce, or family illness, could exacerbate hair loss. Users expressed the challenging reality of living with alopecia, describing a vicious cycle where hair loss induces stress and emotional distress, thereby worsening and accelerating the condition.

Dietary Choices

Users actively discussed the influence of diet on hair loss, suggesting that crash diets and unhealthy eating habits can disrupt hormonal balance and lead to increased hair loss. Some users shared personal experiences of certain diets exacerbating hair loss, whereas others advocated for the benefits of maintaining a healthy and balanced diet to manage alopecia. Personal experiences were also shared, underlining how adopting nutritious eating habits had improved hair health and reduced hair loss for some individuals. Users offered tips and advice on specific foods and supplements that can be beneficial in managing alopecia. It is essential to address and rectify any nutritional deficiencies in patients experiencing hair loss [20].

Hijab Wearing

Users expressed concerns that wearing a hijab, a head covering worn by Muslim women, increases the rate of hair loss in individuals with alopecia. Some users proposed that the lack of exposure to vitamin D due to wearing a hijab could be a contributing factor to hair loss. This discussion sparked conversations about strategies to prevent or reduce hair loss in patients with alopecia who choose to wear a hijab. Conversely, other social media users recommended headscarves as a practical solution for covering bald patches caused by alopecia. They argued that headscarves can offer a way for individuals who are self-conscious about their hair loss to feel more comfortable and confident in public settings.

Environmental Factors

Discussions also frequently referenced environmental contributors, including pollution and water quality, as potential factors associated with hair loss. Users reported experiences suggesting that relocating to different cities influenced the severity of hair loss. Discussions surrounding alopecia and hair loss on social media platforms in the Gulf region may be relevant for pharmaceutical and aesthetic stakeholders seeking to understand public interest and awareness. Water filters were frequently mentioned among products discussed by companies, with associated keywords and hashtags appearing in online conversations about alopecia and hair loss.

Comparison of Gulf Discourse Patterns With Global Studies on Dermatology-Related Social Media Content

The discourse patterns observed in the Gulf region align with broader international research on dermatology-related social media engagement. Several studies have demonstrated that patients with chronic skin conditions such as acne, vitiligo, and psoriasis use social media and online platforms to seek information, share experiences, and influence health care decisions [21-23]. Misinformation, terminology confusion, and commercial influence have also been identified as recurring challenges in digital dermatology conversations worldwide [24].

Cultural Dimension of Alopecia Discourse in the Gulf

The cultural context of the Gulf region may influence how AA is discussed online. Religious, social, and gender norms may appear to shape disclosure behaviors, particularly among women, for whom hair functions as an important marker of social identity and personal meaning. The wearing of the hijab, for example, emerged both as a contributor to hair loss and as a coping strategy for concealing visible symptoms. A study investigated alopecia among Muslim women who self-reported hair loss related to wearing the hijab [25]. The study revealed that a significant number of Muslim women experienced alopecia, and they believed it is associated with hijab wear. More than one-third of the surveyed women were unaware of hijab-associated alopecia. The study discussed the importance of educating Muslim women on ways to mitigate the risks of

XSL•FO
RenderX

alopecia or treat it without compromising their religious beliefs or values [25].

Language use also reflected cultural positioning, with Arabic more common in patient-driven discussions and English predominating in HCP communication. These findings suggest that online discourse around alopecia in the Gulf is embedded within broader sociocultural dynamics rather than being purely biomedical in nature.

Limitations

One significant limitation of our study is the exclusion of data from Facebook and Instagram, as these platforms do not allow data crawling. This restriction clearly impacts the comprehensiveness of our analysis, particularly for social media platforms where data collection is limited. The analysis therefore reflects discussions from publicly accessible, primarily text-based platforms, potentially underrepresenting conversations occurring within private networks or visually driven platforms. The findings were based on a fixed 16-month observation period (2022 - 2023). As social media dynamics evolve rapidly, the identified trends may not fully capture longer-term or ongoing shifts in discourse. Automated language processing and rule-based classification were used in the analysis of bilingual content. Although efforts were made to ensure accuracy, automated processing may introduce translation or categorization bias. The absence of full translation may have resulted in minor inconsistencies in cross-language classification.

Another limitation of this study is that the manual review conducted to contextualize ambiguous posts, especially those in Arabic, was not structured as a formal annotation exercise. It lacked a predefined sampling frame and did not involve dual independent coding. As a result, interrater agreement metrics were not calculated, meaning the findings should be viewed as descriptive rather than indicative of a validated classification model. This limitation may affect the reliability and generalizability of the analysis.

Comparison With Prior Work

A literature review of AA in the Middle East and Africa discussed the prevalence, quality of life, and treatment modalities for AA [2].

A study by Gümüş [5] on social media use in patients with AA using questionnaires showed that 72.6% of the patients used social media to know about AA. The most commonly used platforms were Google (58.5%), Instagram (31.4%), and YouTube (30.5%). In contrast, in our study, users discussed AA across the region on platforms such as Gulf News (4505/8500, 53%) and X (3655/8500, 43%) [5].

An analysis of alopecia-related content on Instagram and TikTok examined popular posts within 10 hair loss–related hashtags on both platforms. Of the 90 posts analyzed within Instagram, nonmedical professional influencers were responsible for 66% of the posts, hair and wig companies for 29%, and medical professionals for only 4%. Of the 100 TikTok posts analyzed, influencers and patients each created 38%, while hair and wig companies contributed 14%, medical professionals accounted for 10%, and none of the top posts were created by board-certified dermatologists [12].

Conclusions

The analysis of online discussions surrounding AA in the Gulf region provided insights into the perceptions and challenges associated with this condition. Throughout the study, we observed a variety of themes, including treatment options and emotional impacts. Key findings included a predominance of discussions originating from the United Arab Emirates. The lack of understanding among patients about AA versus general hair loss leads to delayed diagnosis and inadequate treatment. This study hypothesizes that cultural and societal influences, such as the Oscars incident in 2022, significantly shape perceptions and discussions about AA. Both natural remedies and medical interventions were frequently discussed, but the promotion of nonevidence-based treatments and the lack of standardized treatment protocols highlight the need for evidence-based guidance. The emotional toll of living with AA is significant, affecting an individual's well-being beyond hair loss.

In summary, the study emphasizes the diverse landscape of AA discourse in the Gulf region, encompassing variations in awareness, resource access, and the emotional and social challenges faced by individuals affected by the condition. Moving forward, efforts must focus on enhancing education, promoting evidence-based treatments, and providing holistic support for individuals affected by AA.

Acknowledgments

The authors would like to extend our sincere gratitude to Reshma Susan Reji for her invaluable assistance in preparing this manuscript.

Funding

The study was funded by Pfizer, Inc. Medical writing support was provided by Reshma Susan Reji, NTAM Dubai, and was funded by Pfizer Inc Ltd.

Authors' Contributions

Conceptualization: AR, GAA, Ameen Al Awadhi, Amina Al Awadhi, FE, HMA
Data curation: FE

XSL•FO
RenderX

Formal analysis: FE
Investigation: FE, HMA
Writing—original draft: FE, HMA
Writing—review and editing: AR, GAA, Ameen Al Awadhi, Amina Al Awadhi, FE, HMA

Conflicts of Interest

FE is a former employee of Pfizer Inc., and HM is a full-time employee of Pfizer Inc. The other authors declare no conflicts of interest.

References

1. Pratt CH, King LE Jr, Messenger AG, Christiano AM, Sundberg JP. Alopecia areata. Nat Rev Dis Primers 2017 Mar 16;3:17011. [doi: 10.1038/nrdp.2017.11] [Medline: 28300084]
2. Al Hammadi A, Parmar NV, Aljefri K, et al. Review on alopecia areata in the Middle East and Africa: landscape and unmet needs. Dermatol Ther (Heidelb) 2023 Jul;13(7):1435-1464. [doi: 10.1007/s13555-023-00946-8] [Medline: 37338721]
3. Szeto MD, Mamo A, Afrin A, Militello M, Barber C. Social media in dermatology and an overview of popular social media platforms. Curr Dermatol Rep 2021;10(4):97-104. [doi: 10.1007/s13671-021-00343-4] [Medline: 34692234]
4. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
5. Gümüş S, Uzunçakmak TK, Aşkın Ö. Social media use in patients with alopecia areata. J Turk Acad Dermatol 2023;17(3):62-68. [doi: 10.4274/jtad.galenos.2023.83803]
6. DeBord LC, Patel V, Braun TL, Dao H Jr. Social media in dermatology: clinical relevance, academic value, and trends across platforms. J Dermatolog Treat 2019 Aug;30(5):511-518. [doi: 10.1080/09546634.2018.1530444] [Medline: 30265614]
7. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. J Med Internet Res 2009 Mar 27;11(1):e11. [doi: 10.2196/jmir.1157] [Medline: 19329408]
8. Zachlod C, Samuel O, Ochsner A, Werthmüller S. Analytics of social media data – state of characteristics and application. J Bus Res 2022 May;144(5):1064-1076. [doi: 10.1016/j.jbusres.2022.02.016]
9. Reveilhac M, Steinmetz S, Morselli D. A systematic literature review of how and whether social media data can complement traditional survey data to study public opinion. Multimed Tools Appl 2022;81(7):10107-10142. [doi: 10.1007/s11042-022-12101-0] [Medline: 35194384]
10. Barbarot S, Reguiai Z, Voillot P, Malaab J, Schück S, Tamzali A. The impact of alopecia areata on patients' daily lives: a study using social media in France. JEADV Clin Pract 2024 Dec;3(5):1735-1744. [doi: 10.1002/jvc2.539]
11. Lava CX, Hakimi AA. And the Oscar goes to... alopecia? A Google Trends analysis following the 94th Academy Awards. J Clin Aesthet Dermatol 2022 Dec;15(12):40-41. [Medline: 36569530]
12. Laughter M, Anderson J, Kolla A, et al. An analysis of alopecia related content on Instagram and TikTok. J Drugs Dermatol 2022 Dec 1;21(12):1316-1321. [doi: 10.36849/JDD.6707] [Medline: 36468968]
13. Hosking AM, Juhasz M, Atanaskova Mesinkovska N. Complementary and alternative treatments for alopecia: a comprehensive review. Skin Appendage Disord 2019 Feb;5(2):72-89. [doi: 10.1159/000492035] [Medline: 30815439]
14. Cranwell WC, Lai VW, Photiou L, et al. Treatment of alopecia areata: an Australian expert consensus statement. Australas J Dermatol 2019 May;60(2):163-170. [doi: 10.1111/ajd.12941] [Medline: 30411329]
15. Meah N, Wall D, York K, et al. The Alopecia Areata Consensus of Experts (ACE) study: results of an international expert opinion on treatments for alopecia areata. J Am Acad Dermatol 2020 Jul;83(1):123-130. [doi: 10.1016/j.jaad.2020.03.004] [Medline: 32165196]
16. OLUMIANT® receives Health Canada approval, becoming the second approved treatment for severe alopecia areata in Canada. Canadian Alopecia Areata Foundation. 2024. URL: https://www.canaaf.org/olumiant-receives-health-canada-approval-becoming-the-first-approved-treatment-for-severe-alopecia-areata-in-canada/ [accessed 2026-02-11]
17. South Korea should fund hair loss treatment, says election hopeful in bald bid for power. The Guardian. 2022. URL: https://www.theguardian.com/world/2022/jan/06/south-korean-presidential-hopeful-says-state-should-pay-for-hair-loss-treatment [accessed 2024-04-03]
18. Martin ES, Elewski BE. Tinea capitis in adult women masquerading as bacterial pyoderma. J Am Acad Dermatol 2003 Aug;49(2 Suppl Case Reports):S177-S179. [doi: 10.1067/mjd.2003.327] [Medline: 12894115]
19. Bryant E. How stress causes hair loss. National Institutes of Health. 2021. URL: https://www.nih.gov/news-events/nih-research-matters/how-stress-causes-hair-loss [accessed 2024-04-03]
20. Guo EL, Katta R. Diet and hair loss: effects of nutrient deficiency and supplement use. Dermatol Pract Concept 2017 Jan;7(1):1-10. [doi: 10.5826/dpc.0701a01] [Medline: 28243487]

XSL•FO
RenderX

21. Ertekin SS, Salici NS, Manav Bas V, et al. Influence of social media and internet on treatment decisions in adult female acne patients: a cross-sectional survey study. Dermatol Pract Concept 2024 Jul 1;14(3):e2024156. [doi: 10.5826/dpc.1403a156] [Medline: 39122512]
22. Social media platforms as a resource for vitiligo support. J Drugs Dermatol 2022 Oct 1;21(10):1135-1136. [Medline: 36219048]
23. Aslan Kayıran M, Karadağ AS, Oğuz Topal İ, et al. Habits of using social media and the internet in psoriasis patients. Dermatol Pract Concept 2022 Jul;12(3):e2022143. [doi: 10.5826/dpc.1203a143] [Medline: 36159107]
24. Militello M, Yang RA, Anderson JB, Szeto MD, Presley CL, Laughter MR. Social media and ethical challenges for the dermatologist. Curr Dermatol Rep 2021;10(4):120-127. [doi: 10.1007/s13671-021-00340-7] [Medline: 34540357]
25. Alhanshali L, Bawany F, Buontempo MG, Shapiro J, Lo Sicco K. Understanding perceptions of hair loss in hijab-wearing women: a pilot survey study. Int J Womens Dermatol 2023 Nov;9(4):e115. [doi: 10.1097/JW9.0000000000000115] [Medline: 38028023]

Abbreviations

AA: alopecia areata
HCP: health care professional
JAK: janus kinase
PRP: platelet-rich plasma

Edited by J Meisenheimer; submitted 31.May.2025; peer-reviewed by N Hu; revised version received 26.Jun.2026; accepted 12.Jul.2026; published 13.Aug.2026.

Please cite as:
Reda A, AL Alawi G, Al Awadhi A, Al Obaidli A, Ebeid F, Ahmed HM
Online Discussions About Alopecia Areata in 5 Arabian Gulf Countries: Social Media Listening Study
JMIR Dermatol 2026;9:e78262
URL: https://derma.jmir.org/2026/1/e78262
doi:10.2196/78262

© Ashraf Reda, Ghadear AL Alawi, Ameen Al Awadhi, Amina Al Obaidli, Fadwa Ebeid, Haytham Mohamed Ahmed. Originally published in JMIR Dermatology (http://derma.jmir.org), 13.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing Alopecia Areata Misinformation: Social Media Analysis

Shanika N Francis[1], BS, MHS; Ank A Agarwal[2,3], BA; Aaron H Fanous[2], MD; Dawn H Siegel[3], MD; Roxana Daneshjou[2,3], MD, PhD

[1]Meharry Medical College, School of Medicine, 1005 Dr DB Todd Jr Blvd., Nashville, TN, United States

[2]Department of Biomedical Data Science, Stanford University, 1265 Welch Rd, Stanford, CA, United States

[3]Department of Dermatology, Stanford University, Stanford, CA, United States

Corresponding Author:
Shanika N Francis, BS, MHS
Meharry Medical College, School of Medicine, 1005 Dr DB Todd Jr Blvd., Nashville, TN, United States

Abstract

Background: Alopecia areata (AA) is an autoimmune, hair loss–inducing disease affecting individuals of all genders and ethnic backgrounds worldwide. As social media platforms and artificial intelligence chatbots increasingly influence patient behavior, individuals with AA may turn to these sources for treatment guidance.

Objective: This study evaluated the accuracy of AA-related information on social media and assessed whether patients were exposed to misinformation that could hinder appropriate care. Influencers may recommend treatments applicable to alopecia broadly or promote therapies lacking medical benefit.

Methods: YouTube and TikTok were searched using the hashtags #alopecia and #alopeciatreatment. Posts were selected based on hashtag relevance. Videos were categorized using standard AA treatment guidelines into (1) accurate and medical data–backed (aligned with guideline-supported treatments), (2) limited or anecdotal evidence (unverified treatments without clear misinformation), (3) misinformed and/or false (inaccurate or misleading claims), or (4) other (unrelated to alopecia management).

Results: A total of 98 YouTube and 222 TikTok videos related to #alopecia and #alopeciatreatment were analyzed. Of the YouTube videos analyzed, 76% (n=75) were categorized as “other,” 13% (n=13) as “accurate and medical data–backed,” and 10% (n=10) as “limited or anecdotal evidence”; no videos were classified as “misinformed and/or false.” Of the TikTok videos analyzed, 76% (n=170) were categorized as “other,” 17% (n=38) as “accurate and medical data–backed,” 5% (n=12) as “limited or anecdotal evidence,” and 1% (n=2) as “misinformed and/or false.” Among videos specifically discussing AA treatment, 56% of YouTube videos and 73%of TikTok videos were categorized as “accurate and medical data–backed.” The difference in accurate vs limited or nonmedical YouTube content was statistically significant (P=.002), while misinformation prevalence between platforms was not significantly different (P=.15).

Conclusions: Patients with AA are frequently exposed to generalized alopecia content on social media, which may not consistently offer disease-specific guidance. Health care professionals should help ensure that social media platforms and artificial intelligence tools are leveraged to promote accurate education and proactively combat misinformation surrounding AA treatment.

(JMIR Dermatol 2026;9:e90088) doi:10.2196/90088

KEYWORDS

alopecia areata; social media; digital health; health literacy; health communication; patient education

Introduction

Alopecia areata (AA) is an autoimmune disease prevalent in 2% of the worldwide population that affects both men and women from a range of ethnic backgrounds [1]. AA also affects 1 in 1000 children and adolescents [2]. The physical presentation of AA includes noncicatricial (nonscarring) hair loss—which maintains the potential for regrowth and does not degrade the hair follicle [3]. The disease is also highly impactful on the psychosocial well-being of diagnosed patients; due to negative public associations of hair loss with poor health, many patients with AA experience social discrimination, impairment, and stigma [4].

Social media platforms have become vital components in shaping patient behaviors and trends, especially related to alopecia [5]. Social media influencers use platforms such as Instagram, TikTok, and YouTube to disseminate treatment recommendations to a broad audience, including young adults and children [6]. When experiencing hair loss, patients tend to seek medical advice from social media influencers instead of their health care provider or prior to their first consultation [1]. However, concerns have emerged regarding the accuracy and safety of alopecia-related advice provided by influencers, especially when compared to recommendations from dermatologists [5]. This study aimed to review the accuracy of the information about AA on 2 video-based social media

XSL•FO
RenderX

platforms, YouTube and TikTok, and investigate whether patients may potentially receive misinformation that could either deter or impair appropriate AA treatment from their health care providers [5].

Understanding the information on social media about AA treatment will be critical for guiding physicians on how to better proactively inform their patients with accurate treatment information and to create effective educational materials for correctly managing the disease.

Methods

YouTube and TikTok Data Collection

A web scraping platform, Apify (Apify Technologies sro), was used to extract alopecia-related videos' data from YouTube and TikTok. To identify relevant posts, we selected videos with hashtags #alopecia or #alopeciatreatment in the video caption, description, or title. A total of 98 and 222 posts were collected from YouTube and TikTok respectively between July 17, 2024, and July 24, 2024. Videos retrieved through hashtag searches were manually screened for relevance by 1 reviewer (SF) prior to application of the inclusion and exclusion criteria. Duplicate or reposted videos were excluded from analysis when identified. Included posts were required to meet the following inclusion criteria: (1) posted by dermatologists who explicitly stated their title or had their credentials listed vs other users and (2) focused on hair loss treatment discussions or advice.

For criterion 1, dermatologist status was identified based on self-reported professional titles or board-certified medical doctor credentials included within the caption, title, profile, or video content.

To ensure standardized creator classification, only self-identified US-trained board-certified dermatologists with medical doctor credentials were included. Posts from creators who did not identify as dermatologists were included in the analysis. For criterion 2, posts were excluded if they were unrelated to AA (eg, traction alopecia) or if they were advertisements or non-English content. The relevancy of the video to various hashtags was confirmed by examining post captions for the searched hashtags and relatedness of the content to alopecia to ensure appropriate categorization.

Content Categorization

Data were manually reviewed and categorized based on whether the treatment information provided in the video aligned with medical standards for treating AA. Posts were grouped into four categories: (1) accurate and medical data–backed, (2) limited or anecdotal evidence, (3) misinformed and/or false, and (4) other. Posts classified as "accurate and medical data–backed" referenced treatments aligned with established American Academy of Dermatology or National Alopecia Areata Foundation guidelines, such as corticosteroid injections, topical minoxidil, or platelet-rich plasma therapy. Posts categorized as "limited or anecdotal evidence" included treatments with limited, inconsistent, or primarily anecdotal clinical support, such as herbal oils, supplements, scalp massage, or alternative therapies. While these treatments were not necessarily directly misleading or false, they were not strongly supported by current evidence-based dermatologic guidelines. Limited or anecdotal evidence and "misinformed and/or false" posts included content that directly contradicted established medical guidelines, promoted unsupported curative claims, or recommended treatments without any supporting evidence.

US Food and Drug Administration–approved therapies for AA were categorized as evidence-based treatments. Off-label therapies were evaluated according to consistency with established dermatologic guidelines and recommendations from the American Academy of Dermatology and the National Alopecia Areata Foundation. Posts categorized as "other" included those focused on types of alopecia not related to AA, advertisements, treatment progress updates, or unrelated topics (eg, wig modeling).

Statistical Analysis

Data visualization and analysis were performed using Python 3.3.1 (Python Software Foundation) with Matplotlib: descriptive statistics were used to quantify the proportion of posts falling within each categorization group across both platforms. For instance, of the 98 posts on YouTube, 13% (n=13) were categorized as "accurate and medical data–backed," and 76% (n=75) were classified as "other." A similar approach was used to analyze TikTok posts. Further statistical analysis, including chi-square tests, was performed to assess the difference in the prevalence of misinformation between YouTube and TikTok posts ($P<.05$ was considered statistically significant). These comparisons were performed to evaluate whether misinformation prevalence differed between platforms and to better understand potential platform-specific patterns in AA treatment content.

Ethical Considerations

All necessary permissions were obtained for the use of data scraping tools, and content collection adhered to platform policies [7,8]. Given that no personal data were collected and the study involved publicly available information, institutional review board approval was not required per Stanford University guidelines [9].

Results

YouTube Content Analysis

A total of 98 YouTube videos related to #alopecia and #alopeciatreatment were identified and analyzed. Of these, 76% (n=75) of videos were categorized as "other," as they did not address AA treatment specifically. The remaining 26% (n=23) of videos addressed AA treatment specifically. Among these, 56% (n=13) were categorized as "accurate and medical data–backed," promoting evidence-based treatments such as corticosteroid injections, platelet-rich plasma therapy, and topical or oral minoxidil (Table 1). The remaining 43% (n=10) were classified as "limited or anecdotal evidence," with content recommending alternative treatments such as garlic oil, ginger, rosemary, and scalp massages. Some influencers suggested seeing a dermatologist if no improvement was seen after 3 to 4 months of using these alternative treatments.

XSL•FO
RenderX

Statistical analysis revealed a statistically significant difference in the proportion of accurate content between the YouTube videos categorized as “accurate and medical data–backed” and those classified as “limited or anecdotal evidence” (P=.002). No videos in this sample were categorized as “misinformed and/or false” (Figure 1).

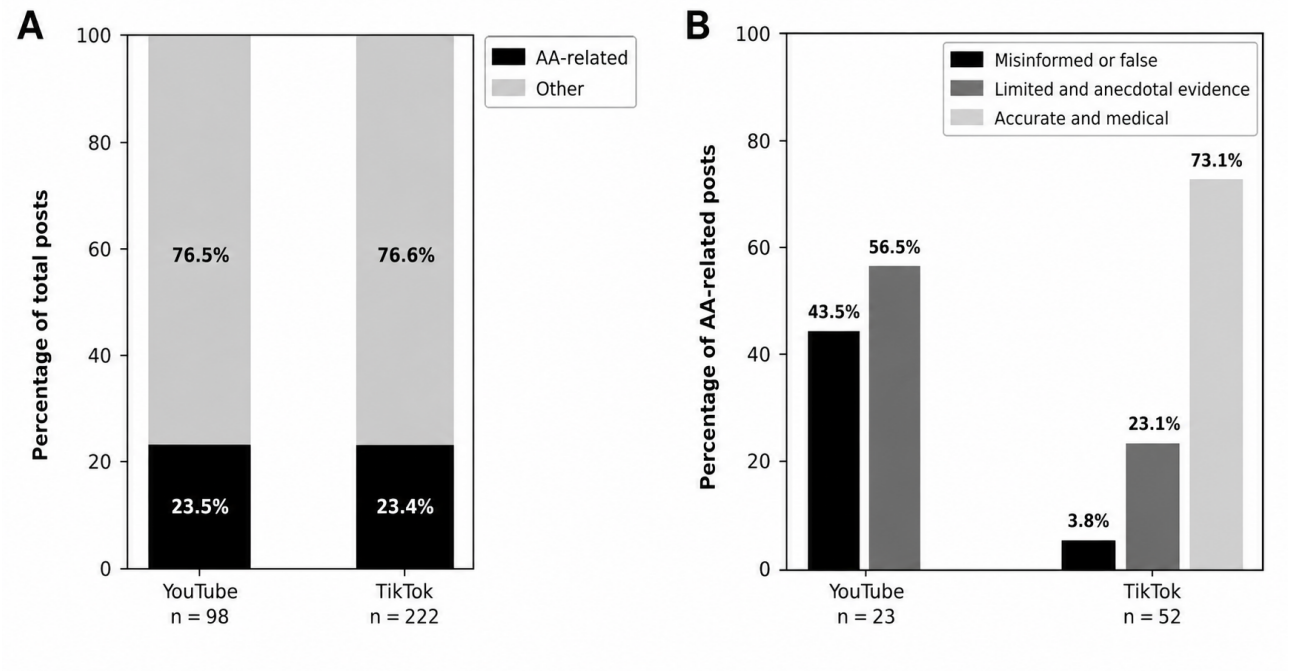
Table . Standard alopecia areata treatments recommended by the American Academy of Dermatology [10].

Condition	Treatment options	Details
1 or 2 bald spots (<1 year)	Wait-and-see approach, corticosteroids, and minoxidil (Rogaine)	Many patients experience spontaneous regrowth. Corticosteroids may stimulate regrowth, while minoxidil may help maintain hair growth.
Children (aged ≤10 years)	Corticosteroids and minoxidil	Corticosteroids are often applied daily. Minoxidil may help maintain regrowth.
Children (aged >10 years)	Contact immunotherapy and JAK[a] inhibitors	Contact immunotherapy typically involves weekly treatments. JAK inhibitors may promote regrowth in severe cases.
Patchy hair loss	Corticosteroid injections, topical corticosteroids, anthralin, and minoxidil	Corticosteroid injections may be effective but painful. Topical therapies may support regrowth.
Extensive hair loss	Contact immunotherapy, JAK inhibitors, and other systemic medications	Systemic therapies and immunotherapy may promote significant regrowth in severe disease.
Hiding hair loss[b]	Styling products; wigs, hairpieces, or scalp prostheses; shaving; eyebrow powder; artificial eyebrows; and microblading	Cosmetic approaches may help improve appearance and confidence.

[a]JAK: Janus kinase.

[b]Supportive or cosmetic management approaches rather than disease-directed alopecia areata treatments.

Figure 1. (A) Distribution of 98 YouTube videos and 222 TikTok posts tagged with #alopecia and/or #alopeciatreatment, showing the proportion that contained alopecia areata (AA) treatment–related content vs other content. (B) Distribution of 23 YouTube videos and 52 TikTok posts specifically addressing AA treatment, showing the proportions categorized as misinformed and/or false, limited or anecdotal, and accurate or medical.



TikTok Content Analysis

A total of 222 TikTok videos related to #alopecia and #alopeciatreatment were reviewed. Of these, 76% (n=170) were classified as “other.” The remaining 23% (n=52) of TikTok posts focused on AA treatment. Of these, 73% (n=38) were categorized as “accurate and medical data–backed,” while 23% (n=12) were classified as “limited or anecdotal evidence.” A small portion of posts (n=2, 3%) were identified as “misinformed and/or false,” and these posts included content promoting unsubstantiated treatments, such as the use of unapproved topical solutions or unproven dietary supplements.

Although TikTok demonstrated a higher prevalence of misinformation-related content, there was no statistically significant difference between platforms (P=.15; Figure 1B).

XSL•FO
RenderX

Discussion

Our study found that a large portion of videos obtained through #alopecia and #alopeciatreatment searches were unrelated to AA treatment or instead focused on other forms of alopecia. Among videos specifically discussing AA, explicit misinformation was limited, although some TikTok videos contained misleading or non–evidence-based treatment recommendations. Additionally, many posts emphasized emotional support and community-building themes, highlighting both the informational and psychosocial roles social media plays for patients with AA.

AA is a visual and psychosociological disruptive disease that may prompt patients to seek online resources outside of the patient-physician relationship. Social media has become a ubiquitous resource for patients to search for health care–related information, including treatments for AA. It is critical that physicians and health care professionals understand the extent and depth of the information that their patients are exposed to on social media to be aware of potential misinformation dissemination and provide correct educational materials on treatment. Previous studies have only sought to understand social media post myths related to alopecia, including untrue hair loss causes and understudied alopecia treatments [5].

The results of this study highlight that a sizeable portion of the video posts centered on #alopecia or #alopeciatreatment searches were related to other types of alopecia (traction alopecia, alopecia universalis, etc), were not understandable, or were unrelated to the treatment of AA. Among the videos related to AA, none of the YouTube videos contained overt misinformation, while 2 TikTok videos included misinformation. However, videos with limited or anecdotal information could prove misleading to patients because social media influencers may be selling their own products or *oils* for profit as accepted treatments along with the standardized treatments such as minoxidil. Although posts containing misinformation represented a minor portion of the dataset, highly viewed or widely disseminated content may still disproportionately influence patient perceptions and treatment choices.

Given that most of the videos were not specifically about AA, another potential pitfall is patients following guidance on another type of alopecia altogether, which would not be appropriate for AA. The results indicate that patients may seek other forms of hair loss solutions that could further stagnate or worsen their hair loss, including the use of wigs, which can damage the hairline, and other oils, which could hamper the growth of hair. Without appropriate guidance on how to use these hair loss supportive measures, patients can worsen their hair loss or cause other diseases, such as scarring alopecia (traction alopecia), related to their scalp health.

There were limitations to the study, including the exclusion of nonvideo social media platforms such as Reddit, which could offer additional insights. The limited hashtag searches could have been broadened to include tags such as #alopeciacure, providing a wider range of posts and enabling a more thorough understanding of the information patients access. Content categorization was performed by a single reviewer, and interrater reliability statistics were not calculated, which may limit reproducibility and introduce potential classification bias. The physician creator inclusion criteria may also limit broader generalizability across international and multidisciplinary social media content environments.

Engagement metrics such as views, likes, comments, shares, and follower counts were not included in our analysis, which limited assessment of the overall reach and potential influence of misinformation-related content. Additionally, this study was limited by the relatively short 1-week data collection period, which may not fully capture the fluctuating nature of social media content and the algorithms of the TikTok and YouTube platforms. Content visibility and viewer engagement may rapidly evolve over time, and the trends observed in AA treatment discussions may have differed over a longer collection period.

Future studies should explore whether patients with AA follow their physician's treatment guidance or turn to social media influencer advice due to barriers in accessing health care or mistrust in the patient-physician relationship. Future studies may further evaluate how creator characteristics and audience engagement influence the spread and impact of AA content on social media. According to Gümüş et al [11], 96% of patients reported that they would trust their physician's advice if it conflicted with social media content on AA and would try to reach their physician through online medical platforms. Gantenbein et al [12] further confirm that the information patients find online does not significantly disrupt their relationship with their physicians. However, the same study also underscores the need for physicians to engage actively on social media to ensure that patients receive correct information upfront.

Our study found that patients both provided and sought out community-building support through video posts, highlighting the importance of physicians addressing the emotional impacts of AA rather than just the physical symptoms. As Aldhouse et al [13] demonstrated, patients often feel sadness, helplessness, social isolation, and a desire for support during treatment, which aligns with the emotional content seen in many posts from our study. It is vital for physicians to adopt a holistic approach in treating AA, offering support for emotional struggles and connecting patients to communities of others experiencing similar challenges.

Health care professionals can counteract misinformation by creating clear and educational social media content centered on evidence-based AA treatments. Meah et al [14] further support the use of evidence-based therapies for AA, including intralesional corticosteroids, topical treatments, and systemic therapies commonly used in clinical practice. These therapies may vary based on disease severity, patient population, and individual treatment considerations, nuances that may not always be fully reflected in social media discussions surrounding AA treatment. Creating precise, informative materials such as digital flyers that can be distributed in clinics gives patients straightforward explanations of their disease and treatment options, supporting ongoing education even after their visit.

As AA remains a hair loss disease of great physical concern and social limitation to patients, it is timely for healthcare

XSL•FO
RenderX

stakeholders to ensure that the social media platforms are used as tools of proactive education to combat any misinformation related to the disease.

Acknowledgments

ChatGPT (OpenAI) was used for basic proofreading, formatting, and language refinement only. All content was reviewed and verified by the authors.

Funding

The authors received no funding for this work.

Data Availability

The data analyzed in this study were derived from publicly available TikTok and YouTube videos. Due to the dynamic nature of social media content, the specific videos analyzed may no longer be available. Details on the search strategy and inclusion criteria are described in the Methods section.

Authors' Contributions

All authors contributed substantially to the conception and design of the project and to drafting and revising the manuscript. SNF led data acquisition, while SNF, AAA, and AHF led analysis and interpretation of data. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Gupta AK, Polla Ravi S, Wang T. Alopecia areata and pattern hair loss (androgenetic alopecia) on social media - current public interest trends and cross-sectional analysis of YouTube and TikTok contents. J Cosmet Dermatol 2023 Feb;22(2):586-592. [doi: 10.1111/jocd.15605] [Medline: 36606397]
2. McKenzie PL, Maltenfort M, Bruckner AL, et al. Evaluation of the prevalence and incidence of pediatric alopecia areata using electronic health record data. JAMA Dermatol 2022 May 1;158(5):547-551. [doi: 10.1001/jamadermatol.2022.0351] [Medline: 35385065]
3. Darwin E, Hirt PA, Fertig R, Doliner B, Delcanto G, Jimenez JJ. Alopecia areata: review of epidemiology, clinical features, pathogenesis, and new treatment options. Int J Trichology 2018;10(2):51-60. [doi: 10.4103/ijt.ijt_99_17] [Medline: 29769777]
4. Paller AS, Rangel SM, Chamlin SL, et al. Stigmatization and mental health impact of chronic pediatric skin disorders. JAMA Dermatol 2024 Jun 1;160(6):621-630. [doi: 10.1001/jamadermatol.2024.0594] [Medline: 38656377]
5. Finnegan P, O'Connor C, Murphy M. BH14 bald-faced lies: a qualitative review of misinformation on alopecia. Br J Dermatol 2024 Jul;191(Supplement_1):i79. [doi: 10.1093/bjd/ljae090.161]
6. Sforza L. Most of Gen Z using TikTok for health advice: survey. The Hill. 2024 Jul 16. URL: https://thehill.com/policy/technology/4774795-tiktok-gen-z-health-advice/ [accessed 2024-07-21]
7. Terms of service. TikTok. URL: https://www.tiktok.com/legal/page/us/terms-of-service/en [accessed 2026-07-03]
8. YouTube API services terms of service. YouTube. URL: https://developers.google.com/youtube/terms/api-services-terms-of-service [accessed 2026-07-03]
9. What is data scraping? Stanford University. URL: https://irb.stanford.edu/for-researchers/for-non-medical-researchers/guidance-specific-topics/data-scraping [accessed 2026-07-3]
10. Hair loss types: alopecia areata diagnosis and treatment. American Academy of Dermatology Association. URL: https://www.aad.org/public/diseases/hair-loss/types/alopecia/treatment [accessed 2026-05-22]
11. Gümüş S, Uzunçakmak TK, Aşkın Ö. Social media use in patients with alopecia areata. J Turk Acad Dermatol 2023 Sep 7;17(3):62-68. [doi: 10.4274/jtad.galenos.2023.83803]
12. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
13. Aldhouse NV, Kitchen H, Knight S, et al. "'You lose your hair, what's the big deal?' I was so embarrassed, I was so self-conscious, I was so depressed:" a qualitative interview study to understand the psychosocial burden of alopecia areata. J Patient Rep Outcomes 2020 Sep 11;4(1):76. [doi: 10.1186/s41687-020-00240-7] [Medline: 32914253]
14. Meah N, Wall D, York K, et al. The Alopecia Areata Consensus of Experts (ACE) study: results of an international expert opinion on treatments for alopecia areata. J Am Acad Dermatol 2020 Jul;83(1):123-130. [doi: 10.1016/j.jaad.2020.03.004] [Medline: 32165196]

XSL•FO
RenderX

Abbreviations

AA: alopecia areata

Edited by M Kirchberger; submitted 21.Dec.2025; peer-reviewed by D Chumachenko, SK Kurosky; revised version received 23.May.2026; accepted 31.May.2026; published 23.Jul.2026.

Please cite as:
Francis SN, Agarwal AA, Fanous AH, Siegel DH, Daneshjou R
Assessing Alopecia Areata Misinformation: Social Media Analysis
JMIR Dermatol 2026;9:e90088
URL: https://derma.jmir.org/2026/1/e90088
doi:10.2196/90088

© Shanika N Francis, Ank A Agarwal, Aaron H Fanous, Dawn H Siegel, Roxana Daneshjou. Originally published in JMIR Dermatology (http://derma.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Climate, Humidity, and Population-Level Interest in Dry Skin: Infodemiology Analysis Using Google Trends Across the United States

Kimiya Aframian[1*], BS; Shaya Naimi[2*], BS; Joy Xu[3], BS; Gordon Bae[1], MD

[1]Department of Dermatology, Stanford Medicine, Stanford, CA, United States

[2]Department of Sciences, CU Boulder Aerospace Engineering Sciences, Boulder, CO, United States

[3]Department of Medicine, David Geffen School of Medicine, 10833 Le Conte Ave, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Joy Xu, BS
Department of Medicine, David Geffen School of Medicine, 10833 Le Conte Ave, Los Angeles, CA, United States

Abstract

Background: Climate and weather factors of temperature and humidity are widely reported to be associated with xerosis (dry skin), a common inflammatory skin condition and frequent driver of pruritus (itchy skin) and reduced quality of life. Growing evidence supports links between environmental conditions and skin barrier function, with extreme climates associated with increased atopic dermatitis–related clinical visits. Mechanistically, temperature and humidity affect the stratum corneum, the skin's primary permeability barrier, with low humidity and high temperature increasing transepidermal water loss and promoting cutaneous inflammation.

Objective: This study examines the relationship between climate, namely temperature and humidity, and the general public's experience in dry skin and moisturizing products, throughout the United States. This study sought to address gaps in traditional epidemiologic approaches by linking climate conditions with population-level online search behavior related to dry skin and moisturizer use across the United States.

Methods: Publicly available climate data were obtained from the National Oceanic and Atmospheric Administration (NOAA), including average temperature and dew point by state over a recent nine-year period (2016 - 2025). Dew point served as a proxy for ambient humidity. Google Trends was used to assess relative search interest for five dry skin– and moisturizer-related terms by state during the same period. Search interest was normalized per million residents, and associations between climate variables and search interest were evaluated using linear regression analyses. Statistical analyses were conducted using R.

Results: Lower average temperatures and lower dew points were associated with higher dry skin–related search interest, while warmer, more humid states showed lower interest. Both temperature and dew point demonstrated significant negative associations with Google search interest. This work was not funded and data collection was performed using publicly available, free databases.

Conclusions: Population-level search behavior related to xerosis reflects national patterns of climate-associated dermatologic burden.

(JMIR Dermatol 2026;9:e93639) doi:10.2196/93639

KEYWORDS

xerosis; dry skin; humidity; dew point; temperature; google trends; infodemiology; United States

Introduction

Climate and weather factors of temperature and humidity are widely reported to be associated with xerosis (dry skin), one of the most common inflammatory skin conditions and frequent drivers of pruritus and reduced quality of life [1]. Although the mechanism is not fully understood, there is a widely growing body of evidence supporting meaningful links between environmental conditions and skin barrier function; regions characterized by extreme environmental conditions (temperature variability, precipitation patterns, sunlight exposure, and airborne pollutants) are associated with higher rates of atopic dermatitis (AD)–related clinical visits [2,3].

Mechanistically, temperature and humidity directly affect the stratum corneum, the skin's primary permeability barrier of highly organized extracellular lipids which limit transepidermal water loss (TEWL) [4]. Factors that increase evaporation from the stratum corneum, such as low humidity and high temperature, impair corneocyte cohesion and promote cutaneous inflammation. Prior studies demonstrate the impacts of dry

XSL•FO
RenderX

environment on increasing TEWL and impaired barrier repair, and high humidity on barrier restoration [5]. Such effects may be amplified in individuals predisposed to xerosis and eczematous dermatoses who already have baseline defects in barrier function [6]. Thus, “cold and dry” conditions may increase AD prevalence and flare risk, and cohort data suggests that climate factors may modulate symptom control in potentially region-specific methods [7].

Moisturizer is the mainstay management for xerosis and AD, so climate variability may plausibly be reflected in variation in skin-care needs on a regional-scale. Traditional epidemiologic analysis of the environment-skin relationship often relies on clinic visits, prescription, and registry records, however these methods do not represent subclinical symptoms and self-management behaviors outside of the formal healthcare environment [8]. Especially with the increase in accessibility to freely-available medical information −83.4% of patients in 2020 report using the Internet prior to their dermatologic appointment - it is critical to consider new metrics of analysis [9].

Google Trends is a freely accessible analytics platform which allows for assessment of Google search query volumes and has increasingly been used as an infodemiologic tool to character population-level health behaviors [10]. Search activity is reflective of self-triage and self-treatment efforts, both before and after clinical visits, offering a potential window into symptom burden captured beyond administrative data [11]. Specifically in dermatologic epidemiology, Google Trends data has been analyzed on a global scale to study temporal and geographic patterns of incidence, symptoms, and treatment of common conditions including eczema, acne, psoriasis, and skin cancer [12].

The aim of this paper is to understand the relationship between climate conditions and population-level interest in dry skin symptoms through linking of NOAA-reported temperature and humidity (dewpoint used as proxy) with Google Trends search interest across each US state for moisturizer-related and dry-skin symptom terms.

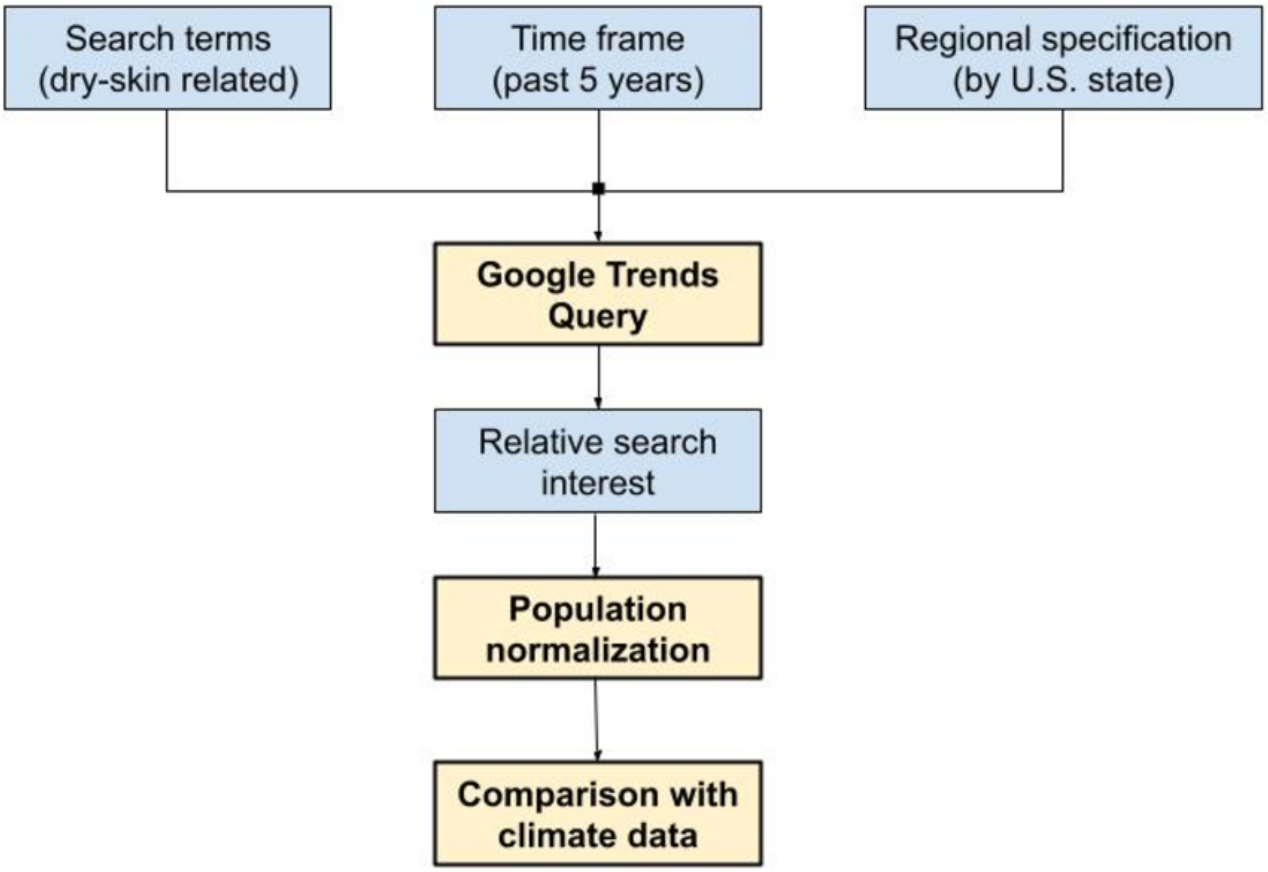
Methods

Google Trends Data

The Google Trends software allows the public to query the search interest of one or more particular search terms over a user-specified period of time and region, and allows the user to compare regions, time frames, and search terms (Figure 1). When queried, the software tool yields relative search interest between regions (eg, US states), normalized between 0 and 100 across the data set, as opposed to absolute search volume. According to this scale, a value of 0 indicates the lowest search popularity and a value of 100 indicates the highest relative search interest amongst the data set. To obtain data for this study, the Google Trends software tool was utilized to query search interest in five search terms related to dry skin and moisturizer: “eczema,” “dry skin,” “moisturizer,” “atopic dermatitis,” and “atopic eczema” over the same 9-year period specified above for climate data and by state. Relative search interest was normalized by each state’s population. Note that this additional step of population normalization is independent of the “normalized” data provided by each Google Trends query; Google Trends provides relative search interest normalized against the highest value in the dataset, which does not involve population. To normalize by state population, each state’s relative interest was divided by its population to obtain “relative search interest per million people” (Figure 1).

XSL•FO
RenderX

Figure 1. Outline of methodology. Google Trends query parameters were gathered, and the output of the query was normalized by state population prior to comparison against climate data.



Climate Data

Publicly available data were obtained from the National Oceanic and Atmospheric Administration (NOAA), a US government agency that studies the earth's oceans, atmosphere, and climate and publishes robust datasets available to the public. In particular, the NOAA offers the Climate Data Online (CDO), a web page that provides access to global historical weather and climate data, including daily, monthly, and hourly measurements of temperature, humidity, wind, and precipitation [13]. For this study, hourly measurements of average temperature and dew point by state were extracted over a recent 9-year period (2016 - 2025). Note that dew point is a widely known excellent and often preferred proxy for humidity, as it measures air moisture more directly and independently of temperature [14]. For each state, these measurements were averaged over the recent 9-year period to yield average temperature and dew point by state over the given time frame.

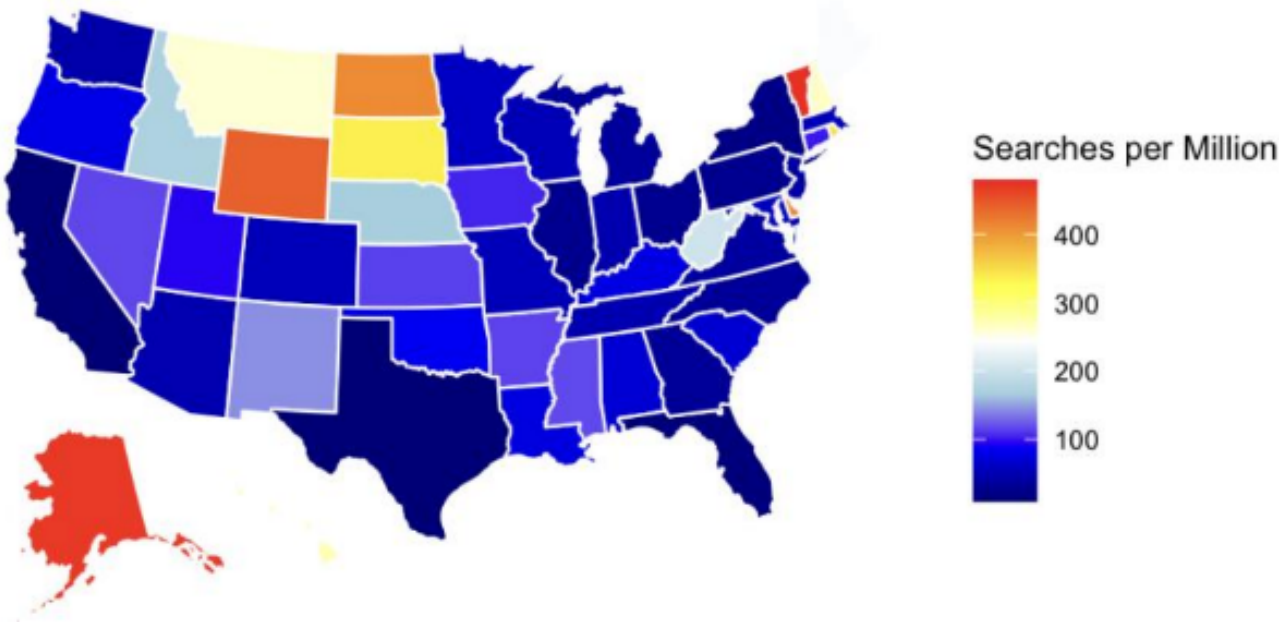
Results

Google Trends Data

Analysis of Google Trends data at the state level demonstrates substantial geographic variability in search interest for moisture-related and dry skin-associated search terms across the U.S. Higher relative search interest was observed in western and southwestern states, whereas lower search interest was observed in northeastern and midwestern states (Figure 2). Note that Google search interest was normalized per million residents, accounting for variability in population per state.

XSL•FO
RenderX

Figure 2. Google Trends Search Interest per capita by US state 2016 - 2025 period. This map was generated using data from the Google Trends software tool.



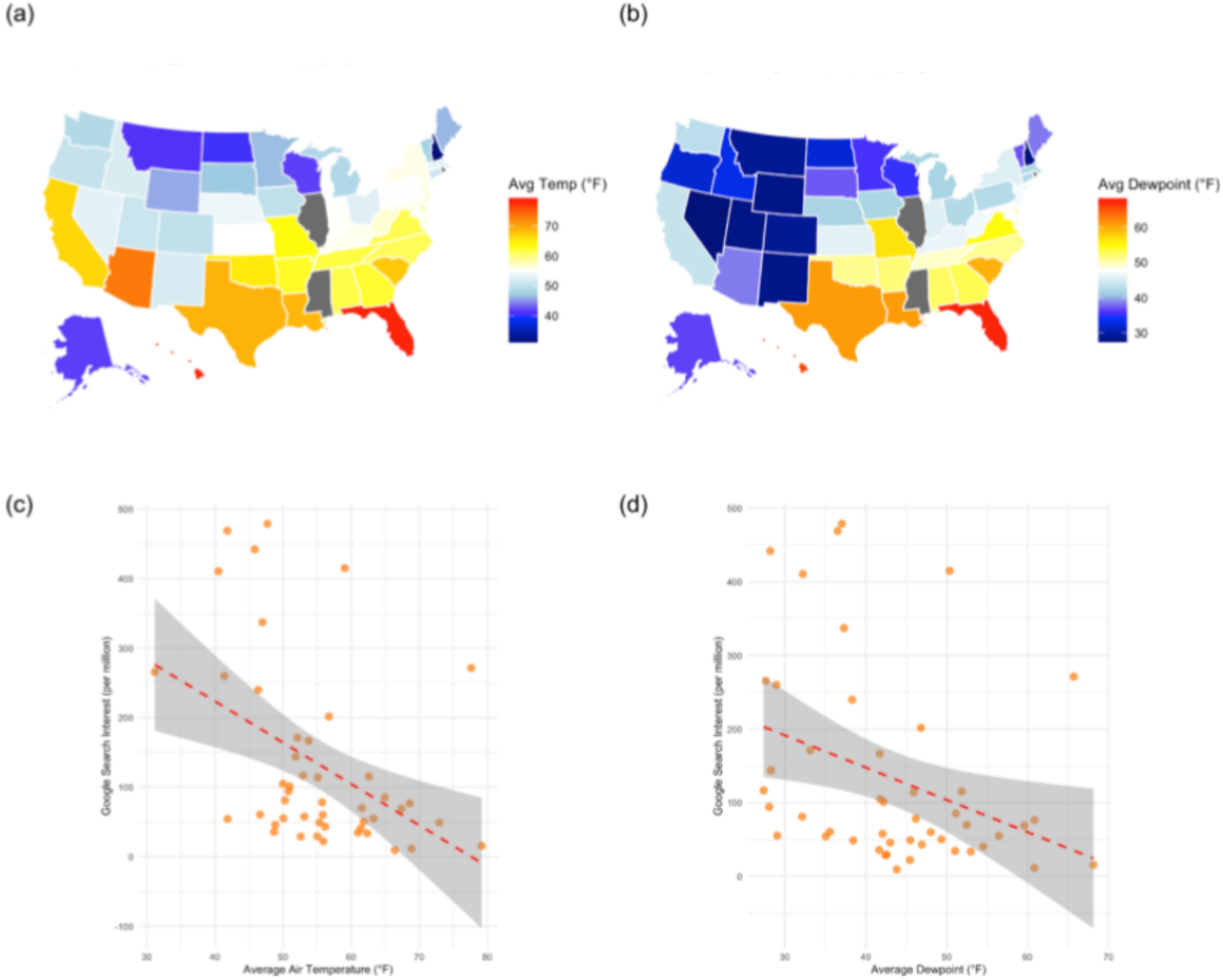
Climate Data

Comparison of the Google Trends data with NOAA-reported climate variables reveal that states such as Vermont and Wyoming, among those with the highest search interest, generally exhibited cooler average temperatures and moderate dew point values, although variation across states was observed. (Figure 3A and C). Examining temperature at the state level, Google search interest (mean 132.79, SD 132.99) was negatively correlated with Average Air Temperature (°F) (mean 55.35, SD 9.8), r(45) = −0.428, 95% CI [−0.637,−0.16], *P*=.002. Statistical analyses were conducted using R, and regression assumptions, including linearity and normality of residuals, were assessed using visual inspection of diagnostic plots.

Average dewpoint is used as a proxy for ambient humidity, showing a similar pattern as temperature on a national level. States with a lower average dewpoint and thus presumably decreased humidity conditions, especially western interior states, exhibited higher relative search interest for moisturizers and dry skin-related symptoms (Figure 3B). Contrarily, states with higher average dewpoints such as the humid southeastern states demonstrated lowest search interest in these search terms. Average dewpoints ranged from 28 degrees Fahrenheit minimum to 68 degrees Fahrenheit maximum. Similarly to temperature, at the state level, Google search interest (M=132.79, SD=132.99) was negatively correlated with Average Dewpoint (°F) (mean 43.5, SD 10.51), r(45) = −0.338, 95% CI [−0.57,−0.056], *P*=.02. (Figure 3D).

XSL•FO
RenderX

Figure 3. (A) Average air temperature by US state over 2016 - 2025 period. (B) Average dewpoint by US state over 2016 - 2025 period. (C) Plot illustrating the relationship between average dry bulb temperature (air temperature) and Google search popularity, averaged over the 2016 - 2025 period, where each point represents one US state. There is a significant downward trend with Pearson *r*=-0.442 and *P*=.002. (D) Plot illustrating the relationship between average dewpoint (proxy for humidity) and Google search popularity, averaged over the 2016 - 2025 period, where each point represents one US state. There is a significant downward trend with Pearson *r*=-0.348, and *P*=.02. All plots were generated using data from the National Oceanic and Atmospheric Administration.



Discussion

Principal Findings

This study found a significant negative correlation between average air temperature and Google search interest in dry skin-related terms per million residents, as well as a significant negative correlation between average dewpoint and Google search interest.

Regional humidity, rather than temperature alone, more closely aligns with patterns of dermatologic search interest according to geographic visualization of climate variables. Western inland states are characterized by lower average dewpoints and higher evaporation demand. These states consistently exhibit elevated Google Trends search interest in our queried terms. In contrast, coastal and southeastern states with high relative humidity and high dewpoints showed lower search interest despite the higher acreage annual temperatures. These findings suggest that atmospheric moisture content, represented by dew point, may be an important contributor to patterns of dermatologic search interest. While these findings demonstrate an association rather than causation, possible mechanisms underlying this relationship have been proposed. For example, temperature and humidity have been shown to influence skin barrier integrity, whereby low humidity may increase evaporation from the skin surface and contribute to skin dryness.

These findings cumulatively represent a robust association between climate conditions and Google search behavior related to moisturizer and dry skin symptoms across the United States. To our knowledge, this study represents an early effort to use Google Trends data as a digital epidemiologic tool to explore geographic variation in dry-skin related search activity across the United States. While search intent does not directly equate to disease prevalence, this approach provides a means of visualizing regional patterns, including clusters of elevated interest that are frequently observed in coastal areas. Google trends has also successfully predicted the outbreak of many viruses [15], indicating its validity as an epidemiological proxy.

XSL•FO
RenderX

Limitations

Limitations of this study include the difference in population per state. Normalization by population in search interest is reasonably assumed to be proportional to population, but states with extremely high or low populations may have resulted in skewed relative search interest per capita. Populations with limited internet access and healthcare utilization as well as lower socioeconomic status may be poorly represented. Demographic differences between states may also account for differences in data. Additionally, Google trends queries return relative search interest normalized between 0 - 100 across the data set instead of absolute search volumes, regardless of the data set being queried. Thus, reported findings are correlational and cannot be assumed as causation. In addition, analysis used multiyear averaged data, which may obscure important seasonal variation in skin dryness and related search behavior. These temporal patterns may not be fully captured in the present approach.

Conclusion

This study finds a strong correlation between NOAA reported temperature and humidity and state-specific Google searches, revealing that online search behavior reflects the general public's experience of the climate across the states. In other words, Americans who live in hotter and drier states are more likely to seek out moisturizer and to experience symptoms of dry skin conditions such as eczema and dermatitis. These findings reveal the potential for Google trends to monitor the general public's awareness of their dermatological response to varying climates, and can ultimately allow physicians to tailor preventative skin-care recommendations and emphasize proactive skin barrier protection in more predisposed environments. These findings highlight the potential for infodemiology tools to complement traditional dermatologic surveillance and may help guide future studies on environmental influences on skin disease.

Furthermore, physicians and medical systems at large could consider prioritization of access to moisturizers and barrier-rapid therapy in higher-risk regions with anticipated increased xerosis burden.

Funding

The authors declared no financial support was received for this work.

Data Availability

Data are available upon request.

Authors' Contributions

KRA & SN conducted manuscript-writing, JX provided critical revisions and edits, GHB oversaw conceptualization and guidance.

Conflicts of Interest

None declared.

References

1. Bylund S, Kobyletzki LB, Svalstedt M, Svensson Å. Prevalence and incidence of atopic dermatitis: a systematic review. Acta Derm Venereol 2020 Jun 9;100(12):adv00160. [doi: 10.2340/00015555-3510] [Medline: 32412646]
2. Hui-Beckman JW, Goleva E, Leung DYM, Kim BE. The impact of temperature on the skin barrier and atopic dermatitis. Ann Allergy Asthma Immunol 2023 Dec;131(6):713-719. [doi: 10.1016/j.anai.2023.08.007] [Medline: 37595740]
3. Park M, Mohsen ST, Katz T, et al. Climate conditions, weather changes, and air pollutants and atopic dermatitis: a meta-analysis. JAMA Dermatol 2025 Aug 1;161(8):828-839. [doi: 10.1001/jamadermatol.2025.1790] [Medline: 40560568]
4. Elias PM. Skin barrier function. Curr Allergy Asthma Rep 2008 Jul;8(4):299-305. [doi: 10.1007/s11882-008-0048-0] [Medline: 18606081]
5. Denda M, Sato J, Masuda Y, et al. Exposure to a dry environment enhances epidermal permeability barrier function. J Invest Dermatol 1998 Nov;111(5):858-863. [doi: 10.1046/j.1523-1747.1998.00333.x] [Medline: 9804350]
6. Cork MJ, Danby SG, Vasilopoulos Y, et al. Epidermal barrier dysfunction in atopic dermatitis. J Invest Dermatol 2009 Aug;129(8):1892-1908. [doi: 10.1038/jid.2009.133] [Medline: 19494826]
7. Engebretsen KA, Johansen JD, Kezic S, Linneberg A, Thyssen JP. The effect of environmental humidity and temperature on skin barrier function and dermatitis. J Eur Acad Dermatol Venereol 2016 Feb;30(2):223-249. [doi: 10.1111/jdv.13301] [Medline: 26449379]
8. Hamann CR, Andersen YMF, Engebretsen KA, et al. The effects of season and weather on healthcare utilization among patients with atopic dermatitis. J Eur Acad Dermatol Venereol 2018 Oct;32(10):1745-1753. [doi: 10.1111/jdv.15023] [Medline: 29706020]
9. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: Search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]

XSL•FO
RenderX

10. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: a scoping review. JMIR Public Health Surveill 2019;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]
11. Sivesind TE, Szeto MD, Kim W, Dellavalle RP. Google Trends in dermatology: scoping review of the literature. JMIR Dermatol 2021 May 25;4(1):e27712. [doi: 10.2196/27712] [Medline: 37632813]
12. Kamiński M, Tizek L, Zink A. Dr. Google, What Is That on My Skin?'-Internet Searches Related to Skin Problems: Google Trends Data from 2004 to 2019. Int J Environ Res Public Health 2021 Mar 4;18(5):2541. [doi: 10.3390/ijerph18052541] [Medline: 33806391]
13. NOAA. URL: https://www.ncei.noaa.gov/cdo-web/ [accessed 2026-03-25]
14. National Weather Service. Analysis information. URL: https://www.weather.gov/mfl/analysisinfo [accessed 2024-12-01]
15. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature New Biol 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]

Abbreviations

AD: atopic dermatitis
NOAA: National Oceanic and Atmospheric Administration
TEWL: transepidermal water loss

Edited by S Subhadarshani; submitted 16.Feb.2026; peer-reviewed by S Adnan; revised version received 13.Apr.2026; accepted 02.May.2026; published 25.May.2026.

Please cite as:
Aframian K, Naimi S, Xu J, Bae G
Climate, Humidity, and Population-Level Interest in Dry Skin: Infodemiology Analysis Using Google Trends Across the United States
JMIR Dermatol 2026;9:e93639
URL: https://derma.jmir.org/2026/1/e93639
doi:10.2196/93639

© Kimiya Aframian, Shaya Naimi, Joy Xu, Gordon Bae. Originally published in JMIR Dermatology (http://derma.jmir.org), 25.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Retraction

Retraction: Assessment and Evaluation of Social Engagement in Dermatology Residency Programs on Instagram: Cross-sectional Study

JMIR Publications Editorial Office[1]

JMIR Publications, Toronto, ON

Corresponding Author:
JMIR Publications
130 Queens Quay East, Suite 1100-1102
Toronto, ON
Phone: 1 4165832040
Email: ed-support@jmir.org

Related Article:

Retraction of: https://derma.jmir.org/2021/2/e32105

(JMIR Dermatol 2026;9:e106666) doi:10.2196/106666

The JMIR Publications Editorial Office is retracting the article "Assessment and Evaluation of Social Engagement in Dermatology Residency Programs on Instagram: Cross-sectional Study" [1] due to concerns regarding the integrity of the peer review process.

Author Chapman Wei responded, stating that there were no conflicts of interest with the reviewers. However, the Editorial Office has determined that there is sufficient evidence indicating that the peer review process was compromised to proceed with a retraction.

Chapman Wei and Nagasai Adusumilli responded but did not explicitly state whether they agreed or disagreed with the decision to retract the article. The remaining authors were contacted but did not respond to communications.

Reference

1. Wei C, Bernstein S, Adusumilli N, Marchitto M, Chen F, Rajpara A. Assessment and evaluation of social engagement in dermatology residency programs on Instagram: cross-sectional study. JMIR Dermatol 2021 Aug 26;4(2):e32105 [FREE Full text] [doi: 10.2196/32105] [Medline: 37632856]

Submitted 09.Jul.2026; this is a non–peer-reviewed article;accepted 13.Jul.2026; published 19.Aug.2026.

Please cite as:
JMIR Publications Editorial Office
Retraction: Assessment and Evaluation of Social Engagement in Dermatology Residency Programs on Instagram: Cross-sectional Study
JMIR Dermatol 2026;9:e106666
URL: https://derma.jmir.org/2026/1/e106666
doi:10.2196/106666
PMID:

©JMIR Publications Editorial Office. Originally published in JMIR Dermatology (http://derma.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX