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 8 (2025) ISSN 2562-0959 Editor in Chief: Robert Dellavalle, MD, PhD, MSPH

Contents

Original Papers

Research Letters

Reviews

Letters to the Editor

Case Reports

Short Paper

Viewpoints

Exploring Attention-Deficit/Hyperactivity Disorder Symptoms in Patients With Atopic Dermatitis by Disease Severity: Cross-Sectional Analysis

Amr Molla[1,2], MD; Raed Jannadi[3], MD, PhD; Dareen Hafez[4], MD; Lujain Alrohaily[5], MD; Ebtesam Abdullah[6], MD; Duha Azouni[6], MD; Muayad Albadrani[3], MD, PhD

[1]Department of Medicine, College of Medicine, Taibah University, P.O.Box 344, Unit 99, Madinah, Saudi Arabia

[2]Department of Dermatology, Saudi German Hospital, Madinah, Saudi Arabia

[3]Department of Family and Community Medicine and Medical Education, College of Medicine, Taibah University, Madinah, Saudi Arabia

[4]Department of Child & Adolescent Psychiatry, King Salman Medical City, Madinah, Saudi Arabia

[5]Department of Pediatric, King Salman Medical City, Madinah, Saudi Arabia

[6]Department of Medicine, King Fahad Hospital, Madinah, Saudi Arabia

Corresponding Author:
Amr Molla, MD
Department of Medicine, College of Medicine, Taibah University, P.O.Box 344, Unit 99, Madinah, Saudi Arabia

Abstract

Background: Atopic dermatitis (AD) is a chronic inflammatory skin condition affecting a significant percentage of the global population. Emerging research suggests a potential link between AD and neurodevelopmental disorders like attention-deficit/hyperactivity disorder (ADHD). However, there is a lack of comprehensive studies within the Saudi Arabian population examining this association.

Objective: This study aims to determine the prevalence of ADHD among patients with AD in Saudi Arabia and to explore potential associations with demographic and clinical factors.

Methods: In this cross-sectional, multicenter study conducted between May and November 2024, 419 patients with AD were recruited from various hospitals in Saudi Arabia. Children were screened for ADHD symptoms using the ADHD Rating Scale-5, while adults were assessed with the Adult Self-Report Scale. Logistic regression was used to evaluate the influence of AD severity, age, gender, nationality, and BMI on the likelihood of ADHD symptoms.

Results: A total of 419 patients with AD were included, of whom 234 (55.8%) were children and 185 (44.2%) were adults; 239 (57%) were female and 360 (85.9%) were Saudi nationals. ADHD symptoms were identified in 84 (20%) patients, with a slightly higher prevalence among children (49/234, 20.9%) compared to adults (35/185, 18.9%; *P*=.61). No significant associations were found between ADHD symptoms and gender, nationality, BMI, or AD severity in either age group. Moderate to severe AD was more common among adults (48/185, 25.9%) than children (42/234, 17.9%; *P*=.048).

Conclusions: This study found that 20% of patients with AD screened positive for ADHD symptoms, with slightly higher rates in children than adults. No significant associations were observed between ADHD symptoms and gender, nationality, BMI, or AD severity. Although no significant clinical predictors were identified, the findings emphasize the need for ADHD screening in patients with AD, particularly in regions with high AD prevalence. Future longitudinal studies should explore underlying mechanisms and assess how managing one condition may influence the other.

(JMIR Dermatol 2025;8:e74126) doi:10.2196/74126

KEYWORDS

atopic dermatitis; ADHD; cross-sectional study; neurodevelopmental disorders; attention-deficit/hyperactivity disorder

Introduction

Atopic dermatitis (AD) is a chronic, relapsing inflammatory condition of multifactorial origin. It primarily affects infants and children but can persist into adulthood. AD is characterized by chronic itching, following a cycle of flare-ups and remission, typically intensifying at night and showing distinct morphology and distribution based on the patient's age [1]. Globally, AD affects approximately 5% of the population [2], whereas in Saudi Arabia, it impacts around 13% of individuals [3].

Patients with AD often have a family history of other atopic conditions, including food allergies, asthma, allergic rhinitis, and allergic conjunctivitis. Clinically, AD presents as excoriated, scaly, eczematous papules and plaques, which may become

overlaid with bacterial infections, predominantly *Staphylococcus aureus*, leading to yellow crusting and exacerbating the condition [1].

Patients with AD often have comorbid atopic conditions, including asthma and allergic rhinitis, and frequently experience disrupted sleep and psychological stress, which can impact cognitive and emotional regulation [1]. Sleep disturbances due to persistent nocturnal pruritus are particularly problematic and may impair concentration, mood, and behavior, potentially mimicking or exacerbating neurodevelopmental symptoms [4]. Furthermore, the chronic nature of AD can negatively affect quality of life and increase the risk of psychiatric comorbidities, including symptoms consistent with attention-deficit/hyperactivity disorder (ADHD) [5,6].

Recent research has identified a significant link between AD and an elevated risk of ADHD, along with other mental comorbidities and reduced quality of life [5,7]. However, no sufficient studies correlate this relation among the Saudi population.

ADHD is a neurodevelopmental disorder affecting both children and adults, characterized by inattention, hyperactivity, and impulsivity, which can disrupt daily functioning in areas like school, work, and social relationships [6]. In Saudi Arabia, ADHD affects 12.4% of children and around 4% of adults [8]. It is more prevalent in men and presents in three types: inattentive, hyperactive-impulsive, and combined. The etiology of ADHD is multifactorial, involving genetic, environmental, and developmental factors such as maternal smoking, socioeconomic status, and perinatal complications [9].

According to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision criteria, ADHD symptoms must appear before age 12, persist in multiple settings, and cause functional impairment. Diagnosis relies on clinical assessment and validated psychometric tools [6,9]. However, ADHD management depends on symptom severity. Mild cases often require psychotherapy, while moderate to severe cases are managed with both pharmacotherapy and psychotherapy. Pharmacotherapy, typically using stimulant and nonstimulant medications, is the first-line treatment [9].

Despite evidence linking AD and ADHD, data in the Saudi population are limited. This study aims to assess the prevalence of ADHD symptoms among adult and pediatric patients with AD in Saudi Arabia and to examine whether AD severity is associated with differences in ADHD symptoms.

Methods

Study Setting and Sampling

This cross-sectional, multicenter study aimed to investigate the relationship between ADHD and AD among the Saudi population. A total of 419 adult and child patients with AD were selected to participate in the study by simple random sampling. This research was conducted from May 2024 to November 2024 at participating hospitals across Saudi Arabia. The inclusion criteria include Saudi and non-Saudi male and female patients of various ages diagnosed by a dermatologist with AD, who presented at the participating hospitals and agreed to participate in the study. The exclusion criteria include children younger than 4 years of age, patients whose ADHD symptoms occur exclusively during the course of schizophrenia or another psychotic disorder, and patients with ADHD symptoms that can be better explained by other mental disorders, such as mood disorders. These criteria are outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, making accurate screening in younger children clinically inappropriate.

Data Collection

Patients with AD with a known diagnosis by a dermatologist, as confirmed from their medical records, were included in the study. Participants were prospectively interviewed and evaluated by a psychiatrist to diagnose ADHD. The ADHD Rating Scale-5 (ADHD-RS-5)-Short Form was used for children, while the Adult Self-Report Scale (ASRS) screening tool was applied for adults. Additional data, including BMI and nationality, were also extracted from medical records to complement the ADHD assessment data. The ADHD screening assessments were administered in both Arabic and English, depending on participant preference and language proficiency. Validated Arabic versions of the ADHD-RS-5 and ASRS tools were used where appropriate.

Study Variables

AD Severity

Diagnoses of AD were confirmed based on historical medical records where the Hanifin and Rajka criteria, one of the earliest and most widely used standards for diagnosing AD, were used by dermatologists [10]. Additionally, these records included Scoring AD (SCORAD) assessments in English, a validated and reliable measure for evaluating AD severity. SCORAD criteria account for extent, intensity, and subjective symptoms, with total scores ranging from 0 to 103. Based on these scores, AD severity is classified into three categories: mild (SCORAD<25), moderate (25≤SCORAD≤50), and severe (SCORAD>50) [11].

ADHD

The ADHD-RS-5-Short Form is a simplified screening tool to identify the likelihood of ADHD in children based on parent or teacher observations [12].

Structure of the ADHD-RS-5-Short Form

The tool consists of 9 items, each representing a core ADHD behavior. Each behavior is scored on a 4-point scale based on frequency: 0=Never or Rarely, 1=Sometimes, 2=Often, and 3=Very Often.

Scoring and Interpretation

The sum of the scores from these 9 items ranges from 0 to 27. A total score of 15 or higher typically indicates a high likelihood of ADHD, suggesting further assessment if other indicators are present. This scoring method serves as an initial screening, with a high score warranting a comprehensive evaluation if ADHD is suspected.

For adults, part A of the Adult Self-Report Scale (ASRS-v1.1) was used, which comprises 6 questions. The first 4 questions

assess inattentive (ADHD-I) symptoms, and the last 2 assess hyperactive (ADHD-H) symptoms. Responses range from 0=never to 4=very often. Patients who selected shaded boxes for 4 or more questions were considered to have symptoms consistent with ADHD, indicating a positive screening result [13]. The ASRS-v1.1 has demonstrated strong diagnostic performance, with a reported sensitivity of 91.4% and specificity of 96% for identifying adult ADHD cases, making it a reliable screening tool in both clinical and research settings [14].

Covariates

The covariates in this study included gender, age, nationality, and BMI. Gender was categorized as male or female. Age groups were defined as children (4 - 17 y) and adults (18 y and older). Nationality was classified as either Saudi or non-Saudi; due to the small number and heterogeneity of non-Saudi participants, they were grouped together for analysis. BMI was classified into three groups: normal weight (18.5 - 24.9 for adults; 15.5 - 21.9 for children), underweight (less than 18.5 for adults; less than 15.5 for children), and overweight or obese (25 or more for adults; 22 or above for children).

Data Analysis

Data were analyzed using SPSS software (version 29; IBM Corp). Bivariate analysis was conducted by applying the chi-square test to examine associations between categorical variables. Logistic regression analysis was used for multiple variable analysis to assess the influence of gender, age, nationality, BMI, and AD severity on ADHD occurrence. A *P* value of less than .05 was considered statistically significant.

Ethical Considerations

Ethical approval for this study was obtained from the Institutional Review Board of Taibah University, Saudi Arabia (TU-24-014), on May 15, 2024. All procedures were conducted in accordance with the Declaration of Helsinki and relevant national regulations. Written informed consent was obtained from all adult participants. For participants younger than 18 years of age, written informed consent was obtained from their parents or legal guardians. The study objectives, procedures, data privacy protections, and estimated duration were clearly explained to all participants prior to enrollment. Participation was entirely voluntary, and participants were informed of their right to withdraw at any time without penalty. All collected data were anonymized to ensure confidentiality and stored securely with restricted access. No financial compensation was provided to participants for their involvement in this study.

Results

Demographic details

The study included 419 patients with AD, of whom 234 (55.8%) were children and 185 (44.2%) were adults. Among all participants, 239 out of 419 (57%) were female and 180 (43%) were male, with no significant gender distribution difference across age groups (*P*=.27). Most participants were Saudi nationals (360/419, 85.9%), while 59 (14.1%) were non-Saudi (*P*=.16). Regarding BMI, 21 (5%) were underweight, 304 (72.6%) had normal weight, and 95 (22.4%) were overweight or obese. These proportions were similar between children and adults (*P*=.37).

ADHD symptoms were likely in 84 out of 419 (20%) patients, with a slightly higher occurrence among children (49/234, 20.9%) than adults (35/185; 18.9%; *P*=.61). ADHD was likely in 84 out of 419 (20%) patients, with a slightly higher occurrence among children (49/234, 20.9%) than adults (35/185, 18.9%; *P*=.61). AD severity also differed significantly by age group. Among children, 192 (82.1%) had mild AD and 42 (17.9%) had moderate to severe AD, while among adults, 137 (74.1%) had mild AD and 48 (25.9%) had moderate to severe AD (*P*=.048; Table 1). While AD severity differences are reported here, they are included to provide clinical context for interpreting ADHD symptom associations, rather than as a primary focus of the study (Table 1).

Table . Descriptive characteristics of patients with AD[a] by age groups.

Variables	Age groups			*P* value[b]
	Study population (N=419), n (%)	Children (n=234), n (%)	Adults (n=185), n (%)	
Sex				.27
Male	180 (43)	95 (40.6)	85 (45.9)	
Female	239 (57)	139 (59.4)	100 (54.1)	
Nationality				.16
Saudi	360 (85.9)	206 (88)	154 (83.2)	
Non-Saudi	59 (14.1)	28 (12)	31 (16.8)	
BMI				.37
Underweight	21 (5)	10 (4.3)	11 (5.9)	
Normal	304 (72.6)	176 (75.2)	128 (69.2)	
Overweight or obese	95 (22.4)	48 (2.5)	46 (24.9)	
ADHD[c]				.61
Likely	84 (20)	49 (20.9)	35 (18.9)	
Unlikely	335 (80)	185 (79.1)	150 (81.1)	
AD severity				.048
Mild	329 (78.5)	192 (82.1)	137 (74.1)	
Moderate to severe	90 (21.5)	42 (17.9)	48 (25.9)	

[a]AD: atopic dermatitis.

[b]*P* values were calculated using the $\chi 2$ test.

[c]ADHD: attention-deficit/hyperactivity disorder.

Association Between ADHD and Demographic and Clinical Variables in Children With AD

Table 2 shows associations between ADHD and study variables among children with AD. Among 95 male children, 23 (24%) were likely to have ADHD, while among 139 female children, 26 (19%) likely had ADHD (*P*=.31). Among Saudi children (n=206), 47 (22.8%) were likely ADHD cases, compared to 2 out of 28 (7%) non-Saudi children (*P*=.06). Regarding BMI, 1 out of 10 (10%) underweight children was likely to have ADHD, compared to 37 out of 176 (21%) normal-weight children and 11 out of 48 (22.9%) overweight or obese children (*P*=.66). Among 192 children with mild AD, 39 (20.3%) were likely ADHD cases, compared to 10 out of 42 (23.8%) children with moderate to severe AD (*P*=.61).

Table . Association between ADHD[a] and study variables among children.

Variables	ADHD			*P* value[b]
	Overall children (n=234), n (%)	Likely (n=49), n (%)	Unlikely (n=185), n (%)	
Gender				.31
Male	95 (40.6)	23 (24.2)	72 (75.8)	
Female	139 (59.4)	26 (18.7)	113 (81.3)	
Nationality				.06
Saudi	206 (88)	47 (22.8)	159 (77.2)	
Non-Saudi	28 (12)	2 (7.1)	26 (92.9)	
BMI				.66
Underweight	10 (4.3)	1 (10)	9 (90)	
Normal	176 (75.2)	37 (21)	139 (79)	
Overweight or obese	48 (20.5)	11 (22.9)	37 (77.1)	
AD[c] severity				.61
Mild	192 (82.1)	39 (20.3)	153 (78.7)	
Moderate to severe	42 (17.9)	10 (23.8)	32 (76.2)	

[a]ADHD: attention-deficit/hyperactivity disorder.

[b]*P* value calculated by using the $\chi 2$ test.

[c]AD: atopic dermatitis.

Association Between ADHD and Demographic and Clinical Variables in Adults With AD

Table 3 presents associations between ADHD and study variables among adult patients with AD. Among 85 male adults with AD, 16 (18.8%) were likely ADHD cases, compared to 19 out of 100 (19%) female adults (*P*=.89). Among 154 Saudi adults, 33 (21.4%) were likely ADHD cases, while only 2 out of 31 (6.5%) non-Saudi adults were likely ADHD cases (*P*=.05). Regarding BMI, 4 out of 11 (36.4%) underweight adults were likely to have ADHD, compared to 25 out of 128 (19.5%) normal-weight adults and 6 out of 46 (13%) overweight or obese adults (*P*=.20). Among 137 adults with mild AD, 29 (21.2%) were likely ADHD cases, while 6 out of 48 (12.5%) adults with moderate to severe AD were likely ADHD cases (*P*=.19).

Table . Association between ADHD[a] and study variables among adults.

Variables	ADHD			*P* value[b]
	Overall adults (n=185), n (%)	Likely (n=35), n (%)	Unlikely (n=150), n (%)	
Sex				.89
Male	85 (45.9)	16 (18.8)	69 (81.2)	
Female	100 (54.1)	19 (19)	81 (81)	
Nationality				.05
Saudi	154 (83.2)	33 (21.4)	121 (78.6)	
Non-Saudi	31 (16.8)	2 (6.5)	29 (93.5)	
BMI				.20
Underweight	11 (5.9)	4 (36.4)	7 (63.6)	
Normal	128 (69.2)	25 (19.5)	103 (80.5)	
Overweight or obese	46 (24.9)	6 (13)	40 (87)	
AD[c] severity				.19
Mild	137 (74.1)	29 (21.2)	108 (78.8)	
Moderate to severe	48 (25.9)	6 (12.5)	42 (87.5)	

[a]ADHD: attention-deficit/hyperactivity disorder.

[b]*P* value calculated by using the $\chi 2$ test.

[c]AD: atopic dermatitis.

Logistic Regression Analysis of ADHD Risk Factors in Children With AD

A logistic regression model was applied to examine the influence of gender, nationality, BMI, and AD severity on the likelihood of ADHD symptoms development among children with AD. Male children with AD were 32% more likely to develop ADHD compared to female children (odds ratio [OR] 1.32, 95% CI 0.69 - 2.54; *P*=.40). Saudi children were 3.45 times more likely than non-Saudi children with AD to develop ADHD (OR 3.45, 95% CI 0.76 - 15.7; *P*=.11). Compared to children with normal BMI, underweight children had a lower probability of developing ADHD symptoms (OR 0.53, 95% CI 0.06 - 5.00), while overweight or obese children had a higher risk of ADHD diagnosis (OR 1.25, 95% CI 0.57 - 2.74; *P*=.58). Furthermore, children with moderate to severe AD were more likely to have ADHD symptoms (OR 1.40, 95% CI 0.6 - 3.25) compared to those with mild AD (*P*=.44; Table 4).

Table . Logistic regression analysis of study variables influencing ADHD[a] onset among childhood patients with AD[b].

Variables	aOR[c] (95% CI)[d]	*P* value
Sex		
Female	1 (reference)	
Male	1.32 (0.69 - 2.54)	.40
Nationality		
Non-Saudi	1 (reference)	
Saudi	3.45 (0.76 - 15.7)	.11
BMI		
Normal	1 (reference)	
Underweight	0.53 (0.06 - 5)	.58
Overweight or obese	1.25 (0.57 - 2.74)	.58
AD severity		
Mild	1 (reference)	
Moderate to severe	1.4 (0.6 - 3.25)	.44

[a]ADHD: attention-deficit/hyperactivity disorder.

[b]AD: atopic dermatitis.

[c]aOR: adjusted odds ratio.

[d]aOR was calculated by including age, gender, nationality, BMI, and AD severity.

Logistic Regression Analysis of ADHD Risk Factors in Adults With AD

Table 5 assesses the effects of gender, nationality, BMI, and AD severity on the development of ADHD symptoms among adult patients with AD. Male and female adults had nearly the same probability of ADHD symptoms (OR 0.97, 95% CI 0.45 - 2.10; *P*=.94). Saudi adults with AD were 3.96 times more likely to have ADHD than non-Saudi adults (OR 3.96, 95% CI 0.89 - 17.7; *P*=.07). Compared to adult patients with AD with normal weight, underweight adults with AD had a higher chance of developing ADHD symptoms (OR 2.40, 95% CI 0.63 - 9.17), while overweight or obese adults had a lower risk (OR 0.58, 95% CI 0.22 - 1.56; *P*=.20 and *P*=.28, respectively). Moreover, adults with moderate to severe AD had a lower probability of having ADHD symptoms (OR 0.54, 95% CI 0.2 - 1.43) compared to those with mild AD (*P*=.21). However, none of the examined predictors showed a significant association with ADHD symptoms development among children and adult patients with AD.

Table . Logistic regression analysis of study variables influencing ADHD[a] onset among adult patients with AD[b].

Variables	aOR[c] (95% CI)[d]	*P* value
Sex		
Female	1 (reference)	
Male	0.97 (0.45 - 2.1)	.94
Nationality		
Non-Saudi	1 (reference)	
Saudi	3.96 (0.89 - 17.7)	.07
BMI		
Normal	1 (reference)	
Underweight	2.4 (0.63 - 9.17)	.20
Overweight or obese	0.58 (0.22 - 1.56)	.28
AD severity		
Mild	1 (reference)	
Moderate to severe	0.54 (0.2 - 1.43)	.21

[a]ADHD: attention-deficit/hyperactivity disorder.

[b]AD: atopic dermatitis.

[c]aOR: adjusted odds ratio.

[d]aOR was calculated by including age, gender, nationality, BMI, and AD severity.

Discussion

Principal Findings

This study aimed to investigate the association between AD and ADHD in the Saudi population. The findings revealed that 20% of patients with AD were likely to have ADHD symptoms, with slightly higher proportions among children (49/234, 20.9%) compared to adults (35/185, 18.9%). Notably, AD severity varied significantly between age groups, with moderate to severe AD observed more frequently in adults (48/185, 25.9%) than in children (42/234, 17.9%). However, logistic regression analyses showed no statistically significant associations between ADHD and gender, nationality, BMI, or AD severity in either age group. These findings align with the study objectives to explore the strength of association between AD and ADHD symptoms while accounting for patient demographics and clinical factors.

This study contributes novel data from the Saudi population, where research on the co-occurrence of AD and ADHD is currently limited. The inclusion of a large, diverse sample and the use of validated tools for assessing both AD severity (SCORAD) and ADHD symptoms (ADHD-RS-5 and ASRS) enhances the reliability of the findings. Additionally, the multicenter design strengthens the study's representativeness and generalizability within Saudi Arabia.

The potential link between AD and ADHD symptoms has been increasingly explored in recent studies, with multiple hypotheses proposed to explain this association. One leading theory suggests that systemic inflammation plays a pivotal role in both conditions. Chronic immune dysregulation in AD, particularly the overactivation of Th2-mediated pathways and elevated levels of cytokines such as IL-4, IL-13, and TNF-α, has been implicated in neuroinflammation, which may contribute to ADHD pathogenesis. Studies have shown that children with AD exhibit higher levels of circulating inflammatory markers, which could influence neurodevelopmental processes and neurotransmitter regulation, particularly in dopaminergic pathways linked to ADHD symptoms [15,16]. Additionally, dysregulated sleep patterns, commonly observed in AD due to persistent pruritus, could further impact cognitive function, emotional regulation, and attentional control, exacerbating ADHD symptoms [4].

Another plausible explanation involves the gut-brain-skin axis, which has gained increasing attention in recent research. Alterations in the gut microbiome composition have been linked to both AD and ADHD, suggesting that microbial dysbiosis may serve as a shared pathophysiological factor. Studies have reported reduced microbial diversity and an imbalance in short-chain fatty acid-producing bacteria in both AD and ADHD populations, which may lead to increased intestinal permeability, systemic inflammation, and neurodevelopmental disturbances [17-19].

The results of this study are consistent with findings from previous research, including a US-based study that identified a higher prevalence of ADHD among children with eczema, particularly those with moderate to severe cases [20]. Another study found that ADHD symptoms were more common among children with eczema, highlighting potential shared inflammatory or neurological pathways [21]. Similarly, a German study emphasized the association between moderate to severe AD and mental health issues, although it focused primarily on adolescents [22]. Finally, a Mendelian randomization analysis suggested potential bidirectional causal relationships between AD and psychiatric disorders, further

supporting the complex interplay between these conditions [23]. These comparisons validate the relevance of our findings while highlighting differences that may be attributed to cultural, genetic, or methodological factors.

While this study provides valuable insights into the Saudi population, its findings may have limited generalizability to other settings due to cultural, environmental, and genetic differences. The notably higher prevalence of AD in Saudi Arabia (13%) compared to the global average (5%) may reflect unique population-specific risk factors [3]. Additionally, the absence of a control group without AD limits our ability to directly compare the prevalence of ADHD symptoms between patients with AD and the general population. The use of validated screening tools rather than formal clinical diagnostic assessments may also result in over- or underestimation of true ADHD prevalence. Furthermore, several key sociodemographic variables—such as income level, employment status, and psychological well-being—were not collected. Their omission limits our ability to examine broader psychosocial influences on ADHD symptom expression in patients with AD. These factors should be addressed in future studies to strengthen interpretability and generalizability.

Future research should focus on prospective longitudinal studies to better establish causal relationships between AD and ADHD, examining how AD progression or treatment interventions influence ADHD symptoms over time. Exploring genetic markers, inflammatory cytokines, and microbiome alterations in affected individuals may also provide deeper mechanistic insights. Furthermore, studies assessing the impact of AD treatment on ADHD symptomatology could help inform integrated management approaches for patients with coexisting dermatological and neurodevelopmental conditions.

Conclusions

This cross-sectional study provides important insights into the potential association between AD and ADHD in the Saudi population. Our findings indicate that 20% of patients with AD are likely to have ADHD, with a slightly higher prevalence among children than adults. While no significant associations were found between ADHD and demographic or clinical factors such as gender, nationality, BMI, or AD severity, the study underscores the need for heightened clinical awareness of ADHD symptoms in patients with AD. Given that AD prevalence in Saudi Arabia exceeds global averages, these findings highlight the importance of integrating neurodevelopmental screening into dermatological care.

The growing evidence linking AD to neuropsychiatric disorders suggests shared inflammatory, neuroimmune, and sleep-related mechanisms that warrant further exploration. Future longitudinal studies should aim to establish causality, assess the impact of AD treatments on ADHD symptoms, and investigate potential biomarkers that may mediate this relationship. Understanding these interactions could pave the way for personalized treatment strategies that optimize both dermatologic and neurodevelopmental outcomes, ultimately improving the quality of life for affected individuals.

Acknowledgments

The authors declare that they received no funding for the research. We extend our sincere gratitude to the medical and administrative teams of Saudi German and Almiqat Hospitals in Madinah for their invaluable support and facilities. Special thanks to Dr. Omer Al-Nozha for contributions to data collection and committee approval, as well as to the nursing staff, support staff, and patients' families for their cooperation, which was vital to the success of this research. This study was conducted without external funding. The authors did not receive financial support or grants for this study.

Data Availability

The data that support the findings of this research are not publicly available due to legal and ethical considerations but are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Wolff K, Goldsmith LA, Katz SI, Gilchrest BA, Paller AS, Leffell DJ. Fitzpatrick’s Dermatology in General Medicine: McGraw-Hill New York; 2008.
2. Raimondo A, Lembo S. Atopic dermatitis: epidemiology and clinical phenotypes. Dermatol Pract Concept 2021 Oct;11(4):e2021146. [doi: 10.5826/dpc.1104a146] [Medline: 35024238]

3. Fatani MI, Al Sheikh AA, Alajlan MA, et al. National Saudi consensus statement on the management of atopic dermatitis (2021). Dermatol Ther (Heidelb) 2022 Jul;12(7):1551-1575. [doi: 10.1007/s13555-022-00762-6] [Medline: 35788543]
4. Bawany F, Northcott CA, Beck LA, Pigeon WR. Sleep disturbances and atopic dermatitis: relationships, methods for assessment, and therapies. J Allergy Clin Immunol Pract 2021 Apr;9(4):1488-1500. [doi: 10.1016/j.jaip.2020.12.007] [Medline: 33321263]
5. Riis JL, Vestergaard C, Deleuran MS, Olsen M. Childhood atopic dermatitis and risk of attention deficit/hyperactivity disorder: a cohort study. J Allergy Clin Immunol 2016 Aug;138(2):608-610. [doi: 10.1016/j.jaci.2016.01.027] [Medline: 27025346]
6. Dulcan MK. Dulcan's Textbook of Child and Adolescent Psychiatry: American Psychiatric Pub; 2021.
7. Horev A, Freud T, Manor I, Cohen AD, Zvulunov A. Risk of attention-deficit/hyperactivity disorder in children with atopic dermatitis. Acta Dermatovenerol Croat 2017 Oct;25(3):210-214. [Medline: 29252173]
8. Aljadani AH, Alshammari TS, Sadaqir RI, et al. Prevalence and risk factors of attention deficit-hyperactivity disorder in the Saudi population: a systematic review and meta-analysis. Saudi J Med Med Sci 2023;11(2):126-134. [doi: 10.4103/sjmms.sjmms_528_22] [Medline: 37252016]
9. American Psychiatric Association. In: Gałecki P, S´wie˛cicki Ł, editors. Desk Reference to the Diagnostic Criteria from DSM-5: Kryteria diagnostyczne z DSM-5, Polish edition; 2013.
10. Joji T. Diagnostic standard for atopic dermatitis. JMAJ 2002;45(11):460-465 [FREE Full text]
11. Oranje AP, Glazenburg EJ, Wolkerstorfer A, de Waard-van der Spek FB. Practical issues on interpretation of scoring atopic dermatitis: the SCORAD index, objective SCORAD and the three-item severity score. Br J Dermatol 2007 Oct;157(4):645-648. [doi: 10.1111/j.1365-2133.2007.08112.x] [Medline: 17714568]
12. Vallejo-Valdivielso M, de Castro-Manglano P, Vidal-Adroher C, Díez-Suárez A, Soutullo CA. Development of a Short Version of the ADHD Rating Scale-IV.es (sADHD-RS-IV.es). J Atten Disord 2024 Mar;28(5):600-607. [doi: 10.1177/10870547241232314] [Medline: 38353419]
13. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. Psychol Med 2005 Feb;35(2):245-256. [doi: 10.1017/s0033291704002892] [Medline: 15841682]
14. Green JG, DeYoung G, Wogan ME, Wolf EJ, Lane KL, Adler LA. Evidence for the reliability and preliminary validity of the Adult ADHD Self-Report Scale v1.1 (ASRS v1.1) screener in an adolescent community sample. Int J Methods Psychiatr Res 2019 Mar;28(1):e1751. [doi: 10.1002/mpr.1751] [Medline: 30407687]
15. Chang SJ, Kuo HC, Chou WJ, Tsai CS, Lee SY, Wang LJ. Cytokine levels and neuropsychological function among patients with attention-deficit/hyperactivity disorder and atopic diseases. J Pers Med 2022 Jul 17;12(7):35887652. [doi: 10.3390/jpm12071155] [Medline: 35887652]
16. Fania L, Moretta G, Antonelli F, et al. Multiple roles for cytokines in atopic dermatitis: from pathogenic mediators to endotype-specific biomarkers to therapeutic targets. Int J Mol Sci 2022 Feb 28;23(5):2684. [doi: 10.3390/ijms23052684] [Medline: 35269828]
17. Molla A. Dietary patterns and their impact on atopic dermatitis: a comprehensive review. TODJ 2024 Jun 6;18(1). [doi: 10.2174/0118743722306189240520075943]
18. Boonchooduang N, Louthrenoo O, Likhitweerawong N, et al. Impact of psychostimulants on microbiota and short-chain fatty acids alterations in children with attention-deficit/hyperactivity disorder. Sci Rep 2025 Jan 24;15(1):3034. [doi: 10.1038/s41598-025-87546-y] [Medline: 39856212]
19. Tang H, Li W, Xu Y, Zhou Y, Hamblin MR, Wen X. Gut microbiota modulation: a key determinant of atopic dermatitis susceptibility in children. Front Microbiol 2025;16(1549895):1549895. [doi: 10.3389/fmicb.2025.1549895] [Medline: 40356648]
20. Strom MA, Fishbein AB, Paller AS, Silverberg JI. Association between atopic dermatitis and attention deficit hyperactivity disorder in U.S. children and adults. Br J Dermatol 2016 Nov;175(5):920-929. [doi: 10.1111/bjd.14697] [Medline: 27105659]
21. Loo EXL, Ooi DSQ, Ong M, et al. Associations between eczema and attention deficit hyperactivity disorder symptoms in children. Front Pediatr 2022;10(837741):837741. [doi: 10.3389/fped.2022.837741] [Medline: 35433544]
22. Paller AS, Gonzalez ME, Barnum S, et al. Attentiveness and mental health in adolescents with moderate-to-severe atopic dermatitis without ADHD. Arch Dermatol Res 2024 Jul 30;316(8):497. [doi: 10.1007/s00403-024-03210-x] [Medline: 39080094]
23. Cao S, Zhang Z, Liu L, et al. Causal relationships between atopic dermatitis and psychiatric disorders: a bidirectional two-sample Mendelian randomization study. BMC Psychiatry 2024 Jan 3;24(1):16. [doi: 10.1186/s12888-023-05478-1] [Medline: 38172785]

Abbreviations

AD: atopic dermatitis
ADHD: attention-deficit/hyperactivity disorder
ADHD-RS-5: Attention-Deficit/Hyperactivity Disorder Rating Scale-5

ASRS: Adult Self-Report Scale
OR: odds ratio
SCORAD: Scoring Atopic Dermatitis

Edited by W Guo; submitted 18.03.25; peer-reviewed by J Meisenheimer, M Gasmi, S Yasamin Parvar; revised version received 03.06.25; accepted 10.06.25; published 15.07.25.

Please cite as:
Molla A, Jannadi R, Hafez D, Alrohaily L, Abdullah E, Azouni D, Albadrani M
Exploring Attention-Deficit/Hyperactivity Disorder Symptoms in Patients With Atopic Dermatitis by Disease Severity: Cross-Sectional Analysis
JMIR Dermatol 2025;8:e74126
URL: https://derma.jmir.org/2025/1/e74126
doi:10.2196/74126

Original Paper

Exploring the Views of Dermatologists, General Practitioners, and Melanographers on the Use of AI Tools in the Context of Good Decision-Making When Detecting Melanoma: Qualitative Interview Study

Brad Partridge[1], PhD; Nicole Gillespie[2,3], PhD; H Peter Soyer[4], MD; Victoria Mar[5,6], MBBS, PhD; Monika Janda[1], PhD

[1]Centre for Health Services Research, University of Queensland, Brisbane, Australia

[2]Melbourne Business School, University of Melbourne, Melbourne, Australia

[3]School of Business, University of Queensland, Brisbane, Australia

[4]Frazer Institute, Dermatology Research Centre, University of Queensland, Brisbane, Australia

[5]Victorian Melanoma Service, Alfred Health, Melbourne, Australia

[6]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

Corresponding Author:
Brad Partridge, PhD
Centre for Health Services Research
University of Queensland
Level 2, Building 33
Princess Alexandra Hospital
Brisbane, 4102
Australia
Phone: 61 7 3176 5530
Email: b.partridge@uq.edu.au

Abstract

Background: Evidence that artificial intelligence (AI) may improve melanoma detection has led to calls for increased human-AI collaboration in clinical workflows. However, AI-based support may entail a wide range of specific functions for AI. To appropriately integrate AI into decision-making processes, it is crucial to understand the precise role that clinicians see AI playing within their clinical deliberations.

Objective: This study aims to provide an in-depth understanding of how a range of clinicians involved in melanoma screening and diagnosis conceptualize the role of AI within their decision-making and what these conceptualizations mean for good decision-making.

Methods: This qualitative exploration used in-depth individual interviews with 30 clinicians, predominantly from Australia and New Zealand (n=26, 87%), who engaged in melanoma detection (n=17, 57% dermatologists; n=6, 20% general practitioners with an interest in skin cancer; and n=7, 23% melanographers). The vast majority of the sample (n=25, 83%) had interacted with or used 2D or 3D skin imaging technologies with AI tools for screening or diagnosis of melanoma, either as part of testing through clinical AI reader studies or within their clinical work.

Results: We constructed the following 5 themes to describe how participants conceptualized the role of AI within decision-making when it comes to melanoma detection: theme 1 (integrative theme)—the importance of good clinical judgment; theme 2—AI as just one tool among many; theme 3—AI as an adjunct after a clinician's decision; theme 4—AI as a second opinion for unresolved decisions; theme 5—AI as an expert guide before decision-making. Participants articulated a major conundrum—AI may benefit inexperienced clinicians when conceptualized as an "expert guide," but overreliance, deskilling, and a failure to recognize AI errors may mean only experienced clinicians should use AI "as a tool." However, experienced clinicians typically relied on their own clinical judgment, and some could be wary of allowing AI to "influence" their deliberations. The benefit of AI was often to reassure decisions once they had been reached by conceptualizing AI as a kind of "checker," "validator," or in a small number of equivocal cases, as a genuine "second opinion." This raised questions about the extent to which experienced clinicians truly seek to "collaborate" with AI or use it to inform decisions.

Conclusions: Clinicians conceptualized AI support in an array of disparate ways that have implications for how AI should be incorporated into clinical workflows. A priority for clinicians is the conservation of good clinical acumen, and our study encourages a more focused engagement with users about the precise way to incorporate AI into the clinical decision-making process for melanoma detection.

(JMIR Dermatol 2025;8:e63923) doi:10.2196/63923

KEYWORDS

artificial intelligence; melanoma; skin cancer; decision-making; decision support; qualitative; attitudes; dermatologists; general practitioners; melanographers; Australia; New Zealand

Introduction

Background

Timely access to accurate, cost-effective melanoma screening and diagnosis is an ongoing area of health care priority, particularly given that early treatment of melanoma is associated with the most favorable patient outcomes. The current detection paradigm relies heavily on clinician examination assisted by dermoscopy; therefore, accuracy is variable depending on the clinician's level of experience and their personal risk threshold for performing a biopsy. Integrating artificial intelligence (AI) with 2D or 3D skin imaging technologies into clinical workflows and decision-making processes may improve melanoma detection in a number of ways. A growing number of studies indicate that, under experimental conditions, AI can correctly identify images of malignant lesions with a level of diagnostic accuracy that exceeds, or is at least comparable to, expert dermatologists [1-4]. AI may help to identify new or changing lesions in patients at high risk when lesion-specific or full-body images are taken sequentially or help to triage patients by identifying suspicious lesions that require more focused clinical assessment by dermatologists [5,6]. While regulatory approval and the integration of AI tools for melanoma diagnosis is not yet a widespread part of dermatological practice, some AI tools are being used in public and private health care settings in several countries to triage patients by identifying suspicious lesions that require more focused clinical assessment by dermatologists. For example, the Deep Ensemble for Recognition of Malignancy (DERM) AI device is approved for use in the United Kingdom and designed to be used for "screening, triage, and assessment of skin lesions" as well as to provide a "suggested diagnosis and referral recommendation" [6]. Using AI for these purposes may improve the efficiency of melanoma detection and increase population access to dermatological assessment, particularly as improvements in the precision of imaging technology and convolutional neural networks allow for more machine autonomy in decision-making, thereby changing or creating new clinical paradigms in melanoma detection.

Calls for human-AI collaboration in melanoma detection are based on the view that "AI-based support of clinical decision-making improves diagnostic accuracy over that of either AI or physicians alone" [7], and the dominant narrative is that AI should ideally be integrated into clinical workflows in a way that "assists" and "supports" clinical decision-making about skin cancer detection [1-3]. However, there is no consensus on how this integration ought to occur nor how or where in the process AI should be used by clinicians to make better decisions. For example, the way AI output is used and regarded may depend on whether it is referred to before or after the clinicians have made their own assessment. Furthermore, the use of AI as part of melanoma screening and diagnosis also strongly depends on clinicians' acceptance of the technology. Therefore, understanding clinician views regarding the incorporation of AI into decision-making is crucial for developing appropriate clinical workflows (eg, a widespread melanoma screening program). To date, a small number of surveys with dermatologists and general practitioners (GPs) across several countries have provided broad level snapshots of overall inclinations toward the use of AI (Al-Ali et al [8], Nelson et al [9], Patrzyk et al [10], Polesie et al [11], Scheetz et al [12], Shen et al [13], Wei et al [14], and Samaran et al [15]). These surveys have generally reported favorable attitudes toward the potential for AI to positively impact dermatological practice in the future [11] as well as a perception that using AI may improve the efficiency of melanoma detection as the precision of imaging technology improves [12]. Studies using qualitative methods, such as focus groups, have allowed for more in-depth descriptions of clinician views on the potential benefits, barriers, and preconditions of using AI for skin cancer detection [16] in ways that elicit more detailed information about the reasoning and beliefs underlying clinician attitudes. Currently, this level of empirical detail is scarce when it comes to understanding the precise role that clinicians see AI playing within their clinical deliberations and how they think AI output ought to be used to inform their own decisions so that any potential benefits of AI can be wholly realized.

The extent to which AI is seen as helpful for making good decisions outside of experimental settings may vary markedly according to the clinical setting, user, and purpose [17,18]. Therefore, a more in-depth investigation of how clinicians conceptualize the specific way AI may be used within their actual decision-making process is needed, as is a more direct assessment that seeks to learn from the actual experiences clinicians may have already had with AI in dermatology. This does not only apply to dermatologists. In countries with high prevalence rates of melanoma, such as Australia and New Zealand [19], skin cancer diagnosis is conducted by dermatologists and GPs, (including those with a special interest in skin cancer), and screening may be conducted by other clinicians such as melanographers (skin imaging technicians who often have a background in nursing).

Objectives

In recognizing the importance of this practice-driven approach, this study uses in-depth individual interviews with a range of clinicians involved in melanoma detection (dermatologists, GPs, and melanographers) to understand how they conceptualize and view the role of AI within their decision-making and what these conceptualizations mean for good decision-making when detecting melanoma. By drawing on the experiences that clinicians have already had with AI where possible, this understanding will help clinical groups, policy makers, and AI developers to respond to the needs of clinicians involved in the detection of melanoma when it comes to AI use.

Methods

Participants and Recruitment

This qualitative study used in-depth individual interviews to explore clinicians' experiences with and views about using AI to detect melanoma. Melanoma is detected by either (1) screening asymptomatic patients for suspicious lesions and identifying new or changing lesions or (2) correctly diagnosing malignant lesions that the patient seeks advice about. We recruited a cohort of clinicians who regularly conduct skin cancer examinations, including dermatologists, GPs with an interest in skin cancer, and melanographers. We sought participants who had familiarity with AI use in the clinical setting or as part of research (but did not exclude those without experience with AI). Participants were recruited through purposive sampling from the authorship group's network of (1) scientific contacts in melanoma research in Australia and New Zealand (including the Australian Centre of Excellence in Melanoma Imaging and Diagnosis, currently one of the largest melanoma cohort studies worldwide using 3D total body imaging for melanoma early detection) and (2) professional clinical contacts in Australia and New Zealand including GPs, melanographers, and fellows and trainees of the Australasian College of Dermatologists (the peak professional body for dermatologists in Australia and New Zealand). An initial group of potential participants was contacted via email, and additional prospective interviewees were contacted on the recommendation of participants themselves through "snowball" recruitment.

The final sample comprised 30 participants: 17 (57%) dermatologists, 6 (20%) GPs with a special interest in skin cancer, and 7 (23%) melanographers (refer to Table 1 for further sample demographics). A total of 26 (87%) participants were from Australia and New Zealand; 4 (13%) participants from Chile, Greece, the United States, and the United Kingdom were recruited through snowball sampling. The vast majority of the sample (n=25, 83%) had interacted with or used 2D or 3D skin imaging technologies with AI tools for screening or diagnosis of melanoma, either as part of experimental testing (or through clinical AI reader studies) or within their clinical work. Dermatologists and GPs in the sample collectively had experience with AI tools for predicting likely diagnosis, detecting lesion change, and screening. The melanographers in the sample had clinical experience with AI tools to screen lesions for melanoma (ie, identifying suspicious lesions in need of referral to a dermatologist). The most commonly nominated AI tools used by participants were part of imaging platforms provided by FotoFinder (FotoFinder systems, Inc), DermEngine (MetaOptima), Canfield (as part of 3D total body photography; Canfield Scientific Inc), and MoleMap.

Table 1. Demographics.

	Dermatologists (n=17)	General practitioners (n=6)	Melanographers (n=7)	Total sample (n=30)
Gender, n (%)				
Men	10 (59)	4 (66)	0 (0)	14 (47)
Women	7 (41)	2 (33)	7 (100)	16 (53)
Age (y), mean (SD; range)	44.5 (7.7; 32-60)	52.1 (12.2; 33-72)	34.7 (10.3; 26-49)	43.7 (6.4; 32-72)
Experience (y), mean (SD; range)	14.1 (7.7; 1-30)	13.8 (5.6; 5-20)	4.6 (4.9; 1-15)	11.4 (2.5; 1-30)
Experience with artificial intelligence, n (%)	14 (82)	4 (66)	7 (100)	25 (83)

Data Collection and Analysis

All interviews were conducted by the first author (BP) via a video link between October 2023 and February 2024, with 2 (7%) of the 30 participants providing written input instead of an interview. The average duration of interviews was approximately 45 minutes, and interviews were digitally recorded with the participant's consent and transcribed. A semistructured interview schedule developed by the research team was used to guide discussion on several areas, including experiences with skin imaging technologies and decision-making with AI tools, expectations for the use of AI as part of decision-making, clinician trust, and potential barriers and enablers of AI use in various clinical workflows.

Using a critical realist stance, we conducted a thematic analysis of transcripts by drawing on the template analysis approach to structure data coding [20,21]. BP read through all interview transcripts for familiarization. BP and NG then started preliminary coding of the data (using a process familiar across all forms of thematic analysis whereby the researcher identifies and "labels" text that may contribute to understanding of the topic; this is described by King et al [21]) by (1) reading through the first 6 interview transcripts together, (2) inductively coding these transcripts for meaning independently, and (3) collaboratively discussing shared interpretations and impressions where diverging views were settled via open discussion. We chose not to make use of a priori themes to guide coding and favored an inductive approach to the development of initial

themes by identifying clusters of shared meaning across interviews. Early in the process of clustering, we identified how conceptualizations of AI were often linked to the way participants talked about "positioning" AI in the decision-making process, and this helped to formulate initial themes. As described in the study by King et al [21], this clustering also allows the development of potential "integrative themes," that is, themes that permeate other clusters of meaning. We generated the integrative theme of "the importance of good clinical judgement" given the way it infused how AI was conceptualized throughout. The first iteration of a coding structure, with initial themes related to good clinical decision-making and the use of AI, was constructed and then applied to the whole dataset, undergoing several refinements, where necessary, to identify how meaningfully it captured the data. At its core, refinement of themes entails repeatedly going back to interview transcripts and "testing" how well the thematic descriptions capture meaning. Here, refinement typically entailed adjusting the wording of themes for clarity in order to ensure precise capture of the data. This process ensured continuous engagement with the data. Construction of themes was then solidified with all authors contributing to the final data interpretation.

Ethical Considerations

This study was approved as low or negligible risk research by the University of Queensland Human Research Ethics Committee (2023/HE001714). All participants provided informed consent to participate and had the ability to opt out. Participants were not provided with compensation or incentives to participate. Data have been deidentified.

Results

Overview

To describe how participants viewed the role of AI within decision-making when it comes to melanoma detection, we constructed the 5 themes. There were 4 main themes and 1 integrative theme (refer to Multimedia Appendix 1 for an outline and additional representative excerpts for thematic context and interpretive validity): theme 1 (integrative theme)—the importance of good clinical judgment; theme 2—AI as just one tool among many within the process of decision-making; theme 3—AI as an adjunct after a clinician's decision; theme 4—AI as a second opinion for unresolved decisions; theme 5—AI as an expert guide before decision-making.

The integrative theme explained how participants conceptualized arriving at good decisions through the development and display of good clinical judgment; in this context, although AI could be accurate, it often lacked contextual awareness. This was the reference point through which participants discussed their views on incorporating AI. It provided a link to the 4 main themes that described the ways participants then conceptualized AI within the decision-making process when screening for or diagnosing melanoma.

Theme 1 (Integrative Theme): The Importance of Good Clinical Judgment

This theme provided a link to the other 4 main themes by encapsulating how participants conceptualized arriving at good decisions when detecting melanoma through the development and display of good clinical judgment, "acumen," or "nous." This was endorsed as one of the hallmarks of being a good clinician that leads to accurate decisions and optimal patient outcomes. Across the accounts of dermatologists, GPs, and melanographers alike were indicators of clinical astuteness, in particular (1) seeking information from multiple sources, (2) being able to see the broader clinical picture and consider contextual factors, (3) synthesizing and balancing the information that has been gleaned, and (4) applying it accurately to individual patients with their benefit in mind. For example, one participant said the following:

> *I think putting all these little clues together is part of the point, and each thing gives you a little bit of incremental information when you have to use your judgment, too, of how you weigh bits of information to make a decision. And then you have to also make decisions about how you rate information in terms of what the patient will consider acceptable practice.* [ID07, dermatologist]

Being a good clinician meant knowing which pieces of information are truly relevant in each specific case and what clinical importance to assign them:

> *...when you have to manage a patient, you don't just see a dermoscopy image and say, "okay, it's melanoma or it's a dysplastic nevus." You take into account so many different factors. There are so many variables that we consider when we have to offer a specific management to the patient.* [ID01, dermatologist]

It was through the practice of repeatedly making decisions for oneself that clinicians developed expert clinical acumen when it comes to melanoma detection, whereas new and inexperienced clinicians tended to "get stuck on one feature" (ID07, dermatologist) and failed to incorporate enough information from multiple sources.

In this context, participants described AI as impressive when arriving at accurate decisions, but the limitations of AI as a "decision-maker" arose because "AI can't put the lesion into the context of the patient" (ID08, GP). Participants often noted "respect" for accurate AI programs and anticipated considerable improvements in accuracy in the future ("when put under the pressure you can see, wow this AI is fast and pretty accurate! So after that, I had a lot of respect." [ID12, dermatologist]) while also highlighting that AI is trained to recognize images of lesions and make decisions in a way that is entirely different from the way clinicians use their acumen to make good decisions with real patients. For example, one of the dermatologists said the following:

> *It's looking at an image. We're not looking at image. This is a completely artificial way of looking at our job...We're looking at a patient who has 100 lesions,*

> *and we're putting so many of their risk factors in the question.* [ID17, dermatologist]

While an accurate AI might come to a correct decision, often the issue for participants was being aware that AI could have little clinical acumen or "nous" and make errors that would be unusual for an expert. This created a need for clinicians to maintain a kind of vigilance against accepting AI output as though it simply reflects the product of good clinical judgment:

> *...there's still a lot of human interaction which goes into clinical decision formulation and management planning, which I think is not yet incorporated in AI appropriately, satisfactorily.* [ID07, dermatologist]

One highly experienced melanographer summed this up by saying the following:

> *...that's something that AI can't do. It can't grab that patient, the background, at this point anyway. How long has it been there? Is it coming and going? Were you aware of it? What does it feel like? From an actual physical touch is it rough, is it raised, is it soft, is it squishy. Then from the patient experience is it sore, is it tender, it is itchy?* [ID16, melanographer]

This impacted the way participants went on to conceptualize and navigate their interactions with AI tools, trust AI's output, and position AI within their own decision-making process:

> *The main question is: will the clinicians become better from using the AI or not?* [ID03, dermatologist]

This conceptualization of good clinical judgment and acumen informed the themes described subsequently.

Theme 2: AI as Just One Tool Among Many Within the Process of Decision-Making

This theme encapsulated a view of AI as a "tool" for consideration when forming decisions rather than AI as a "decision maker" per se. In this light, viewing AI output as simply one piece of information to be synthesized meant that good decision-making on the part of clinicians required knowing *how* or perhaps *whether* to refer to AI output in the course of reaching a decision:

> *Every field of medicine, as technology improves, if you don't use it well then you have to say why not? So this is no different. Just another tool. It's just a more difficult tool to interpret.* [ID13, dermatologist]

But crucially, doing this well was largely dependent on clinicians already having well-developed expertise in order to be judicious in how they interpreted AI output as an "incremental" piece of information ("...a useful tool for an intelligent doctor who knows its limitations" [ID08, GP]). For example, using an AI tool that detects small changes in a lesion over time still requires the clinician to have the expertise to know which change is clinically important for real patients:

> *It [AI] detects change, but it detects every little change...and in the end we manually compare their sequential images and we turn off all the AI, because it circles so many things. It's just a pain.* [ID07, dermatologist]

As such, these clinicians were careful not to assign more weight to AI output than it deserved:

> *I never base solely on AI...So, I never say "AI said this, so I'll do this." No, never.* [ID21, dermatologist]

This served to highlight how the development and implementation of expert clinical acumen was an inherently important part of making good decisions when this involved conceptualizing AI as one tool among many:

> *Look the way I perceive it [AI] would be an incremental step so...but it doesn't replace taking the history, examining the patient, selecting my lesions, it will give me one incremental piece of information in the specific investigations required.* [ID07, dermatologist]

By conceptualizing AI as "simply a tool" that may add one piece of information to be synthesized, often the key question for highly experienced clinicians was whether this tool then added any information beyond what they could already glean through their own good clinical acumen or "nous" to assist in the formation of a decision. As such, many experienced participants in this study felt they did not need the information from an AI tool to help inform their decisions:

> *...de novo interpretation of a particular lesion is where the AI can be helpful, but I don't believe it helps with someone that already has quite a considerable amount of training in lesions.* [ID11, dermatologist]

One GP with considerable familiarity with AI described the experience when receiving information from AI tools currently being developed in research settings:

> *To be honest, I've used this for a number of years, and I don't often find anything new with it...if you're a really good dermoscopist, what we're finding with our data is that the GPs are better than the AI here. We're finding melanomas smaller and earlier than the AI is.* [ID02, GP]

It is not that participants invariably expected AI would be wrong in any given instance (some AI tools could certainly be accurate), rather their wariness was precisely because they knew AI does not incorporate the entire scope of information that expert human clinicians ordinarily make use of. It was because clinicians could not always anticipate being able to relate to AI on the common ground of good clinical acumen (eg, experience, ability to synthesis a wide range of information, and ask questions) that some spoke about their need to better understand exactly how AI arrived at decisions, so they could anticipate how AI output could be incorporated as one piece of information into their deliberations:

> *If someone is coming with an AI, I need to understand where to place it, and the training and testing of the data set, so I can see if my things are the same as what they've been doing. I need to understand the data, how they've been taking their images, how they designed the algorithm, how they've been processing their images, if there was some process, and then of course, I need all the metrics. I need the clinical*

> *validation. I need to have that in my work flow, with my patients, and so I need to know if it has a good sensitivity and specificity, and all the performance metrics...I need to know the limitations, because there are always limitations.* [ID17, dermatologist]

Theme 3: AI as an Adjunct After a Clinician's Decision

Participants often described the value of AI as a source of reinforcement for the decisions they had already made when screening patients for suspicious lesions or diagnosing melanoma. Participants talked about AI by framing it as an "adjunct" or "auxiliary" when it comes to doing skin checks, inspecting individual lesions or diagnosing melanoma. For instance, "I would still be using my clinical acumen and still would be doing skin checks and using this as an adjunct" (ID09, GP). More specifically, participants across clinical groups explained how they were often inclined to position AI *after* they had engaged in their clinical decision-making because this allowed them to first enact their own good clinical acumen and then use output from AI tools to "validate" this. Importantly, this was even when clinicians felt confident in their decision. For instance, two melanographers described the following:

> *...mostly I find it confirms, more than changes, what I'm doing...It's more supporting and confirming.* [ID15, melanographer]

> *I make my decision, and I know in myself "I think it's this." I pop the AI on and then it's a nice confirmation. Hypothetically, maybe it'll catch something I didn't think of, but if you've done the first part right, that shouldn't be happening.* [ID20, melanographer]

Participants described how seeking to use AI for confirmation and reassurance after they had made a decision was different from seeking out information from AI to help *form* their initial decision (as seen in theme 2). When formulated as a "checker," experts were not necessarily seeking out AI's "opinion" to help them make the actual decision (eg, "Is this lesion suspicious?" "Is this lesion change worrying?" and "Should I biopsy this lesion?"), rather, they were seeking out validation for the decisions they had made. For instance, a melanographer who had used an AI tool for screening lesions said the following:

> *I definitely like it as a tool to cross-check my work...Often, even not so much to get AI's opinion. I suppose I use AI to validate and reinforce some of the decisions that I've already made. Often, I'll already have assessed something on the skin.* [ID24, melanographer]

Indeed, participants from all 3 clinical roles described how they enjoyed the feeling of seeing AI tools confirm their good clinical acumen ("I quite like it when AI agrees with me" [ID24]). One dermatologist said the following:

> *Whenever I think something's benign, AI reassures my decision.* [ID21, dermatologist]

They felt it reinforced their good clinical acumen and gave them confidence in their decision-making, particularly in the context of feeling anxious about the prospect of failing to detect a melanoma:

> *If you want me to give you a take home point from my use, it would be just confidence changes, but that's it.* [ID21, dermatologist]

This was also the case for melanographers when examining patients with a large number of unusual looking lesions, where AI was described as being like "a reassuring little friend in clinic" (ID20, melanographer). Participants described how one advantage of incorporating AI as a "checker" was that it helped clinicians to be vigilant when assessing patients. AI was not making decisions before or in lieu of the application of a clinician's own good judgment, rather, when the AI tool "flagged" a lesion, the clinician saw it as an exhortation to ensure they had thoroughly applied their own good clinical judgment in the first place:

> *If I knew the AI would insist on me looking at it, that would be good.* [ID10, dermatologist]

It was an invitation to "double check," confirm their own decision, and be reassured, as described by a dermatologist:

> *...it makes you more vigilant, which is only a good thing. So long as you manage that balance well, where it doesn't turn into paranoia.* [ID22, dermatologist]

Theme 4: AI as a Second Opinion for Unresolved Decisions

In a small number of cases where highly experienced clinicians were still genuinely unsure about a definitive diagnosis after they had applied their clinical judgment, they described seeking out AI for a "second opinion" ("...look in truth, I do look at it quite often if I'm not sure about something, and if I'm vacillating" [ID07, dermatologist]). Participants across all 3 clinical groups likened this specifically to the way they would apply due clinical diligence by seeking out one of their colleagues for a second opinion, which is a familiar practice for clinicians. For instance, one melanographer said the following:

> *I only use it for things that I'm not sure about. Yeah, that's probably where my years of experience come in.* [ID15, melanographer]

Importantly, though, there were varying accounts of whether seeking a second opinion from AI in this way was akin to seeking out a colleague with more expertise, a peer, or perhaps a "less expert colleague." For some, AI could be a second opinion with great decision-making capacity by virtue of it having access to a large training dataset:

> *Having the AI is just like having another person in the room. It's actually swarm intelligence.* [ID08, GP]

Whereas for others, the current state of AI meant that its opinion was perhaps not as authoritative:

> *I can tell you, I've used the AI and then second guessed myself. I think that's what it's there for. It is like having a second opinion from a colleague, if you like, but again, one that's not validated. One that's not ready. A less expert colleague, perhaps.* [ID22, dermatologist]

For some highly experienced clinicians, the fact that their clinical acumen still left them with a degree of uncertainty about a lesion was often reason enough for them to decide on the most cautious course of action (eg, biopsy or excision), regardless of the AI's second opinion:

> *I would use that only for an additional opinion you know the same way I use my colleague's opinion when I am not sure about a lesion of concern. If there's a discordance between us I think I just do what I feel. I would do the same with AI. If the AI says this is benign, and I still feel that this is not benign, or it is something that must be excised, I would go for the excision.* [ID01, dermatologist]

Others thought using AI as a second opinion for equivocal lesions may help them prevent biopsy or excision if it allowed them to incorporate some insights that would lead to a more precise diagnosis rather than simply reverting to caution. This might help reduce the removal of benign lesions when it is not needed and be to the patient's benefit. For instance, one melanographer said the following:

> *...as long as it's accurate, as accurate as it can be, I don't really see any disadvantage of having a second opinion right in front of us. Yeah, I think that it's more thorough and more accurate if you have the person, myself, and then a system to help you as well.* [ID19, melanographer]

Theme 5: AI as an Expert Guide Before Decision-Making

The prospect of patients receiving unbeneficial treatments or melanoma diagnosis being missed was a salient concern:

> *The worst thing is you have an inexperienced practitioner and they cut stuff out that doesn't need to be cut out and they've actually missed the important thing.* [ID04, dermatologist]

With this in mind, a number of participants thought that AI output could potentially act as an "expert guide," for example, when screening for suspicious lesions:

> *For non-dermatology practitioners, so GPs and other health care specialists, AI tools may improve the identification of suspicious lesions or otherwise the reassurance of lesions that are completely benign.* [ID01, dermatologist]

But importantly, the prospect of clinicians referring to AI output *before* they had thoroughly engaged in their own decision-making (or in place of a clinician engaging in decision-making) was also predicted to be "to the detriment of the patient" (ID08, GP) if it inclined clinicians to de-emphasize or set aside the need for "good clinical acumen." It was seen as a "short cut" that could leave clinicians susceptible to error through overreliance on AI and deskilling:

> *...if you are not as good at seeing melanomas, then AI may help you to do that. But it could easily provide false reassurance because we know it doesn't always pick them up. And it could lead to reliance on AI and deskilling, and that's my concern. It's not a substitute for good clinical acumen.* [ID02, GP]

At stake was the distinction between using good clinical judgment to *incorporate* information from AI into one's decision-making ("interpreting" information from AI as part of making one's decision as seen in theme 2) and actually allowing one's clinical judgment to be "influenced" by AI in a way that detracts from displaying good clinical acumen. Given that there was widespread recognition that good clinical acumen involved making a considered judgment based on many sources of information, some explained the value of making a conscious effort *not* to use AI as an expert guide. Some participants described the importance of ensuring their initial decision-making was not influenced by AI at all, so that they were not being "told what to do" by AI, but rather retaining autonomy in their decision-making. For example, one GP with considerable experience with AI tools as part of research said the following:

> *Every patient we reviewed them with a dermatoscope prior to doing any AI analysis...we very, very deliberately take logistical steps [to do that].* [ID02, GP]

That is, it was important to first form one's own assessment before attempting to incorporate information from AI as a tool into the decision-making process:

> *If you click an AI button before you've made your assessment, you won't actually get a real idea what you're thinking what the lesion was. It'll actually adversely influence your decision-making process.* [ID13, dermatologist]

Others similarly cautioned against the temptation to use AI to guide their decision-making (rather than "confirm" or "validate" their decision-making) by likening it to "the ultimate shortcut for people to not learn stuff" (ID12, dermatologist) or "an excuse for not becoming an expert yourself" (ID08, GP). For instance, 3 experienced melanographers noticed the way new trainees were using an AI tool for screening suspicious lesions:

> *I don't know that it's a good idea, in new melanographers. I really think you need to sort of trust your instincts and use your knowledge in the beginning to really know what you're looking at. We've had a couple of new people start with that (AI), and I guess they rely quite heavily on it.* [ID15, melanographer]

> *So we have someone here who is newer and she was saying you could definitely, as a new, person doing this rely on AI without having the training, and the experience. And so it's getting that balance, I suppose, between the experience and not totally relying on it.* [ID16, melanographer]

> *I notice the new nurses, they're using it definitely as their guidance. Whereas I almost use it to just reinforce my decisions. I think at this point I definitely tend to have made my decision up already.* [ID24, melanographer]

Participants thought that if inexperienced clinicians (whatever their clinical role) used AI tools as a guide to melanoma detection without developing and exercising their own clinical acumen, then although they may at times benefit from AI's accuracy, they may also be prone to reinforcing errors arising from AI's lack of acumen. As such, this led some participants to the view that, in the hands of inexperienced clinicians, using AI as a guide was likely to be a mistake, hence perhaps AI tools ought to be considered "expert only" devices:

> *At the beginning we thought that it will be a tool that will be a significant help for non-experienced users, but for experts, okay, it's not so important. I've started to believe it is the opposite. You have to be quite good in order to be able to deal with the strange decision making that the machine makes, and to understand when you should follow and when you should ignore what the machine is saying. This requires quite a lot of confidence and experience. An inexperienced user is prone to fall in all the traps that might happen when you use these tools.* [ID23, dermatologist]

The prospect of using AI as part of screening or triage of patients to filter out benign lesions (the task of melanographers) was alluring for a number of dermatologists and GPs who thought this may increase their efficiency. However, many participants were also skeptical that the current AI tools were sufficiently accurate to wholly defer to as an expert guide in this decision-making. Indeed, several melanographers in this study talked specifically about the dangers of inexperienced clinicians overrelying on AI by using it as though it were an "expert guide" for lesion screening. What is instructive is that some melanographers in this study, with access to AI as part of lesion screening, gave credence to this view:

> *I think I rely on it [AI]. I do. Yeah, I think I relied on it like the most when I first started, but I think as you like learn more and see more skin and see more like diagnoses, I think slowly you start—I don't rely on it as much anymore, but I definitely still rely on it.* [ID19, melanographer]

This also lent support to the view that using AI as an expert guide may "deskill" clinicians, by limiting opportunities to enact good clinical acumen. For instance, one melanographer talked about remaining vigilant and not to overrely on AI in a way that might atrophy her good clinical acumen:

> *What if tomorrow we stopped using it [AI]? Will I lose my clinical skills?... sometimes I think to myself at the end that's why I stay quite cautious.* [ID20, melanographer]

This concern was echoed by other participants who foresaw the prospect that positioning AI as a guide to decision-making may invariably impair their good clinical acumen. Again, this is an acknowledgment that AI tools can be potentially useful or helpful but not if they are used in ways that detract from developing good clinical acumen, as described by a dermatologist:

> *Maybe our human expertise lowers a bit, or gets a bit impaired, that we have the (AI) support. And we've trusted so much that we stopped developing ourselves.* [ID03, dermatologist]

Discussion

Principal Findings

A prominent narrative within the dermatology literature sees AI as supporting clinical decisions [6], with advocates pitching the use of AI to "augment," "assist," and "aid" the detection of melanoma via a human-computer collaboration [22,23]. For instance, Esteva et al [1] pointed to the potential of AI algorithms for "augmenting clinical decision-making for dermatology specialists"; Brinker et al [3] said that "artificial intelligence algorithms may successfully assist dermatologists with melanoma detection in clinical practice," and Haenssle et al [2] said that AI tools may "aid physicians in melanoma detection." However, the notions of augmentation, collaboration, support, and assistance may entail a very wide range of actual functions for AI in the context of decision-making. We find that end users conceptualize these terms in an array of potentially disparate ways that impact AI's incorporation into clinical workflows. These meanings are important for clinical groups, AI developers, and policy makers to understand so that the development of clinical workflows and guidelines for AI use are appropriate and acceptable to end users (in this case, end users being dermatologists, GPs, and melanographers). Indeed, the potential prospects of AI improving decision-making have been tempered more recently by concerns that algorithms with superior performance to human clinicians in research settings do not necessarily translate into better performance with actual patients in the context of real-world decision-making [17,24]. AI algorithms for assessing malignant lesions can often perform less well outside experimental settings [25]. Nevertheless, some AI tools have already been approved for screening, and some skin imaging platforms allow clinicians to access diagnostic AI assessment of lesion images with the caveat that these AI tools have not yet been validated or approved for clinical decision-making.

This study brings to light a timely perspective on how clinicians involved in melanoma detection conceptualize the use of AI within their decision-making processes and explains how they view the role of AI within the context of good decision-making. Our findings can be further contextualized by considering previous assessments of clinician attitudes toward AI in dermatology. The existing literature has been largely characterized by brief surveys exploring broad impressions about the potential impact of AI, and several of those studies have highlighted optimism among clinicians about the prospect that AI may improve melanoma screening [12] and the accuracy of decisions. Surveys from European and Middle Eastern countries as well as China, the United States, and Australia [8-11,13,14,22] commonly reveal expectations that AI-supported decision-making will be beneficial to the field [8,9,11,12], with many of those surveyed seeing AI as having the potential to improve diagnostic accuracy or other decisions [9,10]. Among the biggest expected benefits is the potential for AI to improve patient access to melanoma screening, although inaccurate AI screening or diagnosis is a major concern [12]. In focus groups with Dutch dermatologists [16], greater diagnostic accuracy

was cited as the leading perceived benefit of AI potentially leading to "fewer missed skin cancer diagnoses and less unnecessary biopsies and excisions of benign skin lesions." However, dermatologists also held concerns about the use of AI tools if their accuracy with real patients fell short of the current abilities of expert clinicians.

The use of qualitative methods in our study has extended these snapshots by showing how these beliefs need to be interpreted in line with the way AI is conceptualized by clinicians within the decision-making process, particularly those who have already used AI in some way. Our findings suggest that the potential benefits of AI for improving accuracy in diagnosis or screening depend on where in the clinical decision-making process AI is used, how clinicians engage with it (eg, as a tool, checker, second opinion, or guide), and the level of expertise and experience of the clinicians using it. As such, while clinicians may have an overall view that AI can "improve accuracy," the extent to which they endorse a specific workflow that includes AI is likely to be contingent upon these kinds of details.

Formulating workflows that are sensitive to the positioning of AI has been identified as a critical part of using AI to enhance clinical decisions [26], and our themes could be used in future studies as touchpoints to assess clinician endorsement of various AI workflows. Importantly, our findings show that the importance of these concepts extends beyond dermatologists—they are seen in the accounts of GPs as well as melanographers—and given that melanoma detection is often a multilayered process for patients that encompasses interactions with several clinician groups, it is important for future research to recognize this.

Our main themes support previous evidence that the "confidence" clinicians have in their decisions may be improved with the use of AI. For example, in their 2020 survey with dermatologists from Australia and New Zealand, Scheetz et al [12] found that "improved diagnostic confidence" was one of the most cited potential benefits of AI. However, our study also shows that, for many clinicians, this was the result of using AI as a "validator" for their decisions or as a reassuring adjunct (see theme 3), rather than using AI to help them make the decision in the first place. Similarly, while clinicians see the potential for AI to improve diagnostic accuracy, our results add nuance by showing that many experts conceptualize this as applicable only to less experienced clinicians and mainly when AI is used as simply one "tool" among many or as a "checker" rather than deferred to. This fits with other findings [12] showing that when detecting skin cancer in experimental conditions, experienced clinicians largely ignored AI output if they were confident of their decisions, whereas inexperienced clinicians were more likely to accept AI-output that contradicted their initial decision, so when the accuracy of the AI tool was of a lesser quality, this put decision makers at risk of error.

This understanding of how AI is conceptualized by clinicians allows a better interpretation of clinician views on the acceptability of AI and human-computer collaboration. In determining what the optimal clinician-computer collaboration should look like, it has been pointed out that "the ideal positioning of AI in relation to the clinician also needs to be considered" [26]. Our study shows how participants across clinical roles commonly articulated a major conundrum about the positioning of AI for melanoma detection that encapsulates a number of the potential benefits and drawbacks associated with each conceptualization of AI; in trying to accurately detect melanoma, although AI may be of benefit to inexperienced clinicians when used as an "expert guide" before they engage in their own decision-making, the potential for this to lead to overreliance, deskilling, and a failure to recognize AI errors when they occur may mean that only expert clinicians have the required acumen to use AI properly as "one tool among many" to inform initial decisions. That is, when conceptualized this way, only those with already well-developed clinical judgment are thought to be able to appropriately engage with the limitations of AI (the "traps," as one participant put it), including AI's lack of contextual awareness. However, at the same time, experts in this study often described how they did not necessarily see a strong need for AI to inform their initial decision-making; instead, they preferred to be able to rely on their own clinical judgment when making decisions.

In this study, while some participants conceptualized AI as offering "one piece of information" to be judiciously interpreted within their synthesis, they often did not see any extra information being gleaned from current AI beyond what they could establish about patients through their own good clinical judgment, and they could be wary of allowing AI output to "influence" their initial deliberations in ways that undermined their independence as a decision maker. However, this perhaps raises queries about the extent to which experienced clinicians then truly seek to "collaborate" with AI on decisions or use it to support their decisions in ways that may actually improve accuracy. Participants often thought that AI's main benefit was in reassuring their own decisions after they had been reached, as a kind of "checker," "validator," "confirmation tool," or in a small number of equivocal cases as a genuine "second opinion." As such, this perhaps raises doubts about the extent to which experienced end users always see AI as having the potential to "support" or "assist" within their decision-making process in beneficial ways, again highlighting the importance of understanding what clinicians mean in this regard. Further evidence will be needed to elucidate whether this way of positioning AI in the process will maintain benefits, such as clinician autonomy, without resulting in drawbacks, such as failure to make best use of an accurate AI. Indeed, one recent study of AI use among dermatologists found that despite high confidence in the AI tool, many opted to continue relying on their own decision-making [22], and there is evidence that people often "ignore (AI) recommendations because they do not trust them; or perhaps even worse, follow them blindly, even when the recommendations are wrong" [27]. Our findings extend those from some experimental conditions and surveys showing that experienced clinicians largely ignore AI output when diagnosing melanoma if they are confident of their decisions, whereas inexperienced clinicians are more likely to accept AI-output even when the accuracy of the AI tool was of a lesser quality, thus putting them at risk of error [7]. Still, Tschandl et al [7] found that "faulty AI can mislead the entire spectrum of clinicians, including experts."

With the positioning of AI being very important to the way AI is conceptualized, it will be pertinent to consider practical matters, such as whether clinicians can elect when or whether they see AI output and under what circumstances. Similar challenges in knowing exactly how to incorporate AI in ways that promote good diagnostic decision-making have also been reported by clinicians using AI to detect other types of cancer (eg, radiologists using AI to detect breast and lung cancer) [28]. Interpreting our results through the lens of commonly applied frameworks for assessing user acceptance of technology may yield insights into our participants' perspectives. A recent review [29] found that widely used frameworks such as the technology acceptance model [30,31] and the unified theory of acceptance and use of technology [32] include key factors such as performance expectancy or perceived usefulness among the strongest predictors of behavioral intentions. Regarding our results, constructs such as performance expectancy or perceived usefulness conceivably encompass a range of views expressed by our participants, including (1) the current accuracy, sensitivity, or specificity of AI for melanoma detection, (2) the extent to which AI takes into account the broader patient context, and (3) the perceived need (or lack thereof) for experienced clinicians to rely on information from AI rather than their own clinical discernment, to make good decisions when screening for suspicious lesions, detecting change, or arriving at a diagnosis.

Across our themes, participants described how the performance or perceived usefulness of AI as part of melanoma detection may vary depending on matters such as the kind of role it may play in the decision-making process, the position it occupied in the workflow, or the relative expertise of the clinician. Nevertheless, it is worth noting that the rapid development and unique nature of AI technology has tested the ability of many older technology acceptance models to confidently predict behavior about AI. This is particularly the case in health care settings where there is a complex interplay among social, technical, and organizational structures and with many stakeholders [33]. As such, it is important for inductive work to reveal how AI use within specific cases is conceptualized by stakeholders and what meanings they attach to AI within the scope of their existing values and obligations.

The way participants in this study have conceptualized the role of AI within good decision-making also points to important ethical considerations. The use of AI within dermatology (and indeed health care more broadly) raises many already well-described ethical issues related to data privacy and ownership, transparency, and equitable access [34]. Navigating these issues in an optimal way is likely to require considerable assessment. Our study suggests that ethical obligations to act in accordance with AI may be placed upon clinicians when seeking and receiving information from AI. These may then impact the extent to which they feel accountable for decisions. For instance, if clinicians are to act in their patients' best interests, then it is reasonable to expect that they ought to rely on the best available information when making decisions, that is, there is a prima facie moral obligation within clinical encounters to treat patients based on the best or most accurate available information (at least, it would be unethical for clinicians to *prefer* to rely on information they know is from an inferior or less accurate source). When AI is conceptualized as an "expert guide" then, it is implicitly installed as a kind of epistemic authority in relation to the clinician (AI is the "expert"). This may then create an ethical directive for clinicians to act upon the advice of AI accordingly, because if AI is the acknowledged expert, then it seems hard to justify ignoring its output or making contrary recommendations. However, this also appears to position AI as the accountable party in the workflow. Throughout the study, participants also strongly endorsed a seemingly countervailing imperative for clinicians to "understand when you should follow and when you should ignore what the machine is saying" (ID23), that is, this appears to be a directive to be judicious in accepting AI's decision. In this study, when participants talked about approaching AI output in a judicious rather than deferential way, it appeared to be borne out of two views. First, the current AI tools were not sufficiently accurate to justify an obligation to follow their output without engaging in independent decision-making. second, clinicians are ultimately accountable for their decisions, meaning that "deferring to AI" could be tantamount to recusing oneself from a core ethical responsibility as a clinician (and to patients). This may be why conceptualizing and positioning AI as a "checker" or "second opinion" (themes 3 and 4) was more readily endorsed by many participants; doing so may be seen to preserve their ability to act as the epistemic authority, enact their moral responsibility to promote the welfare of patients, and potentially negate any potential ethical obligation about following the AI. For instance, participants seemed more comfortable about dismissing or "overruling" AI when it was simply consulted as a checker or second opinion. However, the extent to which this is justifiable, or will hold in all potential situations, is unclear. When the accuracy of the AI is known to exceed that of the clinician, then this ability may be left in a perilous ethical state (although it is also important to recognize evidence that AI accuracy in experimental settings is typically far superior to "real-world" AI accuracy [25]). A deeper examination of the scope of ethical obligations raised by AI within specific melanoma screening workflows is certainly warranted to understand how to best implement any future proposals for widespread screening programs.

In doing so, it is worth noting that patient or consumer views toward AI in melanoma detection (or at least, what clinicians believe to be the views of patients) may also in turn impact how clinicians adopt and use AI. Several recent studies have found that most dermatology patients report having few, if any, concerns about AI being used by specialists to diagnose skin cancer as long as diagnostic decisions are not made by AI alone [35], with diagnostic accuracy and explainability as being features of AI that are most important [35-37]. Given that the mere presence of AI within decision-making workflows imbues AI with at least some epistemic legitimacy, this may raise questions for patients about the extent to which AI ought to be deferred to (eg, be used as an expert guide and as second opinion), which may, in turn, impact the extent to which clinicians act accordingly to maintain their patient's trust. Different conceptualizations of the role of AI may present different ways of dealing with issues such as clinician-AI discordance. These issues are likely to be made starkly apparent

in the implementation phase, for instance, in situations where AI output is available to patients and clinicians in real time. It would be pertinent for implementation scientists working on the development of melanoma detection workflows to consider how the different conceptualizations of the role of AI described in this study may accord with the views of patients, and, in turn, impact their trust and acceptance of AI in the process.

The possibility of AI inducing deskilling through overreliance on AI has often been identified in the literature [38-40]. While this is a recurring concern for health care practitioners [39], the potential for AI to hinder learning or erode already-acquired competency permeates more broadly [40]. This study showed that experienced clinicians involved in the detection of melanoma were cognizant of this potential. Some described how they tried to ensure they adopted decision-making workflows that resisted this, for example, by only using AI after they had made an initial decision. Importantly, this study also showed that the prospect of reliance on AI was not merely hypothetical; our interviews uncovered evidence that some newer melanographers who had been trained to use AI to identify suspicious lesions (and refer them for dermatological review) were aware that their anxiety about missing potential melanomas inclined them to regularly rely on AI output as an expert guide, given their initially limited clinical experience. More experienced melanographers were wary of doing this and held concerns that overreliance on AI would facilitate them to "lose their skills" or clinical judgment. This was concerning for participants given that the development and display of good clinical acumen reflected what it meant to be a good clinician; it entailed making decisions in a way that takes into account many pieces of information from the patient, learning how to balance potentially relevant clinical information through experience and reflection, and seeing the broader context of the patient with their interests in mind (refer to the study by Tsang et al [41] for a similar view). Due to this, participants described a primarily AI-led decision-making model for detecting melanoma as one with the potential to stifle the development of good clinical judgment among junior and inexperienced clinicians and atrophy the skills of already experienced operators if it led to good acumen being too regularly bypassed in favor of efficiency. With this in mind, it may be beneficial to investigate how the development of training programs around the use of AI in dermatology as well as clinical guidelines on AI use may take our findings into account. In recent years, clinical groups have published position statements designed to inform dermatologists on the appropriate use of AI. For example, the Australasian College of Dermatologists has outlined out a number of recommendations for AI adopters designed to address commonly seen issues in the application of AI in medical settings (eg, privacy, a desire for transparency in output and training data, and the need for evidence of accuracy and validity) [42]. They recommend dermatologists develop basic knowledge and skills in the use of AI, such as "appropriate use," understand that "output from AI models can produce false-positive and false-negative results," and that their "decision making may be biased by using AI." These very broad-level recommendations could be extended by considering how the different conceptualizations of AI described in this study reveal what clinicians mean by good decision-making in the context of AI use and what our findings indicate, for example, about the concern clinicians have regarding overreliance, deskilling, and maintaining good clinical acumen. In discussing these issues, rather than referring to broad notions of "AI support," it may be more useful to construct more specific recommendations by referring to the constructs we describe here, such as "AI as a tool within decision making," "AI as a checker after decision making or second-opinion on equivocal cases," and "AI used as an expert guide before decision making." This may also improve guidelines on use so that decision-making workflows are sensitive to the desire of clinicians to retain the ability to exercise and develop independent decision-making skills while using AI and also take into account the whole clinical context of the patient.

Given the rapidly evolving nature of AI technologies within dermatology and the health care space more broadly, we suggest several other areas for future research in light of our findings. First, there is a need to evaluate the real-world effects of clinical workflows that position AI in the decision-making process in ways resembling those outlined by participants in this study (eg, "checker," "second opinion," and "expert guide"). For instance, designing and implementing optimal melanoma screening programs will require good validation studies of how human-AI interactions are affected when AI is variously positioned before, after, or during human clinical inspection. Second, it is yet to be determined how the conceptualizations of AI described in this study may translate across melanoma detection workflows that use a range of imaging technologies. For instance, the question arises whether the potential role of AI in good decision-making differs when operating as part of 2D dermoscopic imaging platforms as opposed to 3D total body photography for melanoma screening. The design parameters of some imaging technologies may determine the extent to which some of the AI roles described here are able to be operationalized and how this affects decision-making needs to be better understood. Third, experimental work can provide evidence for how the positioning of AI may impact potential deskilling of experts or possible delayed skill acquisition of novices. This kind of experimental work may also yield insights for developing effective ways to allow a human-AI feedback loop to occur in real-time decision-making as a way of increasing the explainability of decisions. Fourth, there is considerable scope for further qualitative and quantitative research to better understand how the conceptualizations of clinicians described in this study accord with the views of consumers, particularly patients with high risk of melanoma who are likely to be a priority population for melanoma screening and lesion monitoring.

Limitations

While there was diversity in experiences in practice settings among interviewees, overall integration of AI into everyday clinical use remains uncommon. Most of the sample (n=25, 83%) in this study reported having some experience with AI tools for melanoma detection, but currently no AI tool for the diagnosis of melanoma based on dermoscopic images has been approved. Therefore, views on AI for this purpose are based on individual field experience with unapproved tools, testing, or reader studies. Practical experience, testing, and comparison of

various tools in a clinical setting would likely provide additional insights not captured in this study. AI tools are constantly evolving, and some views may be based on early AI tools that are still in development (eg, those with experience through research). Our purposive recruitment was done so that we could elicit, where possible, reflections on the way participants may have already interacted with AI tools when making decisions rather than only form views about hypothetical situations. This is a strength of the study as participants were not reliant on speculating about hypothetical situations; however, we do also acknowledge that the experience of the sample may not translate to all clinicians, particularly those in other health care settings or cultures.

Our sample predominantly comprised clinicians from Australia and New Zealand, raising the question of generalizability of the findings to broader health care systems and cultures. Australia and New Zealand have the highest rates of skin cancer worldwide [18], with populations (eg, in Queensland) having very high rates of sun exposure. Australia has implemented a decades-long public health campaign devoted to fostering sun-protective behaviors. While there is no coordinated widespread melanoma screening program, a shared public-private health care funding model means there is a very wide coverage of consumers seeking screening through opportunistic skin checks conducted by dermatologists, skin cancer clinics, GPs (including those with a special interest in skin cancer), and other clinicians such as melanographers. As such, clinicians working in melanoma detection in Australia and New Zealand (as well as general practice clinicians not specializing in skin cancer) are highly familiar with examining many consumers with highly sun-damaged skin and see many types of skin cancer. This familiarity may mean that a high level of importance is placed on good clinical acumen when it comes to melanoma detection, and they perhaps feel less inclined to rely on AI. The use of AI as an expert guide may be more acceptable to clinicians in other health care systems with less experienced clinicians, lower rates of melanoma, or where consumers have less access to health care.

Research related to AI-clinician collaborations for melanoma detection has, to date, understandably often focused on the decision-making of dermatologists; therefore, a particular strength of this study is the elicitation of views on AI from GPs and melanographers (in addition to dermatologists), given that they too conduct skin checks, identify lesion change, or make diagnostic decisions (GPs). Notably, there was considerable shared meaning across groups, likely due to the shared understanding of the clinical decision-making process and what entailed “good” decision-making. However, further research is needed to draw out potential differences across these groups, particularly in relation to the use of specific AI tools that are developed for use in practice.

Conclusions

Clinicians described their conceptualizations of AI in melanoma detection in ways that prioritize the conservation of good clinical acumen, and this must be a priority when developing and adopting AI into the decision-making process. This has implications for who is likely to be the most appropriate user of AI given its limited contextual awareness, and careful consideration must therefore be given to how (and if) AI is adopted in the clinical setting once AI tools are formally approved by the respective authorities. Our study implores a more focused engagement with users about the precise way, and in what position, they envisage AI being incorporated into their decision-making process for melanoma detection.

Acknowledgments

The authors would like to thank Lara Wild, Tabi Ward, Jake Morrill, and all participants, particularly those who helped them recruit through their own clinical networks. This study was supported by funding from a National Health and Medical Research Council (NHMRC) Synergy Grant titled “Roadmap Options for Melanoma Screening in Australia (Melanoma-ROSA).” MJ is funded by an NHMRC Fellowship (ID 2034422).

Conflicts of Interest

HPS is a shareholder of MoleMap NZ Limited and e-derm-consult GmbH and undertakes regular teledermatological reporting for both companies. HPS is also a medical consultant for Canfield Scientific Inc and Blaze Bioscience Inc and a medical adviser for First Derm. VM has received speaker fees from Novartis, Bristol Myers Squibb, Merck and Janssen, conference travel support from L’Oreal, grant co-funding from MoleMap for a clinical trial, and has participated in Advisory Boards for MSD, L’Oreal and SkylineDx. All other authors declare no conflict of interest.

Multimedia Appendix 1
Outline of themes and additional data excerpts.
[DOCX File , 19 KB - derma_v8i1e63923_app1.docx]

References

1. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Feb 02;542(7639):115-118 [FREE Full text] [doi: 10.1038/nature21056] [Medline: 28117445]

2. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Reader Study Level-I and Level-II Groups, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018 Aug 01;29(8):1836-1842 [FREE Full text] [doi: 10.1093/annonc/mdy166] [Medline: 29846502]
3. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, et al. Deep neural networks are superior to dermatologists in melanoma image classification. Eur J Cancer 2019 Sep;119:11-17 [FREE Full text] [doi: 10.1016/j.ejca.2019.05.023] [Medline: 31401469]
4. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol 2019 Jan 01;155(1):58-65 [FREE Full text] [doi: 10.1001/jamadermatol.2018.4378] [Medline: 30484822]
5. Felmingham C, Pan Y, Kok Y, Kelly J, Gin D, Nguyen J, et al. Improving skin cancer management with ARTificial intelligence: a pre-post intervention trial of an artificial intelligence system used as a diagnostic aid for skin cancer management in a real-world specialist dermatology setting. J Am Acad Dermatol 2023 May;88(5):1138-1142 [FREE Full text] [doi: 10.1016/j.jaad.2022.10.038] [Medline: 36306873]
6. Thomas L, Hyde C, Mullarkey D, Greenhalgh J, Kalsi D, Ko J. Real-world post-deployment performance of a novel machine learning-based digital health technology for skin lesion assessment and suggestions for post-market surveillance. Front Med (Lausanne) 2023 Oct 31;10:1264846 [FREE Full text] [doi: 10.3389/fmed.2023.1264846] [Medline: 38020164]
7. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. Nat Med 2020 Aug 22;26(8):1229-1234 [FREE Full text] [doi: 10.1038/s41591-020-0942-0] [Medline: 32572267]
8. Al-Ali F, Polesie S, Paoli J, Aljasser M, Salah L. Attitudes towards artificial intelligence among dermatologists working in Saudi Arabia. Dermatol Pract Concept 2023 Jan 01;13(1):e2023035 [FREE Full text] [doi: 10.5826/dpc.1301a35] [Medline: 36892376]
9. Nelson CA, Pachauri S, Balk R, Miller J, Theunis R, Ko JM, et al. Dermatologists' perspectives on artificial intelligence and augmented intelligence - a cross-sectional survey. JAMA Dermatol 2021 Jul 01;157(7):871-874 [FREE Full text] [doi: 10.1001/jamadermatol.2021.1685] [Medline: 34037674]
10. Patrzyk S, Bielecki W, Woźniacka A. A study of attitudes among Polish dermatologists and dermatology trainees regarding modern technologies in medicine. Postepy Dermatol Alergol 2022 Jun;39(3):531-537 [FREE Full text] [doi: 10.5114/ada.2022.117738] [Medline: 35950138]
11. Polesie S, Gillstedt M, Kittler H, Lallas A, Tschandl P, Zalaudek I, et al. Attitudes towards artificial intelligence within dermatology: an international online survey. Br J Dermatol 2020 Jul 12;183(1):159-161. [doi: 10.1111/bjd.18875] [Medline: 31953854]
12. Scheetz J, Rothschild P, McGuinness M, Hadoux X, Soyer HP, Janda M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. Sci Rep 2021 Mar 04;11(1):5193 [FREE Full text] [doi: 10.1038/s41598-021-84698-5] [Medline: 33664367]
13. Shen C, Li C, Xu F, Wang Z, Shen X, Gao J, et al. Web-based study on Chinese dermatologists' attitudes towards artificial intelligence. Ann Transl Med 2020 Jun;8(11):698 [FREE Full text] [doi: 10.21037/atm.2019.12.102] [Medline: 32617318]
14. Wei C, Adusumilli NC, Friedman AJ, Patel VA. Perceptions of artificial intelligence integration into dermatology clinical practice: a cross-sectional survey study. J Drugs Dermatol 2022 Feb 01;21(2):135-140. [doi: 10.36849/jdd.6398] [Medline: 35133109]
15. Samaran R, L'Orphelin JM, Dreno B, Rat C, Dompmartin A. Interest in artificial intelligence for the diagnosis of non-melanoma skin cancer: a survey among French general practitioners. Eur J Dermatol 2021 Aug 01;31(4):457-462. [doi: 10.1684/ejd.2021.4090] [Medline: 34642138]
16. Sangers TE, Wakkee M, Moolenburgh FJ, Nijsten T, Lugtenberg M. Towards successful implementation of artificial intelligence in skin cancer care: a qualitative study exploring the views of dermatologists and general practitioners. Arch Dermatol Res 2023 Jul 07;315(5):1187-1195 [FREE Full text] [doi: 10.1007/s00403-022-02492-3] [Medline: 36477587]
17. 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 [FREE Full text] [doi: 10.1016/j.jid.2023.10.004] [Medline: 37978982]
18. Mar VJ, Soyer HP. Artificial intelligence for melanoma diagnosis: how can we deliver on the promise? Ann Oncol 2019 Dec 01;30(12):e1-e3 [FREE Full text] [doi: 10.1093/annonc/mdy191] [Medline: 29790922]
19. Arnold M, Singh D, Laversanne M, Vignat J, Vaccarella S, Meheus F, et al. Global burden of cutaneous melanoma in 2020 and projections to 2040. JAMA Dermatol 2022 May 01;158(5):495-503 [FREE Full text] [doi: 10.1001/jamadermatol.2022.0160] [Medline: 35353115]
20. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. Qual Res Psychol 2015 Apr 03;12(2):202-222 [FREE Full text] [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
21. King N, Brooks J, Tabari S. Template analysis in business and management research. In: Ciesielska M, Jemielniak D, editors. Qualitative Methodologies in Organization Studies. Cham, Switzerland: Palgrave Macmillan; 2017.

22. Goessinger EV, Niederfeilner JC, Cerminara S, Maul JT, Kostner L, Kunz M, et al. Patient and dermatologists' perspectives on augmented intelligence for melanoma screening: a prospective study. J Eur Acad Dermatol Venereol 2024 Dec 27;38(12):2240-2249. [doi: 10.1111/jdv.19905] [Medline: 38411348]
23. Li Y, Rotemberg V. From promise to practice: artificial intelligence in skin cancer screenings. J Eur Acad Dermatol Venereol 2024 Dec 25;38(12):2203-2204. [doi: 10.1111/jdv.20376] [Medline: 39582474]
24. Menzies SW, Sinz C, Menzies M, Lo SN, Yolland W, Lingohr J, et al. Comparison of humans versus mobile phone-powered artificial intelligence for the diagnosis and management of pigmented skin cancer in secondary care: a multicentre, prospective, diagnostic, clinical trial. Lancet Digit Health 2023 Oct;5(10):e679-e691. [doi: 10.1016/s2589-7500(23)00130-9]
25. Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. Lancet Digit Health 2022 May;4(5):e330-e339. [doi: 10.1016/s2589-7500(22)00021-8]
26. Janda M, Soyer HP. Can clinical decision making be enhanced by artificial intelligence? Br J Dermatol 2019 Feb 04;180(2):247-248. [doi: 10.1111/bjd.17110] [Medline: 30714102]
27. Miller T. Explainable AI is dead, long live explainable AI!: hypothesis-driven decision support using evaluative AI. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. 2023 Presented at: FAccT '23; June 12-15, 2023; Chicago, IL. [doi: 10.1145/3593013.3594001]
28. Lebovitz S, Lifshitz-Assaf H, Levina N. To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. Organ Sci 2022 Jan;33(1):126-148. [doi: 10.1287/orsc.2021.1549]
29. Kelly S, Kaye SA, Oviedo-Trespalacios O. What factors contribute to the acceptance of artificial intelligence? A systematic review. Telemat Inform 2023 Feb;77:101925. [doi: 10.1016/j.tele.2022.101925]
30. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
31. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
32. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
33. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e18072 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]
34. Chen M, Zhou AE, Jain N, Gronbeck C, Feng H, Grant-Kels JM. Ethics of artificial intelligence in dermatology. Clin Dermatol 2024 May;42(3):313-316. [doi: 10.1016/j.clindermatol.2024.02.003] [Medline: 38401700]
35. Lim K, Neal-Smith G, Mitchell C, Xerri J, Chuanromanee P. Perceptions of the use of artificial intelligence in the diagnosis of skin cancer: an outpatient survey. Clin Exp Dermatol 2022 Mar 23;47(3):542-546. [doi: 10.1111/ced.14969] [Medline: 34610153]
36. Haggenmüller S, Maron RC, Hekler A, Krieghoff-Henning E, Utikal JS, Gaiser M, et al. Patients' and dermatologists' preferences in artificial intelligence-driven skin cancer diagnostics: a prospective multicentric survey study. J Am Acad Dermatol 2024 Aug;91(2):366-370 [FREE Full text] [doi: 10.1016/j.jaad.2024.04.033] [Medline: 38670313]
37. Wu A, Ngo M, Thomas C. Assessment of patient perceptions of artificial intelligence use in dermatology: a cross-sectional survey. Skin Res Technol 2024 Mar 13;30(3):e13656 [FREE Full text] [doi: 10.1111/srt.13656] [Medline: 38481072]
38. Janda M, Cust AE, Neale RE, Aitken JF, Baade PD, Green AC, et al. Early detection of melanoma: a consensus report from the Australian Skin and Skin Cancer Research Centre Melanoma Screening Summit. Aust N Z J Public Health 2020 Apr;44(2):111-115 [FREE Full text] [doi: 10.1111/1753-6405.12972] [Medline: 32190955]
39. Aquino YS, Rogers WA, Braunack-Mayer A, Frazer H, Win KT, Houssami N, et al. Utopia versus dystopia: professional perspectives on the impact of healthcare artificial intelligence on clinical roles and skills. Int J Med Inform 2023 Jan;169:104903 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104903] [Medline: 36343512]
40. Oschinski M, Crawford A, Wu M. AI and the future of workforce training. Center for Security and Emerging Technology. 2024 Dec. URL: https://cset.georgetown.edu/publication/ai-and-the-future-of-workforce-training/ [accessed 2025-03-07]
41. Tsang M, Martin L, Blissett S, Gauthier S, Ahmed Z, Muhammed D, et al. What do clinicians mean by “good clinical judgment”: a qualitative study. Int Med Educ 2023 Jan 11;2(1):1-10. [doi: 10.3390/ime2010001]
42. Caffery LJ, Janda M, Miller R, Abbott LM, Arnold C, Caccetta T, et al. Informing a position statement on the use of artificial intelligence in dermatology in Australia. Australas J Dermatol 2023 Feb 15;64(1):e11-e20. [doi: 10.1111/ajd.13946] [Medline: 36380357]

Abbreviations

AI: artificial intelligence
DERM: Deep Ensemble for Recognition of Malignancy
GP: general practitioner

Edited by CC Chi; submitted 11.07.24; peer-reviewed by YH Chen, M Popovic, PH Huang; comments to author 23.10.24; revised version received 14.01.25; accepted 03.02.25; published 24.03.25.

Please cite as:
Partridge B, Gillespie N, Soyer HP, Mar V, Janda M
Exploring the Views of Dermatologists, General Practitioners, and Melanographers on the Use of AI Tools in the Context of Good Decision-Making When Detecting Melanoma: Qualitative Interview Study
JMIR Dermatol 2025;8:e63923
URL: https://derma.jmir.org/2025/1/e63923
doi:10.2196/63923
PMID:

Implementing Teledermoscopy to Shorten Doctors' Diagnostic Process for Suspected Skin Cancer: Observational Pilot Study

Rebekka Nordahl Larsen[1], MD; Niels Kvorning Ternov[2], MD, PhD; Kristian Kidholm[3], PhD; Tine Vestergaard[1], MD, PhD

[1]Department of Dermatology and Allergy Centre, Odense University Hospital, Kløvervænget 15, Odense, Denmark

[2]Department of Plastic and Reconstructive Surgery, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark

[3]Department of Clinical Research, Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark

Corresponding Author:
Rebekka Nordahl Larsen, MD
Department of Dermatology and Allergy Centre, Odense University Hospital, Kløvervænget 15, Odense, Denmark

Abstract

Background: Skin cancers are the most frequent types of cancer, and the incidence continues to rise. Teledermoscopy is a promising tool in the diagnostic process of potential skin cancer, and new technologies are constantly being developed. However, little information is available on how teledermoscopy affects physicians' time consumption.

Objective: This study aimed to investigate whether teledermoscopy can shorten the diagnostic process for physicians examining skin lesions suspected of skin cancer.

Methods: We recorded the time primary care providers, dermatologists, plastic surgeons, and pathologists spent examining lesions suspected of skin cancer, with and without teledermoscopy. Furthermore, we looked at five different diagnostic pathways, which reflected the most common ways through the Danish health care system for patients with suspected skin cancer, to estimate the total amount of time physicians spent examining these lesions with and without teledermoscopy.

Results: A total of 118 time recordings were obtained. With teledermoscopy, the diagnostic process was significantly shortened for dermatologists (P=.008) but prolonged for primary care providers (P=.03). While the use of teledermoscopy saved time in one of the diagnostic pathways, it increased the time spent in the four others.

Conclusions: Our research suggests that the implementation of teledermoscopy could save time for dermatologists and potentially plastic surgeons and pathologists, provided that a sufficient number of benign skin lesions can be accurately diagnosed and excluded from further examination and treatment. In contrast, the implementation of teledermoscopy might prolong primary care providers' consultation time.

(JMIR Dermatol 2025;8:e66782) doi:10.2196/66782

KEYWORDS

skin cancers; teledermatology; teledermoscopy; malignant melanoma; nonmelanoma skin cancers; time

Introduction

Background

Skin cancers can be categorized into melanoma and nonmelanoma skin cancers (NMSCs) and are the most frequent types of cancer in Denmark and globally [1,2]. Recent data have shown that NMSC now results in a higher total number of deaths worldwide compared to malignant melanoma (MM), although the individual mortality risk is still higher for MM [3]. For the past decade, there has been a continuous increase in the number of NMSC and MM cases in Denmark [4]. This places a greater burden on a health care system already under pressure from an ageing population, a shortage of health care workers, and the continuous introduction of new, costly treatments [5,6].

In Denmark, primary care providers (PCPs) serve as gatekeepers for all dermatological concerns including skin cancer suspicions. PCPs encounter a broad variety of skin diseases, and studies have suggested that over half of the potentially malignant skin lesions referred to specialists are later diagnosed as benign [7-9]. This could indicate the need for an additional filter function, such as teledermoscopy, to better use the time and resources of health care workers and patients.

Teledermoscopy is a technology within the field of teledermatology where dermoscopic images of skin lesions can be referred to a dermatologist for evaluation [10]. The overall diagnostic accuracy of teledermatology is high, and results from recent studies indicate that accuracy and sensitivity are improving with the emergence of new technology [11]. Some of the advantages of teledermatology include significantly shortened waiting periods for patients and a reduction in the

number of referrals for in-person consultations [11-13]. In addition, studies have found that PCPs, dermatologists, and patients have a general positive approach to the implementation of teledermoscopy [14,15]. Still, a few barriers exist when it comes to the implementation of teledermatology in health care systems. Among other things, physicians have reported concerns about lack of clinical information, technical issues, increased workload, and time consumption [13,16].

Only a few studies have measured the time required for consultations using teledermoscopy or teledermatology, and most only report time spent on completing or evaluating telereferrals as additional data points [12,17-19]. Since the area of teledermoscopy is constantly evolving, we wanted to look into a Danish-developed teledermoscopy platform, in order to see the impact and potential gain of implementation for physicians in terms of time consumption.

Research Objectives

The overall purpose of this study was to investigate whether teledermoscopy can shorten the diagnostic process for physicians when examining skin lesions suspected of skin cancer.

First, we estimated the amount of time spent investigating these lesions with and without teledermoscopy for PCPs and dermatologists. Then time measurements were performed without teledermoscopy for plastic surgeons and pathologists.

Second, in order to compare the total amount of time physicians spend triaging, diagnosing, and treating skin cancer with and without teledermoscopy, we aimed to construct different diagnostic pathways, reflecting the most common ways through the Danish health care system. Time spent by patients and other staff was not included.

Methods

Ethical Considerations

This study did not interfere with the patients' potential treatment for skin cancer. It was conducted as a quality assurance project, which does not need an ethical review by the regional or national research ethics committee [20]. This study was approved by the Danish Data Protection Agency (reference 23/19285).

Clinical Locations

Time recordings were carried out at multiple locations in the Region of Southern Denmark and the Capitol Region of Denmark. Locations included 2 different primary care centers, 2 different private practice dermatologists, the department of pathology at Herlev Hospital, and the departments of dermatology, plastic surgery, and pathology at Odense University Hospital (OUH).

Data Collection

Before data collection, staff members from the primary care centers, dermatologist clinics, and the department of plastic surgery had reviewed their schedules in order to trace patients with potential skin cancer. During the designated data collection period, the investigator (RNL) performed the time recordings on days when these patients were scheduled for examination or removal of the skin lesion. We randomly picked the days for data collection at the department of pathology. The dermatopathologists selected the specimens with potential to be skin cancer.

We intended to collect a minimum of 10 recordings at primary care centers and dermatologists with and without teledermoscopy. Furthermore, we aimed to gather a minimum of 10 recordings at the department of plastic surgery and the department of pathology without teledermoscopy. Data were stored in Microsoft SharePoint.

Inclusion Criteria

At primary care centers, the eligibility criteria included patients presenting at their PCP with concerns about potential skin cancer, regardless of the PCP's final assessment as malignant or benign. In dermatologist clinics, the eligibility criteria included patients referred by PCPs for suspected skin cancer. At the department of plastic surgery, the eligibility criteria included patients referred by PCPs or dermatologists for skin tumor excision. At the department of pathology, the eligibility criteria included patients with skin lesions suspected of skin cancer, referred by either PCPs, dermatologists, or plastic surgeons.

Exclusion Criteria

At primary care centers, the eligibility criteria excluded patients presenting with inflammatory skin diseases or follow-up of previous examined benign lesions. In dermatologist clinics, the eligibility criteria excluded physical referrals of patients with more than one skin tumor. At the department of plastic surgery, the eligibility criteria excluded re-excisions and excisions with skin grafts. At the department of pathology, the eligibility criteria excluded re-excisions and biopsy-verified lesions.

The Teledermoscopy Platform

In this study, teledermoscopy was performed using the teledermoscopy platform Dermloop (Melatech ApS).

Primary Care Providers

PCPs used the Dermloop Capture app (version 2.31; MelaTech IVS on an Apple iPhone 11). Each PCP had a personal account they had to log into upon entering the app. In order to send a telereferral to a dermatologist, the PCP had to state the patient's social security number, answer a short questionnaire about the skin lesion (skin type, recent changes in appearance, tentative clinical diagnosis, risk factors, etc), and take an overview as well as a dermoscopic photograph of the lesion with an attachable dermoscope (Handyscope, Dermlite).

Dermatologists

Dermatologists used Dermloop Desktop (MelaTech) on their computers. To be able to see telereferrals sent by PCPs, dermatologists had to log in to their personal accounts. When logged in to the platform, a list of pending and diagnosed skin lesions appeared on the screen. The dermatologists could then click on the pending skin lesions and examine the image, evaluate the quality of the photograph, assess the difficulty of diagnosing the skin lesions, give a diagnosis, and describe a potential treatment and further plan. After submission, PCPs received an email with the dermatologist´s response.

Time Recordings

Overview of Time Recordings

Time recordings were carried out with a digital stopwatch. The number of skin lesions examined and notes about the circumstances (eg, interruptions) were made to detect potential differences that could have an impact on the time recordings. To make time recordings as uniform as possible, we planned that all recordings would be performed by the same person (RNL). Before this project, physicians had been trained in the use of the teledermoscopy platform as a part of other projects.

Primary Care Provider

Time recordings were carried out at consultations between a PCP and a patient.

Recordings without teledermoscopy included obtaining a medical history about the skin lesion, small talk, an objective clinical examination with or without dermoscopy and informing the patient about the tentative diagnosis, and further plan. Recordings with teledermoscopy included the same tasks plus photography using the teledermoscopy equipment. Before data collection, PCPs using the teledermoscopy equipment had sent between 4 and 19 telereferrals.

Dermatologist

Time recordings without teledermoscopy were carried out at consultations between a dermatologist and a patient referred from a PCP with suspected skin cancer.

Recordings included obtaining a medical history about the skin lesion, small talk, a clinical examination, including dermoscopy, and informing the patient about the tentative diagnosis and further plan.

Time recordings with teledermoscopy only included dermatologists evaluating telereferrals sent from PCPs on the teledermoscopy platform. Recordings included reading about the skin lesion (potential diagnosis, risk factors, etc), examining the lesion based on an overview image and a dermoscopic image, evaluating the quality of the images (clinical and dermoscopic image), and writing a tentative diagnosis and further plan in the teledermoscopy platform.

The Department of Plastic Surgery

The time recordings were carried out in an operating room. The procedure was performed by a plastic surgeon with help from a nurse while the patient was under local anesthesia.

Recordings included obtaining a short medical history about the skin lesion, examining the skin lesion, explaining the procedure, preparing for the procedure (glasses, mask, sterilized gloves, and surgical gown), excision of the skin lesion, stitching, bandaging, and providing the patient with information about precautions and potential side effects of the procedure.

The Department of Pathology

The time recordings were carried out during histopathological evaluations of skin lesions by pathologists at the department of pathology.

The 2 departments included in our study used different types of microscopes in their histopathological evaluation. Tissue sections at Herlev Hospital were examined with regular light microscopes, whereas tissue sections at OUH were examined on a diagnostic computer screen. At OUH, all slides with tissue sections are scanned by “whole slide scanners” and obtained in an “image management system.” The pathologist can access high-resolution microscopic images of the scanned tissue sections in the image management system on their diagnostic screen and examine the skin lesion [21].

Recordings included reading about the removed skin lesion in the referral, examining the skin lesion (regular light microscopy or diagnostic screen), and describing findings.

Depending on the difficulty of diagnosing the skin lesions, additional tasks, such as reviewing the patient’s electronic medical record or discussing findings with a colleague, might be included in the recordings.

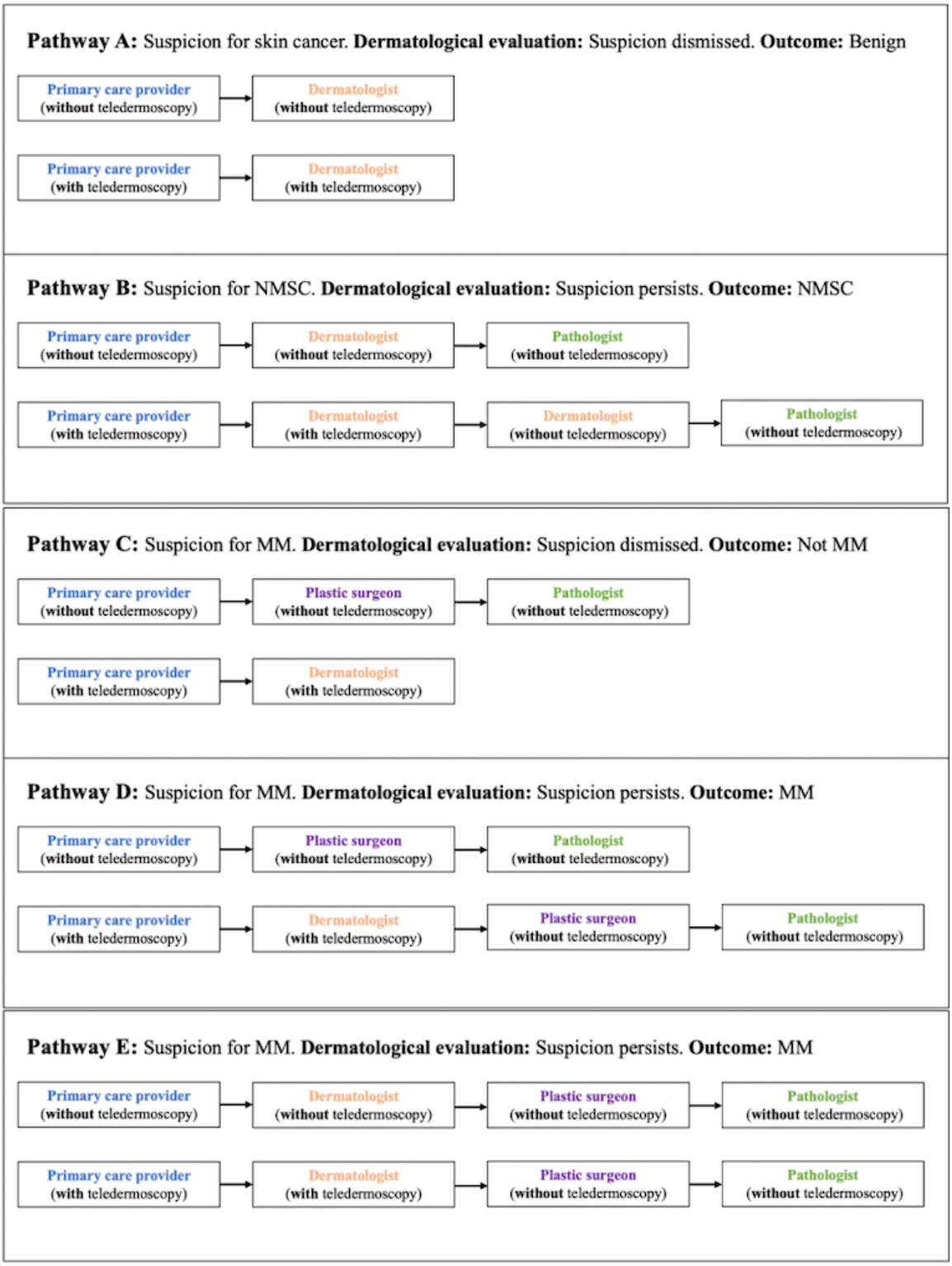
In some cases, only a part of the histopathological evaluation was measured, either because further immunohistochemical staining was required or because the tissue had already gone through this process. Immunohistochemical staining takes approximately one and a half days, which made it difficult to be at the department for both evaluations.

Construction of Diagnostic Pathways

In this study setup, where time and data collection resources were limited and skin cancer consultations at PCPs appeared sporadically, it was a major challenge to identify relevant cases and follow them at every step of the diagnostic process. Thus, it was not possible to measure the total time physicians spend examining potential skin cancer for each patient included.

Consequently, we constructed a variety of diagnostic pathways that aimed to reflect real diagnostic processes. For pathways without the use of teledermoscopy, this was done in accordance with guidelines for physicians regarding referral criteria for lesions suspected of skin cancer, published by the Danish Health Authority [22] and the Region of Southern Denmark [23], where so-called cancer patient pathways (CPPs) are described for both MM and NMSC. Pathways with the use of teledermoscopy were constructed based on expertise from a consultant dermatologist (see Figure 1). It must be emphasized that the constructed pathways are simplified and does not cover all possible scenarios. The total amount of time physicians spend examining skin lesions in each of the constructed pathways was calculated by adding together the median amount of time PCPs, dermatologists, plastic surgeons, and pathologists spend on the individual parts of the process.

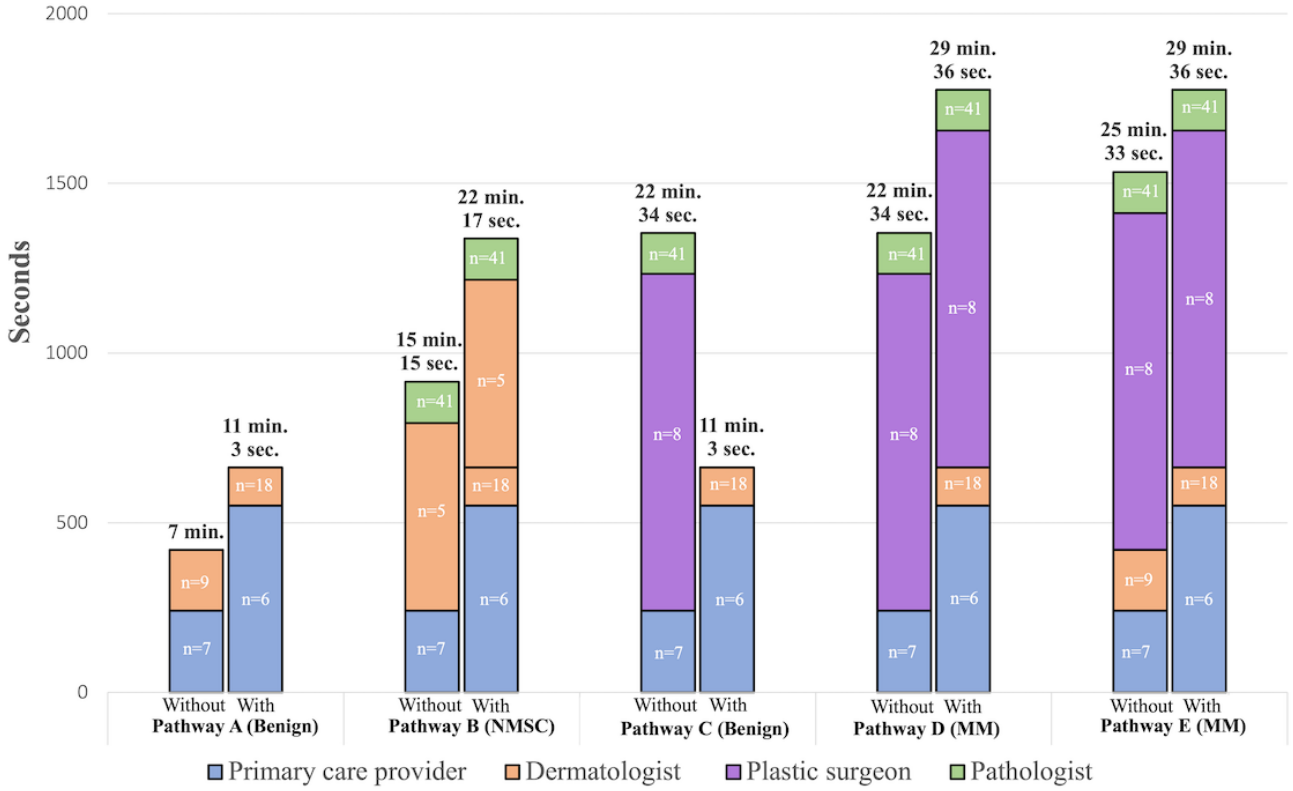
Figure 1. Constructed diagnostic pathways with and without teledermoscopy for patients with suspected skin cancer. Diagnostic pathway A: potential skin cancer. The suspicion is dismissed by the dermatologist, with or without the use of teledermoscopy. Outcome: benign lesion. Diagnostic pathway B: suspected NMSC. Without teledermoscopy: the PCP refers the patient to an in-person consultation with a dermatologist. The dermatologist removes the skin lesion and sends it to a pathologist for evaluation. With teledermoscopy: the suspicion persists, and the patient is referred to an in-person consultation with a dermatologist. The dermatologist removes the skin lesion and sends it to a pathologist for evaluation. Outcome: NMSC. Diagnostic pathway C: suspected MM. Without teledermoscopy: the patient is referred to the department of plastic surgery, where the lesion is removed and sent to a pathologist for evaluation. The pathologist dismisses the suspicion. With teledermoscopy: the suspicion is dismissed by the teledermatologist. Outcome: benign lesion. Diagnostic pathway D: suspected MM. Without teledermoscopy: the patient is referred to the department of plastic surgery, where the lesion is removed and sent to a pathologist for evaluation. The pathologist confirms the suspicion. With teledermoscopy: the suspicion persists, and the patient is referred to the department of plastic surgery, where the lesion is removed and sent to a pathologist for evaluation. The pathologist confirms the suspicion. Outcome: MM. Diagnostic pathway E: suspected MM. Without teledermoscopy: the patient is referred to an in-person consultation at the dermatologist. With teledermoscopy: images of the lesion are sent for evaluation by the teledermatologist. The suspicion persists, and the patient is referred to the department of plastic surgery, where the lesion is removed and sent to a pathologist for evaluation. Outcome: MM. MM: malignant melanoma; NMSC: nonmelanoma skin cancer.



Data Analysis

Stata/BE 18.0 was used to calculate all statistical analyses. For each individual part of the process, the Mann-Whitney U test was used to compare time consumption with and without teledermoscopy. $P<.05$ was considered significant. Data are presented as median (IQR). All time statements are presented in minutes (min) and seconds (s), that is, min:s. For each of the constructed diagnostic pathways (see Figure 1), the calculated median for every individual part of the pathway was added together (see Figure 2). The absolute difference between the constructed pathways with and without teledermoscopy was calculated.

Figure 2. Stacked bar chart showing the amount of time various physicians spend investigating suspected skin cancer for the constructed diagnostic pathways A-E (see legends for Figure 1). The total (median) time physicians spend in each pathway is shown at the top of each column. The height of each segment represents the (median) time physicians in each group spend investigating skin lesions. The diagnostic outcome is shown in brackets. The number of time recordings is shown in each segment. MM: malignant melanoma; NMSC: nonmelanoma skin cancer; min: minutes; sec: seconds; With: with teledermoscopy; Without: without teledermoscopy.



Results

Baseline Characteristics

Data were collected from April 2023 to December 2023. The dataset comprised 118 time recordings and included 17 physicians examining lesions suspected of skin cancer. When asked if it was acceptable for us to perform time recordings during examinations or treatment for potential skin cancer, no patients declined.

Upon evaluation of the data collected, a total of 24 recordings were excluded due to data not fulfilling the inclusion criteria or to various interruptions or errors during the time measurements. The exact number of time recordings excluded at each department was 5, 3, 8, and 8 at the PCP, dermatologists, department of plastic surgery, and department of pathology, respectively. Thus, a total of 94 time recordings were included in this study.

Tasks included in the procedure at the department of plastic surgery, dermatology, and pathology were repeated per lesion. Consequently, time recordings including more than 1 lesion were divided, resulting in each time recording comprising 1 skin lesion.

In order to make time recordings as uniform as possible, the majority of measurements were carried out by the same person (RNL). Due to time constraints, the physician (TV) at the department of dermatology at OUH made some of the recordings on her own (6.8%), following instructions.

Individual Parts of the Process

Table 1 shows the data for the individual parts of the diagnostic processes with and without teledermoscopy for physicians. PCPs spend significantly longer time with teledermoscopy, whereas the diagnostic process was significantly shortened for dermatologists. The absolute median differences with and without teledermoscopy was 05:09 for PCPs and 01:06 for dermatologists.

Table . The amount of time primary care providers, dermatologists, plastic surgeons, and pathologists spend investigating skin lesions suspected of skin cancer with and without teledermoscopy.

Characteristic	Time recordings without teledermoscopy (n=70)	Time recordings with teledermoscopy (n=24)	*P* value
PCP[a]			
Time recordings, n	7	6	—[b]
Patients, n	7	6	—
Skin lesions, n	11	7	—
Number of PCPs involved in the recordings, n	5	5	—
Time, median (IQR)	04:01 (02:41-04:55)[c]	09:10 (08:19-11:59)	.03
Dermatologists (without treatment)			
Time recordings, n	9	18	—
Patients, n	9	18	—
Skin lesions, n	9	18	—
Number of dermatologists involved in the recordings, n	2	2	—
Time, median (IQR)	02:59 (02:05-04:58)	01:53 (01:14-02:56)	.008
Dermatologists (with treatment)			
Time recordings, n	5	—	—
Patients, n	5	—	—
Skin lesions, n	5	—	—
Number of dermatologists involved in the recordings, n	2	—	—
Time, median (IQR)	09:13 (07:13-10:14)	—	—
The department of plastic surgery			
Time recordings, n	8	—	—
Patients, n	8	—	—
Skin lesions, n	8	—	—
Number of plastic surgeons involved in the recordings, n	2	—	—
Time, median (IQR)	16:32 (13:33-18:20)	—	—
The department of pathology			
Time recordings, n	41	—	—
Patients, n	30	—	—
Skin lesions, n	41	—	—
Number of pathologists involved in the recordings, n	4	—	—
Time, median (IQR)	02:01 (01:31-03:11)	—	—

[a]PCP: primary care provider.

[b]Not available.

[c]Presented as min:s, ie, minutes:seconds.

Constructed Diagnostic Pathways

Figure 2 shows the total amount of time physicians consumed at each of the constructed diagnostic pathways. Teledermoscopy shortened diagnostic pathway C with 11:31, whereas pathways A, B, D, and E were prolonged with 04:03, 07:02, 07:02, and 04:03, respectively.

Post Hoc Analyses

Despite the use of different types of microscopes at the two departments of pathology involved, no significant difference between the time recordings at Herlev Hospital and OUH was found (P=.27).

In Denmark, 400.000 dermatological referrals are made per year, out of which about 30% (120,000) deal with suspicion of skin cancer [24,25]. Assuming 75% (90,000) of these referrals are benign and 25% (30,000) are NMSC [7], the amount of time physicians spend in pathway A and B with teledermoscopy can be estimated as shown in Table 2.

Table . Calculated total time consumption in hours per year with and without teledermoscopy for primary care providers, dermatologists, plastic surgeons and pathologists, for all constructed pathways.

		Primary care providers (hours per year)	Dermatologists (hours per year)	Plastic surgeons (hours per year)	Pathologists (hours per year)
Pathway A					
	Without TDS[a]	6025	4475	—[b]	—
	With TDS	13,750	2825	—	—
	Sum	+7725	−1650	—	—
Pathway B					
	Without TDS	2008	4608	—	1008
	With TDS	4583	5550	—	1008
	Sum	+2575	+942	—	0
Pathway C					
	Without TDS	281	—	1156	141
	With TDS	641	132	—	—
	Sum	+360	+132	−1156	−141
Pathway D					
	Without TDS	203	—	837	102
	With TDS	464	95	837	102
	Sum	+261	+95	0	0
Pathway E					
	Without TDS	203	151	837	102
	With TDS	464	95	837	102
	Sum	+261	−56	0	0
All pathways					
	Total amount of time spend with TDS per year	+10,921	–481 to –632	–1156	–141

[a]TDS: teledermoscopy.

[b]Not available.

The most recent report from the Danish Health Authority regarding CPPs for skin lesions suspected of MM stated that in 58% of the cases, the suspicion was dismissed at some point in the diagnostic pathway [9]. The number of patients diagnosed with MM in 2022 was 3038 [4], which corresponds to a total of 7233 CPPs for suspected MM, with 4195 (58%) dismissed cases in 2022 [9]. The amount of time physicians spend in pathway C, D and E with teledermoscopy can be estimated as displayed in Table 2.

Discussion

Principal Findings

The use of teledermoscopy by PCPs prolonged the diagnostic process, whereas the process was significantly shortened for dermatologists. The constructed pathway C (benign lesion) was shortened for physicians with teledermoscopy, while the diagnostic pathways A (benign lesion), B (NMSC), D (MM), and E (MM) were prolonged by 16% - 58%.

Comparison With Previous Work

Primary Care Providers

Only a few studies have looked at the duration of consultations with teledermoscopy at PCPs. Similar to our findings, Berghout et al [17] and Nami et al [18] both reported prolonged consultation times compared to consultations without teledermoscopy. They found an average duration of consultations with teledermoscopy of 11:32 and 19:00, respectively [17,18]. This is longer than the median consultation time of 09:10 (IQR 08:19-11:59) in our study, but a direct comparison is challenging because of variations in the many components forming a typical consultation and the use of different types of teledermoscopy technologies.

Time stamps specifying how much time PCPs spent exclusively on teledermoscopy were not registered during time recordings in our study, yet the time can be estimated to be 05:09, based on the median time used for consultations with (09:10) and without (04:01) teledermoscopy. Nami et al [18] and van Sinderen et al [12] reported a mean and median time consumption solely for teledermoscopy usage of 04:00 and 05:24, respectively. Furthermore, based on the data provided by Berghout et al [17], it can be calculated that the time spent only on teledermoscopy was 06:49 (11:32 minus 04:43). Again, a statistical comparison of these results with our estimations is difficult due to variations in method, technical equipment, and the absence of statistical data, such as SD and range in some of the studies.

In the randomized controlled trial by Berghout et al [17], the authors discovered that in the first consultations, both the duration of the consultation and the task of filling out the telereferral were significantly prolonged compared to the later consultations. This tendency was also observed by van Sinderen et al [12]. During our data collection, 3 out of 5 PCPs included in the teledermoscopy group expressed a lack of confidence in using the teledermoscopy equipment because they had not used it for some time. This may have affected our results and prolonged some of the recordings. In 2 out of 6 of our time recordings with teledermoscopy, the PCPs had to get the teledermoscopy device from a colleague during the consultation, thereby extending the time recordings. This aspect however, might just reflect the current conditions in a real work setting where a number of PCPs in a primary care center share the teledermoscopy equipment.

Despite recordings with teledermoscopy being longer than those without, the duration was still within the time limit of 15 minutes that most PCPs allocate for consultations. On the other hand, it is not uncommon for patients to raise concerns about skin issues as additional topics during consultations, which may necessitate reducing the amount of time spent with teledermoscopy. In our study, all images were obtained by the PCPs themselves. Since the beginning of this project, teledermoscopy has been implemented at numerous primary care centers where the process of obtaining the photos has been assigned to other staff (eg, nurses), which could solve this issue.

Dermatologists

In contrast to the prolonged consultations at PCPs, teledermoscopy significantly shortened the examination of the skin lesions for dermatologists. Other studies reported that evaluation of telereferral by dermatologists lasted between 01:05 and 02:30 [12,18,19], which is consistent with our results. None of the studies compared their measurements with the duration of in-person consultations with a dermatologist.

Pathologists

Given that we only measured a part of the histopathological evaluation, in a number of cases, it is likely that the median time pathologists spend in our study is shorter than the time they typically spend examining these lesions. However, since we used the same calculated median for pathologists in all the constructed diagnostic pathways with and without teledermoscopy, it does not influence the absolute time difference in these pathways.

In Denmark lesions suspected of melanoma are analyzed within 7 days [26], while carcinomas and other skin biopsies may take longer. This aspect is not accounted for in the time recordings.

Constructed Diagnostic Pathways

Alternative diagnostic pathways could have been constructed, which might have changed the outcome. For instance, the simplification of the pathways could overlook larger time differences in more complicated patient courses. However, we argue that the displayed pathways in this study were the most common based on previous experience and the pathways described in the national guidelines.

According to our calculations, dermatologists, plastic surgeons, and pathologists could overall save time with teledermoscopy, while PCPs would spend more time using the teledermoscopy equipment. To provide perspective, PCPs constitute the largest specialty group in Denmark, including about 3500 physicians [27]. Consequently, each PCP would spend slightly more than 3 hours per year or less than 1 minute a day on average, using teledermoscopy. Enhancing PCPs' skills might reduce the amount of time PCPs spends with teledermoscopy further.

In pathway B, the duration of the physical consultation at the dermatologist might be reduced, as the dermatologist had previously examined the lesion using teledermoscopy, and therefore, only needed to treat the lesion. Similarly, information regarding the size and placement of the skin lesion obtained with teledermoscopy could be beneficial in the planning of the surgical procedure at the department of plastic surgery, thereby making a physical preliminary examination of the skin lesion unnecessary in pathways D and E.

It is important to notice, that the focus of this study was solely on physicians "hands-on" time consumption. Hence, additional time required for tasks, such as preparation time and tissue slide collection, was not included. Furthermore, because most diagnostic pathways involve secretaries, nurses, and medical laboratory technicians, preventing benign skin lesions from unnecessary treatment might save time for other medical staff as well.

Patient resources are also crucial when considering the implementation of teledermoscopy. While teledermoscopy might save time for specific groups of physicians, we did not consider the impact on patients time or the overall duration of the diagnostic process for patients. Fortunately, other studies have looked into this. A review by Jones and Oakley [11] found that the majority of studies investigating time outcomes for teledermatology reported reduced time from referrals to biopsy and treatment with teledermatology. The study by van Sinderen et al [12] reviewing 11 years of teledermoscopy in the Netherlands, found that the median response time for teledermatologists was 2.4 hours compared to an average waiting time of 2.8 weeks for in-person consultations. Furthermore, several studies reported that the number of face-to-face consultations with dermatologists was reduced with teledermoscopy [12,28,29]. Consequently, fewer patients and accompanying persons might have to take time off from work to attend in-person consultations [8].

Strengths and Limitations

This study has several limitations. First, due to limited resources and different geographic locations, it was not possible to follow each patient throughout the diagnostic process. This resulted in fragmented measurements of physician's time consumption, involving independent groups of patients. Second, identifying relevant cases was difficult, especially at the PCPs, because patients have a tendency to bring up skin problems as supplementary subjects during consultations. Consequently, we were unable to achieve the intended number of patients and most time recordings were performed on patients with benign skin tumors. Third, our data are very heterogeneous due to the nature of the different departments included and the large spectrum of potential diagnoses. Fourth, patients with multiple potential skin cancers were not included in this study. Finally, this study was mainly based on structures in the Danish health care system. Some of our results might not apply directly to other health care systems.

This is one of few studies measuring clinicians' time, as suggested by a recent review [30]. A strength of the study is that all time recordings were obtained in a real work setting, which gives the results high external validity. Furthermore, the same person performed the majority of the recordings.

Further Research

More time recordings would be required to obtain a more accurate picture of physicians' time consumption, for example, by including more departments and primary care centers, increasing the number of data collectors, or involving medical staff at the departments to a greater extent. In addition, it would increase data transparency if time stamps stating how much time PCPs and pathologists spend on each task were registered during the recordings. Altogether, several modifications to this pilot study's design would be required to make it feasible for a larger-scale investigation.

Conclusion

Teledermoscopy significantly shortened the diagnostic process for dermatologists but prolonged it for PCPs. Previous studies have concluded that teledermoscopy reduces the number of physical referrals as well as surgical procedures, and our time study indicates that this would altogether result in time saving for dermatologists, plastic surgeons, and pathologists. Part of the time savings attained might be reallocated to other areas of the health care system, such as consultations at PCPs.

Acknowledgments

We would like to thank Jonas Strunge, Christian Tørring, Rune Lindberg-Larsen, Jakob Torp Madsen, Vibeke Vestergaard, Birgitte Bols, and Vivi Bakholdt for making the data collection possible. We would also like to thank all participating patients, doctors, and other staff at the involved hospital departments and clinics.

Conflicts of Interest

NKT is the cofounder and chief executive officer of the company called Melatech that provides the eConsult platform, Dermloop, which was used by clinicians during the time studies in this paper. None declared by other authors.

References

1. Skin cancer statistics. World Cancer Research Fund International. URL: https://www.wcrf.org/cancer-trends/skin-cancer-statistics [accessed 2024-09-01]
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. Global cancer observatory: cancer today. International Agency for Research on Cancer. 2024. URL: https://gco.iarc.who.int/today [accessed 2025-01-29]
4. Nye kræfttilfælde i Danmark 2022 [Article in Danish]. The Danish Health Data Authority. 2023. URL: https://sundhedsdatastyrelsen.dk/-/media/sds/filer/find-tal-og-analyser/sygdomme-og-behandlinger/kraeft/kraeft_nye_tilfaelde_aarsrapporter/kraefttilfaelde-2022.pdf [accessed 2024-01-11]
5. Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Ioannides D. Epidemiological trends in skin cancer. Dermatol Pract Concept 2017 Apr;7(2):1-6. [doi: 10.5826/dpc.0702a01] [Medline: 28515985]

6. Steenberger A. Sundhedsøkonom: ingen partier har løsningen på sundhedsvæsenets største problem [Article in Danish]. Ugeskrift for Læger. 2022. URL: https://ugeskriftet.dk/nyhed/sundhedsokonom-ingen-partier-har-losningen-pa-sundhedsvaesenets-storste-problem [accessed 2025-05-23]
7. Vestergaard T, Prasad SC, Schuster A, Laurinaviciene R, Andersen MK, Bygum A. Diagnostic accuracy and interobserver concordance: teledermoscopy of 600 suspicious skin lesions in Southern Denmark. J Eur Acad Dermatol Venereol 2020 Jul;34(7):1601-1608. [doi: 10.1111/jdv.16275] [Medline: 32031277]
8. Vestergaard T, Andersen MK, Kidholm K. A cost-minimization analysis comparing teledermoscopy and face-to-face evaluations of suspicious skin lesions in Southern Denmark. J Telemed Telecare 2024 May;30(4):661-667. [doi: 10.1177/1357633X221077864] [Medline: 35139672]
9. Monitorering af forløbstider pa˚ kræftomra˚det - a˚rsopgørelse for 2022 [Article in Danish]. Danish Health Authority. 2023. URL: https://www.sundhedsstyrelsen.dk/-/media/Udgivelser/2023/Kraeft/Monitorering-af-forloebstider-AArsopgoerelse-for-2022-_endelig_.ashx?sc_lang=da&hash=6BC4B6EB48D61E5C1C75EEE51299339D [accessed 2024-01-04]
10. Deda LC, Goldberg RH, Jamerson TA, Lee I, Tejasvi T. Dermoscopy practice guidelines for use in telemedicine. NPJ Digit Med 2022 Apr 27;5(1):55. [doi: 10.1038/s41746-022-00587-9] [Medline: 35477979]
11. Jones LK, Oakley A. Store-and-forward teledermatology for assessing skin cancer in 2023: literature review. JMIR Dermatol 2023 May 17;6:e43395. [doi: 10.2196/43395] [Medline: 37632914]
12. van Sinderen F, Tensen E, Lansink RA, Jaspers MW, Peute LW. Eleven years of teledermoscopy in the Netherlands: a retrospective quality and performance analysis of 18,738 consultations. J Telemed Telecare 2024 Jul;30(6):1037-1046. [doi: 10.1177/1357633X221122113] [Medline: 36052405]
13. Lee KJ, Finnane A, Soyer HP. Recent trends in teledermatology and teledermoscopy. Dermatol Pract Concept 2018 Jul;8(3):214-223. [doi: 10.5826/dpc.0803a13] [Medline: 30116667]
14. Vestergaard T, Andersen MK, Bygum A. Acceptance of teledermoscopy by general practitioners and dermatologists in Denmark. Dermatol Pract Concept 2021 Mar;11(2):e2021033. [doi: 10.5826/dpc.1102a33] [Medline: 33954016]
15. Kong F, Horsham C, Rayner J, et al. Consumer preferences for skin cancer screening using mobile teledermoscopy: a qualitative study. Dermatology (Basel) 2020;236(2):97-104. [doi: 10.1159/000505620] [Medline: 32126557]
16. Dovigi E, Kwok EYL, English 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]
17. Berghout RM, Eminović N, de Keizer NF, Birnie E. Evaluation of general practitioner's time investment during a store-and-forward teledermatology consultation. Int J Med Inform 2007 Dec;76 Suppl 3:S384-S391. [doi: 10.1016/j.ijmedinf.2007.04.004] [Medline: 17532256]
18. Nami N, Massone C, Rubegni P, Cevenini G, Fimiani M, Hofmann-Wellenhof R. Concordance and time estimation of store-and-forward mobile teledermatology compared to classical face-to-face consultation. Acta Derm Venereol 2015 Jan;95(1):35-39. [doi: 10.2340/00015555-1876] [Medline: 24889827]
19. Ferrándiz L, Ojeda-Vila T, Corrales A, et al. Internet-based skin cancer screening using clinical images alone or in conjunction with dermoscopic images: a randomized teledermoscopy trial. J Am Acad Dermatol 2017 Apr;76(4):676-682. [doi: 10.1016/j.jaad.2016.10.041] [Medline: 28089728]
20. Hvad skal anmeldes? [Article in Danish]. De videnskabsetiske Komiteer for Region Syddanmark. URL: https://komite.regionsyddanmark.dk/i-tvivl-om-anmeldelse [accessed 2023-12-30]
21. Detlefsen S, Hansen S, Waldstrøm M, Marcussen N, Korsgaard N, Green TM. Digital pathology. Ugeskr Laeger 2022 Jun 20;184(25):V01220044. [Medline: 35781361]
22. Pakkeforløb for modermærkekræft i huden [Article in Danish]. Danish Health Authority. 2020. URL: https://www.sst.dk/-/media/Udgivelser/2020/Modermaerkekraeft/Pakkeforloeb-for-modermaerkekraeft-i-huden.ashx [accessed 2024-01-09]
23. Visitation af nonmelanom hudkræft (BCC og PCC) [Article in Danish]. The Region of Southern Denmark. 2023. URL: https://www.sundhed.dk/sundhedsfaglig/information-til-praksis/syddanmark/almen-praksis/patientforloeb/visitation/ouh/visitation-nonmelanom-hudkraeft-bcc-og-pcc [accessed 2023-12-28]
24. Olsen KR, Sørensen TH, Vedsted P, Gyrd-Hansen D, Bech M. Analyse Af Henvisningsmønstret i Dansk Almen Praksis: Delanalyse 2 – En Registerundersøgelse: Dansk Sundhedsinstitut; 2008. URL: https://pure.vive.dk/ws/files/2054729/dsi-2432.pdf [accessed 2025-05-21]
25. Dobry A, Begaj T, Mengistu K, et al. Implementation and impact of a store-and-forward teledermatology platform in an urban academic safety-net health care system. Telemed J E Health 2021 Mar;27(3):308-315. [doi: 10.1089/tmj.2020.0069] [Medline: 32522105]
26. Pakkeforløb og opfølgningsprogrammer [Article in Danish]. Danish Health Authority. 2018. URL: https://www.sst.dk/-/media/Udgivelser/2019/Pakkeforloeb-og-opfoelgningsprogrammer-Begreber-forloebstider-og-monitorering.ashx [accessed 2024-06-13]
27. Kühnau A, Duedahl M. Danske regioner til praktiserende læger: vi vil styrke almen praksis og lægedækningen [Article in Danish]. Danske Regioner. 2024. URL: https://www.regioner.dk/services/nyheder/2024/februar/danske-regioner-til-praktiserende-laeger-vi-vil-styrke-almen-praksis-og-laegedaekningen [accessed 2024-08-22]

28. Teoh NSC, Oakley A. A 9-year teledermoscopy service in New Zealand: retrospective service review. JMIR Dermatol 2022 Oct 6;5(4):e36351. [doi: 10.2196/36351] [Medline: 37632888]
29. Lowe A, Atwan A, Mills C. Teledermoscopy as a community based diagnostic test in the era of COVID-19? Clin Exp Dermatol 2021 Jan;46(1):173-174. [doi: 10.1111/ced.14399] [Medline: 33405282]
30. Kidholm K, Jensen LK, Johansson M, Montori VM. Telemedicine and the assessment of clinician time: a scoping review. Int J Technol Assess Health Care 2023 Dec 15;40(1):e3. [doi: 10.1017/S0266462323002830] [Medline: 38099431]

Abbreviations

CPP: cancer patient pathway
MM: malignant melanoma
NMSC: nonmelanoma skin cancer
OUH: Odense University Hospital
PCP: primary care provider

Edited by R Dellavalle; submitted 24.09.24; peer-reviewed by J Meisenheimer, T Dicker; revised version received 03.03.25; accepted 04.03.25; published 30.05.25.

Please cite as:
Nordahl Larsen R, Kvorning Ternov N, Kidholm K, Vestergaard T
Implementing Teledermoscopy to Shorten Doctors' Diagnostic Process for Suspected Skin Cancer: Observational Pilot Study
JMIR Dermatol 2025;8:e66782
URL: https://derma.jmir.org/2025/1/e66782
doi:10.2196/66782

Provider Perspectives on Pediatric Store-and-Forward Teledermatology at Boston Medical Center: Cross-Sectional Survey

Mason McDowell[1], BS; Sara Stulac[2], MD; Margaret S Lee[1,2,3], MD, PhD

[1]Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

[2]Department of Pediatrics, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

[3]Department of Dermatology, Boston University Chobanian & Avedisian School of Medicine, 609 Albany St, Boston, MA, United States

Corresponding Author:
Margaret S Lee, MD, PhD
Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

Abstract

This cross-sectional survey of pediatric dermatology and primary care pediatric providers found that store-and-forward teledermatology is an efficient and educational means of delivering care to a safety-net pediatric patient population.

(JMIR Dermatol 2025;8:e67728) doi:10.2196/67728

KEYWORDS

teledermatology; residency education; pediatric dermatology access; eConsult; telemedicine; pediatric dermatology

Introduction

Access to pediatric dermatologists is limited by prolonged waiting times, limited appointments, and uneven geographic availability [1]. Provider-to-provider store-and-forward teledermatology (SAFTD), which allows referring providers to send images and clinical information to dermatologists for asynchronous evaluation, triage, and recommendations, alleviates these barriers [2,3]. In 2020, Boston Medical Center (BMC) launched an Epic-based SAFTD service for pediatric providers (PPs), with dermatology residents and board-certified pediatric dermatologists responding to requests within 72 hours. We hypothesized that SAFTD is particularly helpful to patients and providers in safety-net hospital (SNH) systems like BMC, where language, transportation, and financial challenges play heightened roles in care delivery [4,5].

Methods

Study Design

A cross-sectional REDCap (Research Electronic Data Capture [Vanderbilt University]) survey was emailed to all BMC pediatric SAFTD users, including pediatric dermatologists, dermatology residents, pediatricians, pediatric residents, and pediatric nurse practitioners. Completed responses were collected between June 29 and August 7, 2023. Surveys included qualitative and Likert scale (range 0-5) data, which were examined via thematic and univariate analyses, respectively.

Ethical Considerations

The survey was anonymous and was approved as exempt by the BMC institutional review board (H-43783). No renumeration was provided.

Results

Among 15 (58%) responses obtained from 26 PPs, the mean satisfaction score was 4.93 (SD 0.29), with 93% (14/15) reporting they were very satisfied. All PPs preferred SAFTD over traditional referral methods (Table 1), primarily citing decreased time to intervention and saving patients resources (eg, time and cost of travel to clinic). Satisfaction with the response time (mean 4.8, SD 0.4), digital template (mean 4.33, SD 1.07), and time to face-to-face visit (mean 4.26, SD 0.93) was high. Recommendations were communicated via phone for 66% (10/15) of PPs, with 33% (5/15) using the patient portal. Barriers to using SAFTD included difficulty capturing high-quality photographs, providers' limited time for contacting patients regarding recommendations, and challenges with uploading photographs. PPs appreciated the opportunity to learn dermatology in real time, and 93% (14/15) reported changing their subsequent patient management practices after using SAFTD.

Table . Pediatric provider responses to multiple-choice questions. The total number of pediatric provider respondents was 15.

Questions and respondents' selections	Respondents, n (%)
How satisfied are you with the Pediatric Dermatology eConsult[a] service?[b]	
4	1 (7)
5	14 (93)
How satisfied are you with the response time for pediatric dermatology eConsults?[b]	
4	3 (20)
5	12 (80)
How satisfied are you with the template used to provide content in the pediatric dermatology eConsults?[b]	
1	1 (7)
3	1 (7)
4	4 (27)
5	9 (60)
How satisfied are you with the referral time to in-person evaluation, if indicated, by pediatric dermatology eConsult?[b]	
2	1 (7)
3	2 (13)
4	4 (27)
5	8 (53)
What proportion of eConsult diagnoses are concordant with your initial evaluation?[c]	
1	1 (7)
2	1 (7)
3	4 (27)
4	7 (47)
5	2 (13)
Has the information provided in the pediatric dermatology eConsult(s) changed the way you manage other patients?	
No	1 (7)
Yes	14 (93)
What impact has the information provided in the pediatric dermatology eConsults had on the number of patients you refer to dermatology?	
Decreased	4 (27)
No change	3 (20)
Increased	8 (53)
Do you prefer the eConsult or traditional referral system?	
eConsult	15 (100)
With what method do you inform patients of recommendations made via eConsult?	
Patient portal	5 (33)
Phone call	10 (67)

[a]"eConsult" is synonymous with provider-to-provider store-and-forward teledermatology.

[b]Likert scale: 1=not satisfied at all; 5=the most satisfied.

[c]Likert scale: 1=always discordant; 5=always concordant.

Among 7 (41%) responses obtained from 17 dermatology providers (DPs), the mean satisfaction score was 4 (SD 0.53), and 86% (6/7) preferred SAFTD over traditional referral methods (Table 2), citing improved triage and decreased time to intervention. Additional SAFTD benefits included increased collaboration between dermatology and pediatrics departments and decreased language barriers. DP priorities for photograph improvement included focus, lighting, and an adequate number of photos and views. Patients' and guardians' comprehension of teledermatology recommendations at follow-up in-person

appointments was assessed by DPs as moderate overall (mean 2.67, SD 0.51). One respondent preferred traditional referrals, stating that in-person assessment is necessary for adequate diagnoses and that SAFTD may delay appropriate care.

Table . Dermatology provider responses to multiple-choice questions. The total number of dermatology provider respondents was 7.

Questions and respondents' selections	Respondents, n (%)
How satisfied are you with the Pediatric Dermatology eConsult[a] service?[b]	
3	1 (14)
4	5 (71)
5	1 (14)
To what extent do eConsults include adequate quality photos?[c]	
2	1 (14)
3	3 (43)
4	3 (43)
Rate the importance of the below actions for eConsult quality	
Checking that images are in focus[d]	
5	7 (100)
Making sure images have clear orientation and location[d]	
3	3 (43)
4	1 (14)
5	3 (43)
Including images of symmetrical contralateral skin for rashes[d]	
2	2 (29)
3	3 (43)
4	1 (14)
5	1 (14)
Ensuring adequate lighting[d]	
4	3 (43)
5	4 (57)
Verifying adequate number of photos/views[d]	
3	1 (14)
4	4 (57)
5	2 (29)
Making sure hair does not obscure condition[d]	
3	2 (29)
4	4 (57)
5	1 (14)
Do eConsults include adequate clinical information?[c]	
2	1 (14)
3	4 (57)
4	2 (29)
What is your overall level of comfort diagnosing conditions via eConsult?[e]	
3	3 (43)
4	4 (57)
To what extent do eConsults reduce unnecessary face-to-face dermatology appointments?[f]	
3	5 (71)

Questions and respondents' selections	Respondents, n (%)
4	2 (29)

[a]"eConsult" is synonymous with provider-to-provider store-and-forward teledermatology.

[b]Likert scale: 1=not satisfied at all; 5=the most satisfied.

[c]Likert scale: 1=none of the time; 5=all the time.

[d]Likert scale: 1=not important; 5=very important.

[e]Likert scale: 1=unable to make diagnoses via eConsult; 5=equally comfortable diagnosing as compared to face-to-face visits.

[f]Likert scale: 1=not at all; 5=greatly.

Discussion

BMC PPs and DPs are highly satisfied with SAFTD, particularly with its ability to facilitate prompt treatment, superior triage, and reduced barriers to care. Providers highlighted that it decreases language and transportation barriers, which disproportionately affect the diverse and low-income populations served by BMC [4,5]. The proportion of BMC PPs "very satisfied" with SAFTD surpasses those reported in similar studies conducted in non-SNH settings [6,7]. Although participation bias may have contributed, PPs' high satisfaction with SAFTD suggests heightened value in SNHs. These findings support literature showing that SAFTD increases access to care [2,3,8].

Previous studies noted photograph quality as an SAFTD limitation [3,6]. Our results provide more details for improving pediatric photography support or training to ensure images are in focus, are taken with adequate illumination, and provide a comprehensive range of clinical views to improve SAFTD quality.

Our study shows that SAFTD can be used to train pediatric and dermatology residents, as well as PPs with less dermatology experience. PPs valued how SAFTD facilitated real-time learning in dermatology; 93% affirmed that the service improved their management practices for dermatological conditions. This shows how learning co-occurs with routine patient care, supporting quiz-based data and survey studies suggesting that PPs learn from SAFTD over time [8,9]. In the context of variable dermatological training for pediatricians [10] and pediatric advanced practice providers, SAFTD presents an opportunity to incorporate dermatology education into day-to-day primary care. With respect to mitigating board-certified pediatric dermatologist workforce shortages [1], increasing PPs' knowledge of common dermatological conditions allows DPs to focus on complex and severe conditions.

We provide evidence that SAFTD may be particularly helpful for pediatric residency programs and SNHs serving resource-limited populations. Addressing PPs' barriers to SAFTD use by providing photography support and allocating protected time for dermatologic recommendations would strengthen SAFTD's benefits. Efforts to optimize SAFTD for primary care, dermatology education, and expanded pediatric dermatology access hold significant promise.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files (Multimedia Appendix 1).

Authors' Contributions

MM conceptualized and designed the study, designed data collection instruments, participated in data analysis, drafted the initial manuscript, and critically reviewed and revised the manuscript. SS designed data collection instruments, participated in data collection and analysis, and critically reviewed and revised the manuscript. MSL conceptualized and designed the study, designed data collection instruments, participated in data analysis, drafted the initial manuscript, and critically reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data and information files.
[XLSX File, 16 KB - derma_v8i1e67728_app1.xlsx]

References

1. 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]

2. Wang RF, Trinidad J, Lawrence J, et al. Improved patient access and outcomes with the integration of an eConsult program (teledermatology) within a large academic medical center. J Am Acad Dermatol 2020 Dec;83(6):1633-1638. [doi: 10.1016/j.jaad.2019.10.053] [Medline: 31678336]
3. Seiger K, Hawryluk EB, Kroshinsky D, Kvedar JC, Das S. Pediatric dermatology eConsults: reduced wait times and dermatology office visits. Pediatr Dermatol 2020 Sep;37(5):804-810. [doi: 10.1111/pde.14187] [Medline: 32544276]
4. Health Research & Educational Trust. Social determinants of health series: transportation and the role of hospitals. American Hospital Association. 2017 Nov. URL: https://www.aha.org/system/files/hpoe/Reports-HPOE/2017/sdoh-transportation-role-of-hospitals.pdf [accessed 2025-05-20]
5. Al Shamsi H, Almutairi AG, Al Mashrafi S, Al Kalbani T. Implications of language barriers for healthcare: a systematic review. Oman Med J 2020 Apr 30;35(2):e122. [doi: 10.5001/omj.2020.40] [Medline: 32411417]
6. Pahalyants V, Murphy WS, Gunasekera NS, Das S, Hawryluk EB, Kroshinsky D. Evaluation of electronic consults for outpatient pediatric patients with dermatologic complaints. Pediatr Dermatol 2021 Sep;38(5):1210-1218. [doi: 10.1111/pde.14719] [Medline: 34467570]
7. Kohn LL, Pickett K, Day JA, et al. When is synchronous telehealth acceptable for pediatric dermatology? Pediatr Dermatol 2022 Mar;39(2):236-242. [doi: 10.1111/pde.14919] [Medline: 35178735]
8. Crockett JL, Cordoro KM. Pediatric dermatology eConsultation: insights to reduce barriers to utilization and increase access to care. Pediatr Dermatol 2024;41(6):1114-1116. [doi: 10.1111/pde.15698] [Medline: 39078081]
9. Mohan GC, Molina GE, Stavert R. Store and forward teledermatology improves dermatology knowledge among referring primary care providers: a survey-based cohort study. J Am Acad Dermatol 2018 Nov;79(5):960-961. [doi: 10.1016/j.jaad.2018.05.006] [Medline: 29753059]
10. Khorsand K, Brandling-Bennett HA. Deficiencies in dermatologic training in pediatric residency: perspective of pediatric residency program directors. Pediatr Dermatol 2015;32(6):819-824. [doi: 10.1111/pde.12662] [Medline: 26337718]

Abbreviations:

BMC: Boston Medical Center
DP: dermatology provider
PP: pediatric provider
REDCap: Research Electronic Data Capture
SAFTD: store-and-forward teledermatology
SNH: safety-net hospital

Edited by E Parker; submitted 19.10.24; peer-reviewed by M Mars, S Albers; revised version received 01.04.25; accepted 19.04.25; published 02.06.25.

Please cite as:
McDowell M, Stulac S, Lee MS
Provider Perspectives on Pediatric Store-and-Forward Teledermatology at Boston Medical Center: Cross-Sectional Survey
JMIR Dermatol 2025;8:e67728
URL: https://derma.jmir.org/2025/1/e67728
doi:10.2196/67728

Mobile Health App as an Auxiliary Tool in Management of Atopic Dermatitis in Children: Randomized Controlled Trial

Alex Zvulunov[1], MHA, MD; Stepan Lenevich[2], PhD; Natalia Migacheva[3], MD, PhD

[1]Sheba Medical Center and the Recanati School of Medicine, Reichman University, Tel Hashomer, Ramat-Gan, Israel

[2]AvantaTrading Ltd, Moscow, Russian Federation

[3]Department of Pediatrics, Samara State Medical University, Samara, Russian Federation

Corresponding Author:
Alex Zvulunov, MHA, MD
Sheba Medical Center and the Recanati School of Medicine, Reichman University, Tel Hashomer, Ramat-Gan, Israel

Abstract

Background: Mobile health apps can boost treatment adherence and support disease management at home. The Atopic App and web-based Atopic School patient education program offer a chance to enhance adherence to atopic dermatitis (AD) management.

Objective: We aim to evaluate the feasibility, acceptability, and preliminary efficacy of the Atopic App mobile health intervention in the managing of AD in children.

Methods: A randomized controlled study in children with AD divided participants into 3 groups: a control group (no app), an observational group with the app, and an interventional group with investigator supervision. Patients were examined at screening and follow-up visits 1 and 2 at 3-month intervals. Outcome measures included SCORAD (Scoring Atopic Dermatitis) for objective severity and Patient-Oriented Eczema Measure (POEM) for subjective effectiveness. Statistical analysis used paired *t* tests (2-tailed), the Mann-Whitney U test, and multiple regression.

Results: Fifty-eight participants entered this study (38 boys and 20 girls): group 1 (control) comprised 17 patients, while experimental groups 2 and 3 consisted of 20 and 21 patients, respectively. The rates of missed appointments were similar and statistically insignificant across the groups. All groups showed a significant decrease in SCORAD and POEM scores ($P<.05$). Usage of the app for ≥8 days showed a more significant decrease in severity scores compared to those who used it for ≤7 days, or did not use it at all. Participants who used the app for ≥8 days had a median SCORAD of 6.25 (95% CI 4.6 - 14.1; IQR 4-16.3) at visit 1, significantly lower than nonusers (17.9, 95% CI 13.9 - 24.0; IQR 13.9-24; $P=.03$) and those using it ≤7 days (13, 95% CI 9.35 - 27; IQR 7.2-27; $P=.04$). Their median POEM of 2 (95% CI 1.0 - 4.5; IQR 1-5.3) was also significantly lower than those using the app ≤7 days (9, 95% CI 2 - 12; IQR 2-12; $P=.04$) and lower, though not significantly, than nonusers (7, 95% CI 1 - 9; IQR 1-9; $P=.14$). Additionally, using the Atopic App for ≥8 days after the screening visit strongly predicted a decrease in both SCORAD and POEM scores ($P=.01$ and $P=.04$, respectively). The time since the screening visit significantly predicted increased outcome scores, while prescriptions of topical calcineurin inhibitors, oral antihistamines, and oral antibiotics were weak and insignificant predictors of score changes.

Conclusions: Our findings indicate that the Atopic App is helpful tool in managing AD in children, and they underscore the potential of mobile health interventions in the disease management.

Trial Registration: ClinicalTrials.gov NCT06412094; https://clinicaltrials.gov/study/NCT06412094

(JMIR Dermatol 2025;8:e60479) doi:10.2196/60479

KEYWORDS

atopic dermatitis; skin; disease management; children; pediatric; feasibility; mHealth; mobile health; app; eczema; Atopic App; dermatology

Introduction

Poor medication adherence is a major barrier to treatment success in atopic dermatitis (AD), due to various underlying causes, including forgetfulness, medication side effects, complex dosing regimens, cost barriers, etc [1]. Mobile health apps may improve treatment adherence [2]. There is a growing list of AD-related mobile apps. A recent systematic review and meta-analysis of mobile health applications for AD reported a significant improvement in patients' quality of life (assessed by Dermatology Life Quality Index) and self-management (assessed by Patient-Oriented Eczema Measure [POEM]) but no significant impact on AD severity (assessed by SCORAD [Scoring Atopic Dermatitis]) [3]. POEM is a subjective measure

completed by patients, capturing their experience of AD severity (scores range from 0‐28, higher scores indicate worse symptoms). SCORAD is a clinician-administered tool that combines objective assessment of disease signs with patient-reported symptoms (scores range from 0‐103, higher scores indicate worse severity). The most important feature required for development of mobile apps for caregivers of children with AD is an educational functionality including knowledge of the disease, management of symptoms, medication usage, and triggers [4]. Most available apps for AD primarily assess disease severity, lacking educational functionality or bidirectional communication. Few have been scientifically studied, mainly demonstrating feasibility [3-5].

The Atopic App is a free to download app on the App Store (Apple Inc) and Google Play (Google LLC) that address these shortcomings by offering the following features: (1) a chatbot-directed instruction on app use and targeted education to enhance user understanding and adherence to treatment plans, (2) online patient-education program beyond basic disease information, (3) automatic artificial intelligence (AI)-powered severity assessment for efficient self-monitoring by patients and caregivers, (4) integration of personalized action plans prescribed by health care providers, and (5) a tool for identification of personal trigger factors to reduce flare-ups. The engagement process with the app includes completion of the POEM questionnaire, acquisition of clinical photographs and numerical rating scale for severity of itch, transcription of action plans prescribed by a treating physician, documentation of suspected triggers of exacerbations, and the patient education Atopic School program, while an integrated AI tool automatically calculates severity scores using the Eczema Area and Severity Index method based on photographs taken by users [6]. Recently, we reported on feasibility and impact of the Atopic App that provided a real-world data on severity dynamics, treatment patterns and exacerbation-trigger correlations, indicating the tool's potential impact on health care engagement of AD in children [6].

The purpose of this randomized controlled study is to evaluate the impact of the Atopic App mobile health app as an auxiliary management tool for children with AD.

Methods

Study Design

Study participants were children with AD aged 4 months till 16 years and their parents, consecutively recruited from our dermatology clinic regardless of gender, or disease severity. Informed consent was obtained from all participating parents.

This study used a parallel, 3-arm randomized controlled trial design with a 1:1:1 allocation ratio. Participants were randomized to 1 of 3 groups: a control group that did not use the Atopic App (group 1), an experimental observational group provided with the mobile app without supervision by the investigators (group 2), and an experimental interventional group provided with the mobile app with potential supervision by the investigators (group 3). To ensure allocation concealment and minimize selection bias, a sequential allocation to a study group was used. Upon study enrollment, participants received recommended treatment plans and instructions for contacting the doctor via messenger for any questions during treatment. Furthermore, participants in groups 2 and 3 were instructed to download the mobile app within 1 day post screening, while those in group 3 were also informed about the doctors' virtual oversight, including registration status and regularity of use of the app. Patients aged older than 14 years were permitted to use the app. Beyond this, no specific instructions or recommendations were given regarding usage of the Atopic App, allowing participants to access its features and functionalities as needed to manage their children's AD. Push notifications served as reminders to submit a POEM form, initially after 7 days and then daily until completion of participation in this study.

The investigators did not initiate communication or reminders regarding usage of the app, but could refer to usage data reports during patients' visits or queries. The WhatsApp (Meta Platforms) application was used for patient-doctor communication. Reasons for communications and for interim visits were registered by the following categories: exacerbation, lack of improvement, or clarification questions. Whenever adjustments to therapy were deemed necessary due to ineffectiveness or exacerbations, patients were invited for an in-person interim visit.

Patient-parent pairs were excluded if they had previous experience with the Atopic App or participation in affiliated online Atopic School program or presence of concomitant skin disease or pathological conditions that may affect the assessment of effectiveness (severe somatic diseases, mental disorders, oncologic or acute infectious diseases, etc) or, regarding participants in groups 2 or 3, avoidance from registration during consecutively 5 days following the screening visit. In addition, participants were excluded from the final analysis, if the time gap between their visits deviated by more than 30 days from the scheduled dates, either by occurring more than 30 days before or exceeding 30 days after the planned follow-up visit. This exclusion criterion resulted in variations in the number of patients across different stages of this study within the groups.

The intended duration of this study was 6 months with 3 months intervals between visits.

The outcome end points included objective severity assessment using the SCORAD scale, and subjective assessment of effectiveness using the POEM scale.

To evaluate the significance of pairwise differences among the groups under consideration, the Mann-Whitney U test was used. Additionally, multiple regression analysis was used to explore relationships between POEM and SCORAD score changes and independent variables such as the prescription of different types of medications at the previous visit, the period of time since the screening visit, and whether the patient engaged with the Atopic App for 8 and more days following the screening visit.

To evaluate the impact of usage of the Atopic App on AD severity dynamics, participants were stratified into 3 engagement groups. Group A included control group participants who did not install the app, group B included participants from both

experimental groups who used the app 7 days or less between screening and visit 1 or 2, and group C included participants from both experimental groups who used the app 8 or more days between screening and visits 1 or 2.

Ethical Considerations

This study was approved by the Samara State Medical University Ethics Committee (review 242). Written informed consent was obtained, detailing data usage, potential risks, and the right to withdraw at any time. Comprehensive safety and security procedures were implemented to protect participant privacy and reduce harm. Data transmission was encrypted using HTTPS, and participant data were anonymized with unique codes. Staff were trained on data security protocols and privacy regulations. No compensation was provided to participants. Due to the pilot nature of this study and its limited sample size, registration in a World Health Organization–accredited registry was not conducted before this study. However, it was registered retrospectively (ClinicalTrials.gov NCT06412094).

Results

During the period from March 2022 till June 2022, a total of 66 children with AD and their parents were recruited for this study. Seven patients from experimental groups 2 and 3 were excluded after the screening: those who did not install the app on time, those who did not provide their email address, or those who cancelled their participation after the screening. Moreover, 1 participant from the control group 1 was excluded due to installation of the app during this study. So, 58 participants entered this study (38 boys and 20 girls): group 1 (control) comprised 17 patients, while experimental groups 2 and 3 consisted of 20 and 21 patients, respectively. Flow of participants through each stage of the study (enrollment, intervention allocation, follow-up, and analysis) is depicted in the Figure 1. Baseline demographic data and clinical features in each group are presented in the Table 1.

Figure 1. CONSORT flow diagram. *Patients who missed follow-up appointments beyond the ±30 day window were excluded from the analysis. AD: atopic dermatitis; CONSORT: Consolidated Standards of Reporting Trials.

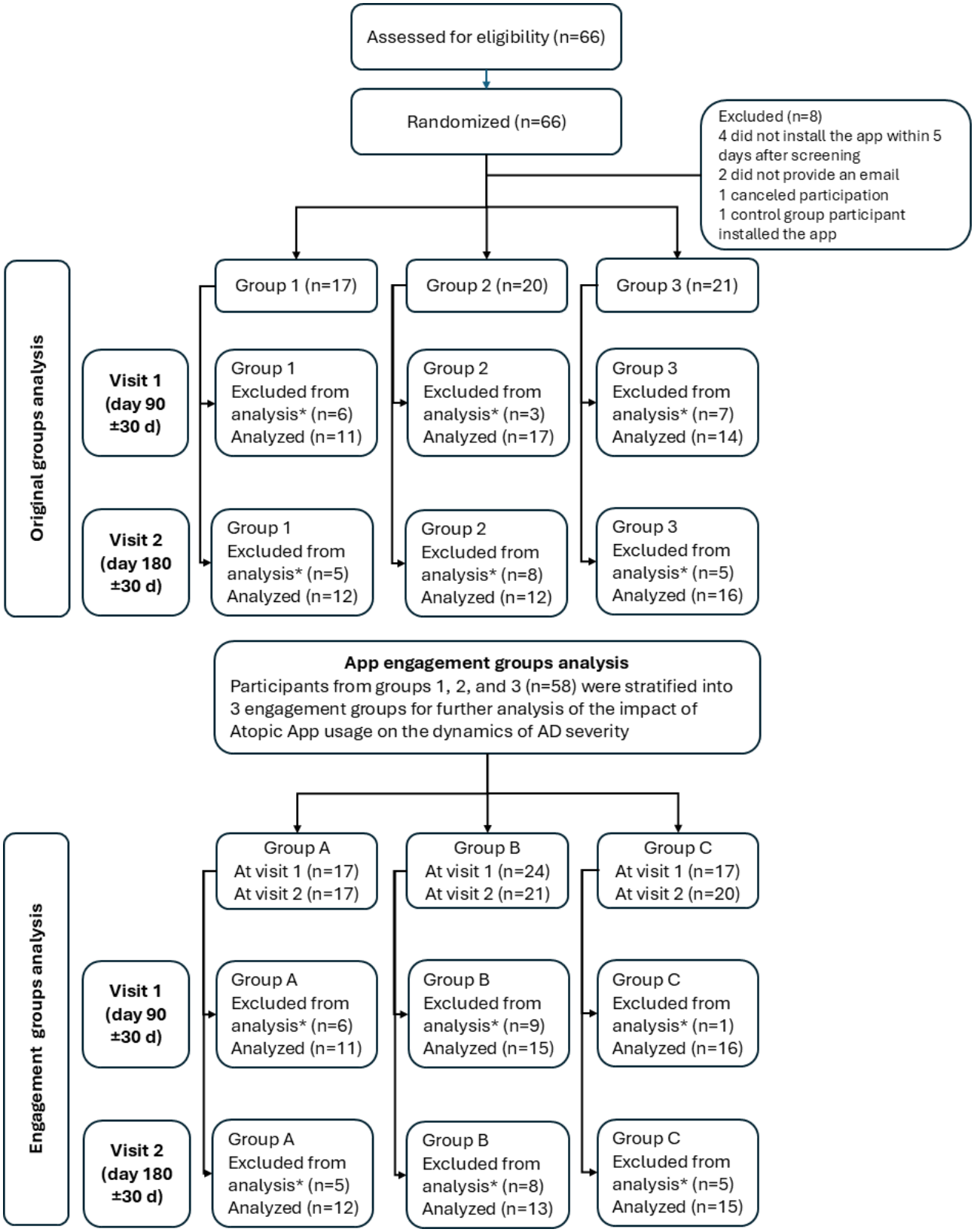


Table . Baseline demographic data and clinical features in each study group.

		Group 1 (n=17)	Group 2 (n=20)	Group 3 (n=21)
Age (years)				
	Median (IQR)	4.4 (1.1 - 6.2)	2.1 (0.6 - 8)	7.2 (1.5 - 11.3)
	Range	0.5 - 15	0.3 - 16.3	0.5 - 14.1
Male, n		13	13	12
Female, n		4	7	9
SCORAD[a]				
	Median (IQR)	31.8 (24.5 - 38.35)	31.4 (24.5 - 38.2)	34 (22.7 - 43.6)
	Range	8 - 61	12 - 50.9	9.4 - 49
POEM[b]				
	Median (IQR)	14 (8.5 - 18.5)	12.5 (8.25 - 16)	13 (9-16)
	Range	2 - 23	5 - 26	3 - 20
Topical corticosteroids, n (%)		11 (65)	14 (70)	10 (48)
Topical calcineurin inhibitor, n (%)		10 (59)	16 (80)	9 (43)
Oral antihistamines, n (%)		11 (65)	16 (80)	11 (52)
Oral antibiotics, n (%)		1 (6)	2 (10)	0 (0)

[a]SCORAD: Scoring Atopic Dermatitis.

[b]POEM: Patient-Oriented Eczema Measure.

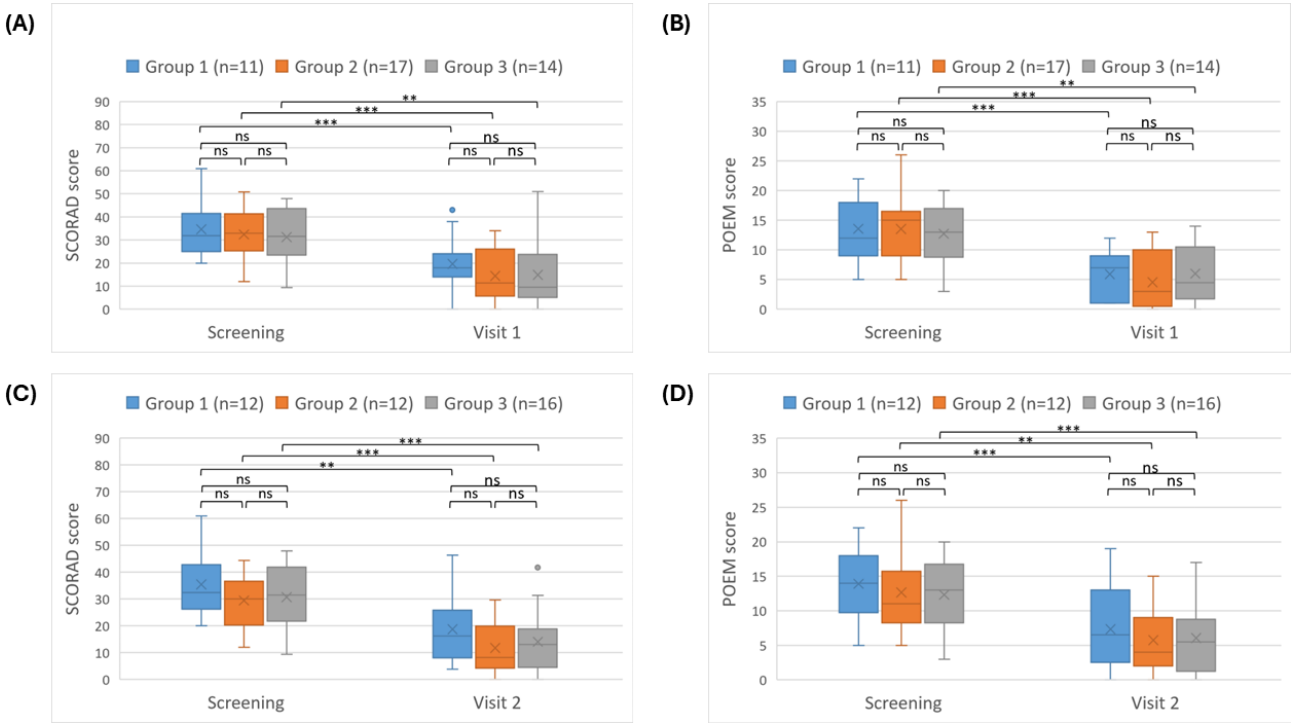
The vast majority of this study's cohort (48/58, 83%) had a documented history of allergic diseases: 66% (38/58) exhibited concurrent food allergies, bronchial asthma, and allergic rhinitis or a combination of these. Overall, 29% (17/58) exhibited mild AD, while 66% (38/58) presented with moderate to severe AD, and the remaining 5% (3/58) had severe AD. No statistically significant differences were found between patients of different groups by gender, age, and severity of the disease at the time of inclusion in this study.

At the screening visit, group 3 had higher median AD severity scores (POEM and SCORAD) compared to groups 1 and 2. The difference in SCORAD and POEM between the groups at screening was not statistically significant (Table 1).

Patients who missed follow-up appointments outside the ±30 day window were excluded from the analysis, resulting in variations in the number of participants between visits. The rates of missed appointments were similar and statistically insignificant across the groups.

Throughout the observation period, all groups demonstrated a significant decrease in the values of SCORAD and POEM scores at visit 1 and visit 2 (Figure 2) compared to corresponding scores at the screening visit (P<.05)

Figure 2. Distribution of AD severity scores across groups 1, 2, and 3. The error bars represent the minimum and maximum values excluding outliers. Above the bars, the intergroup statistical analysis is reported as follows: * *P*<.05, ** *P*<.01, *** *P*<.001. (A) SCORAD at screening and visit 1 (day 90±30 d). (B) POEM at screening and visit 1 (day 90±30 d). (C) SCORAD at screening and visit 2 (day 180±30 d). (D) POEM at screening and visit 2 (day 180±30 d). AD: atopic dermatitis; ns: not significant; POEM: Patient-Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis.



At visit 1, the median SCORAD scores for groups 2 and 3 were lower than group 1, at 11.3 (95% CI 6 - 25) and 9.5 (95% CI 5.5 - 20), respectively, compared to 17.9 (95% CI 13.9 - 24) for group 1. At visit 2, this pattern continued, with medians of 8.20 (95% CI 4.85 - 19.15) and 13 (95% CI 4.5 - 18) for groups 2 and 3, respectively, versus 16.3 (95% CI 10.2 - 24.9) for group 1. POEM scores followed a similar pattern. At visit 1, the median POEM scores for groups 2 and 3 were 3 (95% CI 1 - 8) and 4.5 (95% CI 2 - 9.5), respectively, compared to 7 (95% CI 1 - 9) for group 1. At visit 2, the medians were 4 (95% CI 2 - 9) for group 2 and 5.5 (95% CI 2 - 8) for group 3, versus 6.5 (95% CI 3 - 12) for group 1.

The reduction of SCORAD and POEM scores was more prominent and consistent for patients in groups 2 and 3 than in group 1, although the differences between the groups were not statistically significant (Table 2 and Figure 3).

Table . Median values of SCORAD[a] and POEM[b] scores across groups 1, 2, and 3 and visits.

	SCORAD, median (IQR)		POEM, median (IQR)	
	Screening	Visit	Screening	Visit
Visit 1: day 90 (±30 d)				
Group 1[c]	31.8 (25 - 41.5)	17.9 (13.9 - 24)	12 (9-18)	7 (1-9)
Group 2[d]	33 (25.3 - 41.3)	11.3 (5.75 - 26)	15 (9 - 16.5)	3 (0.5 - 10)
Group 3[e]	31.5 (23.4 - 43.6)	9.5 (5.05 - 23.8)	13 (8.75 - 17)	4.5 (1.75 - 10.5)
Visit 2: day 180 (±30 d)				
Group 1	32.4 (26.2 - 42.8)	16.3 (8.1 - 25.9)	14 (9.8 - 18)	6.5 (2.5 - 13)
Group 2	30 (20.3 - 36.6)	8.2 (4.3 - 19.9)	11 (8.3 - 15.8)	4 (2-9)
Group 3	31.5 (21.8 - 41.9)	13 (4.5 - 18.9)	13 (8.3 - 16.8)	5.5 (1.3 - 8.8)

[a]SCORAD: Scoring Atopic Dermatitis.

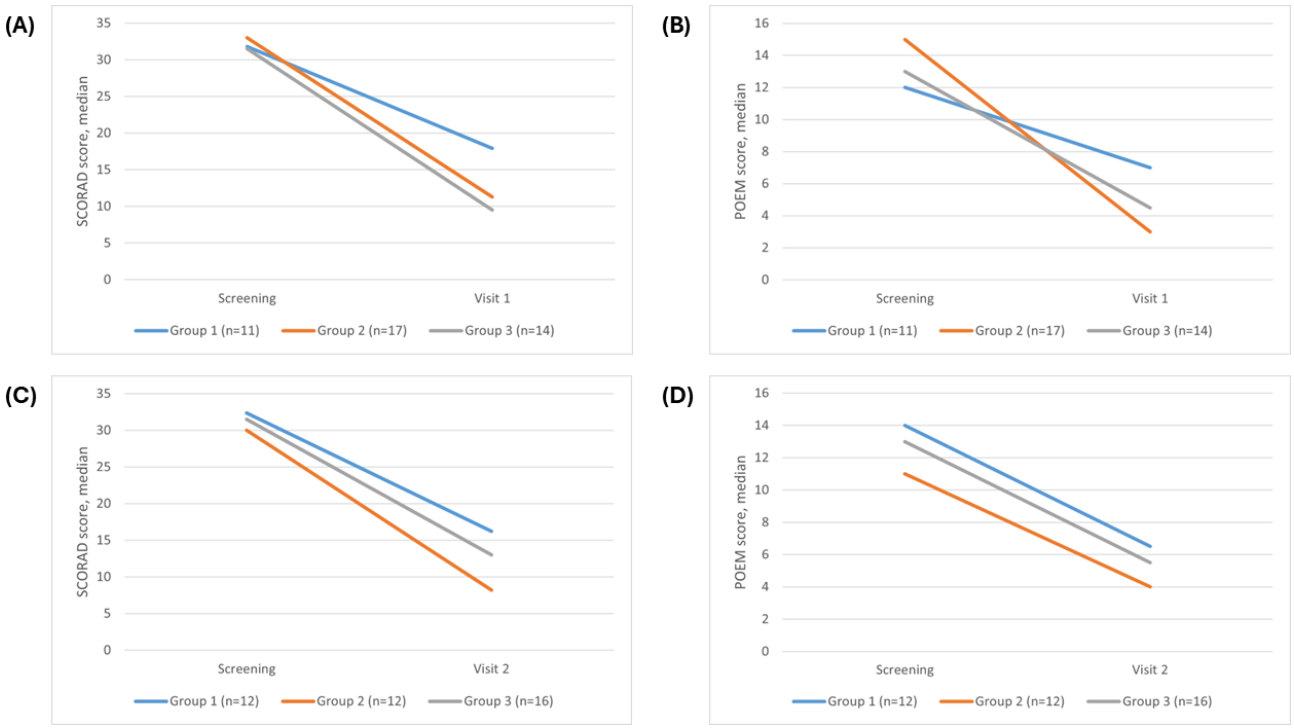
[b]POEM: Patient-Oriented Eczema Measure.

[c]Group 1: control group that did not use the Atopic App.

[d]Group 2: experimental observational group provided with the mobile app without supervision by the investigators.

[e]Group 3: experimental interventional group provided with the mobile app with potential supervision by the investigators.

Figure 3. Trends in median AD severity scores across groups 1, 2, and 3. (A) Median SCORAD scores from screening to visit 1 (day 90±30 d). (B) Median POEM scores from screening to visit 1 (day 90±30 d). (C) Median SCORAD scores from screening to visit 2 (day 180±30 d). (D) Median POEM scores from screening to visit 2 (day 180±30 d). AD: atopic dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis.



To further evaluate the effectiveness of using the Atopic App mobile health app to monitor the course of AD, an analysis was carried out considering the influence of engagement with the mobile application on treatment outcomes.

At visit 1, the median SCORAD score of group C (6.25, 95% CI 4.6 - 14.1) was significantly lower than that of both group A (17.9, 95% CI 13.9 - 24; *P*=.03, *r*=0.418) and group B (13, 95% CI 9.35 - 27; *P*=.04, *r*=0.369). The median POEM score of group C (2, 95% CI 1 - 4.5) was significantly lower than that of group B (9, 95% CI 2 - 12; *P*=.04, *r*=0.369) and lower than that of group A (7, 95% CI 1 - 9), although this difference was not statistically significant (*P*=.14, *r*=0.285). At visit 2, despite a more pronounced reduction trend in group C (Figure 4; SCORAD 9, 95% CI 6 - 18; POEM 5, 95% CI 2 - 9), there were no statistically significant differences compared to group A (SCORAD 16.25, 95% CI 10.5 - 24.9; POEM 6.5, 95% CI 3 - 12) or group B (SCORAD 9, 95% CI 3.5 - 17.5; POEM 5, 95% CI 1 - 8).

The decrease in SCORAD and POEM scores was more significant in patients who used the app for 8 or more days as compared to those who used it for 7 days or less or did not use the app at all (Figures 4 and 5; Table 3).

Figure 4. Trends in median AD severity scores across groups A, B, and C. (A) Median SCORAD scores from screening to visit 1 (day 90±30 d). (B) Median POEM scores from screening to visit 1 (day 90±30 d). (C) Median SCORAD scores from screening to visit 2 (180±30 d). (D). Median POEM scores from screening to visit 2 (day 180±30 d). AD: atopic dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis.

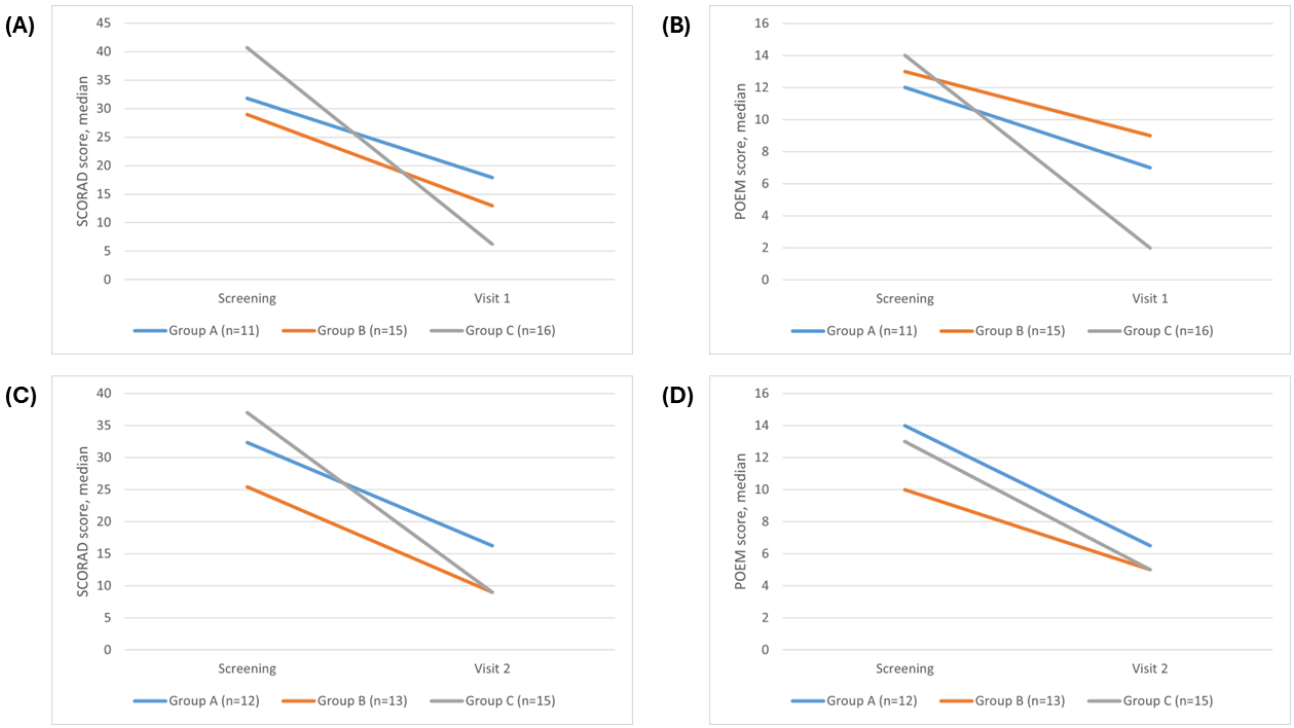


Figure 5. Distribution of AD severity scores across groups A, B, and C. The error bars represent the minimum and maximum values excluding outliers. Above the bars, the intergroup statistical analysis is reported as follows: * P<.05, ** P<.01, *** P<.001. (A) SCORAD at screening and visit 1 (day 90±30 d). (B) POEM at screening and visit 1 (day 90±30 d). (C) SCORAD at screening and visit 2 (day 180±30 d). (D) POEM at screening and visit 2 (day 180±30 d). AD: atopic dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis.

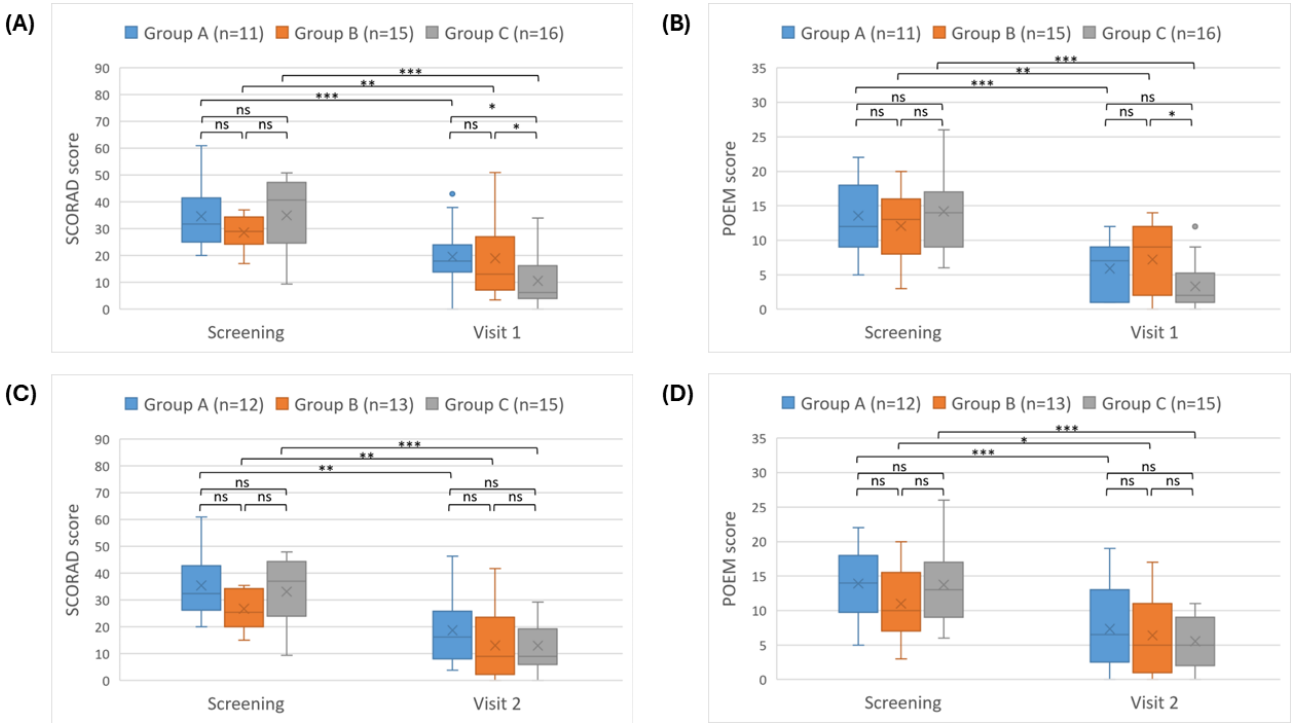


Table . Median values of SCORAD[a] and POEM[b] scores.

	SCORAD, median (IQR)		POEM, median (IQR)	
	Screening	Visit	Screening	Visit
Visit 1: day 90 (±30 d)				
Group A[c]	31.8 (25 - 41.5)	17.9 (13.9 - 24)	12 (9-18)	7 (1-9)
Group B[d]	29 (24.2 - 34.4)	13 (7.2 - 27)	13 (8-16)	9 (2-12)
Group C[e]	40.7 (24.6 - 47.3)	6.25 (4 - 16.3)	14 (9-17)	2 (1 - 5.3)
Visit 2: day 180 (±30 d)				
Group A	32.4 (26.3 - 42.8)	16.3 (8.1 - 25.9)	14 (9.8 - 18)	6.5 (2.5 - 13)
Group B	25.4 (20 - 34.2)	9 (2.3 - 23.6)	10 (7 - 15.5)	5 (1-11)
Group C	37 (24 - 44.4)	9 (6 - 19.2)	13 (9-17)	5 (2-9)

[a]SCORAD: Scoring Atopic Dermatitis.

[b]POEM: Patient-Oriented Eczema Measure.

[c]Group A: participants who did not receive the app.

[d]Group B: participants who engaged with the application on fewer than 8 days.

[e]Group C: participants who engaged with the app on 8 or more days.

Multiple regression analysis aimed to explore effects of various independent variables on changes in SCORAD and POEM scores between visits. The independent variables included the prescription of different types of medications at the previous visit, interval since the screening visit, and use of the Atopic App for 8 days or more following the screening visit.

An *F* test was used as a measure of the models' accuracy on the dataset:

- For the model with SCORAD score as the dependent variable: multiple R=0.6282, R^2=0.3947, $F_{6,75}$=8.1497, P<.001.
- For the model with POEM score as the dependent variable: multiple R=0.6731, R^2=0.4531, $F_{6,75}$=10.3548, P<.001.

The explored models identified statistically significant predictors of SCORAD and POEM score changes:

1. The prescription of topical corticosteroids at the previous visit was a strong predictor of a decrease in both SCORAD and POEM scores (P<.001 and P<.001, respectively).
2. Engagement with the Atopic App for 8 days or more following the screening visit strongly predicted a decrease in both SCORAD and POEM scores (P=.01 and P=.04, respectively).
3. The time elapsed since the screening visit significantly predicted an increase in both SCORAD and POEM scores (P<.001 and P<.001, respectively).
4. Prescriptions of topical calcineurin inhibitors, oral antihistamines, and oral antibiotics were weak and insignificant predictors of SCORAD and POEM score changes.

Most participants experienced seamless use of the installed app, with only four inquiries pertaining to its use arising during the initial week post installation.

Discussion

Principal Findings

The number of AD-related mobile apps is increasing [3]. However, only a few of these apps are designed to enable bidirectional communication between patients or caregivers and the app [7]. Most studies on apps usage for AD focus primarily on the acceptability and feasibility of using AD-related mobile apps to assess disease severity or burden of the disease [5,8,9] and all lack supporting evidence, input from clinicians and dermatologists [10]. However, this is a randomized controlled study that provides evidence for the potential impact of engaging caregivers of children with AD on the dynamic changes in AD severity, using an app developed with the aid of 20 dermatologists, allergists, adult patients with AD, and parents of children with AD, and aimed to identify difficulties in management of AD in home settings [6]. The differences in methodology of these studies from this randomized controlled trial study preclude meaningful comparison of our findings to other studies on mobile health apps for AD.

In their analysis of studies on the effectiveness of mobile phone apps in influencing health-related behavior change, Zhao et al [11] concluded that apps' use for treatment reminders leads to increased overall adherence. More recently, Joergensen et al [12] investigated the effects of requested self-reporting on treatment adherence using memory buttons, with or without a mobile app, and reported improved outcomes in the group that used the mobile app. A recent meta-analysis [3] demonstrated that while mHealth applications significantly improve patients' quality of life and self-management, they show no significant impact on AD severity (SCORAD). This suggests that the reviewed apps fail to offer comprehensive tools for tracking and addressing clinical symptoms, which may be attributed to limited personalization of treatment plans and lack of real-time communication with health care providers. In a qualitative study, caregivers and health care professionals highlighted key

shortcomings in existing apps, such as confusion over treatment, lack of empowerment through education, and limited emotional support [4]. These factors contribute to poor user engagement and adherence, which aligns with findings from a feasibility study [5], where declining weekly interactions with app features over 6 weeks was noted despite initially high engagement with medication reminders and educational content. Maintz et al [13] highlighted a common limitation of AD apps: inadequate interoperability, data exchange, and personalized care. Lack of bidirectional communication remains a key challenge in many existing tools, as they often function as static symptom trackers without the capacity to facilitate ongoing interaction between patients and health care professionals.

The Atopic App addresses the identified limitations in current AD management, such as AI-powered severity assessment, incorporation of personalized action plans, bidirectional communication and sustained engagement strategies.

The minimal clinically important differences (MCIDs) for SCORAD and POEM scores in AD have been established to represent the smallest changes in scores that are perceived as meaningful by patients: the MCID for SCORAD is about 8.7 points and for the POEM about 3.4 points [14]. The reductions in SCORAD and POEM scores observed in our study significantly exceed the established MCID thresholds, reinforcing the effectiveness of the Atopic App as a valuable adjunct tool in managing AD in children. While all groups exhibited reduced disease severity over time, patients who engaged more with the app—specifically, those who used it for 8 days or more between visits—demonstrated a more significant reduction in severity scores. This suggests that repeated exposure to patient-education content and the app's feedback may lead to better adherence to treatment plans and improved disease management [6]. On the other hand, participants who actively engaged with the app may have been more likely to adhere to treatment plans and recommendations, potentially influencing their improved outcomes. To strengthen causal inferences between app usage and improved outcomes in AD management, future research could use random assignment of app access, independent of participants' initial engagement levels that would isolate the app's true impact on treatment adherence and clinical outcomes.

While the reductions in AD severity observed across all study groups could be attributed to various factors such as prescribed treatments, the natural course of the disease, increased awareness and monitoring, or a placebo effect, the results of the multiple regression analyses indicate that these possibilities are highly unlikely. The analyses showed a significant correlation between higher SCORAD and POEM scores and the time period from the screening visit, which suggests that factors other than those mentioned above are at play. Additionally, the prescription of topical corticosteroids at a preceding visit was identified as a strong predictor of decreased SCORAD and POEM scores, highlighting the well-established effectiveness of this treatment. However, the time elapsed since the screening visit was found to be a significant predictor of increased disease severity, emphasizing the need for regular monitoring and intervention. Notably, sustained app engagement was linked to lower disease scores, emphasizing its potential role in AD management. While capturing disease state at exacerbation onset would offer deeper insights, this study primarily focused on longitudinal changes.

Limitations of this Study

This study's population exhibited a gender imbalance with a higher proportion of boys compared to girls. While this does not reflect the general population, this pilot study prioritized feasibility and initial impact assessment of the Atopic App. To comprehensively evaluate the app's effectiveness, a larger study with a balanced gender distribution is warranted.

Possibly, patients who engaged more with the app may have been more proactive in seeking medical advice or adjusting their treatment based on the app's recommendations, leading to better outcomes.

The predominance of female caregivers in our study highlights the importance of considering gender as a contextual factor, as recommended by the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) guidelines. Due to our limited sample size, we were unable to perform a formal subgroup analysis by gender. Nevertheless, the high proportion of female participants likely reduces the risk of significant gender-related bias in our findings. Future research involving larger and more diverse populations will be crucial to further examine the influence of gender and other contextual factors on the effectiveness of the app.

Conclusion

The Atopic App represents a significant advancement in the digital management of AD by offering AI-powered, personalized solutions to the common shortcomings of existing apps, including lack of bidirectional communication, low engagement, and inadequate personalization. By addressing these gaps, the Atopic App provides a more effective, user-centered approach to improving treatment adherence and clinical outcomes in AD, positioning it as a valuable contribution to the field.

Future research directions should explore: (1) sustained use of the app over extended periods to assess its effect on long-term patient outcomes, treatment adherence, and disease control; (2) economic benefits of integrating the Atopic App into clinical workflows, including potential reductions in health care visits and treatment costs; and (3) potential uses of the app's technology and personalized approach could be adapted for managing other chronic skin conditions.

Conflicts of Interest

SL is an employee of AvantaTrading Ltd.

Checklist 1

CONSORT-eHEALTH checklist (V 1.6.1). CONSORT: Consolidated Standards of Reporting Trials
[PDF File, 11352 KB - derma_v8i1e60479_app1.pdf]

References

1. Choi JY, Dawe R, Ibbotson S, Fleming C, Doney A, Foerster J. Quantitative analysis of topical treatments in atopic dermatitis: unexpectedly low use of emollients and strong correlation of topical corticosteroid use both with depression and concurrent asthma. Br J Dermatol 2020 Apr;182(4):1017-1025. [doi: 10.1111/bjd.18265] [Medline: 31257575]
2. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile apps for increasing treatment adherence: systematic review. J Med Internet Res 2019 Jun 18;21(6):e12505. [doi: 10.2196/12505] [Medline: 31215517]
3. 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]
4. Xu X, Griva K, Koh M, et al. Creating a smartphone app for caregivers of children with atopic dermatitis with caregivers, health care professionals, and digital health experts: participatory co-design. JMIR mHealth uHealth 2020 Oct 29;8(10):e16898. [doi: 10.2196/16898] [Medline: 33118949]
5. Gudmundsdóttir SL, Ballarini T, Ámundadóttir ML, et al. Engagement, retention, and acceptability in a digital health program for atopic dermatitis: prospective interventional study. JMIR Form Res 2023 Jun 14;7:e41227. [doi: 10.2196/41227] [Medline: 36975050]
6. Zvulunov A, Lenevich S, Migacheva N. A mobile health app for facilitating disease management in children with atopic dermatitis: feasibility and impact study. JMIR Dermatol 2023 Dec 13;6:e49278. [doi: 10.2196/49278] [Medline: 38090787]
7. Santer M, Muller I, Becque T, et al. Eczema Care Online behavioural interventions to support self-care for children and young people: two independent, pragmatic, randomised controlled trials. BMJ 2022 Dec 7;379:e072007. [doi: 10.1136/bmj-2022-072007] [Medline: 36740888]
8. Ragamin A, Schappin R, Tan Nguyen N, et al. Remote severity assessment in atopic dermatitis: validity and reliability of the remote Eczema Area and Severity Index and Self-Administered Eczema Area and Severity Index. JAAD Int 2023 Dec;13:184-191. [doi: 10.1016/j.jdin.2023.07.019] [Medline: 38025299]
9. Shah S, Kemp JM, Kvedar JC, Gracey LE. A feasibility study of the burden of disease of atopic dermatitis using a smartphone research application, myEczema. Int J Womens Dermatol 2020 Dec;6(5):424-428. [doi: 10.1016/j.ijwd.2020.08.001] [Medline: 33898712]
10. 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.2024.0468] [Medline: 38452263]
11. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287. [doi: 10.2196/jmir.5692] [Medline: 27806926]
12. Joergensen KM, Vestergaard C, Joergensen MS, et al. Memory buttons in combination with mobile application-induced objective and subjective effects in patients with atopic dermatitis. Dermatol Res Pract 2020;2020:8915893. [doi: 10.1155/2020/8915893] [Medline: 32099541]
13. Maintz L, Bieber T, Bissonnette R, Jack C. Measuring atopic dermatitis disease severity: the potential for electronic tools to benefit clinical care. J Allergy Clin Immunol Pract 2021 Apr;9(4):1473-1486. [doi: 10.1016/j.jaip.2021.02.027] [Medline: 33838840]
14. Schram ME, Spuls PI, Leeflang MMG, Lindeboom R, Bos JD, Schmitt J. EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. Allergy 2012 Jan;67(1):99-106. [doi: 10.1111/j.1398-9995.2011.02719.x] [Medline: 21951293]

Abbreviations

AD: atopic dermatitis
AI: artificial intelligence
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
MCID: minimal clinically important difference
POEM: Patient-Oriented Eczema Measure
SCORAD: Scoring Atopic Dermatitis

Edited by R Dellavalle; submitted 13.05.24; peer-reviewed by GK Gupta, S Gulliver; revised version received 04.10.24; accepted 22.10.24; published 22.01.25.

Please cite as:
Zvulunov A, Lenevich S, Migacheva N
Mobile Health App as an Auxiliary Tool in Management of Atopic Dermatitis in Children: Randomized Controlled Trial
JMIR Dermatol 2025;8:e60479
URL: https://derma.jmir.org/2025/1/e60479
doi:10.2196/60479

Characterization of Reddit Posts About Xylazine-Associated Wounds: Qualitative Study

Anthony Spadaro[1], MPH, MD; JaMor Hairston[2], MS; Sahithi Lakamana[2], MS; Rachel Wightman[3], MD; Jennifer Love[4], MD; Jeanmarie Perrone[5], MD; Abeed Sarker[2], PhD

[1]Department of Emergency Medicine, Rutgers New Jersey Medical School, 140 Bergen St, Newark, NJ, United States

[2]Department of Biomedical Informatics, Emory University, Atlanta, GA, United States

[3]Department of Emergency Medicine, Brown University, Providence, RI, United States

[4]Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[5]Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Anthony Spadaro, MPH, MD
Department of Emergency Medicine, Rutgers New Jersey Medical School, 140 Bergen St, Newark, NJ, United States

Abstract

Background: Xylazine has been associated with skin wounds. The rising prevalence of xylazine and its debated role in wound causation have sparked concerns among public health professionals, medical experts, and people who use drugs.

Objective: This study used a qualitative evaluation of Reddit posts to understand the experiences of people who use drugs concerning xylazine-associated wounds.

Methods: This study explored xylazine discussions on Reddit. Data were collected from 930+ drug-related subreddits via the PRAW Python application programming interface, and natural language processing methods were employed to identify posts that mentioned xylazine and wound-related keywords. Retrieved posts were manually coded for thematic analysis, and a term frequency–inverse document frequency analysis was performed per theme to obtain additional insights.

Results: The manual classification of 286 posts revealed predominant themes related to the pathophysiology of xylazine, wound locations on the body, and management strategies. The 3 most frequent xylazine wound-related themes were “Mechanisms of xylazine-associated wounds” (84 posts, 29.4%), “Geographic region” (67, 23.4%), and “Location of wounds on the body” (56, 19.6%). The analysis showed xylazine’s presence in the discussions among Reddit’s drug-using communities, with a notable focus on wound management and geographic trends. The term frequency–inverse document frequency analysis revealed prominent lexical markers within each theme.

Conclusions: The findings suggest that social media platforms such as Reddit can serve as valuable resources for understanding emerging health issues such as xylazine-associated wounds. The study’s findings highlight patterns of use, the characteristics of wounds on people who use drugs, and discussions about wound management. This study adds to a growing body of literature using social media to understand the consequences of emerging drugs on human health.

(JMIR Dermatol 2025;8:e70329) doi:10.2196/70329

KEYWORDS

xylazine; social media; substance-related disorders; opioid-related disorders; wounds

Introduction

Xylazine is an α -adrenergic agonist that is increasingly prevalent in the unregulated opioid supply [1]. In the northeast United States, xylazine may be present in more than 90% of the fentanyl supply in some localities [2]. As the prevalence of xylazine has increased, there has been growing concern voiced by people who use drugs, the media, and medical providers that xylazine use is associated with skin wounds [3,4]. The etiology of wounds associated with xylazine exposure, the relationship to route of drug use, the location on the body where the wounds develop, and the optimal treatment of wounds are not known [1-3]. Theories proposed include vasoconstrictive effects, tissue hypoxia, cytotoxicity, impaired glucose control, prolonged sedation, pressure-related injury, and other factors associated with drug use practices, including access to sterile supplies, proper nutrition, and clean water [5,6]. Despite these unknowns, there is significant concern from public health and government regulatory agencies about the association between xylazine and skin wound development [6].

Xylazine-associated wounds are extensively discussed among people who use drugs, including on social media [7]. Reddit is a popular social network that allows for anonymous posting by subscribers (redditors), with more than 1.2 billion monthly

active subscribers [8]. Because of its anonymity, Reddit has become a forum for people to discuss sensitive or stigmatized topics such as drug use [9]. For these reasons, social media in general, and Reddit specifically, has been analyzed to understand various aspects of drug use [7,10]. Various social media sites such as Reddit, X, or TikTok may be used to study attitudes toward drugs; however, each site attracts different demographics, may restrict discussion of drug-related topics, and has different rules for researchers to extract data [10]. Reddit has been used for pharmacovigilance of emerging substances, to explore adverse effects from prescribed medications in stigmatized communities, and to explore the general public's perceptions of drug-related topics [7,11,12]. Others have used Reddit to study perceptions of xylazine in general and noted concerns about wounds as an adverse effect, although this study did not focus specifically on issues related to xylazine-associated wounds [13].

In prior work, we leveraged natural language processing (NLP) to identify social media posts on drug-related topics of interest to medical toxicology and addiction medicine researchers and used these social media posts to identify potential adverse effects of xylazine use [7]. For this study, we sought to specifically explore an adverse effect that was identified in our prior work, xylazine-associated wounds. Our primary aim was to perform a thematic analysis of Reddit posts specifically related to xylazine and wounds in order to identify associated factors and salient issues xylazine-associated wounds may be causing for people who use drugs.

Methods

Ethical Considerations

This study was approved by Emory University Institutional Review Board (number STUDY00002458). We combined NLP and expert-driven qualitative analysis to thematically characterize a set of Reddit posts mentioning xylazine and wounds. Posts were not modified from their original form in order to not alter their meaning. All data were publicly available and anonymous at the time of collection. Usernames were not included in the data analysis to further protect the privacy and anonymity of the Reddit users. The authors had no contact with individual Reddit users and no compensation was provided to any Reddit users.

Data Collection

We collected data from 961 drug-related subreddits (Appendix 1 in Multimedia Appendix 1) via the PRAW (Python Reddit API Wrapper) application programming interface (API) (version 7.7.1; GitHub). We used Python (version 3.12; Python Software Foundation) for PRAW and for filtering the data. Data filtering was conducted using the Natural Language Toolkit for tokenization and stop word removal, and regular expressions (via Python's remodule) to identify and exclude noninformative or low-quality posts [7]. The API enables the collection of publicly available data from chosen subreddits in a secure manner following authorization. Posts that are removed by moderators or the original posters are not available via the API. No subreddit that had protected data and required joining prior to viewing data was included in this study. First, we filtered all the collected data using the keywords "xylazine," "tranq," and other possible lexical variants, including street names and common phonetic and typographic misspellings (Appendix 2 in Multimedia Appendix 1) [14]. Then, we filtered the data using a predetermined set of keywords (Appendix 2 in Multimedia Appendix 1) potentially related to wounds (eg, mentions of wounds, necrosis, and skin ulcers). Thus, posts mentioning xylazine and potentially wound-related information were included for further thematic analysis.

Data Analyses

A sample of 15 posts identified by NLP as having content related to xylazine and wounds was reviewed by 2 authors with expertise in toxicology (ASp and JP) to identify relevant categories (codes) into which the posts could be classified [7]. Codes were developed using a content analysis approach [15]. ASp and JP prepared a guidebook to drive the coding of the social media posts (Table S1 in Multimedia Appendix 1). An initial randomized set of 50 posts was independently assigned codes by 2 authors (ASp and JH). The larger author group reviewed the coding assignments, and the codes were iteratively defined through consensus until agreement was reached regarding the definitions of the codes and their adequacy in describing the themes in the posts. This process ensured that codes that were too inclusive or poorly defined could be eliminated, and additional codes could be created. This iteration generated 13 primary codes, including 2 codes: "non-relevant post about xylazine" and "not about xylazine at all" for posts that lack relevance to human xylazine use (eg, posts purely about veterinary use of xylazine) and posts that lack relevance to xylazine completely, respectively. Each code represented a potential theme that could be present in a post, and each post could be assigned multiple codes if the post contained multiple themes [15]. Table 1 shows the 13 codes and their definitions.

Table . Themes from posts extracted by natural language processing with example quotes, distribution, and indicative n-grams detected via term frequency–inverse document frequency.

Theme	Example quotes	Number of posts coded	Percentage of posts with theme (out of 286)	Illustrative top TF-IDF[a] terms and values
Hypothesized mechanisms of xylazine-associated wounds	"it's a Vasoconstrictor and Will Greatly Reduce the Ability of Oxygen-Rich Blood to Get Out into All the Small Veins. If You Repeatedly Inject into a Single Site, You're Counting on Your Circulatory System to Repair the Damage and Catch Any Bacterial Infection."	84	29.4	Top terms include xylazine (184.91) itself, krokodil (29.38) for comparison, xylazine induced (19.09) for direct effect, skin necrosis (15.22), skin ulceration (16.70), and references to pmc articles (10.45).
Geographic region	"This shit is RAVAGING South Jersey."	67	23.4	Prominent terms include philly (21.74), harm reduction (11.24), care supplies (11.17), and https www (10.16) referring to web-based sources. Xylazine skin necrosis (20.65) is noted as a health consequence by locale.
Locations of wounds on the body	"It ate my skin on both arms. They looked horrible, talking down the bone."	56	19.5	Top terms for this theme include xylazine (122.35) itself, alongside skin (52.44), oxygenation (20.99), skin ulceration (16.70), and skin necrosis (13.04), indicating the severe nature of wounds.
Management of wounds	"Other than OP's suggestion to have good hygiene (which I think meant using clean supplies & cleaning IV sites) I personally use saline nasal spray to clean my nose out about 15 - 20 min after snorting."	34	11.8	Illustrative terms include xylazine (99.36), oxygenation (20.99) as critical for healing, skin ulceration (16.70), impaired healing wounds (7.16), and whole wrap (12.00) possibly for bandaging.
Posts about specific xylazine withdrawal symptoms	"I'm wondering if the withdrawal anxiety (I don't rly get any other withdrawal symptoms it's like the methadone covers any sickness and other parts of withdrawal except for the anxiety which makes me think the rebound is not from the fent but from the cut…possibly xylazine) so yeah I was saying I wonder if the "withdrawal" or rebound anxiety is rly from sniffing a shit ton of xylazine."	29	10.1	Examples reflecting withdrawal symptoms include withdrawals (14.13), organs hurting (6.97), zombifying bodies (7.16), and every hour (10.49). Tranq dope (17.00) indicates polysubstance context.
Stigma related to xylazine wounds	"Tranq will actually keep your wounds from healing, and they are calling it a zombie drug."	23	8.0	Examples reflecting stigma include terms such as flesh eating (11.93), skin rotting (9.37) for graphic descriptions, zombie (12.65) for dehumanization, dirty (11.17) for uncleanliness, and opioid deaths (9.54) for links to crises.

Theme	Example quotes	Number of posts coded	Percentage of posts with theme (out of 286)	Illustrative top TF-IDF[a] terms and values
Other drug use habits	"Makes me think it's the common analog in benzos called Etizolam."	20	6.9	Top terms illustrate polysubstance use such as xylazine (46.92) with morphine (24.38) and fent (fentanyl). Phrases such as fent mixed xylazine (4.20), every hour (10.49) for frequency, and methadone clinic (5.58) are also key.
Xylazine use habits	"I got lucky with my IV use, never got hep or anything but I did get a bad infection from a muscle injection."	19	6.6	Illustrative top terms include nose (18.48) for intranasal use, xylazine injectable (10.0 and 8.38) for injection, every hours (10.49) for frequency, fent (11.0) for polysubstance use, and skin (15.64) for side effects.
Posts about MOUD[b]	"Now I have methadone in me yet it isn't making a difference now that's probably cause from using on top of the done I jacked my tolerance so high I no longer feel the methadone at all it's mine as we'll be water I need to increase my dose, but it seems like something else is going on."	19	6.6	Examples include xylazine (65.32) caused with MOUDs such as methadone (17.39), phrases such as suboxone kill cravings (4.77), terms such as fentanyl (21.16), and community queries such as quick question anyone (4.22).
Non-MOUD management of withdrawal	"Yes clonidine is the best drug I'd say for coming off tranq."	11	3.8	Examples include xylazine (19.32) and benzos (7.73) for self-medication, 3-mg clonidine (2.81), terms such as noradrenaline new (3.48), and indicators of injection risks such as iv warned (2.81).
Ability to get into rehabilitation clinic or addiction treatment	"I can't get wounds healed & rehabs won't take me with open festering wounds."	10	3.5	Key terms include rehab (16.70), xylazine (11.04) as a complicating factor, phrases such as hospitals around area (4.77), save someone's life (4.77) for motivation, and insurance xylazine imagine (3.48) as a barrier.
Nonrelevant posts about xylazine	"I run a wildlife hospital and we use xylazine in a lot of our cases. Because it's not a scheduled drug it's easy for us to keep on hand for euthanasia, sedation for fractious animals needing wound care, exams, etc."	37	12.9	N/A[c]

Theme	Example quotes	Number of posts coded	Percentage of posts with theme (out of 286)	Illustrative top TF-IDF[a] terms and values
Not about xylazine at all	"If I'm remembering right wouldn't him jumping into a gross jungle river soon after losing it cause some serious infection, especially in the 60 s?"	70	24.4	N/A

[a]TF-IDF: term frequency–inverse document frequency.

[b]MOUD: medications for opioid use disorder.

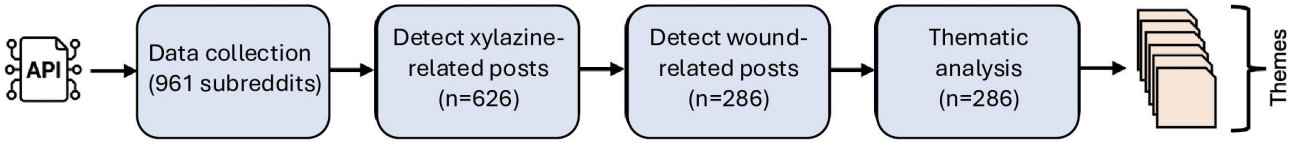
[c]N/A: not applicable.

Using the guidebook, author JH manually coded the remaining posts. All the coded posts were reviewed by authors ASp and JP, and any disagreement about the assigned codes was discussed as a group and resolved by consensus. Following the thematic categorization, word n-gram (n=1 - 3)–level term frequency–inverse document frequency (TF-IDF), where each document included all the posts within a specific theme, was computed to identify n-grams uniquely indicative of each theme. An n-gram is a sequence of n adjacent symbols in a particular order—in this context, a contiguous sequence of words. The top n-grams for each theme were processed with a large language model (LLM) (Google Gemini 2.5) for summarized explanation and interpretation. The 2 noninformative themes were excluded from this analysis. The TF-IDF analysis of the contents associated with each theme and their LLM-assisted explanations revealed some key topics associated with each theme.

Results

We retrieved 5373 posts from 961 subreddits from January 2019 to March 2023. A total of 626 posts were detected to potentially mention xylazine. Within these posts, 286 posts were detected via NLP to contain potentially wound-related keywords, and all of these posts were manually reviewed for thematic analysis (Figure 1). Of these 286 posts, 37 (12.9%) of posts were about xylazine but determined to be about nonhuman use, such as veterinary use. An additional 70 of the 286 posts (24.4%) were determined to be not about xylazine at all and were either incorrectly selected by NLP or were nonrelevant comments on posts about xylazine and wounds. In total, 179 of the 286 (62.6%) posts extracted by NLP were about xylazine and wounds. The posts described several important aspects of the experiences people shared on Reddit about xylazine-associated wounds.

Figure 1. Flow diagram of data collection to thematic analysis. API: application programming interface.



The 3 most frequent xylazine wound-related themes were "Hypothesized mechanisms of xylazine-associated wounds" (84/286 posts, 29.4%), "Geographic region" (67/286, 23.4%), and "Location of wounds on the body" (56/286, 19.6%). Table 1 shows the frequencies of posts for 13 xylazine wound-related themes, along with representative examples. Less prominent themes were "Management of wounds" (34/286, 11.8%), "Posts about specific xylazine withdrawal symptoms" (29/286, 10.1%), and "Stigma related to xylazine wounds" (23/286, 8.0%), which contained important information about managing wounds, their impacts, stigma, and withdrawal.

Posts described how the redditors were using xylazine and speculated on the relationship between how xylazine was being used and the development of wounds. For example, "People are getting wounds on their noses and mouths. Sniffing and smoking is giving people sores." Other posts discussed specific mechanisms and the pathophysiology of xylazine and how it may cause wounds, for example, "It causes severe vasoconstriction which is why it leads to wounds." Some sought help in managing wounds and described their wounds, for example, "How do you treat these skin wounds? I have them on my chest, arms, and legs. How do I know if I need to see a doctor? Any info what to do would be great." The discussions of the pathophysiology and mechanisms of xylazine were often connected to discussions on the locations of the wounds and how xylazine was being used, seemingly to provide explanations for why certain wounds were developing. For example, "Apparently it adds longevity to the high. However the side effect is that it's much harder on the blood vessels. As blood vessels break down, infection risk skyrockets. Users are forced to find new veins more frequently, injection sites don't heal well and this is the result. Infection and destroyed blood vessels causes strain to organs, and the extremities are starved for blood."

Some of the posts expressed that having xylazine-associated wounds made getting addiction treatment more difficult; for example, one post described someone's experience trying to get into addiction treatment: "She wants to, but they won't take her into rehab with the wound and without rehab she'll likely keep injecting." Many of the posts about wounds also discussed issues with attempts at stopping the use of xylazine and unregulated opioids. Posts described both the attempts at quitting "cold turkey" and using medications for opioid use disorder (MOUD), "Was hooked on Philly tranq bad from Kensington

for years. It ate my skin on both arms. They looked horrible, talking down to the bone. The withdrawals would be so bad. Nonstop throwing up, literally every couple mins. Even when nothing is left, you'll just puke bile or dry heave. It's terrible. I was stuck on it for years. Was on 200 mgs of methadone on top of it which did absolutely nothing." Posts that commented on these themes of getting into addiction treatment, complicated withdrawal treatment, and xylazine wounds were sometimes connected. These posts would describe how xylazine made withdrawal more complex to manage and that made wounds more difficult to treat. For example, "There isn't a lot of good hospitals around that area that actually seem to know what they are doing for xylazine/fentanyl withdrawal and that one I mentioned does an excellent job and if some one reads it, and can save someone's life I don't see the harm in posting that. My fiancé died last year because we both thought the hospitals around that area would not know how to treat her withdrawal for an infection. And the infection spread into her blood and she died."

Several posts disclosed the geographic location of where the poster was, where they were buying drugs, or where they were seeing people with wounds. Many of the posts used stigmatizing language to describe wounds and the people who get them, "I do heroin regularly and never had any issues. These people are dirty junkies who don't take care of themselves and won't even swab their skin with an alcohol pad but it's definitely not a combination of unsanitary conditions and dirty drugs full of flesh eating bacteria that can be filtered out with cheap micron filters." Many of the posts also referred to xylazine as the "zombie-drug" and that people who used it were "zombies." Many posts with stigmatizing language were directed toward others rather than the person making the post. As posts were anonymous, it was not known in all instances whether the redditor making the post was a person who uses drugs. However, some posts with stigmatizing language did seem to come from health care providers, and some of the stigma seemed to be coming from the non–people who use drugs community. For example, one post from a health care provider described their first encounter with a patient who used xylazine, "I met my first patient with tranq wounds yesterday. Both their legs look absolutely awful, and the smell is just unspeakable."

The TF-IDF analysis revealed some key topics associated with each theme, such as "Philly" as a prominently mentioned location, "wraps" or bandages for wound management, and graphic descriptions (eg, "flesh eating" and "zombie") expressing stigma associated with xylazine-inflicted wounds. Further examples and explanations are shown in Table 1.

Discussion

Principal Findings

This study uses social media data and NLP to explore a critical clinical complication of the emerging drug xylazine. In this study, social media posts described purported mechanisms behind toxicity from xylazine, experiences with xylazine wounds including how people managed wounds, and stigma related to xylazine use. While the association between xylazine and the development of distinct wounds has been described in case reports and the media, there is limited understanding of the pathophysiology, the optimal management of these wounds, and individual experiences with xylazine-associated wounds [3-6,16-18]. There have been many postulated mechanisms by which xylazine could cause wounds. Xylazine may cause vasoconstriction through its α-adrenergic agonist effect, which could cause tissue hypoxia and injury [6]. Xylazine may also have systemic cytotoxic or hypoxic effects, which have been proposed as a possible explanation for wounds distant from the site of drug use [6]. There are reports of people developing wounds at sites where they do not inject and in those who report only intranasal use [3]. Some experts are skeptical about whether these wounds are a distinct clinical entity separate from skin and soft tissue infections that can occur with intravenous drug use [5]. The results of this study show that people self-report developing wounds after using xylazine. Many of the posts commented on the development of wounds in sites where they do not inject and in association with intranasal use. This has been previously reported, for example, at a wound care clinic in Philadelphia, patients reported developing wounds on their extremities at sites where they do not inject and in those who reported only intranasal use [3]. The development of wounds in the nose and oropharynx is not well described but may be clinically relevant, and it is notable that it is mentioned on Reddit before it is described in the medical literature. These findings support that xylazine-associated wounds can develop without intravenous use. Additionally, many posts speculated on the pathophysiology of xylazine and how it might lead to wound development. While this does not prove an association between xylazine and wound development, it emphasizes the concern about xylazine-associated wounds among Reddit users who are posting on drug-related subreddits and the need for more research into causation.

Many of the posts also used stigmatizing language and employed terms such as "zombie" to discuss the wounds and the people who have them. While this language may reflect the language used in popular media around xylazine use, it is of concern that xylazine-associated wounds may confer additional stigma on a population that already faces a significant amount of stigma [16-20]. Although the posts were anonymous, people occasionally self-disclosed drug use, and there were some notable examples of stigmatizing language used by people who use drugs without xylazine-associated wounds toward people who use drugs with xylazine-associated wounds. Stigma toward people who use drugs with xylazine-associated wounds from popular media and health care institutions has been described, but less is known about stigma within the people who use drugs community around xylazine wounds [20].

Published experiences from low-barrier wound clinics have reported that patients with xylazine-associated wounds face significant structural barriers and fear of withdrawal that may prevent them from getting adequate wound care [4]. Several of the posts discussed withdrawal symptoms that were attributed to xylazine and inadequate relief of their withdrawal symptoms with MOUD. Withdrawal symptoms and issues with MOUD were often brought up in the context of wanting to quit using xylazine so that wounds would heal. Some posts also discussed difficulty with getting into inpatient detox, residential, and

rehabilitation treatment facilities due to the presence of wounds. These posts highlight that xylazine may complicate the treatment of opioid use disorder if people are having difficulty getting into treatment programs or staying on MOUD [3,4].

Limitations

There are several limitations to this study. The posts were self-reported experiences without confirmation that xylazine was present in the drugs the redditors were reportedly using. Redditors who experience more negative effects of drug use may be more likely to post about it. Additionally, this analysis was limited to publicly available posts, and thus private posts were not included, and the posts analyzed may not be representative of everyone's experiences. Individuals who post on Reddit may not be representative of the larger community of people who use drugs, particularly those without phone or internet access. Another limitation is that only keyword-matched posts were included in the thematic analyses; thus, potentially relevant information in posts not containing the required keywords was excluded. The filtering, however, was a necessary step since it is not feasible to review the large volume of content generated on Reddit manually. Future work could expand on the relatively small number of posts analyzed here by looking at other social media sites, which may have different user bases and find a broader range of themes. As the posts were anonymous, we do not have demographic information on the redditors to know in which ways they might be different from other populations. More research is needed to understand the real-world impact of these issues being discussed on social media. Future studies using social media analysis should include exploring other related emerging drugs such as medetomidine, nitazenes, and tianeptine. Future work will include exploring machine learning methodologies such as LLMs to automate the classification performed in this study. Automating this process will increase the feasibility of scaling it for much larger datasets of social media posts and allow for more rapid analysis of studies on emerging drugs. This study has been informed by the clinical experience and expertise of some of the authors; however, there may be delays between when a novel drug emerges and when a clinician may encounter it in a health care setting. Future work should explore the ability to identify emerging drugs of abuse being discussed on social media before expert clinicians even begin to encounter them in the health care setting, which could allow for health care systems to proactively prepare for emerging drugs.

Conclusions

Reddit posts revealed discussion behind the pathophysiology of xylazine-associated wounds and discussed the impact of wounds on treatment and service access. As the unregulated drug market changes and new drugs emerge, social media may continue to be a valuable resource to study adverse effects from novel drugs and their impact on the people who use them.

Acknowledgments

Research reported in this publication was supported by the National Institutes of Health (NIH) under award number R01DA057599. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

AS participated in conceptualization, methodology, writing—original draft, writing—reviewing and editing, supervision, and funding acquisition. JMH contributed to software, validation, data curation, writing—original draft, writing—reviewing, and editing. SL contributed to software, validation, data curation, and writing—reviewing and editing. RW participated in conceptualization, writing—reviewing, and editing. JL participated in conceptualization, writing—reviewing, and editing. JP participated in conceptualization, methodology, writing—reviewing and editing, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of subreddits used, lexical variants searched, and definition of themes.
[DOCX File, 27 KB - derma_v8i1e70329_app1.docx]

References

1. Friedman J, Montero F, Bourgois P, et al. Xylazine spreads across the US: a growing component of the increasingly synthetic and polysubstance overdose crisis. Drug Alcohol Depend 2022 Apr 1;233:109380. [doi: 10.1016/j.drugalcdep.2022.109380] [Medline: 35247724]
2. Zhu DT. Public health impact and harm reduction implications of xylazine-involved overdoses: a narrative review. Harm Reduct J 2023 Sep 12;20(1):131. [doi: 10.1186/s12954-023-00867-x] [Medline: 37700329]
3. McFadden R, Wallace-Keeshen S, Petrillo Straub K, et al. Xylazine-associated wounds: clinical experience from a low-barrier wound care clinic in Philadelphia. J Addict Med 2024;18(1):9-12. [doi: 10.1097/ADM.0000000000001245] [Medline: 38019592]

4. Reed MK, Imperato NS, Bowles JM, Salcedo VJ, Guth A, Rising KL. Perspectives of people in Philadelphia who use fentanyl/heroin adulterated with the animal tranquilizer xylazine; making a case for xylazine test strips. Drug Alcohol Depend Rep 2022 Sep;4:100074. [doi: 10.1016/j.dadr.2022.100074] [Medline: 36846574]
5. Hoffman RS. Closing the xylazine knowledge gap. Clin Toxicol (Phila) 2023 Dec 2;61(12):1013-1016. [doi: 10.1080/15563650.2023.2294619]
6. D'Orazio J, Nelson L, Perrone J, Wightman R, Haroz R. Xylazine adulteration of the heroin-fentanyl drug supply : a narrative review. Ann Intern Med 2023 Oct;176(10):1370-1376. [doi: 10.7326/M23-2001] [Medline: 37812779]
7. Spadaro A, O'Connor K, Lakamana S, et al. Self-reported xylazine experiences: a mixed-methods study of Reddit subscribers. J Addict Med 2023;17(6):691-694. [doi: 10.1097/ADM.0000000000001216] [Medline: 37934533]
8. Dean B. Reddit user and growth stats. Backlinko. 2024 Apr 8. URL: https://backlinko.com/reddit-users#reddit-monthly-active-users [accessed 2025-08-26]
9. Chi Y, Chen HY. Investigating substance use via Reddit: systematic scoping review. J Med Internet Res 2023;25:e48905. [doi: 10.2196/48905]
10. Carpenter KA, Nguyen AT, Smith DA, et al. Which social media platforms facilitate monitoring the opioid crisis? PLOS Digit Health 2025 Apr;4(4):e0000842. [doi: 10.1371/journal.pdig.0000842] [Medline: 40293990]
11. Zhu J, Zhang X, Jin R, Jiang H, Kenne DR. Probing public perceptions of antidepressants on social media: mixed methods study. JMIR Form Res 2025 Feb 26;9:e62680. [doi: 10.2196/62680] [Medline: 40009806]
12. Gilbert MK, Daughton AR, Chilcoat HD, Laffont CM, Strafford S, DeVeaugh-Geiss AM. Social listening for patient experiences with stopping extended-release buprenorphine: content analysis of Reddit messages. J Med Internet Res 2025 Apr 25;27:e71245. [doi: 10.2196/71245] [Medline: 40279635]
13. Heidari O, Sugarman OK, Winiker AK, et al. Personal experiences with xylazine and behavior change: a qualitative content analysis of Reddit posts. J Addict Med 2025;19(2):135-142. [doi: 10.1097/ADM.0000000000001383] [Medline: 39329377]
14. Sarker A. LexExp: a system for automatically expanding concept lexicons for noisy biomedical texts. Bioinformatics 2021 Aug 25;37(16):2499-2501. [doi: 10.1093/bioinformatics/btaa995] [Medline: 33244602]
15. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048]
16. Hoffman J. Tranq dope: animal sedative mixed with fentanyl brings fresh horror to U.S. drug zones. The New York Times. 2023. URL: https://www.nytimes.com/2023/01/07/health/fentanyl-xylazine-drug.html [accessed 2025-08-26]
17. Sue KL, Hawk K. Clinical considerations for the management of xylazine overdoses and xylazine-related wounds. Addiction 2024 Apr;119(4):606-608. [doi: 10.1111/add.16388] [Medline: 37939387]
18. Jawa R, Ismail S, Shang M, et al. Drug use practices and wound care experiences in the age of xylazine adulteration. Drug Alcohol Depend 2024 Oct 1;263:112390. [doi: 10.1016/j.drugalcdep.2024.112390] [Medline: 39173221]
19. Bowles JM, Copulsky EC, Reed MK. Media framing xylazine as a "zombie drug" is amplifying stigma onto people who use drugs. Int J Drug Policy 2024 Mar;125:104338. [doi: 10.1016/j.drugpo.2024.104338] [Medline: 38364357]
20. Chen AT, Wang LC, Johnny S, et al. Stigma and behavior change techniques in substance use recovery: qualitative study of social media narratives. JMIR Form Res 2025 Mar 26;9:e57468. [doi: 10.2196/57468] [Medline: 40138682]

Abbreviations

API: application programming interface
LLM: large language model
MOUD: medications for opioid use disorder
NLP: natural language processing
PWUD: people who use drugs

Edited by A Kristensen-Cabrera; submitted 01.01.25; peer-reviewed by C Zagorski, F Lamy; revised version received 28.06.25; accepted 24.07.25; published 12.09.25.

Please cite as:
Spadaro A, Hairston J, Lakamana S, Wightman R, Love J, Perrone J, Sarker A
Characterization of Reddit Posts About Xylazine-Associated Wounds: Qualitative Study
JMIR Dermatol 2025;8:e70329
URL: https://derma.jmir.org/2025/1/e70329
doi:10.2196/70329

Modern Digital Query Analytics of Patient Education Materials on Acanthosis Nigricans: Systematic Search and Content Analysis

Kevin Johny Varghese[1], MD; Som Singh[2], MD; Emily Kamali[2], BLA; Fahad Qureshi[3], MD; Aleena Jamal[4], BLA; Fawad Qureshi[5], MD

[1]Division of Dermatology, University of Kansas Medical Center, 2060 W 39th Ave, Kansas City, KS, United States

[2]Department of Biomedical Sciences, University of Missouri-Kansas City School of Medicine, Kansas City, MO, United States

[3]Department of Interventional Radiology, Loma Linda University, Loma Linda, CA, United States

[4]Department of Internal Medicine, Sidney Kimmel Medical College, Philadelphia, PA, United States

[5]Division of Nephrology and Hypertension, Department of Medicine, Mayo Clinic, Rochester, MN, United States

Corresponding Author:
Kevin Johny Varghese, MD
Division of Dermatology, University of Kansas Medical Center, 2060 W 39th Ave, Kansas City, KS, United States

Abstract

Background: Online digital materials are integral to patient education and health care outcomes in dermatology. Acanthosis nigricans (AN) is a common condition, often associated with underlying diseases such as insulin resistance. Patients frequently search the internet for information related to this cutaneous finding. To our knowledge, the quality of online educational materials for AN has not been systematically examined.

Objective: The primary objective of this study was to profile the readability and quality of the content of publicly available digital educational materials on AN and identify questions frequently asked by patients.

Methods: This study analyzed publicly available internet sources to identify the most frequent questions searched by patients regarding AN using the Google Rankbrain algorithm. Furthermore, available articles on AN were evaluated for quality and reading level using metrics such as the Brief DISCERN score, and readability was determined using three specific scales including the Flesch-Kincaid score, Gunning Fog index, and the Coleman-Liau index, based on literature.

Results: Patients most frequently accessed facts on AN from government sources, which comprised 30% (n=15) of the analyzed sources. The available articles did not meet quality standards and were at a reading level not appropriate for the general public. The majority of articles (n=29/50, 58%) had substandard Brief DISCERN scores, failing to meet the criteria for *good quality*.

Conclusions: Clinicians should be aware of the paucity of valuable online educational material on AN and educate their patients accordingly.

(JMIR Dermatol 2025;8:e60210) doi:10.2196/60210

KEYWORDS

acanthosis nigricans; dermatology; patient education; public health; skin; readability; information resource; DISCERN; general public; reading level; information seeking; information behavior

Introduction

Online digital materials are increasingly central to patient education [1]. Internet-based resources such as websites, telehealth platforms, and mobile health apps, are tools that patients interact with, in order develop health care literacy [2-4]. Furthermore, health care literacy is associated with outcomes which determine the patient's experience [5]. Patients with poor health literacy are more likely to suffer suboptimal health care outcomes [6,7]. Digital patient education plays an important role in improving outcomes in fields such as dermatology [8].

Acanthosis nigricans (AN) is a common cutaneous disorder characterized by symmetric and velvety hyperpigmented plaques [9]. These are often found in intertriginous areas such as the posterior neck, axillae, or the inguinal or inframammary regions [9]. The prevalence of AN is as high as 74% in some populations, and its incidence increases with age [10]. Obese individuals are at higher risk for AN [10]. Furthermore, a higher prevalence of AN is observed among the Native American, African American, and Hispanic populations [11]. AN typically indicates underlying insulin resistance or other endocrinological pathologies, including malignancy [12,13]. It may also be associated with other findings such as metabolic syndrome, acrochordons, hyperandrogenism, or diabetes mellitus [14]. The characteristic hyperpigmented plaques occur due to increased levels of insulin and insulin-like growth factor 1, which stimulate keratinocyte proliferation [14]. Rarely, AN may be

induced by drugs such as nicotinic acid or insulin [15]. Furthermore, some cases of AN are inherited through familial mutations in genes such as fibroblast growth factor receptor 3 (FGFR3) [16,17]. Lastly, AN can be a paraneoplastic manifestation of malignancies such as gastric adenocarcinoma [18].

Given that AN may be a manifestation of highly prevalent cardiovascular conditions such as diabetes and insulin resistance, we sought to evaluate the current online resources available to patients [12]. Additionally, patients searching online resources may encounter associations with diseases such as inherited mutations or malignancy. To our knowledge, the quality of online educational materials for AN has not been evaluated. The primary objective of this study was to assess the readability and quality of the content of publicly available digital educational materials on AN and determine the most frequently searched questions by patients.

Methods

Study design

In March 2024, a digital search was performed to extract 50 unique frequently asked questions on AN generated by the Google Rankbrain algorithm. The reviewers evaluated only materials in English. To reduce the impact of tracking cookies associated with the digital search, this search was performed using a newly installed internet browser. The digital articles associated with each question were examined for further health literacy analysis.The questions and digital articles that were extracted were then reviewed by 3 reviewers based on specific inclusion criteria as follows: (1) the article pertained to AN, (2) the article was publicly available without the requirement of a paid subscription, and (3) the content of the article consisted of at least 150 words and was written in English.

The extracted questions underwent evaluation using Rothwell's classification of questions and were categorized as either fact, policy, or value. Questions were then sorted according to their category [19-22]. For each digital article associated with a question, reviewers initially categorized each article's source as one of the following: academic institution, commercial, medical practice, government source, media outlet, or other. Following source classification, the reviewers of this study subsequently evaluated the content of each article for quality using the Brief DISCERN score [23]. The cutoff score for this instrument was established as ≥16 out of 30 for a digital article to be considered good-quality content [23]. Following quality assessment, the text of each digital article was extracted onto a plain text document and evaluated for readability. Moreover, readability was determined using 3 specific scales based on prior literature: Flesch-Kincaid score, Gunning Fog index, and the Coleman-Liau index [24-27]. This study established grade reading level recommendations for content to be approximately at the 6th-grade reading level based on previous literature [28].

Ethical Considerations

This study did not involve human subjects, and according to University of Missouri-Kansas City Institutional Review Board, under one of the categories identified in 45 CFR 46.101 (b)(4), simple observational studies of public behavior that do not involve human subjects are exempt from institutional board approval since there is no intervention involved and the behavior is not private. The data was both anonymized and deidentified.

Results

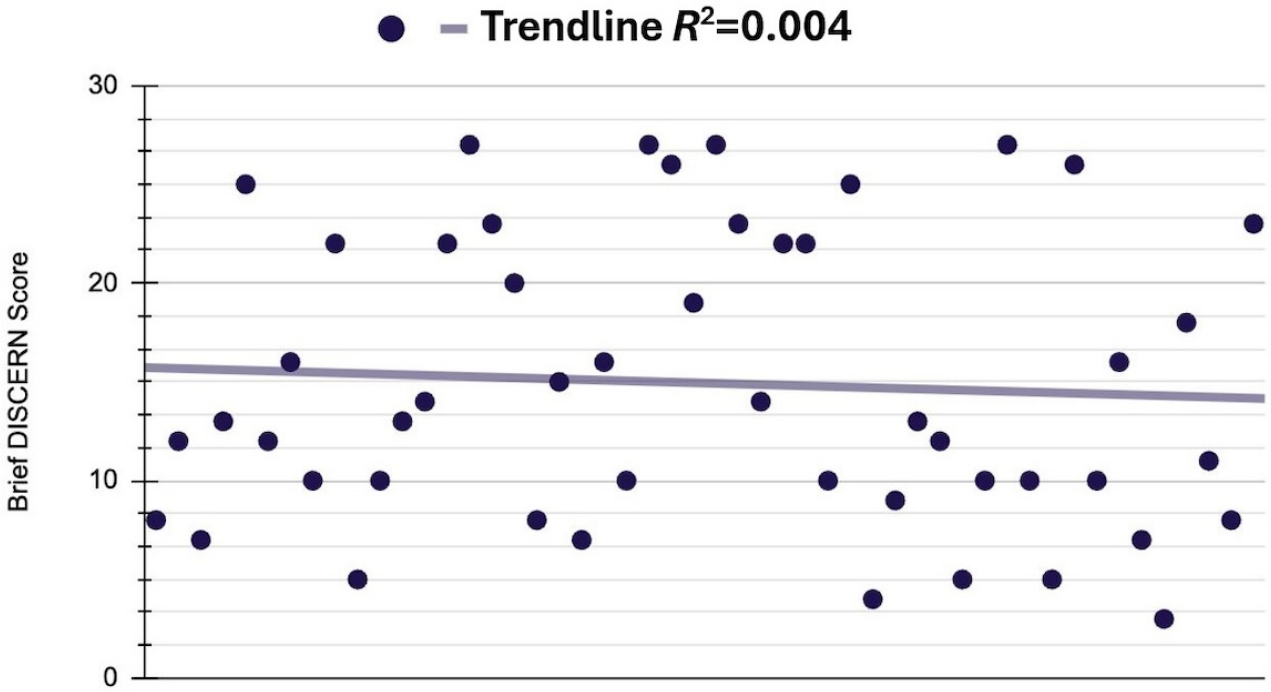
Among the 50 questions and associated digital articles extracted for this study, 15 (30%) of the sources originated from the government, followed by 13 (26%) from academic sources, 11 (22%) from commercial sources, and 5 (10%) from media outlets (Multimedia Appendix 1). Most questions (n=27, 54%) were classified as *fact* using Rothwell's classification of questions. This was followed by *policy* (n=14, 28%) and *value* (n=9, 18%) (Multimedia Appendix 2).

The mean readability of digital articles on AN did not meet grade reading level recommendations across all 3 readability metrics (Table 1). The mean Flesch-Kincaid score of the digital articles was 11.0 (SD 3.4; range 1.7-18.2). The mean Gunning Fog score was 14.7 (SD 3.5; range 7.4-21.4). and the mean Coleman-Liau index was 13.1 (SD 3.5; range 6.0-26.7). Brief DISCERN scores for articles included in this study did not meet the recommended criteria (≥16), to be considered *good quality*; the mean brief DISCERN score was 14.9 (SD 7.3; range 3.0-27). Additionally, most (29/50; 58%) articles were substandard and did not meet *good quality* (Figure 1).

Table . Mean readability scores of available articles on acanthosis nigricans.

Variable	Tools used for analysis		
	Flesch-Kincaid score	Gunning Fog score	Coleman-Liau index
Readability scores, mean (SD)	11.0 (3.4)	14.7 (3.6)	13.1 (3.5)

Figure 1. Distribution of articles on acanthosis nigricans based on the Brief DISCERN scores



Discussion

There is a paucity of literature exploring innovations in patient education on AN [29,30]. To the best of our knowledge, this study is among the first to evaluate the quality of digital educational materials for AN using the Brief DISCERN score. Government sources emerged as the most frequent contributors to public digital education materials. The findings of this study effectively demonstrate that most articles on AN do not meet established quality standards. Furthermore, the mean readability and grade reading level of these articles are often more complex than recommended guidelines [31]. This suggests a need for improvements in the publicly available digital resources.

In addition, most of the frequently searched questions by patients on AN were classified as *facts*. his suggested that patients are investigating on AN and building their knowledge base. Furthermore, it may indicate that patients require further education from their health care providers regarding this diagnosis, as they will encounter online materials of varying quality. For example, providers should consider incorporating standardized AN educational material into clinic check-out or discharge paperwork.

A key strength of this study is the use of standardized content quality assessment tools, including readability and Brief DISCERN [23]. Readability is a well-established concept in patient education that provides a metric for reviewing educational materials qualitatively based on word count, syllables, sentence structure, etc [32,33]. This is particularly important in the field of dermatology, where prior research has found that most patient educational materials do not meet recommended reading guidelines [34]. Readability has been evaluated in literature on dermatology over the years, and most articles are at a recommended grade reading level.

However, readability alone does not appropriately and effectively evaluate content quality. As a result, this study also employed the Brief DISCERN tool. Our findings indicate that the mean Brief DISCERN score of the available articles is below the minimum quality threshold. The Brief DISCERN tool provides high reliability when evaluating online articles [23]. Investigations in the future should re-evaluate the Brief DISCERN scores of articles on AN to determine the potential effect of changes over time on the quality of these materials. A key limitation of our study is the subjective nature of Rothwell's classification of questions, which may introduce potential bias in raters' scoring [35,36].

In conclusion, AN is often associated with chronic diseases, such as insulin resistance, which significantly contribute to morbidity and mortality. Patients with AN frequently use the internet for education purposes. This study implemented tools established in the literature to analyze the quality, readability, and content of patient educational materials on AN. The overall quality of these materials is poor and did not meet recommended readability standards in the United States. Clinicians can improve patient outcomes by educating patients directly regarding thi condition. Consequently, patients will have access to reliable sources of information. Future studies should examine whether the quality of online materials changes over time.

Authors' Contributions

Conceptualization: KV and FQ
Data curation: EK, AJ, SP

Formal analysis: KV and FQ
Supervision: FQ
Writing – original draft: FQ, KV
Writing – review and editing: AJ, EK, SP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Source distribution of articles.
[PNG File, 30 KB - derma_v8i1e60210_app1.png]

Multimedia Appendix 2
Rothwell's classification of questions.
[PNG File, 25 KB - derma_v8i1e60210_app2.png]

References

1. Kataria S, Ravindran V. Digital health: a new dimension in rheumatology patient care. Rheumatol Int 2018 Nov;38(11):1949-1957. [doi: 10.1007/s00296-018-4037-x] [Medline: 29713795]
2. Holland Brown TM, Bewick M. Digital health education: the need for a digitally ready workforce. Arch Dis Child Educ Pract Ed 2023 Jun;108(3):214-217. [doi: 10.1136/archdischild-2021-322022] [Medline: 35697475]
3. Kuwabara A, Su S, Krauss J. Utilizing digital health technologies for patient education in lifestyle medicine. Am J Lifestyle Med 2020;14(2):137-142. [doi: 10.1177/1559827619892547] [Medline: 32231478]
4. Conard S. Best practices in digital health literacy. Int J Cardiol 2019 Oct 1;292:277-279. [doi: 10.1016/j.ijcard.2019.05.070] [Medline: 31230937]
5. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
6. Shahid R, Shoker M, Chu LM, Frehlick R, Ward H, Pahwa P. Impact of low health literacy on patients' health outcomes: a multicenter cohort study. BMC Health Serv Res 2022 Sep 12;22(1):1148. [doi: 10.1186/s12913-022-08527-9] [Medline: 36096793]
7. Mancuso CA, Rincon M. Impact of health literacy on longitudinal asthma outcomes. J Gen Intern Med 2006 Aug;21(8):813-817. [doi: 10.1111/j.1525-1497.2006.00528.x] [Medline: 16881939]
8. Zirwas MJ, Holder JL. Patient education strategies in dermatology: part 2: methods. J Clin Aesthet Dermatol 2009 Dec;2(12):28-34. [Medline: 20725582]
9. Das A, Datta D, Kassir M, et al. Acanthosis nigricans: a review. J Cosmet Dermatol 2020 Aug;19(8):1857-1865. [doi: 10.1111/jocd.13544] [Medline: 32516476]
10. Radu AM, Carsote M, Dumitrascu MC, Sandru F. Acanthosis nigricans: pointer of endocrine entities. Diagnostics (Basel) 2022 Oct 17;12(10):2519. [doi: 10.3390/diagnostics12102519] [Medline: 36292208]
11. Stuart CA, Driscoll MS, Lundquist KF, Gilkison CR, Shaheb S, Smith MM. Acanthosis nigricans. J Basic Clin Physiol Pharmacol 1998;9(2-4):407-418. [doi: 10.1515/JBCPP.1998.9.2-4.407] [Medline: 10212845]
12. Leung AKC, Lam JM, Barankin B, Leong KF, Hon KL. Acanthosis nigricans: an updated review. Curr Pediatr Rev 2022;19(1):68-82. [doi: 10.2174/1573396318666220429085231] [Medline: 36698243]
13. Higgins SP, Freemark M, Prose NS. Acanthosis nigricans: a practical approach to evaluation and management. Dermatol Online J 2008 Sep 15;14(9):19061584. [Medline: 19061584]
14. Hermanns-Lê T, Scheen A, Piérard GE. Acanthosis nigricans associated with insulin resistance: pathophysiology and management. Am J Clin Dermatol 2004;5(3):199-203. [doi: 10.2165/00128071-200405030-00008] [Medline: 15186199]
15. Mourad AI, Haber RM. Drug-induced acanthosis nigricans: a systematic review and new classification. Dermatol Ther 2021 Mar;34(2):e14794. [doi: 10.1111/dth.14794] [Medline: 33480113]
16. Fukuchi K, Tatsuno K, Matsushita K, Kubo A, Ito T, Tokura Y. Familial acanthosis nigricans with p.K650T FGFR3 mutation. J Dermatol 2018 Feb;45(2):207-210. [doi: 10.1111/1346-8138.14107] [Medline: 29068064]
17. Fu J, Zhao Y, Wang T, Zhang Q, Xiao X. Acanthosis nigricans in a Chinese girl with FGFR3 K650 T mutation: a case report and literature review. BMC Med Genet 2019 Jan 11;20(1):8. [doi: 10.1186/s12881-019-0748-4] [Medline: 30635042]
18. Rizwan M, Iftikhar N, Sarfraz T, Ullah O. Malignant acanthosis nigricans: an indicator of internal malignancy. J Coll Physicians Surg Pak 2019 Sep;29(9):888-890. [doi: 10.29271/jcpsp.2019.09.888] [Medline: 31455489]
19. Kanthawala S, Vermeesch A, Given B, Huh J. Answers to health questions: internet search results versus online health community responses. J Med Internet Res 2016 Apr 28;18(4):e95. [doi: 10.2196/jmir.5369] [Medline: 27125622]

20. Varghese KJ, Singh SP, Qureshi FM, Shreekumar S, Ramprasad A, Qureshi F. Digital patient education on xanthelasma palpebrarum: a content analysis. Clin Pract 2023 Sep 29;13(5):1207-1214. [doi: 10.3390/clinpract13050108] [Medline: 37887084]
21. Sajjadi NB, Shepard S, Ottwell R, et al. Examining the public's most frequently asked questions regarding COVID-19 vaccines using search engine analytics in the United States: observational study. JMIR Infodemiology 2021;1(1):e28740. [doi: 10.2196/28740] [Medline: 34458683]
22. Singh SP, Ramprasad A, Luu A, Zaidi R, Siddiqui Z, Pham T. Health literacy analytics of accessible patient resources in cardiovascular medicine: What are patients wanting to know? Kans J Med 2023;16:309-315. [doi: 10.17161/kjm.vol16.20554] [Medline: 38298385]
23. Khazaal Y, Chatton A, Cochand S, et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Educ Couns 2009 Oct;77(1):33-37. [doi: 10.1016/j.pec.2009.02.016] [Medline: 19372023]
24. Kher A, Johnson S, Griffith R. Readability assessment of online patient education material on congestive heart failure. Adv Prev Med 2017;2017:9780317. [doi: 10.1155/2017/9780317] [Medline: 28656111]
25. Singh SP, Qureshi FM, Borthwick KG, Singh S, Menon S, Barthel B. Comprehension profile of patient education materials in endocrine care. Kans J Med 2022;15:247-252. [doi: 10.17161/kjm.vol15.16529] [Medline: 35899057]
26. Sudah SY, Faccone RD, Manzi JE, et al. Most patient education materials on shoulder conditions from the American Academy of Orthopaedic Surgeons exceed recommended readability levels. JSES Int 2023 Jan;7(1):126-131. [doi: 10.1016/j.jseint.2022.09.004] [Medline: 36820439]
27. Michel C, Dijanic C, Abdelmalek G, et al. Readability assessment of patient educational materials for pediatric spinal deformity from top academic orthopedic institutions. Spine Deform 2022 Nov;10(6):1315-1321. [doi: 10.1007/s43390-022-00545-1] [Medline: 35819724]
28. Eltorai AEM, Ghanian S, Adams CA Jr, Born CT, Daniels AH. Readability of patient education materials on the American Association for Surgery of Trauma website. Arch Trauma Res 2014 Jun;3(2):e18161. [doi: 10.5812/atr.18161] [Medline: 25147778]
29. Popa ML, Popa AC, Tanase C, Gheorghisan-Galateanu AA. Acanthosis nigricans: to be or not to be afraid. Oncol Lett 2019 May;17(5):4133-4138. [doi: 10.3892/ol.2018.9736] [Medline: 30944606]
30. Kong AS, Williams RL, Smith M, et al. Acanthosis nigricans and diabetes risk factors: prevalence in young persons seen in southwestern US primary care practices. Ann Fam Med 2007;5(3):202-208. [doi: 10.1370/afm.678] [Medline: 17548847]
31. Rooney MK, Santiago G, Perni S, et al. Readability of patient education materials from high-impact medical journals: A 20-year analysis. J Patient Exp 2021;8:2374373521998847. [doi: 10.1177/2374373521998847] [Medline: 34179407]
32. Fontaine G, Zagury-Orly I, Maheu-Cadotte MA, et al. A meta-analysis of the effect of paper versus digital reading on reading comprehension in health professional education. Am J Pharm Educ 2021 Nov;85(10):8525. [doi: 10.5688/ajpe8525] [Medline: 34301544]
33. Daraz L, Morrow AS, Ponce OJ, et al. Readability of online health information: a meta-narrative systematic review. Am J Med Qual 2018;33(5):487-492. [doi: 10.1177/1062860617751639] [Medline: 29345143]
34. Hunt WTN, McGrath EJ. Evaluation of the readability of dermatological postoperative patient information leaflets across England. Dermatol Surg 2016 Jun;42(6):757-763. [doi: 10.1097/DSS.0000000000000720] [Medline: 27176870]
35. Smith SR, Hodakowski A, McCormick JR, et al. Patient-directed online education for carpal tunnel syndrome and release: analysis of what patients ask and quality of resources. J Hand Surg Glob Online 2023 Nov;5(6):818-822. [doi: 10.1016/j.jhsg.2023.07.015] [Medline: 38106941]
36. Ferrell SC, Ferrell MC, Claassen A, Balogun SA, Vassar M. Frequently asked questions about mobility devices among older adults. Eur Geriatr Med 2023 Oct;14(5):1075-1081. [doi: 10.1007/s41999-023-00815-9] [Medline: 37505403]

Abbreviations

AN: acanthosis nigricans

Edited by R Dellavalle; submitted 04.05.24; peer-reviewed by A Michenko, T Skrzypczak; revised version received 05.11.24; accepted 06.11.24; published 06.01.25.

Please cite as:
Varghese KJ, Singh S, Kamali E, Qureshi F, Jamal A, Qureshi F
Modern Digital Query Analytics of Patient Education Materials on Acanthosis Nigricans: Systematic Search and Content Analysis
JMIR Dermatol 2025;8:e60210
URL: https://derma.jmir.org/2025/1/e60210
doi:10.2196/60210

Long-Term Efficacy and Safety of755-nm Alexandrite Laser for Axillary Hair Removal: A Comparative Analysis of Single vs. DualFlash lamp Systems

Kawthar Shurrab, PhD; Manal Asad Nassr, MSc

Biomedical Photonics Lab, Higher Institute for Laser Research and Applications, Damascus University, Airport Highway, Damascus, Syrian Arab Republic

Corresponding Author:
Kawthar Shurrab, PhD
Biomedical Photonics Lab, Higher Institute for Laser Research and Applications, Damascus University, Airport Highway, Damascus, Syrian Arab Republic

Abstract

Background: Laser hair removal is a noninvasive cosmetic procedure that targets melanin in hair follicles through selective photothermolysis.

Objective: This study aims to evaluate the long-term efficacy and safety of the 755-nm alexandrite laser for axillary hair removal by comparing single- and dual-flash lamp systems.

Methods: In total, 40 women aged 20 to 35 years with Fitzpatrick skin types II and III participated in a study on laser hair removal for their axillae. Participants underwent 3 treatment sessions, each spaced 4 weeks apart, from January to April 2024. They were divided into 2 groups, both receiving treatments with an alexandrite laser (755 nm, 14 mm spot size). The first group was treated with a dose of 8 J/cm² using a single-flash lamp device operating at 5 Hz, with a pulse duration of 10 ms. The second group received doses between 9 and 11 J/cm² from a dual-flash lamp device operating at 2.5 Hz, with pulse durations ranging from 10 to 15 ms. Photographs and hair counts were taken at baseline and 1 month after the final session. A 2-tailed *t* test was used to assess statistical significance, and regression analysis evaluated treatment effects. Pain scores and side effects were assessed using a visual analog scale in a satisfaction questionnaire.

Results: The dual-flash lamp laser achieved an overall hair reduction of 94%, while the single-flash lamp laser resulted in a 91% reduction in the axilla. The difference was not statistically significant (P=.14). No serious adverse effects were reported with either device, indicating effective safety features.

Conclusions: The outcomes show that both systems provide similar results in terms of efficacy and safety, with no reported side effects, and results were maintained even after 6 and 12 months of follow-up.

(JMIR Dermatol 2025;8:e76523) doi:10.2196/76523

KEYWORDS

hair removal; 755-nm laser; single flash lamp; dual flash lamp; photothermolysis; axillary hair reduction; alexandrite laser

Introduction

Laser hair removal (LHR) is a widely used and effective approach for long-term hair reduction [1,2]. The technique is grounded in the principle of selective photothermolysis, wherein laser pulses target melanin in hair follicles to induce photothermal destruction while preserving surrounding tissue [3-5]. Wavelengths between 600 and 1100 nm are particularly effective for this purpose, as they efficiently absorb melanin, ensuring effective follicular stem cell disruption. Common devices for LHR include long-pulsed lasers (eg, ruby, alexandrite, diode, and neodymium-doped yttrium aluminum garnet) and intense pulsed light systems [6-9]. Among these, the 755-nm alexandrite laser is considered the gold standard due to its superior melanin absorption and efficacy, particularly for skin types I to III [10-12]. Its long pulses maximize follicle energy absorption while reducing skin damage risk [13-15].

Lasers operate through flashlamp pumping, which uses gas-filled tubes, typically xenon or krypton, to generate high-intensity light that excites the laser medium [16,17]. This technology has significantly advanced LHR devices. Many systems use a single flashlamp to target hair follicles while protecting surrounding tissues, offering reliability and low maintenance. In contrast, dual-flashlamp–pumped systems, which feature 2 flashlamps adjacent to the laser crystal, provide uniform pumping and enhanced stability, enabling rapid pulses and higher energy fluence [18]. This enhances treatment efficiency and reduces session durations, making it ideal for larger areas. The effectiveness of alexandrite systems depends on their energy delivery, while single-flashlamp devices use low fluence with

higher repetition rates, allowing gradual thermal accumulation in hair follicles. Dual-flashlamp configurations offer higher fluence and more uniform energy output, enhancing treatment efficiency for larger areas. Recent research indicates that treatment outcomes are influenced more by the energy delivery method than by absolute energy levels, affecting follicular thermal thresholds and hair reduction durability [19,20]. In addition to hair removal, alexandrite lasers have shown promising applications in other dermatologic contexts. A recent split-body clinical trial demonstrated the efficacy and safety of these systems in treating keratosis pilaris, confirming that optimized fluence and pulse duration settings can extend their therapeutic applications beyond traditional photoepilation [15] to include vascular and pigmented lesion treatments [21]. Moreover, novel approaches in LHR emphasize expanding safety profiles across all skin types, highlighting the role of advanced epidermal cooling and energy modulation in minimizing side effects [22]. Although emerging platforms combining multiple wavelengths (810, 940, and 1060 nm) are increasingly adopted for treating fine or less pigmented hairs, the alexandrite laser remains the benchmark for comparative evaluation, offering clinicians robust evidence to optimize treatment protocols [23]. This study represents the first direct comparison between single- and dual-flashlamp alexandrite laser systems for hair removal using the 755-nm wavelength. Unlike previous research that focused solely on single-flash devices and short-term outcomes, our work provides a detailed analysis of treatment parameters, fluence, frequency, pulse duration, and the number of passes per session and includes a comprehensive 12-month follow-up, which enhances the reproducibility and clinical utility of our findings. By integrating patient-centered considerations such as session cost, treatment duration, and side effect profiles, this study offers novel insights that support personalized decision-making in clinical dermatology and contribute to optimizing long-term hair removal protocols.

Methods

Ethical Considerations

The comparative study methods and protocols were approved by the Ethics Committee of Damascus University (approval ID: HILRA-261124-372). All participants voluntarily took part in the study and provided written informed consent prior to enrollment. No financial or material compensation was provided, and participants did not receive any free treatments or benefits in exchange for participation. Before enrollment, all participants were informed about the complete treatment protocol and provided written consent for data collection and the publication of their images. They were also informed that participation was entirely voluntary and that they could withdraw at any time without penalty. All identifying information was kept confidential, with only anonymized reference numbers used throughout data collection, storage, and analysis.

Clinical Data

The study was conducted at a private laser clinic in Damascus, Syria, from January to April 2024. The participants in the study were aged between 20 and 35 years, with a mean age of 31.3 (SD 4.3) years for one group and mean age of 30.1 (SD 4.2) years for another. None of the participants had previously undergone laser treatments in the axillary area, and they were classified as Fitzpatrick skin types II and III. All participants had dark terminal hairs. We screened interested female participants with unwanted axillary hair who were first-time laser users and met the inclusion criteria for participation. Exclusion criteria included individuals who were suntanned, those with contraindications to laser treatment, and those who were not willing or committed to following the required precare and postcare procedures. Female participants were chosen for the study to focus on one sex and to minimize variables such as hormonal differences between sexes [24]. The study excluded individuals who did not meet certain eligibility criteria, including those with blonde, red, or light-colored hair in the axillary region; those with tanned or sun-exposed skin; and those who were pregnant or breastfeeding. Additional exclusions are applied to individuals with a history of seizures, prior laser treatments in the underarm area, skin infections, or those prone to hypertrophic scarring or keloids.

After establishing the study sample, each participant was assigned a randomly generated reference number to allocate them to a specific laser machine [25].

Participants on medications such as isotretinoin, antibiotics, or anticoagulants, as well as those with tattoos in the treatment area, joint pain during gold therapy, contagious diseases, or diabetes, were also disqualified. Other criteria included those with suspicious pigmented lesions, users of photosensitive medications, and patients undergoing radiation or chemotherapy.

Research Design

All 40 participants were randomly assigned to 2 groups, A and B, simultaneously. A number generator was used to ensure that the allocation to the treatment groups was concealed, thereby reducing bias and enhancing the reliability of the results [25,26]. Participants were assigned to receive hair removal treatments using either the DEKA Motus Axe 755-nm alexandrite laser (DEKA Laser), which features a 20 mm spot size and uses single-flashlamp technology, along with a Moveo handpiece that allows for continuous motion delivery instead of traditional single pulses and provides epidermal cooling, or the DEKA Again 755-nm alexandrite laser (DEKA Laser). The latter also has a 20 mm spot size and uses dual-flashlamp technology combined with air cooling down to −20 °C, minimizing pain and thermal damage.

Participants in the study received 3 treatment sessions, scheduled 4 weeks apart. No topical anesthetics or medications were used during these sessions, and patients reported no discomfort or redness. They were instructed to avoid all other methods of hair removal, except for shaving. They were allowed to shave the area 2 weeks after the laser session if needed.

Each participant received aftercare and treatment instructions, including guidelines for shaving to ensure uniformity. They were asked to shave their underarm hair with a razor 3 days before their treatment. Additionally, they were advised to use a broad-spectrum sunscreen and to avoid heat, humidity,

sweating, friction, rubbing, cosmetics, and salon procedures for 3 to 5 days after each laser session.

A single technician conducted all sessions. After completing the 3 sessions, treatment was paused for evaluation. Posttreatment evaluations were performed at 3 time points: 1 month after the final session, 6 months thereafter, and at the 12-month follow-up.

The participants were divided into 2 groups: group A received 755-nm alexandrite treatments using the DEKA Motus AX device, with a dose of 8 J/cm², a frequency of 5 Hz, and a pulse duration of 10 milliseconds, and group B underwent 755-nm alexandrite treatments using the DEKA Again device, with a dose ranging from 9 to 11 J/cm², a frequency of 2.5 Hz, and a pulse duration between 10 and 15 milliseconds. The entire area was treated by completing 1 pass horizontally per session.

Evaluation and Outcome Measures

Clinical photographs were taken before the first treatment session, 1 month after the last session, and at a final follow-up appointment scheduled 12 months after treatment. This was done using a mobile device to assess hair counts and thickness [27-29]. Hair counts were analyzed using HowMany AI (YesChat AI), a free object-counting software tool whose accuracy was validated by comparison with manual counting [30].

Patients were instructed to measure their pain intensity during the LHR sessions using a 10-cm Visual Analog Scale. Pain levels were systematically classified according to the following criteria: no pain (0), mild pain (1-3), moderate pain (4-6), and severe pain (7-10) [7-10,31,32].

After completing the treatment, patients received a questionnaire to evaluate their satisfaction with the level of improvement or any adverse effects. The ratings were categorized into 4 groups: mild (<25% improvement), moderate (25% to <50% improvement), good (50% to <75% improvement), and very good (>75% improvement) [33].

Data Analysis

All collected data were recorded and analyzed using the SPSS software (version 26.0; IBM Corp). Various descriptive and inferential statistical techniques were used. Statistical descriptors, such as tables, means, and percentages, were used to analyze patterns within the collected data. Additionally, Microsoft Excel was used to enhance visual representation through the generation of relevant charts and graphs.

During statistical analysis, data were assessed using either parametric or nonparametric tests to determine the normality of the variables based on their distribution. A *t* test was used to compare differences within groups, while an analysis of covariance (ANCOVA) regression analysis was used to determine the impact of these differences. A *P* value <.05 was considered statistically significant. To evaluate the differences between the 2 technologies, we compared the percentage of hair reduction achieved with the dual-flashlamp technology to that obtained with the single-flashlamp technology, using the following equation [34]:

$$\Delta\text{Reduction} = [H_{b,\text{dual}} - H_{p,\text{dual}} H_{b,\text{dual}} - H_{b,\text{single}} - H_{p,\text{single}} H_{b,\text{single}}] \times 100$$

Where $H_{b,\text{dual}}$ and $H_{p,\text{dual}}$ represent the baseline and posttreatment hair counts for the dual-flashlamp group, and $H_{b,\text{single}}$ and $H_{p,\text{single}}$ represent the baseline and posttreatment counts for the single-flashlamp group.

Results

Assessment Scores

A total of 40 Syrian female patients, aged between 20 and 35 years (with a mean age of 31.3, SD 4.3, years in one group and 30.1, SD 4.2, years in another), participated in this study. According to the Fitzpatrick skin type classification, 17 (43%) patients were classified as skin phototype II, while 23 (58%) patients were classified as skin phototype III.

For the treatment, the initial fluence parameters were set as follows: for the single-flashlamp laser group, the settings were 8 J/cm², a frequency of 5 Hz, and a pulse duration of 10 milliseconds. For the dual-flashlamp laser group, the parameters ranged from 9 to 11 J/cm², with a frequency of 2.5 Hz and a pulse duration between 10 and 15 milliseconds.

The findings indicated significant hair reduction in both groups when comparing baseline hair counts to measurements taken 1 month after the final 3 treatment sessions. Statistically significant differences were observed in both groups (P< .001), highlighting the effectiveness of the treatment.

Visual documentation in Figure 1 and Table 1 supports the quantitative findings by showcasing representative baseline and posttreatment photographs of 4 patients from both groups, illustrating consistent axillary hair reduction across both skin phototypes and treatment modalities.

Figure 1. Clinical photographs taken at baseline and 1 month after the final session of 3 treatments show a significant decrease in hair count. (A) A 26-year-old female patient with skin phototype II and (B) a 31-year-old female patient with skin phototype III from group A were treated with single-flashlamp technology, while (C) a 27-year-old female patient with skin phototype II and (D) a 30-year-old female patient with skin phototype III from group B were treated with dual-flashlamp technology.

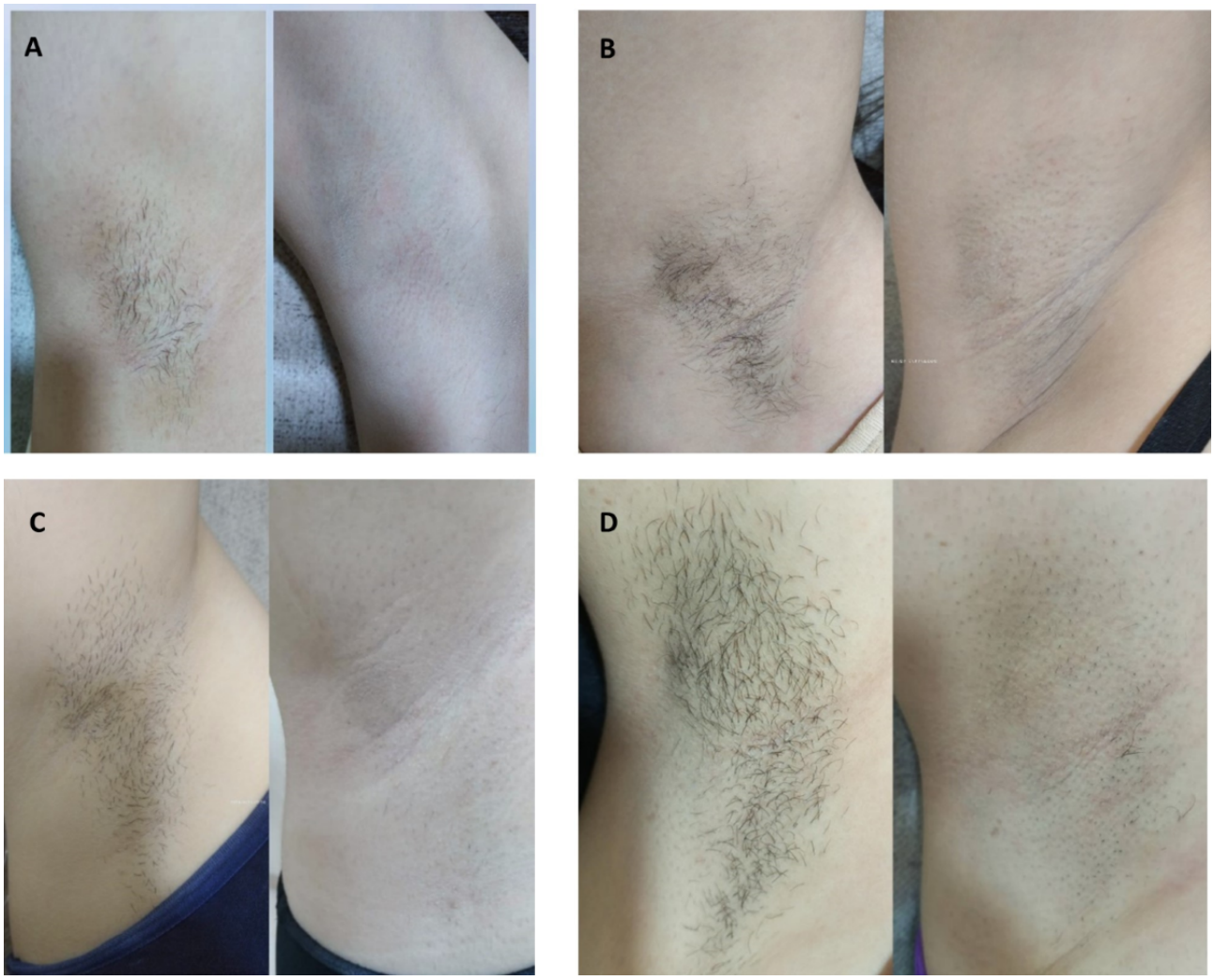


Table . Baseline hair counts, mean reduction, and 95% CIs after 3 treatment sessions in single- and dual-flashlamp laser systems.

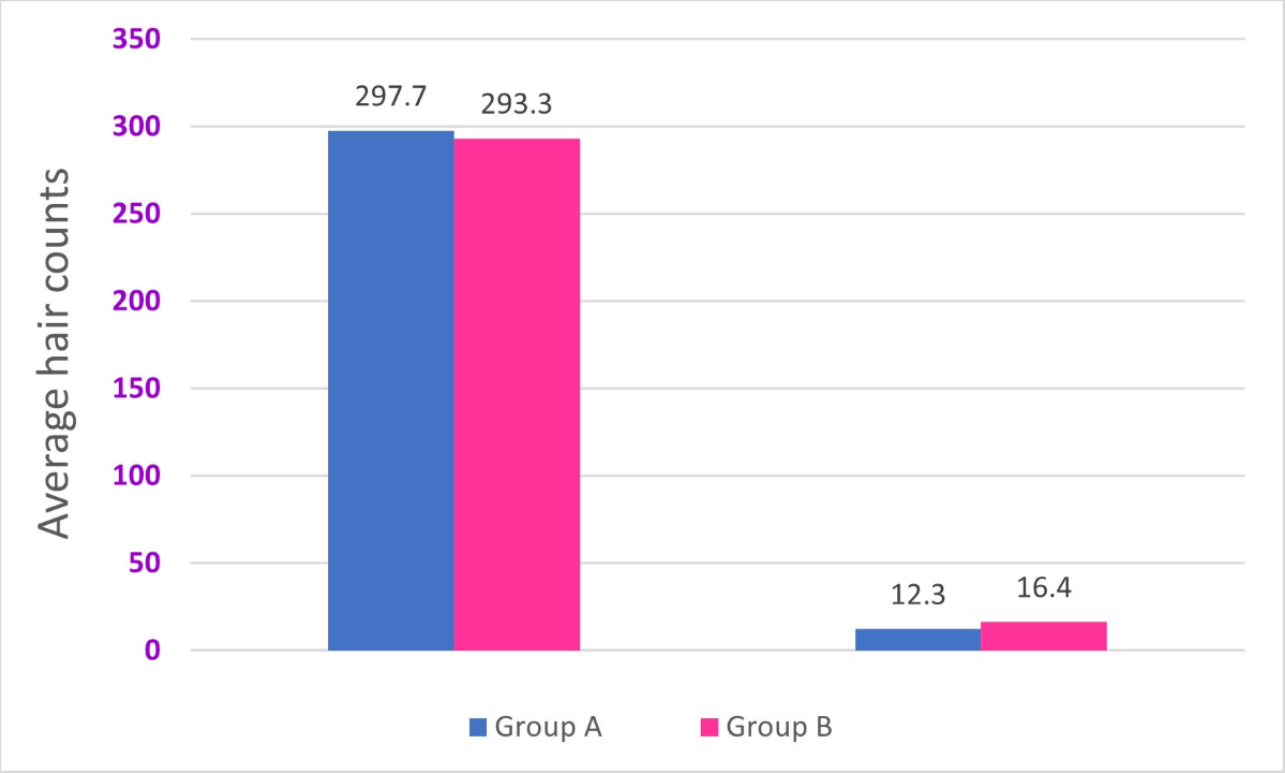
	Baseline, mean (SD)	After 3 treatments, mean (SD)	Mean reduction (95% CI)	Reduction[a] (95% CI)
Group A	293.7 (5.9)	17.6 (2.2)	268.5 (262.8 to 274.2)	91.4 (88.8 to 99.5)
Group B	297.7 (4.9)	16.6 (3.7)	279.2 (276.5 to 281.9)	94 (91.9 to 97.0)
Between-group difference	N/A[b]	N/A	10.7 (4.5 to 16.9)	2.6 (–0.5 to 5.7)

[a]The values are presented as percentages.

[b]N/A: not applicable; because the between-group comparison reflects a mean difference with a 95% CI rather than raw measurements required for mean (SD) values.

The comparison of the 2 groups showed a significant reduction in hair counts after treatment as shown in Figure 2. In group A, the average hair count dropped from 293.7 (SD 5.9) at baseline to 17.3 (SD 2.2) after 3 treatments (t_{19}=209.2; P<.001), yielding a mean reduction of 268.47 (95% CI 262.76-274.18) hairs, equivalent to 91.4% (95% CI 88.8%-99.5%) reduction. Similarly, group B showed a decrease from 297.7 (SD 4.9) at baseline to 16.6 (3.7) after 3 treatments (t_{19}=107.5; P<.001), with a mean reduction of 279.20 (95% CI 276.49-281.91) hairs, equivalent to 94% (95% CI 91.9%-97.0%) reduction.

Figure 2. Comparative analysis of groups A and B using an independent samples test (91.4% vs 94%, with a difference of 2.7%).



The between-group difference in mean reduction was 10.73 (95% CI 4.54-16.92) hairs, favoring the dual flashlamp. However, the posttreatment mean difference in residual hair count was −0.65 (95% CI −2.6 to 1.3) hairs, indicating that both technologies achieved highly comparable and clinically significant efficacy, with no statistically or clinically meaningful difference. These results align with the ANCOVA model, which confirmed no significant group effect after baseline adjustment (P=.14).

To evaluate the differences between the 2 laser treatments, an ANCOVA test was performed, and the results are as follows.

The ANCOVA showed no statistically significant differences between the groups (single- vs dual-flashlamp technology) for the posttreatment hair count after adjusting for baseline measurements ($F_{1,37}$=2.23; P=.14; η^2=0.06). Additionally, the baseline measurement did not significantly impact the outcome ($F_{1,37}$=0.718; P=.40; η^2=0.02). The corrected model accounted for 9% of the variance in posttreatment outcomes (R^2=0.09), with an adjusted R^2 value of 0.04, as presented in Table 2.

Table . The analysis of covariance results comparing the single- and dual-flashlamp laser groups after 3 treatments.

Source	Type III sum of squares (*df*)	Mean square	*F* test (*df*)	*P* value	Partial η^2
Corrected model[a]	15.04 (2)	7.52	1.84 (2.37)	.17	0.09
Intercept	12.99 (1)	12.99	3.17 (1.37)	.08	0.08
Baseline (before)	2.94 (1)	2.94	0.72 (1.37)	.40	0.02
Group (A vs B)	9.14 (1)	9.14	2.24 (1.37)	.14	0.06
Error	151.36 (37)	4.09	—[b]	—	—
Total	11,322.00 (40)	—	—	—	—
Corrected total	166.40 (39)	—	—	—	—

[a]Model fit: R^2=0.09 (adjusted R^2=0.04).

[b]Not available.

Side Effects and Patient Satisfaction

Mild side effects were reported during the initial treatment session in both groups, but these effects diminished in subsequent sessions. Approximately 65% (13/20) of participants who received a single-flashlamp laser treatment reported experiencing pain rated as mild to moderate, and 25% (5/20) noticed slight redness or mild erythema. In contrast, 70% (14/20) of participants who underwent dual-flashlamp laser treatment reported similar pain levels, while 20% (4/20) experienced redness that only appeared during the treatment (Table 3). All these side effects disappeared within minutes after treatment.

Axillary hyperhidrosis was observed following LHR in both groups. In group A, 30% (6/20) of patients reported this condition, whereas group B exhibited a higher incidence at 60% (12/20). While some participants regained normal perspiration spontaneously, others continued to experience moderate hyperhidrosis for up to 12 months. The effects were moderate for all participants who received treatment. There were no unexpected or severe adverse events recorded during any of the follow-up periods at 3, 6, and 12 months after the last treatment. Additionally, no cases of burns, hyperpigmentation, depigmentation, or paradoxical hair growth were reported in either group.

Table . Comparison of side effects between single- and dual-flashlamp groups with statistical analysis

Side effect	Single-flashlamp laser (group A), n (%)	Dual-flashlamp laser (group B), n (%)	Chi-square (*df*)	*P* value[a]	95% CI for difference
Pain (mild- to moderate)	13 (65)	14 (70)	0.1 (1)	.74	−28.2 to 38.2
Redness or erythema	5 (25)	4 (20)	0.1 (1)	.72	−23.9 to 33.9
Axillary hyperhidrosis	6 (30)	12 (60)	3.3 (1)	.07	−2.5 to 62.5

[a]*P* value corresponds to the Fisher exact test.

Despite the specific guidelines about shaving frequency, none of the participants in either group needed to shave between sessions.

To assess the reliability of observed differences in side effects, 95% CIs were calculated. For pain and erythema, the 95% CIs were −28.2% to 38.2% and −23.9% to 33.9%, respectively. Both intervals included 0, indicating no statistically significant differences between the 2 laser modalities in terms of tolerability. However, the 95% CI for axillary hyperhidrosis (−2.5% to 62.5%) was notably wide, suggesting a potential trend toward increased incidence in the dual-flashlamp group. Although this finding did not reach statistical significance (*P*=.07), it warrants further investigation in larger cohorts to determine whether higher-energy delivery systems contribute to this side effect.

Patient satisfaction was high for both groups. Of the total 40 patients, 6 (15%) reported their improvement as good, while the remaining 34 (85%) rated it as very good. After just 3 sessions, all participants experienced a significant reduction in hair counts, and these results remained consistent for 12 months.

Discussion

Principal Findings

This study presents the first direct comparison of single- and dual-flashlamp alexandrite laser systems for axillary hair removal, demonstrating substantial and sustained hair reduction in both groups, with no statistically significant difference in efficacy or safety.

Hair removal is a widely practiced cosmetic procedure among men and women, with a notably higher demand among women. Motivations for this practice include esthetic enhancement, personal comfort, and hygiene considerations. LHR has emerged as a preferred method due to its efficacy in achieving smooth skin, reducing ingrown hairs, and enhancing self-confidence [33].

Despite extensive research on laser technologies using various wavelengths [5,28,35-37], limited studies have examined the comparative efficacy of different laser systems.

Both modalities demonstrated substantial and sustained hair reduction across Fitzpatrick skin types II and III, with results maintained at 6- and 12-month follow-ups. Group A (single-flash) achieved a 91.4% (95% CI 88.8%-99.5%) reduction, while group B (dual-flash) achieved 94% (95% CI 91.9%-97.0%) reduction. Although the dual-flash system yielded a slightly higher mean reduction (10.7, 95% CI 4.54-16.9 hairs), the adjusted ANCOVA analysis confirmed no statistically significant difference (*P*=.14), reinforcing the clinical equivalence of both techniques.

The results suggest that LHR efficacy is primarily determined by energy and energy density delivery to the tissue rather than energy output or the number of flashlamps used. Low fluence, high repetition rate lasers typically found in single-flashlamp systems deliver energy in multiple passes, facilitating gradual thermal accumulation within hair follicles [19,20]. Conversely, although the dual-flashlamp technique generates higher overall energy output, hair follicles may reach their thermal damage threshold using a single-flashlamp laser, rendering additional energy ineffective in enhancing treatment outcomes. Thus, both methods produce comparable levels of follicular damage, yielding similar hair reduction results.

Importantly, our outcomes enhance reproducibility and clinical applicability of the results by detailing all treatment parameters, fluence, frequency, pulse duration, and number of passes and by incorporating a rigorous 12-month follow-up. Unlike many previous reports that lacked consistent technical documentation, our findings establish evidence-based guidance for optimizing treatment protocols [66].

Patients undergoing LHR may experience a range of side effects, from mild discomfort to more severe complications [36]. Arsiwala and Majid [37] highlighted that laser treatment efficacy is influenced by both the laser device and the expertise of the practitioner. Serious adverse effects can occur when procedures are performed by untrained individuals.

Regarding safety, both systems exhibited favorable profiles. Mild to moderate pain was reported by 65% (13/20) of group A and 70% (14/20) of group B participants, with transient erythema observed in 25% (5/20) and 20% (4/20) , respectively. These side effects resolved within minutes, likely due to

adherence to professional laser protocols, including optimized treatment parameters and adequate epidermal cooling before, during, and after the procedure, and were not statistically significant (pain 95% CI −28.2% to 38.2%; erythema 95% CI −23.9% to 33.9%). However, axillary hyperhidrosis was more prevalent in group B (12/20, 60%) than in group A (6/20, 30%), with a wide 95% CI −2.5% to 62.5%, suggesting a potential trend that warrants further investigation. No cases of burns, pigmentary changes, or paradoxical hair growth were reported.

Patient satisfaction was uniformly high, with 85% (34/40) rating their improvement as very good and 15% (6/40) as good. Notably, none of the participants required shaving between sessions, indicating robust follicular suppression.

From a practical standpoint, the comparable efficacy of both systems allows clinicians to prioritize other factors in device selection. Single-flash systems may offer cost advantages and slightly reduced side effects, while dual-flash systems may provide faster treatment times and enhanced device longevity. These considerations are especially relevant in resource-limited settings or high-throughput clinics.

Limitations and Strengths

Several limitations need to be acknowledged. First, the relatively small sample size (n=40) may have limited the statistical power to detect minor differences between the 2 treatment modalities. Second, the inclusion of only female participants restricts the generalizability of our findings to male populations, in whom hormonal and physiological differences may affect treatment outcomes; they may require a separate study. Third, our study population consisted exclusively of individuals with Fitzpatrick skin types II and III, which limits the applicability of our results to patients with darker skin types, where safety and efficacy profiles of alexandrite lasers may differ. Fourth, although we used the HowMany AI–based hair counting tool, which demonstrated consistency with manual counts, the possibility of measurement bias or inaccuracy cannot be entirely excluded. Finally, no formal a priori or post hoc power analysis was conducted, and this should be considered when interpreting the findings. Future research with larger, more diverse cohorts, including both sexes and a wider range of skin types, is warranted. Incorporating multivariate models with additional covariates such as age, baseline hair density, and hair thickness may further improve adjustment for potential confounders and enhance the generalizability of results.

Although there are some limitations to the study, it also has significant strengths. In particular, the inclusion of a 12-month follow-up period provides strong evidence of the sustainability and validity of treatment outcomes, which is often lacking in comparative studies. Additionally, reporting detailed treatment parameters enhances the reproducibility of the results and offers practical guidance for clinicians seeking to optimize their alexandrite LHR protocols.

Conclusions

This study compared single- and dual-flashlamp laser systems for hair removal at a wavelength of 755 nm in individuals with Fitzpatrick skin types II and III. Findings indicated that the method of energy delivery, shaped by flashlamp configuration, rather than the total energy output, significantly influenced treatment efficacy, with both systems achieving comparable hair reduction. While dual-flashlamp systems may enhance device longevity and performance consistency and reduce session time, single-flashlamp systems offer greater cost-effectiveness and are associated with fewer side effects. Further research with larger sample sizes is warranted to refine treatment guidelines and optimize outcomes in LHR.

Acknowledgments

The authors would like to thank the participants of the study. All authors declare 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.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: KS
Data curation: KS
Formal analysis: KS
Investigation: KS
Methodology: KS
Project administration: KS
Resources: KS, MAN
Software: KS
Supervision: KS
Validation: KS
Visualization: KS
Writing – original draft: KS

Writing – review & editing: KS

Conflicts of Interest

None declared.

References

1. Eremia S, Li CY, Umar SH, Newman N. Laser hair removal: long-term results with a 755 nm alexandrite laser. Dermatol Surg 2001 Nov;27(11):920-924. [doi: 10.1046/j.1524-4725.2001.01074.x] [Medline: 11737124]
2. Gan SD, Graber EM. Laser hair removal: a review. Dermatol Surg 2013 Jun;39(6):823-838. [doi: 10.1111/dsu.12116] [Medline: 23332016]
3. Altshuler GB, Anderson RR, Manstein D, Zenzie HH, Smirnov MZ. Extended theory of selective photothermolysis. Lasers Surg Med 2001;29(5):416-432. [doi: 10.1002/lsm.1136] [Medline: 11891730]
4. Thomas MM, Houreld NN. The “in’s and outs” of laser hair removal: a mini review. J Cosmet Laser Ther 2019;21(6):316-322. [doi: 10.1080/14764172.2019.1605449] [Medline: 31018716]
5. Hendricks K, Nxumalo CT, Makgobole MU, Ghuman S, Jacobs D, Mpofana N. Evaluating the effectiveness of laser hair reduction using a home use laser in comparison to a Diode laser. PLoS One 2023;18(5):e0286162. [doi: 10.1371/journal.pone.0286162] [Medline: 37235570]
6. Haedersdal M, Wulf HC. Evidence-based review of hair removal using lasers and light sources. J Eur Acad Dermatol Venereol 2006 Jan;20(1):9-20. [doi: 10.1111/j.1468-3083.2005.01327.x] [Medline: 16405602]
7. Elman M, Klein A, Slatkine M. Dark skin tissue reaction in laser assisted hair removal with a long-pulse ruby laser. J Cutan Laser Ther 2000 Mar;2(1):17-20. [doi: 10.1080/14628830050516551] [Medline: 11446086]
8. Amin SP, Goldberg DJ. Clinical comparison of four hair removal lasers and light sources. J Cosmet Laser Ther 2006 Jun;8(2):65-68. [doi: 10.1080/14764170600717902] [Medline: 16766483]
9. Eremia S, Li C, Newman N. Laser hair removal with alexandrite versus diode laser using four treatment sessions: 1-year results. Dermatol Surg 2001 Nov;27(11):925-929. [doi: 10.1046/j.1524-4725.2001.01073.x] [Medline: 11737125]
10. Russe E, Purschke M, Herold M, Sakamoto FH, Wechselberger G, Russe-Wilflingseder K. Evaluation of safety and efficacy of laser hair removal with the long-pulsed 755 nm wavelength laser: a two-center study with 948 patients. Lasers Surg Med 2020 Jan;52(1):77-83. [doi: 10.1002/lsm.23160] [Medline: 31579971]
11. Bernstein EF, Basilavecchio L, Plugis J. Bilateral axilla hair removal comparing a single wavelength alexandrite laser with combined multiplexed alexandrite and Nd:YAG laser treatment from a single laser platform. J Drugs Dermatol 2012 Feb;11(2):185-190. [Medline: 22270200]
12. Khoury JG, Saluja R, Goldman MP. Comparative evaluation of long-pulse alexandrite and long-pulse Nd:YAG laser systems used individually and in combination for axillary hair removal. Dermatol Surg 2008 May;34(5):665-670. [doi: 10.1111/j.1524-4725.2007.34125.x] [Medline: 18261098]
13. Tatlidede S, Egemen O, Saltat A, Turgut G, Karasoy A, Kuran I. Hair removal with the long-pulse alexandrite laser. Aesthet Surg J 2005;25(2):138-143. [doi: 10.1016/j.asj.2005.02.003] [Medline: 19338804]
14. Garcia C, Alamoudi H, Nakib M, Zimmo S. Alexandrite laser hair removal is safe for Fitzpatrick skin types IV-VI. Dermatol Surg 2000 Feb;26(2):130-134. [doi: 10.1046/j.1524-4725.2000.99185.x] [Medline: 10691941]
15. Li M, Bai Y, Duan Z, et al. Efficacy and safety of long-pulsed 755-nm alexandrite laser for keratosis pilaris: a split-body randomized clinical trial. Dermatol Ther (Heidelb) 2022 Aug;12(8):1897-1906. [doi: 10.1007/s13555-022-00771-5] [Medline: 35838863]
16. Choubey A, Jain RK, Vishwakarma SC, Upadhyaya BN, Oak SM. Performance improvement of long pulse Nd:YAG laser using advanced diffuse ceramic reflectors. Mater Focus 2013 Oct 1;2(5):362-368. [doi: 10.1166/mat.2013.1103]
17. Koechner W. Solid State Laser Engineering, 6th edition: Springer; 2006.
18. Choubey A, Vishwakarma SC, Misra P, et al. A highly efficient and compact long pulse Nd:YAG rod laser with 540 J of pulse energy for welding application. Rev Sci Instrum 2013 Jul;84(7):073108. [doi: 10.1063/1.4812635] [Medline: 23902045]
19. Braun M. Permanent laser hair removal with low fluence high repetition rate versus high fluence low repetition rate 810 nm diode laser--a split leg comparison study. J Drugs Dermatol 2009 Nov;8(11 Suppl):s14-s17. [Medline: 19916262]
20. Koo B, Ball K, Tremaine AM, Zachary CB. A comparison of two 810 diode lasers for hair removal: low fluence, multiple pass versus a high fluence, single pass technique. Lasers Surg Med 2014 Apr;46(4):270-274. [doi: 10.1002/lsm.22226] [Medline: 24752608]
21. Chen C, Ke Y. Applications of long-pulse alexandrite laser in cosmetic dermatology: a review. Clin Cosmet Investig Dermatol 2023;16:3349-3357. [doi: 10.2147/CCID.S441169] [Medline: 38021435]
22. Gold MH, Weiss E, Biron J. Novel laser hair removal in all skin types. J Cosmetic Dermatol 2023 Apr;22(4):1261-1265. [doi: 10.1111/jocd.15674] [Medline: 36756716]
23. Pall A, Viera-Mármol G. Triple wavelength and 810 nm diode lasers for hair removal: a clinical and in Silico comparative study on Indian skin. J Cosmet Dermatol Sci Appl 2022;12(4):164-173. [doi: 10.4236/jcdsa.2022.124014]
24. Grymowicz M, Rudnicka E, Podfigurna A, et al. Hormonal effects on hair follicles. Int J Mol Sci 2020 Jul 28;21(15):5342. [doi: 10.3390/ijms21155342] [Medline: 32731328]

25. Bikos A, Nastou PE, Petroudis G, Stamatiou YC. Random number generators: principles and applications. Cryptography 2023;7(4):54. [doi: 10.3390/cryptography7040054]
26. L'Ecuyer P. Random number generators. In: Gass SI, Harris CM, editors. Encyclopedia of Operations Research and Management Science: Springer; 2001.
27. 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]
28. Thaysen-Petersen D, Barbet-Pfeilsticker M, Beerwerth F, et al. Quantitative assessment of growing hair counts, thickness and colour during and after treatments with a low-fluence, home-device laser: a randomized controlled trial. Br J Dermatol 2015 Jan;172(1):151-159. [doi: 10.1111/bjd.13254] [Medline: 25039260]
29. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012 Jun 28;9(7):676-682. [doi: 10.1038/nmeth.2019] [Medline: 22743772]
30. HowMany - object counting AI tool. YESCHAT AI. 2024. URL: https://www.yeschat.ai/gpts-2OToXUmaUZ-HowMany [accessed 2025-11-11]
31. Myles PS, Myles DB, Galagher W, et al. Measuring acute postoperative pain using the visual analog scale: the minimal clinically important difference and patient acceptable symptom state. Br J Anaesth 2017 Mar;118(3):424-429. [doi: 10.1093/bja/aew466]
32. Delgado DA, Lambert BS, Boutris N, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based visual analog scale in adults. J Am Acad Orthop Surg Glob Res Rev 2018 Mar;2(3):e088. [doi: 10.5435/JAAOSGlobal-D-17-00088] [Medline: 30211382]
33. Pall A, Mármol GV. Study on the efficacy and safety of a high-power triple wavelength diode laser (810, 940, 1060 nm) for removing fine and less pigmented facial hair on Asian skin. J Cosmet Dermatol 2024 Apr;23(4):1282-1290. [doi: 10.1111/jocd.16129] [Medline: 38348571]
34. Modena DAO, Miranda ACG, Grecco C, Liebano RE, Cordeiro RCT, Guidi RM. Efficacy and safety of ND:YAG 1064 nm lasers for photoepilation: a systematic review. Lasers Med Sci 2020 Jun;35(4):797-806. [doi: 10.1007/s10103-019-02939-6]
35. Al-Niaimi F. Laser and energy-based devices' complications in dermatology. J Cosmet Laser Ther 2016;18(1):25-30. [doi: 10.3109/14764172.2015.1052511] [Medline: 26052809]
36. Relf I, Chow R, Pirotta M. Blinding techniques in randomized controlled trials of laser therapy: an overview and possible solution. Evid Based Complement Alternat Med 2008 Jan;5(4):383-389. [doi: 10.1093/ecam/nem085]
37. Arsiwala SZ, Majid IM. Methods to overcome poor responses and challenges of laser hair removal in dark skin. Indian J Dermatol Venereol Leprol 2019;85(1):3-9. [doi: 10.4103/ijdvl.IJDVL_1103_16] [Medline: 30027915]

Abbreviations

ANCOVA: analysis of covariance
LHR: laser hair removal

Edited by J Meisenheimer; submitted 25.04.25; peer-reviewed by D Puspitosari, M Jaradat; revised version received 22.09.25; accepted 25.09.25; published 21.11.25.

Please cite as:
Shurrab K, Nassr MA
Long-Term Efficacy and Safety of755-nm Alexandrite Laser for Axillary Hair Removal: A Comparative Analysis of Single vs. DualFlash lamp Systems
JMIR Dermatol 2025;8:e76523
URL: https://derma.jmir.org/2025/1/e76523
doi:10.2196/76523

Effectiveness of a Machine Learning-Enabled Skincare Recommendation for Mild-to-Moderate Acne Vulgaris: 8-Week Evaluator-Blinded Randomized Controlled Trial

Misbah Noshela Ghazanfar[1], MD, PhD; Ali Al-Mousawi[1], MD; Christian Riemer[2], MSc; Benóný Þór Björnsson[2], MSc; Charlotte Boissard[2], MSc; Ivy Lee[3], MD; Zarqa Ali[1], MD, PhD; Simon Francis Thomsen[1,4], MD, PhD

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

[2]NØIE, Copenhagen, Denmark

[3]Division of Dermatology, University of California Los Angeles, Los Angeles, CA, United States

[4]Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Misbah Noshela Ghazanfar, MD, PhD
Department of Dermatology, Bispebjerg Hospital, Bispebjerg bakke 23, Copenhagen, Denmark

Abstract

Background: Acne vulgaris (AV) is one of the most common skin disorders, with a peak incidence in adolescence and early adulthood. Topical treatments are usually used for mild to moderate AV; however, a lack of adherence to topical treatment is seen in patients due to various reasons. Therefore, personalized skincare recommendations may be beneficial for treating mild-to-moderate AV.

Objective: This study aimed to evaluate the effectiveness of a novel machine learning approach in predicting the optimal treatment for mild-to-moderate AV based on self-assessment and objective measures.

Methods: A randomized, evaluator-blinded, parallel-group study was conducted on 100 patients recruited from an internet-based database and randomized in a 1:1 ratio (groups A and B) based on their consent form submission. Groups A and B received customized product recommendations using a Bayesian machine learning model and self-selected treatments, respectively. The patients submitted self-assessed disease scores and photographs after the 8-week treatment. The primary and secondary outcomes were photograph evaluation by two board-certified dermatologists using the Investigator Global Assessment (IGA) scores and quality of life (QoL) measured using the Dermatology Life Quality Index (DLQI), respectively.

Results: Overall, 99 patients were screened, and 68 patients (mean age: 27 years, SD 4.56 years) were randomized into groups A (customized) and B (self-selected). IGA scores significantly improved after treatment in group A but not in group B (mean difference in IGA score; group A=0.32, P=.04 vs group B=0.09, P=.54). The DLQI significantly improved in group A from 7.75 at baseline to 3.5 (P<.001) after treatment but reduced in group B from 7.53 to 5.3 (P>.05). IGA scores and the DLQI were significantly correlated in group A, but not in group B. A total of 3 patients reported adverse reactions in group B, but none in group A.

Conclusions: Using a machine learning model for personalized skincare recommendations significantly reduced symptoms and improved severity and overall QoL of patients with mild-to-moderate AV, supporting the potential of machine learning-based personalized treatment options in dermatology.

(JMIR Dermatol 2025;8:e60883) doi:10.2196/60883

KEYWORDS

machine learning; personalised skincare; acne vulgaris; dermatology; skincare

Introduction

Acne vulgaris (AV) is one of the most common skin disorders, with a peak incidence in adolescence and early adulthood, affecting approximately 85% of individuals aged 12 - 24 years. Although acne is most prevalent in teenagers, it can emerge at any age. Approximately 25% and 12% of women and men in their 40s, respectively, report experiencing acne, usually accompanied by a high degree of stigma and impairment in quality of life (QoL) [1].

Topical treatments, such as retinoids, antibiotics, and combinations of antibiotics and benzoyl peroxide, are usually used for mild-to-moderate AV [2]; however, there is a lack of adherence to topical treatments among patients with AV [3]. A wide range of skincare products are available at beauty stores, pharmacies, and web-based shops. Many of these products are

inefficient, and users usually lack knowledge about which ingredients are effective and beneficial for their skin condition.

A Danish skincare company (NØIE) has developed a method for customizing skincare products based on in-depth phenotyping and direct feedback loops from over 80,000 patients with a skin condition by collecting clinical data on skin characteristics during an web-based survey and combining it with dermatological knowledge, feedback from users, and statistical modeling. In 2019, after 3 years of development, the project successfully launched a data model that modeled personalized skincare solutions based on an individual's specific skin and personal needs.

Therefore, this study aimed to evaluate the effectiveness of a novel machine learning approach for predicting the optimal treatment for mild-to-moderate AV based on subjective patient self-assessment and objective measures, including the physician-rated Investigator Global Assessment (IGA) and the patient-rated Dermatology Life Quality Index (DLQI).

Methods

Study Design and Participants

This evaluator-blinded, randomized controlled parallel-group trial included 100 patients who were randomized into groups A and B on a 1:1 ratio based on their submission of consent forms. It was conducted in accordance with the CONSORT (Consolidated Standards of Reporting Trials) statement (Checklist 1) [4]. Each group was assessed and assigned treatment based on the self-assessed reporting and image or real-life face-to-face interaction: group A used NØIE's Bayesian machine learning model and group B found and chose products themselves. The patients submitted self-assessed disease scores and a standardized set of facial images, known as a collected dataset, as the skin profile after 8 weeks of adherence to treatment. Subsequently, these assessments and images were scored by 2 independent board-certified dermatologists to evaluate the effectiveness of the 2 methods.

Recruitment Process

Patients were recruited through several channels affiliated with NØIE and e-mail campaigns targeting NØIE's database. NØIE contacted respondents to determine eligibility via a short survey before the screening process. The eligibility criteria were patients aged 18 - 40 years with a known diagnosis of AV who were interested in participating in the study from mid-February onwards for an 8-week duration and resided in Denmark, Germany, the Netherlands, Sweden, the United Kingdom, Switzerland, Belgium, or Austria.

All patients were required to submit a high-resolution image of their facial acne, which an employee of NØIE objectively assessed to confirm the AV and ensure that the disease severity was mild-to-moderate based on lesion counts. Mild AV was categorized as mostly whiteheads and blackheads with a few papules and pustules, whereas moderate acne was categorized as multiple papules and pustules.

The inclusion criteria were healthy female or male patients aged 18 - 40 years with mild-to-moderate AV and who had not previously used NØIE products. In contrast, the exclusion criteria were pregnancy, breastfeeding, or any changes in birth control during the intervention period since these would cause fluctuations in the hormonal impact on disease severity. Patients who used prescription medical treatments for acne treatment were also excluded.

Bayesian Model Guidance for Product Development

NØIE is established upon substantial data collection from over 65,000 individuals with various skin diseases, including skin reactions to well-categorized products commercially sold across Europe. These data were reverse-engineered into a matrix of stratified user segmentation, with an underlying layer of products developed by NØIE to support the optimal needs of individual user's skin. In addition, this precision medicine approach has only been used in oncology to date; however, NØIE has incorporated Bayesian modeling to stratify users based on differences in their epigenetic features, lifestyle, and personal preferences and their response to the products being used, based on the collected data, in-house product development, and direct contact with each user. The feedback from real-world data not only trains the model for precision but also guides skincare product development simultaneously to better meet the diverse needs of patients.

Precision medicine is an emerging approach in clinical research and patient care that focuses on understanding and treating diseases by integrating multimodal or multiomics data from each patient to make personalized treatment decisions [5]. In addition, dealing with the large and intricate datasets generated by precision medicine diagnostics requires the development of innovative techniques to process and interpret this complex information. Concurrently, rapid advancements in computer science have enabled the storage, processing, and analysis of these intricate datasets—a task that traditional statistics and early computing technologies could not accomplish. Therefore, the Bayesian modeling approach provides a means to formalize previous beliefs and combine them with available observations, aiming to derive rational criteria for optimal decision-making and measure the outcomes of these decisions [5].

This approach forms the foundational core of NØIE, aiming to identify intricate patterns in data for making predictions and classifications and conducting advanced exploratory data analysis on new, unseen data to guide their product development and distribution for better and safer treatments for various phenotypes of a given skin disease.

The fundamental input for the Bayesian model was the patient undergoing a skin test, which is a cumbersome in-depth web-based survey capturing 31 parameters relevant to the underlying skin condition or disease. These different data points are stored in real-time, forming the patient's skin profile, which the Bayesian models activate to recommend the ideal skin products for providing alleviation. Since all patients are asked to provide feedback on the effectiveness of their given treatment, it establishes a closed loop in the modeling process where results are automatically considered, thereby continuously strengthening the model's precision.

Product Description

Since NØIE's products are classified as cosmetics and medical device class I, the ingredients distinguish themselves from classic active pharmaceutical ingredients known in the pharmaceutical industry, although they rely on synergistic effects between conventionally used cosmetic ingredients and innovative modified peptides. Conventional ingredients include salicylic acid, retinol, and niacinamide, whereas innovative ingredients include *Curcuma longa* callus lysate, *Morinda citrifolia* callus culture, and Lactiplantibacillus fermentation lysate. Both groups' routine skincare regimens comprised a face cleanser and cream.

Group A: Customized Skincare

All patients in this group underwent a skin test to create a skin profile. Using the Bayesian model, they received product recommendations including a personalized face cleanser and cream. These products were shipped by NØIE.

Group B: Self-Selected Skincare

Similar to group A, patients in group B filled out using the same skin profile as their starting point. However, compared with group A, it did not activate NØIE's machine learning endpoint for a recommendation. All patients in this group were instructed to select a face cream and cleanser that they believed would improve their acne symptoms over 8 weeks. The choice of product was completely at the patients' discretion; they were permitted to seek advice from family, friends, pharmacists, and doctors but were not allowed medical treatment. NØIE purchased and shipped the self-selected products to patients, or patients purchased the products themselves and NØIE refunded the invoice.

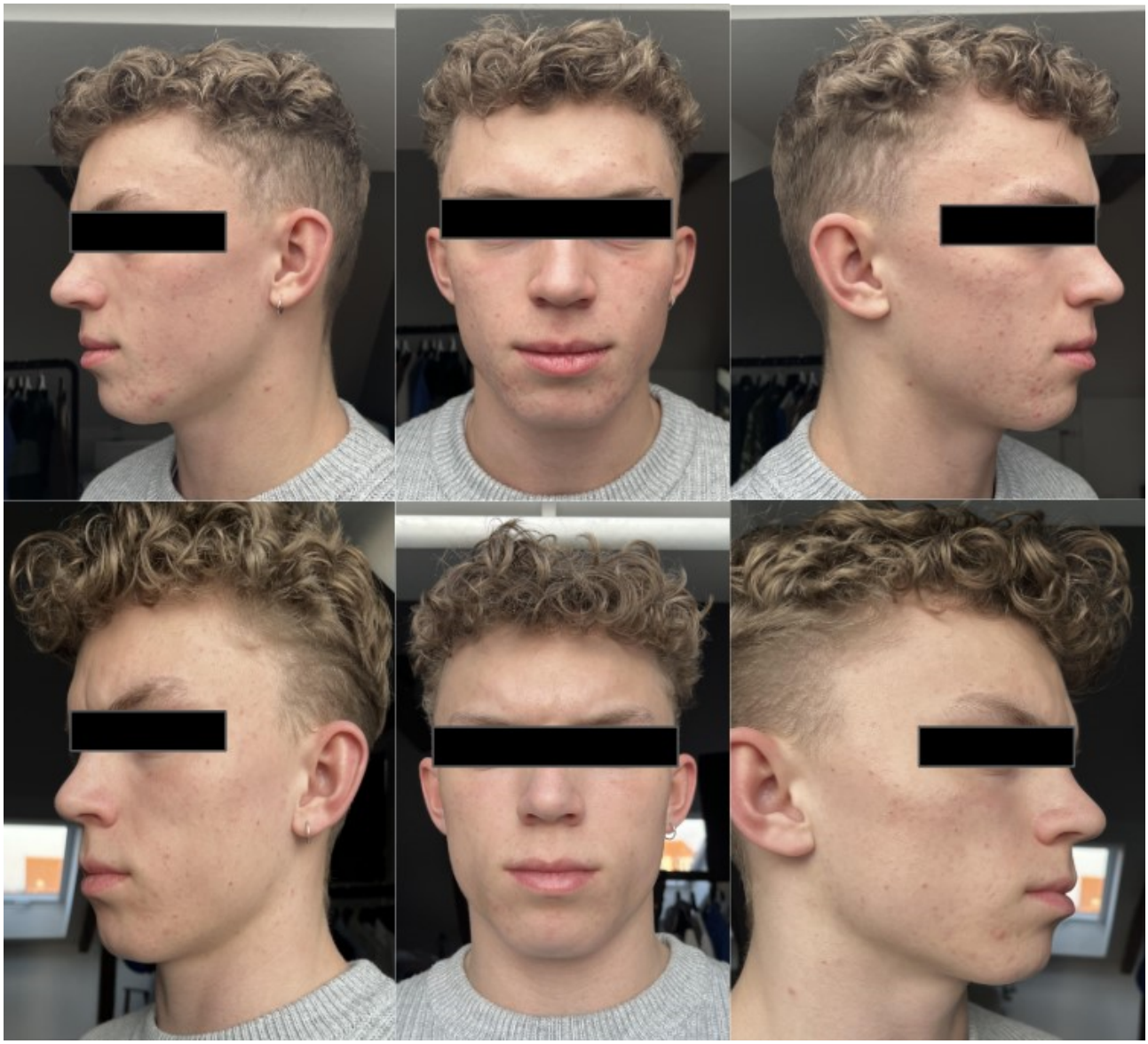
Adverse Reactions

Patients were instructed to contact NØIE immediately in the event of any adverse reactions. In such cases, patients were given the option to either substitute the product or were automatically assigned an alternative formulation by NØIE's underlying machine learning model.

Dermatological Assessment

A dermatological assessment was conducted based on three images provided by the patients on days 0 and 56. The 3 images included each cheek side and a frontal profile captured in high resolution, with balanced lighting and without makeup (Figure 1).

Figure 1. Patient photographs for assessment by dermatologists.



The two assessors were licensed dermatologists located in Denmark and the United States of America, respectively, and unaffiliated with NØIE in any sense. In addition, the images were rated according to the IGA [6], and the assessors were blinded to each other's ratings and the group origin of the assessed participant alongside the three images on days 0 and 56. Therefore, biasing the outcomes of the assessments was impossible.

The triplets of before-and-after images from each participant receiving personalized intervention were randomized so that the same participant's datasets were not evaluated consecutively. In addition, the order of the images (before and after) was randomized. This ensured that the dermatologist assessed each triplet without bias, as if conducting individual consultations, rather than being influenced by prior comparisons.

After the complete assessment of 72 image triplets, the results were organized in the correct before-and-after sequence for each participant. This allowed for the calculation of the intervention's development and effectiveness for each individual.

Primary and Secondary Outcomes

The primary outcome measure was the evaluation of changes in acne severity based on IGA scores by blinded assessors, whereas the secondary outcome was QoL measured using the DLQI. IGA scores were graded from 0 to 4 (from clear 0 to severe acne 4), and only facial acne was included in the assessment. Chest, back, and shoulder acne was not considered in this study. Both inflammatory and noninflammatory lesion counts have been reported [7]. Furthermore, the DLQI is a 10-item retrospective questionnaire that assesses the QoL of patients with skin conditions. Each question has scores ranging from 0 to 3, with a maximum total score of 30. A higher score indicates a greater impairment in QoL, and the minimal clinically important difference for DLQI is 3.10 points [8].

Self-Assessment

The self-assessment conducted during the study covered five lesion types for the patients. In addition, the question "Do you deal with symptom X" comprises a response scale ranging from "Not at all" to "Extremely". Subsequently, the positioning on the sliding scale was automatically converted to a numerical

grade in NØIE's backend, where direct "Not at all" and "Extremely" are equivalent to 0 and 10, respectively. This implies that patients could score their acne severity from 0 to 50 across all five lesion types. The five acne lesion types assessed were papules, pimples, blackheads, whiteheads, and oiliness, excluding nodules or cysts as they mainly belong to the severe state of AV.

Changes in Hypothetical Medical Intervention From Blinded Dermatologists

A third-party unaffiliated dermatologist with decades of experience in treating AV across all severities was randomly assigned to the group receiving customized treatment (group A) according to the sequence and the order of the intervention. Based on the 3 photographs, we assessed which treatment to prescribe, if any, as a hypothetical action which could be considered a development caused by the study intervention, as all other parameters remained unchanged. Multimedia Appendix 1 presents the 6 interventions performed by a blinded dermatologist.

Intervention severity was scored from 1 (least severe) to 6 (most severe). The mean severity was calculated at baseline and after the 8-week intervention, as well as the mean difference for each patient. The higher the mean score, the more severe the intervention required. The difference was calculated as the intervention score after the 8-week treatment minus the intervention score at baseline, where a negative score reflects a downgrade in the severity of the intervention required.

A *t* test was performed to determine the difference between the interventions at baseline and after 8 weeks. Statistical significance was set at an α value of .05.

Statistical Analysis

Differences in mean values within and between groups were analyzed using the *t* test where normalcy and continuity of the samples have been verified and with the nonparametric Mann-Whiteny *U* test where those requirements were not met, along with comparing percentage differences. Correlations were calculated using the Pearson correlation coefficient where applicable, and using Spearman correlation where the normalcy requirement was not met for the Pearson correlation. Statistical analysis, correlation calculations, and graphing were performed using Python version 3.7.12, Pandas library version 1.3.5, Numpy library version 1.21.6, Seaborn library version 0.12.1, and Scipy library version1.11.2.

Cohen Test to Ensure Adequate Sample Sizes

A power calculation using the Cohen test was conducted before initiating the recruitment process to estimate the sample size required for each group for determining the likelihood of detecting an effect in the experiment if it truly existed. First, the effect size (ES) was calculated by substituting the proportions of patients expected to be improved by each treatment, p_1=0.86 and p_2=0.50 (p_1=0.86 refers to the product satisfaction rate obtained by NØIE within the last 2 years; p_2=0.5 was set as a high estimate for the self-selected group since competent guidance was expected), and the overall proportion, *P*=.68 (ie, 0.86 [0.50] / 2):

$$ES=|p1-p2|p(1-p)=|0.86-0.5|0.68(1-0.68)=0.7717$$

Subsequently, we calculated the sample size (n_i) for each group (i=1, 2). ES was used in the equation for two independent samples, along with the confidence level, ensuring that a two-sided test with a 5% level of significance (ie, α=.05) and 80% power to detect the estimated response difference between the 2 groups. The equation is presented as follows:

$$ni=2(Z1-\alpha 2+Z1-\beta ES)2=2(1.96+0.840.7717)2=26.32$$

The use of n_1=26 and n_2=26 ensures that the test of the hypothesis has 80% power to detect a product effect. We anticipated patient dropouts due to strict inclusion criteria and noncompliance with product usage. Therefore, we recruited 50 patients in each group (n_1=50 and n_2=50).

Ethical Considerations

All patients provided written informed consent before enrollment. Approval from the scientific ethics committee was not required as no medical interventions were conducted in the study.

This study investigates the efficacy of skincare products through secondary analysis of consumer data. It is important to note that no medical interventions or treatments were involved in the study. The data used were collected for nonmedical purposes, and all participants had provided informed written consent for their identifiable data to be shared with researchers.

Ethical approval is not needed for this type of study in Denmark, where the intervention is a behavioral change without a medicinal product involved. This study did not involve any clinical treatments, procedures, or health interventions, and since the analysis was based on existing consumer data with appropriate safeguards in place, the study did not meet the criteria for requiring ethical approval from a health research ethics committee, in accordance with the Danish National Committee on Health Research Ethics (NVK) guidelines [6]. The NVK generally requires ethical approval for research involving medical interventions or when sensitive health data is being processed [6].

Furthermore, the data used for publication have been fully anonymized, ensuring the privacy and confidentiality of the participants. The General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) [9] and the Danish Data Protection Act [10] govern the use of personal data in research and allow for the processing of personal data when explicit consent is obtained and anonymization techniques are applied to protect privacy [10].

Participant enrollment alongside the collected data during the study interventions relied on secondary analysis conducted under a voluntary, waived consent framework. Accordingly, the consent form for study participation included detailed information about both primary data collection and the secondary analysis of the collected data, ensuring comprehensive informed consent for participants regarding the dual-purpose use of their data. The study participants were informed in writing about the study and consented to participate and with an option to contact NØIE in case of queries.

All the data collected on enrolled participants were stored on a secure, compliant server in an anonymized format, with participant identities protected through a unique ID system. Deciphering individual participants required conversion via a separate platform to maintain confidentiality. Access to this encrypted data was restricted solely to NØIE, and only aggregated data, without identifiers, was shared through a compliant system with external assessors. Access was disabled upon completion of the assessment, and NØIE permanently deleted all collected identifiable data exactly one year after collection, effected in May 2024.

None of the enrolled participants received or were offered financial compensation for their participation in the study. However, all participants were uniformly provided with complimentary products corresponding to their respective cohorts throughout the 8-week intervention period.

The publication does not contain any identifiable information about individual participants; all data are aggregated by cohort and lack identifiable elements, with the exception of images.

The appendix includes images of one study participant, anonymized with a censor bar covering the eyes. An additional signed consent form was obtained from this participant, stating:

> *I, the undersigned, agree to allow NØIE and a team of unaffiliated doctors to showcase my before and after images obtained during the 8 weeks of study involvement as part of summarising the learnings and insights obtained in the study. The intention is for readers of the publication to assess the image quality, and not intended to address any changes seen during the 8 weeks of intervention.*
>
> *I, the undersigned understand and hereby authorise and agree to NØIE utilising my photos, for scientific and awareness purposes to support aggregated numbers and results published in a physical dermatological journal as well as its corresponding online version. The material with me will likely be used in an intended publication to showcase results of a novel methodology like the AI driven recommendation carried out by NØIE as part of the study.*

All images will be eye censored to hinder identification.

Furthermore, no compensation was offered or given to the participant for the consent to use the images in the publication. This study involved no medical intervention; therefore, approval from the ethics committee was not required. Written consent was obtained from all participants before participation in the study.

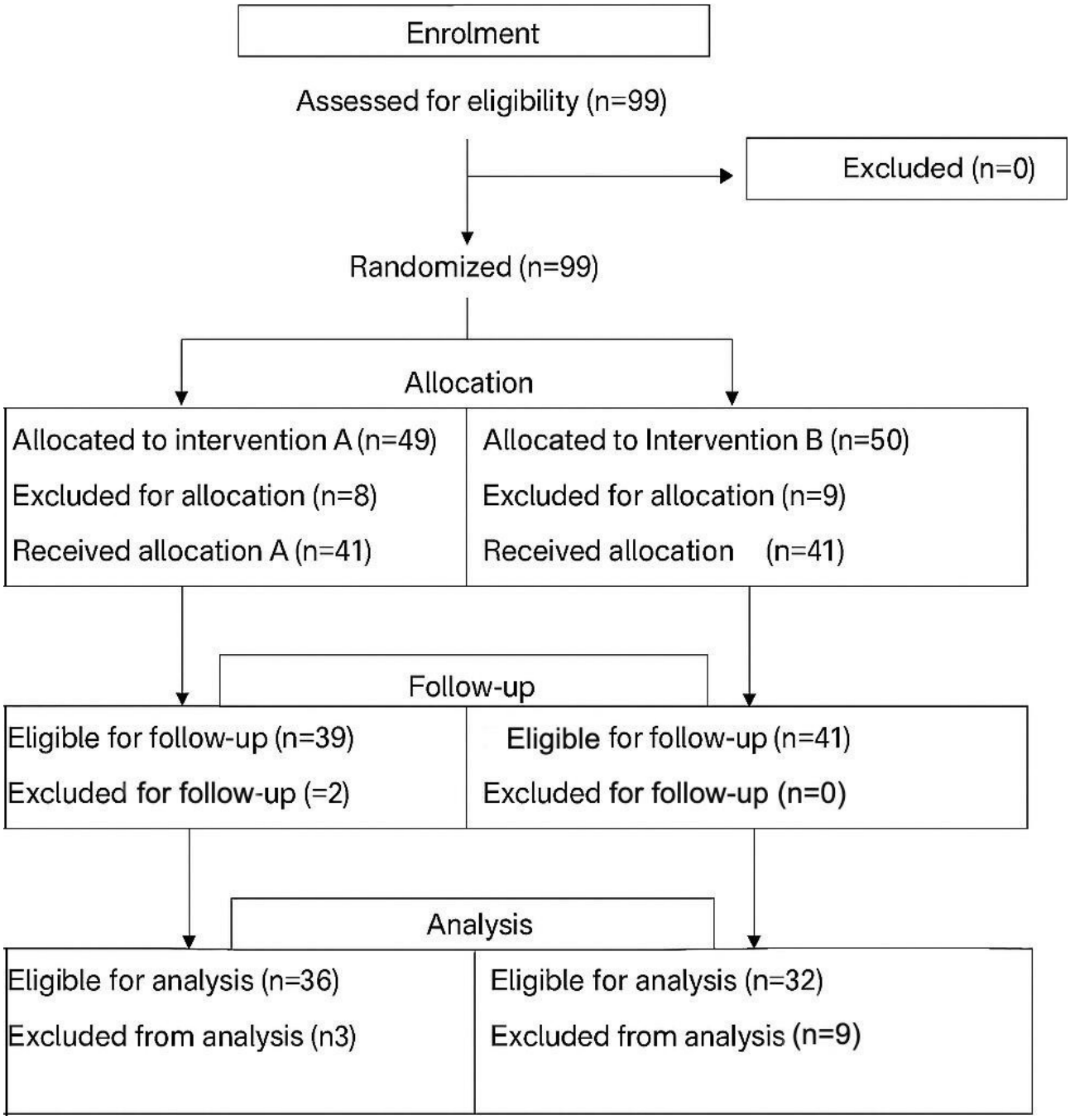
Results

Randomization and Screening

A total of 99 patients were screened for randomization. The failure to obtain the intended 100 patients was attributed to a strict deadline, given the coordination required for product shipment and the need to minimize the period between the initial screening of AV severity and the first day of the intervention.

Of the 99 patients, 17 were excluded before randomization because of the inclusion criteria. A total of 5 patients from group A, and 9 patients group B were excluded from data analysis due to noncompliance with usage guidelines (self-reported) or intervention, as shown in Figure 2.

Figure 2. The consolidated standards of reporting trials diagram. IGA: Investigator Global Assessment.



The patients excluded after filling out the initial dataset did not exhibit any specific characteristics regarding age, sex, or disease severity compared with those analyzed in either of the 2 groups. Although modafinil and lamotrigine are not prescribed for AV, they are well-known for their ability to affect the skin and were among the medications causing exclusion due to their potential interference with the study outcome.

In total, 68 patients were randomized into group A (customized), which included 36 patients (34 females and 2 males) with a mean age of 27.1 years, SD 4.76, and group B (self-selected) comprising 32 patients (30 women and 2 men) with a mean age of 26.3 years, SD 4.36 (Table 1). The duration of AV was 5 and 6 years in groups A and B, respectively. In addition, 58.3% and 62.5% of the patients in groups A and B, respectively, had previously undergone medical treatment. Baseline DLQI were 7.75 and 7.53 in groups A and B, respectively, indicating a moderate effect on QoL in both groups.

Table . Patient characteristics.

Characteristics	Group A (customized)	Group B (self-selected)
Age (years), mean (SD; range)	27.1 (4.76; 20–38)	26.3 (4.36; 19–35)
Sex, n (%)		
Female	34 (94.4)	30 (93.7)
Male	2 (5.5)	2 (6.3)
Duration of AV[a] (years), mean (SD)	4.94 (3.75)	5.98 (4.84)
Previous use of acne medication, n (%)	19 (58.3)	23 (62.5

[a]AV: acne vulgaris.

Primary Outcome: IGA Assessments

The mean (SD) IGA score improved from 1.53 (0.83) at baseline to 1.21 (0.76) after the 8-week treatment in group A (mean difference in IGA score=0.32, *P*=.04) and improved from 1.42 (0.42) at baseline to 1.33 (0.87) after the 8-week treatment in group B (mean difference in IGA score=0.09, *P*=.54) (Table 2).

Table . Primary and secondary outcomes at baseline and the 8-week follow-up.

Outcomes	Baseline	Follow-up at 8 weeks	*P* value
IGA[a], mean (SD)			
Group A	1.53 (0.83)	1.21 (0.76)	.03
Group B	1.42 (0.42)	1.33 (0.87)	.53
DLQI[b], mean (SD)			
Group A	7.75 (5.03)	3.5 (4.1)	<.001
Group B	7.53 (6.16)	5.3 (4.7)	>.05
Self-assessment, mean (SD)			
Group A	26.6 (4.9)	20.3 (7.6)	<.001
Group B	24.9 (7.6)	20.6 (8.44)	.03

[a]IGA: Investigator Global Assessment.

[b]DLQI: Dermatology Life Quality Index.

Table 3 illustrates the agreement between the 2 assessors' IGA scores. The 2 assessors disagreed in 36 cases but agreed upon no change (n=16), improvement (n=13), and regression (n=3) in 32 cases. Significant correlations were observed between the 2 assessors at baseline (*r*=0.452, 95% CI=0.24 - 0.62, *P*<.001) and after the 8-week treatment (*r*=0.54, 95% CI 0.35 - 0.69; *P*<.001).

Table . Matrix showing inter-agreements plotted as aggregate numbers for the 2 groups.

Assessor evaluation	Assessor #2 IGA[a] score reduced	Assessor #2 IGA score remained unchanged	Assessor #2 IGA score increased
Assessor #1 IGA score reduced	13[b]	5[c]	3[d]
Assessor #1 IGA score remained unchanged	9[c]	16[b]	7[c]
Assessor #1 IGA score increased	6[d]	6[c]	3[b]

[a]IGA: Investigator Global Assessment.

[b]Consensus between 2 assessors.

[c]Minor disagreement.

[d]Contradictory assessments.

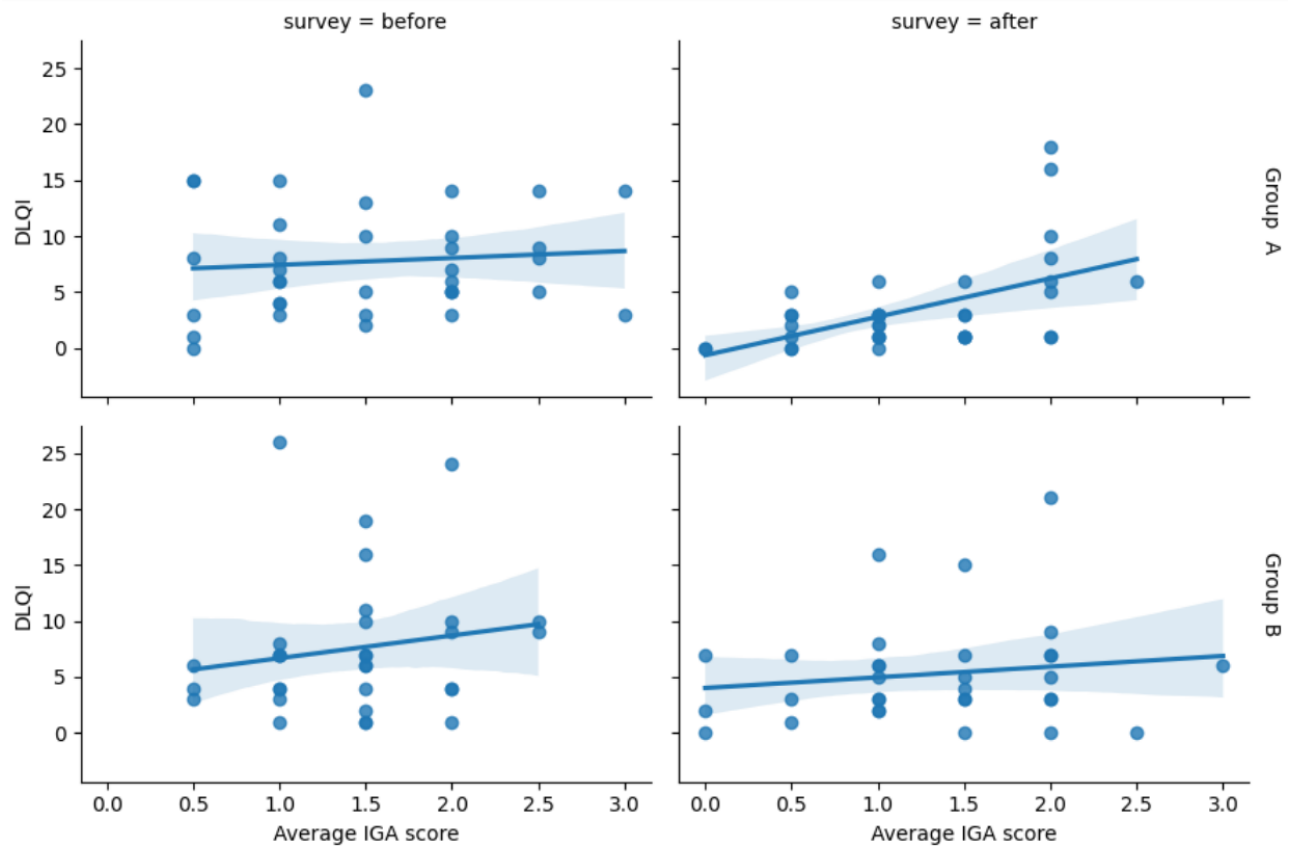
Secondary Outcome: QoL

This parameter measured using the DLQI significantly improved in group A from 7.75 (5.03) at baseline to 3.5 (4.1) after the 8-week treatment (mean difference in DLQI=4.3, *P*<.001). In contrast, a reduction in DLQI from 7.53 (6.16) at baseline to 5.3 (4.7) after the 8-week treatment was observed in group B (mean difference in DLQI=2.3), although this was not significant (*P*>.05; Table 2).

Correlation Between the IGA Score and the DLQI

Weak and moderate correlations were observed between the IGA score and the DLQI at baseline (r=0.11, 95% CI −0.23 to 0.42, P>.05) and after the 8-week treatment (r=0.516, 95% CI 0.23-0.72, P<.001) in group A, respectively. However, a weak correlation was observed between acne severity and the DLQI at baseline (r=0.242, 95% CI −0.12 to 0.54, P>.05) and after the 8-week treatment (r=0.136, 95% CI −0.22 to 0.47, P>.05) in group B (Figure 3).

Figure 3. Correlation between acne severity (IGA score) and DLQI before and after intervention in groups A and B. DLQI: Dermatology Life Quality Index; IGA: Investigator Global Assessment.



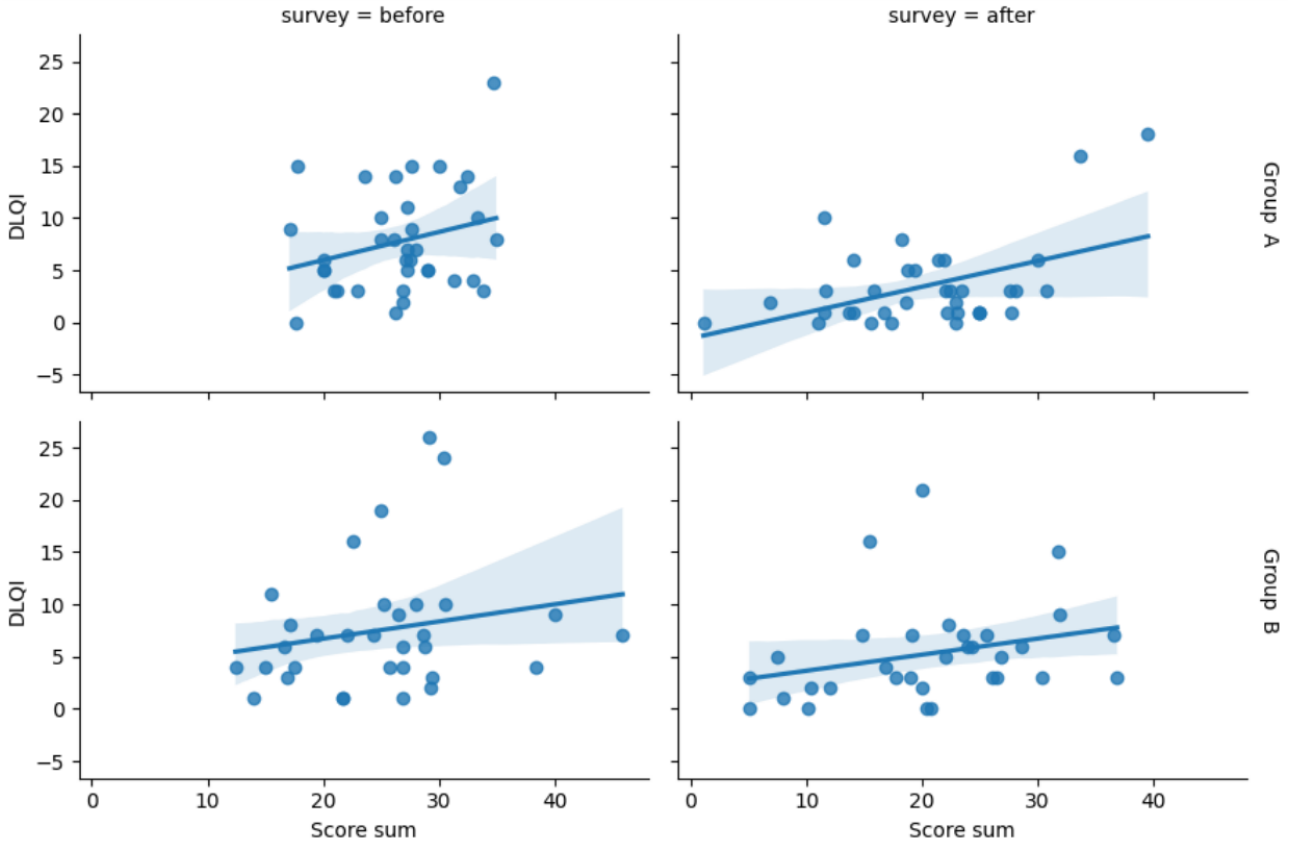
Self-Assessed Symptoms

Group A had a total mean (SD) symptom score of 26.6 (4.9) at baseline, which improved to 20.3 (7.6) after the 8-week treatment (mean difference=6.3, 95% CI 3.22 - 9.32, P<.001). In contrast, group B had a total mean (SD) symptom score of 24.9 (7.6) at baseline, which improved to 20.6 (8.44) after the 8-week treatment (mean difference=4.31, 95% CI=0.26 - 8.37, P=.03) (Table 2). Furthermore, significant reductions in the proportions of blackheads (25%, 8/32), whiteheads (31%, 10/32), pimples (26%, 8/32), and skin oiliness (23%, 7/32) were observed in group A after the 8-week follow-up. In total, 61% (20/32) and 56% (20/36) of the patients in groups A and B, respectively, reported that their skin appeared healthier after using the customized skincare routine and the 8-week treatment, respectively. Greater reductions in the proportions of blackheads (53%, 17/32 vs 35%, 13/36) and pimples (67%, 22/32 vs 34%, 12/36) were observed in group A than in group B. Moreover, a significant visual improvement in the skin was observed in group A compared with that in group B (67%, 22/32 vs 37%, 13/36, P=.008).

Correlation Between Self-Assessed Symptoms and the DLQI

Weak and moderate correlations were observed between acne severity (self-assessed) and the DLQI at baseline (r=0.218, 95% CI −0.12 to 0.51; P>.05) and after the 8-week treatment (r=0.279, 95% CI −0.05 to 0.56; P>.05) in group A, respectively. In group B, a weak correlation was observed between acne severity and the DLQI at baseline (r=0.244, 95% CI −0.11 to 0.55; P>.05) and at the 8-week follow-up (r=0.388, 95% CI 0.05-0.65; P=.03; Figure 4).

Figure 4. Correlation between acne severity (self-assessed) and DLQI before and after intervention in groups A and B. Abbreviations DLQI, Dermatology Life Quality Index.



Adverse Reactions

Overall, 3 patients in group B reported minor adverse events in the form of blushing, stinging, and itchy sensations upon the first application on day 1. They contacted NØIE and were instructed to discontinue the products and find alternative products the same day. Notably, the alternative products were well tolerated. However, none of the patients in group A reported any adverse events.

Changes in Hypothetical Medical Intervention From the Blinded Dermatologists

The mean (SD) severity score in group A at baseline was 2.97 (1.77), which improved to 2.22 (1.66) after 8 weeks of intervention (mean difference 0.75 [1.84]; *P*=.05). After the 8-week intervention, 42.9% of the patients (14/32) were administered only dermo-cosmetics. Multimedia Appendix 2 illustrates the need for a hypothetical medically prescribed treatment for AV at baseline and after the 8-week follow-up for each intervention.

Discussion

Precision medicine is an emerging approach in the field of dermatology, and recent clinical research has shown that the impact and management of skin diseases differ among patients. This has resulted in an increased focus on the development of personalized treatment approaches to optimize treatment response, minimize adverse reactions, and improve the overall QoL of patients [11,12]. However, the knowledge of optimal personalized treatment approaches in clinical practice remains limited.

Main Findings

This 8-week, evaluator-blinded, randomized controlled trial, based on patient-taken photographs, evaluated the effectiveness of a novel machine learning approach for predicting the optimal treatment for mild-to-moderate AV, using dermatological evaluation combined with self-assessment. We found that IGA scores, as assessed by board-certified dermatologists, demonstrated a significant reduction in acne severity in group A compared with that in group B, indicating that personalized product recommendations generated by a machine learning model were more effective in improving acne severity than a self-selected treatment approach.

Furthermore, a significant improvement in acne severity (self-assessed) was observed after the 8-week treatment in both groups. Significant reductions in acne symptoms, such as blackheads, whiteheads, pimples, and skin oiliness, were observed in group A compared with that in group B. Therefore, this finding adds to the potential benefits of personalized treatment approaches for patients with AV.

Interpretation

Previous studies have reported that even mild acne significantly impacts the psychological well-being and QoL of patients [12,13]. In our study, a significant improvement was observed in the QoL of patients in group A after the 8-week treatment but not in group B. This suggests that the improvement in QoL was greater with personalized product recommendations than

with self-selected products. Furthermore, a weak correlation was found between acne severity and the DLQI in both groups, suggesting that the impact of acne on QoL is weakly correlated with its severity, which highlights the complexity of clinical assessments and the psychological impact of acne. These findings align with those of previous reports showing no correlation between acne severity and QoL impairment [14,15].

Minor adverse reactions, including blushing, stinging, and itchiness, alone were reported from the use of self-selected products. However, no minor or severe adverse reactions were reported from the use of customized treatment. Therefore, this finding indicates that personalized treatments may have a better safety profile than self-selected products.

Medication should ideally be the last resort in treatment given the potential side effects, intolerance, and restrictions on long-term use. Consequently, treatment approaches should prioritize resolving conditions through the least invasive methods when possible. For mild-to-moderate AV, dermo-cosmetic treatments should be considered the first line of treatment, emphasizing gentler options before resorting to medication.

Strengths and Limitations

The strengths of this study include its randomized design which ensured the assigning of unbiased treatment and the relatively large group of patients. Furthermore, two board-certified dermatologists assessed disease severity using the IGA score. A substantial agreement was noted between the IGA assessments of the two board-certified dermatologists regarding the dynamics of acne evolution throughout the trial. However, IGA scores are limited to the face and do not cover disease activity on the back or chest. Changes in the IGA score are small, making it difficult to be used in clinical settings with interventions [5]. The Global Acne Grading System, another score that might be more suitable for research purposes, is a broader scoring system and provides a more detailed picture of disease severity; however, it is time-consuming and difficult to use in clinics with limited time [16]. This study's limitations include a short follow-up time (8 week), a primary focus on mild-to-moderate AV, and a self-assessment of acne severity. Furthermore, self-assessment methods have been reported to be unreliable because it is difficult for patients to objectively score the severity of their acne [17]. Further research is needed in personalized skincare and machine learning models with longer follow-up times.

Conclusions

The use of a machine learning model for personalized skincare recommendations significantly improved the severity of acne, reduced symptoms, and improved the overall QoL of patients with mild-to-moderate AV. Therefore, these findings support the potential of machine learning-based personalized treatment options in dermatology.

Acknowledgments

The study is sponsored by NØIE (see conflict of interest).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

ZA and CR conceived of and planned the study. SFT and IL performed dermatological assessments. CR, BB, and CB analyzed the results. MNG, AA, CR, BB, CB, ZA, and SFT contributed to the interpretation of results and models. MNG took the lead in writing the manuscript and was supervised by ZA, CR, and SFT. All authors discussed the results and contributed to the final manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

CB, BB, and CR are employed by NØIE. None of the other authors have any affiliation with NØIE.
With no relation to the present manuscript, SFT has received research support from Janssen, LEO Pharma, Novartis, Sanofi, and UCB and has been a speaker/consultant for AbbVie, Almirall, Eli Lilly, Galderma, Janssen, LEO Pharma, Novartis, Pfizer, Sanofi, Symphogen, UCB, and Union Therapeutics. MNG, AA, ZA, and IL have no conflict of interest.

Multimedia Appendix 1
Supplementary Table 1. Overview of the six interventions.
[DOCX File, 17 KB - derma_v8i1e60883_app1.docx]

Multimedia Appendix 2
Supplementary Figure 1. Need for medically prescribed interventions.
[DOCX File, 151 KB - derma_v8i1e60883_app2.docx]

Checklist 1

CONSORT 2010 checklist
[PDF File, 68 KB - derma_v8i1e60883_app3.pdf]

References

1. Oge' LK, Broussard A, Marshall MD. Acne vulgaris: diagnosis and treatment. Am Fam Physician 2019 Oct 15;100(8):475-484. [Medline: 31613567]
2. 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]
3. Sevimli Dikicier B. Topical treatment of acne vulgaris: efficiency, side effects, and adherence rate. J Int Med Res 2019 Jul;47(7):2987-2992. [doi: 10.1177/0300060519847367] [Medline: 31122106]
4. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8:18. [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
5. Balaneshinkordan S, Kotov A. Bayesian approach to incorporating different types of biomedical knowledge bases into information retrieval systems for clinical decision support in precision medicine. J Biomed Inform 2019 Oct;98:103238. [doi: 10.1016/j.jbi.2019.103238] [Medline: 31301513]
6. The Danish national committee on health research ethics. Danish National Committee on Health Research Ethics (NVK). 2024. URL: https://researchethics.dk/research-ethics-themes/about-the-scientific-ethical-assessment [accessed 2025-05-28]
7. Alsulaimani H, Kokandi A, Khawandanh S, Hamad R. Severity of acne vulgaris: comparison of two assessment methods. Clin Cosmet Investig Dermatol 2020;13:711-716. [doi: 10.2147/CCID.S266320] [Medline: 33061511]
8. Basra MKA, Fenech R, Gatt RM, Salek MS, Finlay AY. The dermatology life quality index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol 2008 Nov;159(5):997-1035. [doi: 10.1111/j.1365-2133.2008.08832.x] [Medline: 18795920]
9. Art. 24 GDPR – responsibility of the controller. GDPR-info.eu. URL: https://gdpr-info.eu/art-24-gdpr/ [accessed 2025-05-28]
10. Act on supplementary provisions to the regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (the data protection act). Danish Data Protection Act. URL: https://www.datatilsynet.dk/media/7753/danish-data-protection-act.pdf [accessed 2025-05-28]
11. Šitum M, Franceschi D, Franceschi N. Challenges and strategies in dermatologic therapy-personalized medicine, patient safety, and pharmacoeconomics. Dermatol Ther 2019 Jul;32(4):e13011. [doi: 10.1111/dth.13011] [Medline: 31237068]
12. Nie D, Tegtmeyer K, Zhao J, Lio PA. Developing patient-specific adverse effect profiles: the next frontier for precision medicine in dermatology. J Dermatolog Treat 2020 May;31(3):211-212. [doi: 10.1080/09546634.2019.1687810] [Medline: 31663793]
13. Gieler U, Gieler T, Kupfer JP. Acne and quality of life – impact and management. Acad Dermatol Venereol 2015 Jun;29(S4):12-14 [FREE Full text] [doi: 10.1111/jdv.13191]
14. Gupta A, Sharma YK, Dash KN, Chaudhari ND, Jethani S. Quality of life in acne vulgaris: relationship to clinical severity and demographic data. Indian J Dermatol Venereol Leprol 2016;82(3):292-297. [doi: 10.4103/0378-6323.173593] [Medline: 27088931]
15. Ilgen E, Derya A. There is no correlation between acne severity and AQOLS/DLQI scores. J Dermatol 2005 Sep;32(9):705-710. [doi: 10.1111/j.1346-8138.2005.tb00829.x] [Medline: 16361712]
16. Hacivelioglu S, Gungor ANC, Gencer M, et al. Acne severity and the global acne grading system in polycystic ovary syndrome. Intl J Gynecology & Obste 2013 Oct;123(1):33-36. [doi: 10.1016/j.ijgo.2013.05.005]
17. Adah R, Yusufu H, Otene QAV. Epidemiology and perception of acne among adolescents in Jos, Nigeria: cross-sectional school-based study. JMIR Dermatol 2023 Jul 12;6(GDPR):e44441. [doi: 10.2196/44441] [Medline: 37632928]

Abbreviations

AV: acne vulgaris
CONSORT: Consolidated Standards of Reporting Trials
DLQI: Dermatology Life Quality Index
IGA: Investigator Global Assessment scores
QoL: quality of life

Edited by R Dellavalle; submitted 24.05.24; peer-reviewed by U Sinha, V Astha; revised version received 07.05.25; accepted 08.05.25; published 16.07.25.

Please cite as:
Ghazanfar MN, Al-Mousawi A, Riemer C, Björnsson BÞ, Boissard C, Lee I, Ali Z, Thomsen SF
Effectiveness of a Machine Learning-Enabled Skincare Recommendation for Mild-to-Moderate Acne Vulgaris: 8-Week Evaluator-Blinded Randomized Controlled Trial
JMIR Dermatol 2025;8:e60883
URL: https://derma.jmir.org/2025/1/e60883
doi:10.2196/60883

Perception, Quality, and Accuracy of Sunscreen Content on TikTok: SkinMedia Cross-Sectional Content Analysis

Jaclyn Roland-McGowan[1,2*], BS; Kyra Diehl[2,3*], BS; Tayler Tobey[2], BS, BA; Autumn Shafer[4], PhD; Paige Clement[5], MSc; Oliver J Wisco[5], DO; Alex G Ortega-Loayza[2], MCR, MD; Sancy Leachman[2,6], MD, PhD

[1]School of Medicine, Oregon Health and Science University, Portland, OR, United States

[2]Department of Dermatology, Oregon Health and Science University, 3303 SW Bond Ave, CHH1 16th Floor, Portland, OR, United States

[3]College of Osteopathic Medicine, Western University of Health Sciences, Pomona, CA, United States

[4]School of Journalism and Communication, University of Oregon, Eugene, OR, United States

[5]Department of Dermatology, Warren Alpert Medical School of Brown University, Providence, RI, United States

[6]Department of Dermatology, University of Utah, Salt Lake City, UT, United States

[*]these authors contributed equally

Corresponding Author:
Alex G Ortega-Loayza, MCR, MD
Department of Dermatology, Oregon Health and Science University, 3303 SW Bond Ave, CHH1 16th Floor, Portland, OR, United States

Abstract

Background: TikTok, with more than 2 billion users worldwide, has become an influential venue for health information, including dermatologic advice. However, concerns remain about the accuracy and impact of sunscreen-related content.

Objective: This study aimed to assess the quality, accuracy, and themes of popular TikTok videos about sunscreen; evaluate associations with creator credentials and promotional content; and identify implications for public health.

Methods: We conducted a cross-sectional content analysis of the 100 most-liked English-language TikTok videos generated by the search term "sunscreen." Metadata, creator characteristics, Global Quality Score (GQS), accuracy, attitudes, promotional disclosures, and reference use were extracted using a structured codebook. Thematic and statistical analyses (ie, Pearson correlations, χ^2, 2-tailed *t* tests, and ANOVA) were conducted, with significance defined as $P<.05$.

Results: Of the top 100 videos, 74 (74%) expressed a positive attitude toward sunscreen, 35 (35%) were accurate, 57 (57%) were opinion based, and 6 (6%) were inaccurate. None of the videos cited references. GQS ratings were low: 40 (40%) videos were rated poor (score=1), 31 (31%) below average (score=2), and only 2 (2%) excellent (score=5). Promotional content appeared in 27 (27%) videos. Accuracy was negatively correlated with likes (r=–0.229; P=.02) and views (r=–0.242; P=.02), while GQS correlated positively with accuracy (r=0.270; P=.007) but not with engagement. Likes and views were strongly correlated (r=0.726; P<.001).

Conclusions: Despite broadly positive sentiment toward sunscreen, misinformation and promotional bias are common in highly engaged TikTok videos, and user engagement is often unrelated to accuracy or educational quality. Dermatologists and public health experts must proactively engage on social platforms to counter misinformation and promote reliable skin health information.

(JMIR Dermatol 2025;8:e70010) doi:10.2196/70010

KEYWORDS

TikTok; social media; health information; misinformation; Global Quality Scale; GQS; dermatology; sunscreen; sun safety; sun protection factor; SPF

Introduction

Social media is an increasingly important source of health information, and TikTok, one of the world's fastest-growing video-sharing platforms, has rapidly become a key venue for health messaging [1]. With its short-form video style and algorithmic feed, the platform shapes health behaviors, particularly among younger audiences. Within dermatology, TikTok content is both vast and highly visible, with dermatology-related hashtags collectively accumulating more than 18 billion views [2]. Sunscreen, a cornerstone of skin cancer prevention and photoaging mitigation, has emerged as a recurring theme across the platform, with *#sunscreen* alone garnering billions of views [3]. However, within the same feed that promotes sun protection, viral hashtags such as *#AntiSunscreen* circulate misinformation, propagating doubt about established guidance and exposing users considered

vulnerable to confusion about one of the most important preventive health behaviors [4,5].

Existing research highlights the complexity of sunscreen discourse on TikTok (ByteDance Ltd). Haff et al [6] reported that much of dermatology content is low quality, with an average DISCERN score of 1.58 out of 5. Pediatric analyses similarly found that only 26% of daytime skincare regimens included sunscreen, despite frequent use of sun-sensitizing ingredients [7]. At the other extreme, misinformation has proliferated: Kearney et al [4] showed that 69% of videos under *sunscreen conspiracy* searches were antisunscreen, and Nikookam et al [5] emphasized the viral reach of the *#AntiSunscreen* movement. Even outside conspirative spaces, video quality is limited. Khan et al [3] found that the most popular sunscreen videos averaged a DISCERN score of 2.68 out of 5. In addition, promotional content is common: Ranpariya et al [8] showed that dermatology influencers frequently post sponsored material, raising questions about transparency and commercial influence.

Despite these insights, no study has systematically examined whether the most-engaged sunscreen content aligns with evidence-based dermatology guidance or how popularity favors accuracy. Our study addresses this gap through a systematic evaluation of the 100 most-liked *#sunscreen* TikTok videos, assessing accuracy, Global Quality Scale (GQS) score, sentiment, references, and promotional content and analyzing their relationship to engagement metrics.

Methods

Ethical Considerations

As this study relied exclusively on publicly available data without human participants, it was reviewed by the Oregon Health & Science University Institutional Review Board and deemed not to be human research (00027957).

Study Design

This cross-sectional content analysis aimed to evaluate the quality, accuracy, and attitudes of popular sunscreen content on TikTok.

Preliminary Screening and Search Strategy

To establish a search strategy, 2 reviewers (KD and JR-M) first conducted exploratory searches using targeted terms such as “toxic sunscreen,” “sunscreen risks,” and “sunscreen myths.” On the basis of these findings, we selected the broad search term “sunscreen,” limiting our final dataset to a single, neutral term to minimize search bias and ensure broad inclusion of relevant content. This approach enabled capturing a wide range of perspectives within a single organic search.

Sample Selection

The final search was conducted on August 14, 2024, using newly created TikTok accounts to reduce algorithmic personalization. We applied TikTok’s “most-liked” filter and identified the top 106 most-liked videos returned by the search term “sunscreen.” For each video, we recorded metadata (URL, creator username, date posted, number of likes, views, and creator follower count) in Microsoft Excel (Microsoft Corporation) at the time of search.

Inclusion and Exclusion Criteria

The top 100 most-liked videos were screened for inclusion. Eligible videos were required to (1) be in English or contain English subtitles and (2) be relevant to sunscreen use, safety, or recommendations. Duplicates and irrelevant videos were excluded. Videos ranked 101 to 106 were used exclusively for coder calibration and interrater reliability (IRR) testing. A sample size of 100 was selected to balance feasibility of detailed manual coding with representativeness. Although larger-scale TikTok analyses exist, such as the study by Steinke et al [9], our aim was to balance depth with representativeness.

Variables and Coding Definitions

Videos were coded for the following variables:

- Content and creator characteristics included the date posted, creator username, number of likes, number of views, and follower count (recorded at the time of data collection).
- Content quality or usability was assessed using the 5-point GQS described by Bernard et al [10] (1=poor to 5=excellent); in our study, GQS anchors emphasized viewer-facing usefulness and clarity (refer to “Codebook Development and IRR”).
- Accuracy was categorized as accurate, inaccurate, mixed, or opinion-based, according to concordance with the American Academy of Dermatology sunscreen guidelines [11].
- Reference use was coded based on presence or absence of identifiable sources such as peer-reviewed articles.
- Attitudes toward sunscreen were categorized as positive, negative, or neutral.
- Promotional content was identified by the presence of TikTok “Paid Partnership” labels, affiliate links, discount codes, or clearly disclosed sponsorships in captions or on-screen text.

Codebook Development and IRR

Two reviewers (KD and JR-M) independently coded 6 calibration videos (videos 101-106) to refine the structured codebook (Multimedia Appendix 1) and assess IRR. IRR for categorical variables was evaluated using percent agreement, Scott pi, Cohen kappa, and Krippendorff alpha. Using the original 1 - 5 GQS wording from prior literature [10,12-14], initial GQS agreement was 60% despite more than 85% agreement on other variables.

To improve consistency, coders received additional training and in-person calibration meetings, clarified the construct to reflect how lay viewers interpret scientific information, and iteratively revised scale anchors: 1=poor (largely missing information or unhelpful for viewer use), 2=below average (minimal content with major omissions), 3=moderate (some helpful information but incomplete or unclear for viewers), 4=good (covers main points understandably, though not comprehensive), and 5=excellent (clear, well structured, comprehensive, and highly useful).

Reviewers completed several rounds of independent recoding and consensus review over a few weeks. Disagreements were

resolved by consensus using the written anchors and comparison to exemplar videos.

Calibration continued until IRR exceeded 0.90 across all variables and GQS reliability surpassed 0.75. After this, full extraction was conducted, with 1 reviewer coding all 100 videos and the second reviewer independently coding a blinded 10% (10/100) subset. Final IRR for this subset remained high (agreement=97.9%; Scott pi=0.954; Cohen kappa=0.954; Krippendorff alpha=0.956; Multimedia Appendix 2).

Statistical Analysis

Descriptive statistics summarized video characteristics, themes, and creator demographics. Inferential tests included independent samples *t* tests and one-way ANOVA for comparing continuous variables across groups, Pearson correlations to examine relationships between engagement metrics and quality indicators (ie, accuracy and GQS), and χ^2 tests for associations between categorical variables. Statistical significance was defined as $P<.05$. All analyses were conducted in Microsoft Excel and R (R Foundation for Statistical Computing). No videos were excluded due to missing data.

Decile-Based Engagement Analysis

Because engagement metrics (ie, likes and views) were highly skewed, we performed a decile-based stratification. Videos were grouped by engagement deciles, and accuracy patterns were examined across strata to identify nonparametric trends that might have been obscured by extreme outliers.

Results

Descriptive Overview

Analysis of the 100 most-liked TikTok videos generated using the search term “sunscreen” revealed a mean like count of 227,671 (SD 351,902; range 26,700 - 2,400,000) and a mean view count of 4,203,243 (SD 5,775,935; range 309,800 - 34,300,000). The average number of followers per creator was 1,033,824 (SD 2,977,532; range 272 - 18,100,000). Of the 100 content creators, only 18 (18%) identified as health care professionals, including 13 (13%) dermatologists, 2 (2%) medical aestheticians, 1 (1%) pharmacist, 1 (1%) esthetic nurse practitioner, and 1 (1%) cosmetic chemist. Ingredient mentions appeared in 13 (13%) videos, and toxicity concerns in 10 (10%) videos. Positive sunscreen attitudes were expressed in 74 (74%) videos.

Out of the 100 videos, 35 (35%) were classified as accurate, 6 (6%) as inaccurate, 2 (2%) as mixed, and 57 (57%) were opinion-based or not applicable for factual accuracy evaluation. Positive attitudes toward sunscreen were expressed in 74 (74%) videos, with the remainder split between 17 (17%) neutral and 9 (9%) negative sentiment videos.

Sunscreen product recommendations appeared in 85 (85%) videos, and 27 (27%) videos contained explicit promotional content such as affiliate links, brand sponsorships, or paid advertisements (Table 1).

Table . Descriptive characteristics of the top 100 most-liked TikTok videos related to sunscreen, including content accuracy, Global Quality Scores (GQS), thematic content, and creator background.

Characteristics	TikTok videos, n (%)
Attitude	
Positive	74 (74)
Negative	7 (7)
Neutral	9 (9)
Accuracy	
Opinion-based	57 (57)
Accurate	35 (35)
Inaccurate	6 (6)
Mixed	2 (2)
Used references	0 (0)
GQS	
1 (Poor quality)	40 (40)
2 (Below average)	31 (31)
3 (Moderate)	23 (23)
4 (Good)	4 (4)
5 (Excellent)	2 (2)
Content type	
Non-white skin–specific content (eg, white cast, blending, or sunscreen considerations for darker skin tones)	14 (14)
Sun protection factor	30 (30)
Sun safety education	31 (31)
Specific ingredients	13 (13)
Toxicity	10 (10)
Content creator background	
Board-certified dermatologist	13 (13)
Medical esthetician	2 (2)
Cosmetic chemist	1 (1)
Pharmacist	1 (1)
Nurse practitioner	1 (1)
Promotional content	
Creators with advertisements or promotion	27 (27)

Engagement Metrics by Accuracy

Although inaccurate videos were fewer in number, they received disproportionately high engagement. On average, the 6 (6%) inaccurate videos received 2,141,683 (SD 2,235,616) views and 119,383 (SD 194,427) likes and the 35 (35%) accurate videos received 2,732,594 (SD 3,067,518) views and 135,711 (SD 131,773) likes.

A strong positive correlation was observed between the number of likes and views (r=0.726; df=98; P<.001). In contrast, accuracy was negatively correlated with both likes (r=–0.229; df=98; P=.02) and views (r=–0.242; df=98; P=.02), indicating that less accurate content may attract more user engagement. Refer to Table 2 for all Pearson correlation coefficients among key video characteristics.

Table . Correlation coefficients (Pearson *r*) among key variables in sunscreen-related TikTok videos.

Comparison	*r*	*P* value
Likes vs views	0.73[a]	<.001
Likes vs attitude	−0.19	.06
Likes vs promotional	−0.16	.11
Likes vs accuracy	−0.23	.02
Likes vs GQS[b]	0.12	.24
Views vs attitude	−0.15	.13
Views vs promotional	−0.04	.67
Views vs accuracy	−0.24	.02
Views vs GQS	0.19	.05
Attitude vs promotional	0.05	.63
Attitude vs accuracy	0.1	.32

[a]Statistically significant correlations.

[b]GQS: Global Quality Score.

To account for the skewed distribution of engagement metrics, a decile-based analysis was conducted. This analysis revealed a strong, statistically significant negative correlation between accuracy and views (r=–0.720; df=8; P=.02) and a moderate, nonsignificant negative correlation between accuracy and likes (r=–0.440; df=8; P=.21). These findings suggest that less accurate videos may be algorithmically favored or more widely shared, resulting in greater visibility despite lower informational quality (Figures 1 and 2).

Figure 1. Relationship between accurate sunscreen content and views. The 100 most-liked videos were grouped into deciles based on their view counts. The percentage of videos within each decile that contain accurate information is shown on the y-axis, while the average views per decile are shown on the x-axis.

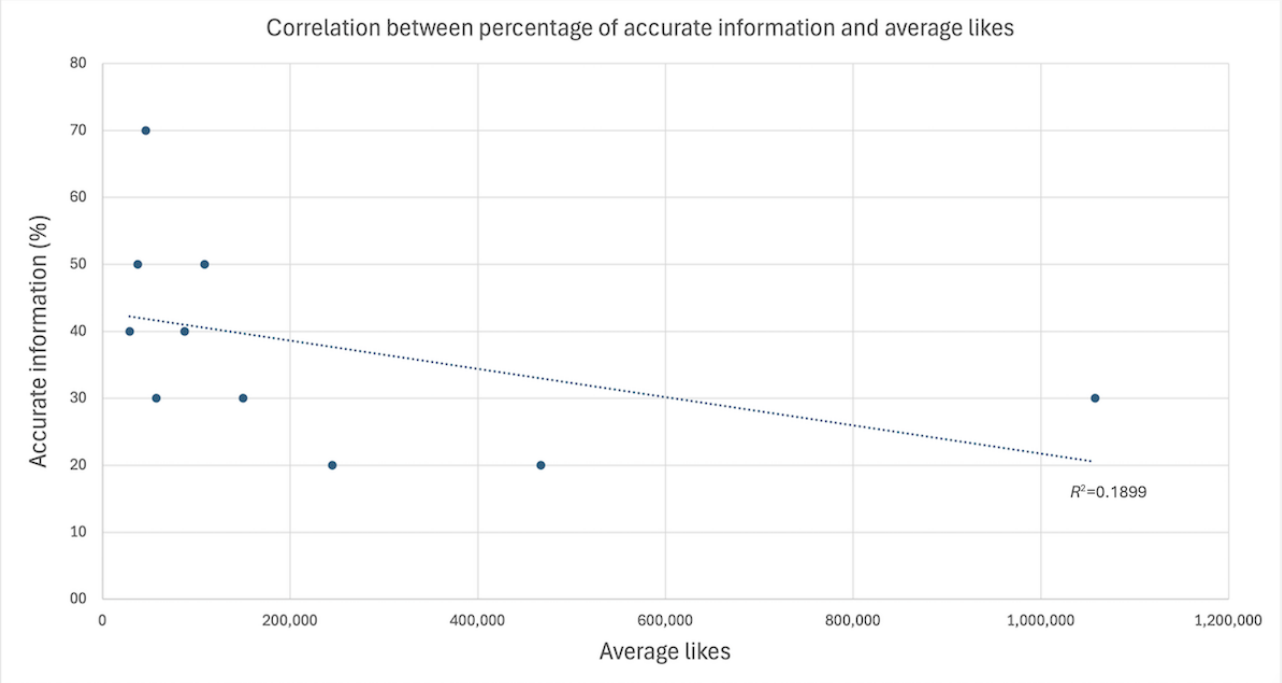
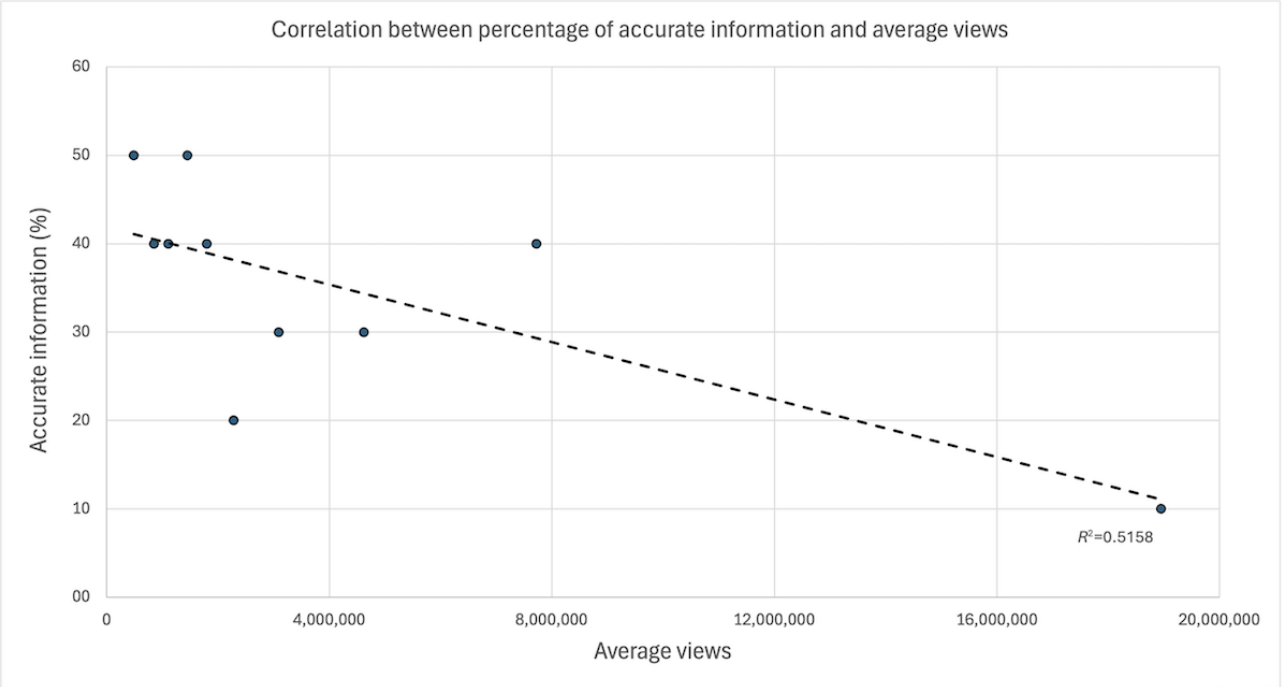


Figure 2. Relationship between accurate sunscreen content and likes. The 100 most-liked videos were grouped into deciles based on their like counts. The percentage of videos within each decile that contain accurate information is shown on the y-axis, while the average likes per decile are shown on the x-axis.



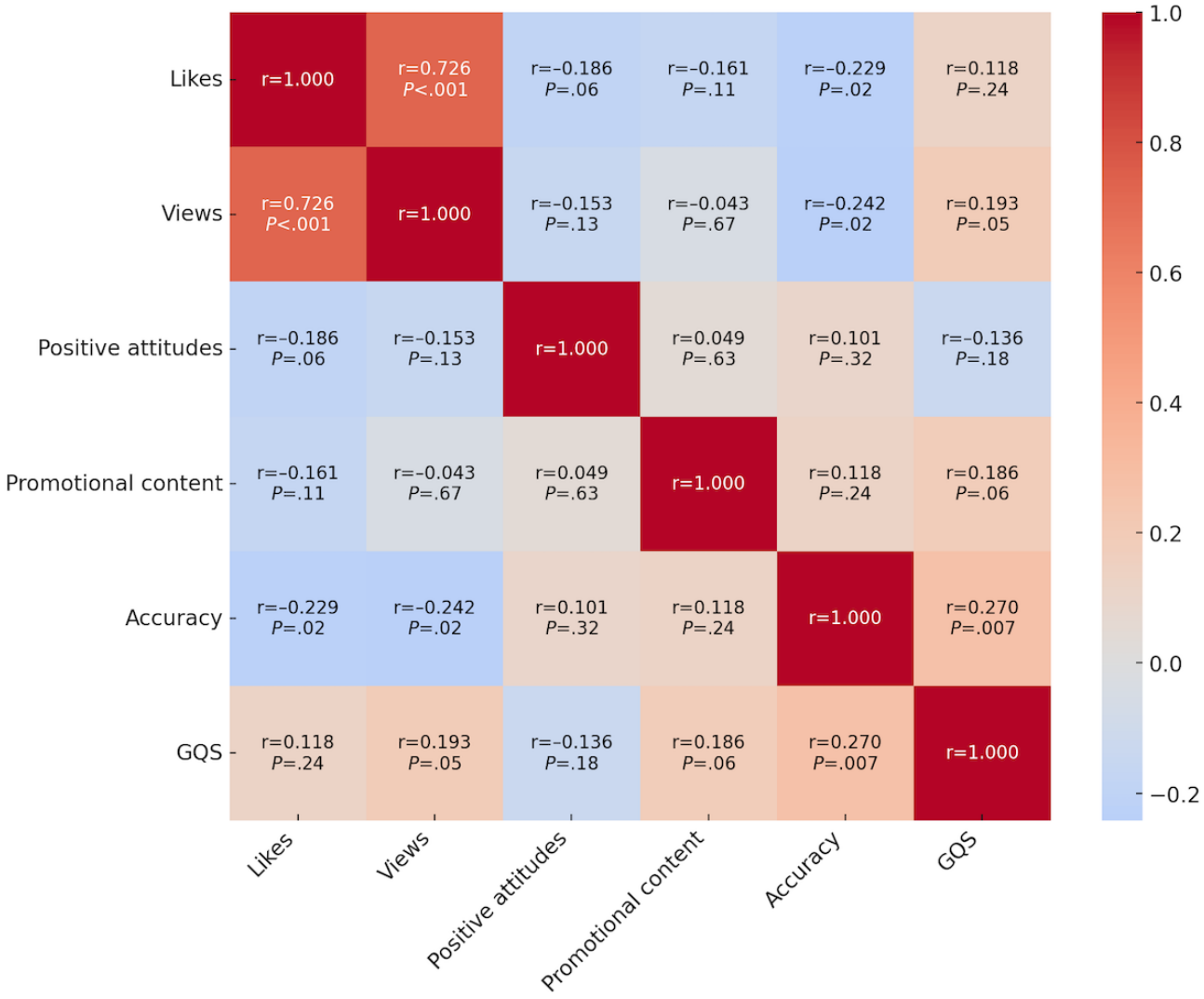
GQS and Accuracy

Content quality was generally low across the top 100 most-liked sunscreen-related TikTok videos. The majority (71/100, 71%) of videos received GQS of 1 (poor quality) or 2 (below average), with only 2 (2%) scoring a 5 (excellent quality). GQS was positively correlated with accuracy (r=0.270; df=98; P=.007), indicating that videos deemed more accurate were also rated as higher quality. However, GQS showed no significant correlation with engagement metrics such as likes (r=0.118; df=98; P=.24) or views (r=0.193; df=98; P=.05), suggesting that content quality did not predict popularity on the platform. These findings reflect a disconnect between perceived usefulness and user engagement.

Promotional Content

Promotional content was present in 27 (27%) of the top 100 most-liked sunscreen-related TikTok videos. However, the presence of promotional material was not significantly correlated with either likes (r=–0.161; df=98; P=.11) or views (r=–0.043; df=98; P=.67), suggesting that promotional status alone did not drive engagement. Similarly, promotional content showed only weak, nonsignificant correlations with GQS (r=0.186; df=98; P=.06) and accuracy (r=0.118; df=98; P=.24). These findings indicate that promotional content was common but not necessarily more engaging or accurate than nonpromotional content. Figure 3 visualizes these correlations across 6 content variables.

Figure 3. Heat map depicting Pearson correlations among key variables across the top 100 most-liked TikTok videos about sunscreen. Variables include Global Quality Score (GQS), accuracy, likes, views, promotional content, and positive attitudes. Warmer colors indicate positive correlations; cooler colors indicate negative correlations.



Discussion

Principal Findings

Our analysis of trending TikTok content on sunscreen revealed an average of 4,203,243 views and 227,671 likes per video, highlighting the platform's massive reach. Among the 100 most-liked videos, 74 (74%) expressed a positive attitude toward sunscreen, 35 (35%) contained accurate information, and only 6 (6%) contained inaccurate content, per the American Academy of Dermatology guidelines. Notably, 57 (57%) videos were purely opinion-based. Our study is the first to demonstrate that sunscreen content on TikTok is predominantly positive but frequently lacks factual grounding or educational value. These findings expand on prior reports of variable quality in dermatology TikToks by specifically examining sunscreen content, an area of high public interest and misinformation.

Comparison With Prior Work

Consistent with prior literature, we found that videos deemed accurate were rated significantly higher in quality, as measured by the GQS, aligning with established findings that content accuracy enhances educational utility [3-6]. However, more than half of the videos (57/100, 57%) were categorized as opinion-based and lacked references, reflecting a shift toward anecdotal rather than evidence-based guidance.

Promotional content was present in 27 (27%) videos but was not significantly associated with engagement metrics or content quality. It showed only weak, nonsignificant correlations with GQS (r=0.186; df=98; P=.06), accuracy (r=0.118; df=98; P=.24), likes (r=–0.161; df=98; P=.11), and views (r=–0.043; df=98; P=.67), suggesting that the presence of commercial messaging neither enhanced nor substantially degraded video quality or popularity. Although Ranpariya et al [8] documented the prevalence of promotional content among dermatology influencers, our analysis suggests that such content does not necessarily drive engagement or correlate with accuracy.

Despite comprising only 6 (6%) videos, inaccurate content attracted disproportionately high engagement, receiving a mean of 2,141,683 views and 119,383 likes, which was comparable to that of the 35 (35%) accurate videos (mean views 2,732,594; mean likes 135,711). A strong negative correlation was found between accuracy and views in decile-based analysis (r=–0.72; df=8; P=.02), and a moderate but nonsignificant negative

correlation was found between accuracy and likes (r=–0.44; df=8; P=.21). This reinforces prior concerns that social media virality may be driven more by sensationalism or esthetics than by informational accuracy [11-13].

Implications

The low overall GQS (mean 1.97) emphasizes the poor educational utility of the majority of sunscreen-related content on TikTok. Despite their reach, most videos failed to meet standards of quality, evidence, or clarity. Content with promotional framing was particularly vulnerable to low-quality scores, even when expressing positive sunscreen messaging. Given that accurate content did not consistently outperform anecdotal or promotional posts in engagement, our findings reinforce the need for dermatologists to adapt communication strategies to the realities of algorithm-driven platforms.

To address these gaps, we recommend that health care providers stay informed about dermatology trends on social media, discuss misinformation and content sources with patients, and engage on platforms such as TikTok to provide evidence-based perspectives.

Encouragingly, 18% (18/100) of the content creators in our dataset were health care providers, and 1 board-certified dermatologist had more than 18 million followers, demonstrating the potential impact medical professionals can have. To further promote quality content, we advocate for standardized disclosure of conflicts of interest, citation of sources, and increased collaboration with platform moderators.

SkinMedia Initiative

Our team is committed to continuing this work through the SkinMedia series. This initiative aims to systematically track dermatologic misinformation and translate findings into publicly accessible content. This work directly addresses the gap identified in the current literature by moving beyond analysis to the development of interventions that amplify accurate dermatologic messaging. By partnering with dermatologists, educators, and digital platforms, we aim to promote scientifically grounded messaging, particularly for younger audiences. Future work will include the dissemination of open-access educational resources.

Limitations

This study is limited by its cross-sectional design and the selection of only the 100 most-liked videos, which may not reflect less popular but potentially higher-quality content. The large percentage of promotional content (27/100, 27%) may also have skewed findings. IRR for most variables was high; however, agreement on GQS was lower (κ=0.712), consistent with its known subjectivity. Additionally, most content creators were nonhealth care professionals, which may explain the infrequent use of citations.

Conclusions

This study was one of the first to systematically analyze the most-liked sunscreen videos on TikTok, demonstrating an abundance of opinion-based, low-quality content despite an overall positive stance toward sunscreen use. We urge dermatologists and other health professionals to expand their digital presence to ensure that younger populations have access to engaging, reliable, and evidence-based skin health information. Future efforts should explore collaborative strategies with platforms and influencers to amplify high-quality dermatology content and counteract misinformation.

Acknowledgments

The authors would like to thank the SkinMedia team for their contributions to developing a framework to identify and counter dermatology-related misinformation on social media platforms.

Authors' Contributions

Conceptualization: JR-M (lead), KD (equal), AS (supporting), SL (supporting)
Data curation: JR-M (lead), KD (equal)
Formal analysis: JR-M (lead), KD (equal), TT (supporting), PC (supporting)
Funding acquisition: AGO-L (lead), SL (supporting)
Methodology: JR-M (lead), KD (equal), AS (supporting), PC (supporting), SL (supporting),
Project administration: JR-M (lead), KD (equal), SL (supporting)
Supervision: SL (lead), AGO-L (supporting)
Validation: PC (lead), TT (supporting)
Visualization: JR-M (lead), KD (equal)
Writing – original draft: JR-M (lead), KD (equal)
Writing – review & editing: TT (equal), AS (equal), PC (equal), OJW (supporting), AGO-L (supporting), SL (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example of the codebook framework for video content extraction.
[DOCX File, 16 KB - derma_v8i1e70010_app1.docx]

Multimedia Appendix 2
Interrater reliability for key coded variables in a 10% blinded subset of videos, assessed by 2 independent reviewers. Percent agreement, Scott pi, Cohen kappa, and Krippendorff alpha are reported. “Nan” indicates invariant responses where chance-adjusted metrics could not be calculated due to lack of variability in responses.
[DOCX File, 17 KB - derma_v8i1e70010_app2.docx]

References

1. 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]
2. 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]
3. Khan S, Shaik M, Khan N, Potts G. Sunscreen on screen: cross-sectional study of TikTok content on sunscreen. Skin Res Technol 2024 Oct;30(10):e70061. [doi: 10.1111/srt.70061] [Medline: 39366913]
4. Kearney N, Leahy M, Stanciu M, O’Flynn A, Walsh MC, Laing M. A cross-sectional analysis of TikTok content relating to sunscreen conspiracy theories and correlation with published evidence on sunscreen risks. Clin Exp Dermatol 2025 Mar 26;50(4):867-869. [doi: 10.1093/ced/llae456] [Medline: 39440937]
5. Nikookam Y, Potluru A, Guckian J. Comment on “A cross-sectional analysis of TikTok content relating to sunscreen conspiracy theories and correlation with published evidence on sunscreen risks”: debunking sunscreen myths and trends - lessons for dermatology educators. Clin Exp Dermatol 2025 Apr 24;50(5):1025-1027. [doi: 10.1093/ced/llae492] [Medline: 39549268]
6. Haff PL, Jacobson A, Taylor MM, et al. The new media landscape and its effects on skin cancer diagnostics, prognostics, and prevention: scoping review. JMIR Dermatol 2024 Apr 8;7:e53373. [doi: 10.2196/53373] [Medline: 38587890]
7. Hales M, Rigali S, Paller A, Liszewski W, Lagu T. Pediatric skin care regimens on TikTok. Pediatrics 2025 Jul 1;156(1):e2024070309. [doi: 10.1542/peds.2024-070309] [Medline: 40484399]
8. Ranpariya VK, Fathy R, Chu B, Wang S, Lipoff JB. Patterns of promotional content by dermatology influencers on TikTok. JMIR Dermatol 2022 Mar 30;5(1):e34935. [doi: 10.2196/34935] [Medline: 37632857]
9. Steinke J, Gilbert C, Coletti A, Levin SH, Suk J, Oeldorf-Hirsch A. Women in STEM on TikTok: advancing visibility and voice through STEM identity expression. Soc Media Soc 2024 Jul;10(3). [doi: 10.1177/20563051241274675]
10. 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]
11. Sunscreen FAQs. American Academy of Dermatology Association. URL: https://www.aad.org/media/stats-sunscreen [accessed 2025-09-22]
12. Diehl K, Barber C, Simpson EL, et al. TikTok misinformation on the role of diet in atopic dermatitis management. Dermatitis 2025;36(4):e399-e401. [doi: 10.1089/derm.2024.0123] [Medline: 39109445]
13. Guan JL, Xia SH, Zhao K, et al. Videos in short-video sharing platforms as sources of information on colorectal polyps: cross-sectional content analysis study. J Med Internet Res 2024;26:e51655. [doi: 10.2196/51655]
14. Chen AY, Azizi B, Borba AJ, Armstrong AW. Rosacea videos on social media: a comparison of accuracy, quality, and viewer engagement. Dermatol Online J 2021 Feb 15;27(2):13030/qt55c0g9wz. [Medline: 33818977]

Abbreviations

GQS: Global Quality Scale
IRR: interrater reliability

Edited by I Brooks; submitted 13.12.24; peer-reviewed by C Gilbert, M Buller, P Siuda, S Grantham; revised version received 30.09.25; accepted 29.10.25; published 01.12.25.

Please cite as:
Roland-McGowan J, Diehl K, Tobey T, Shafer A, Clement P, Wisco OJ, Ortega-Loayza AG, Leachman S
Perception, Quality, and Accuracy of Sunscreen Content on TikTok: SkinMedia Cross-Sectional Content Analysis
JMIR Dermatol 2025;8:e70010
URL: https://derma.jmir.org/2025/1/e70010
doi:10.2196/70010

The Prevalence of Dermoscopy Use Among Dermatology Residents in Riyadh, Saudi Arabia: Cross-Sectional Study

Abdullah Almeziny[1], MBBS; Rahaf Almutairi[1], MBBS; Amal Altamimi[1], MBBS; Khloud Alshehri[1], MBBS; Latifah Almehaideb[1], MBBS; Asem Shadid[2], MBBS; Mohammed Al Mashali[3], MBBS

[1]College of Medicine, Al-Imam Muhammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

[2]Department of Dermatology, King Fahad Medical City, Riyadh, Saudi Arabia

[3]Department of Dermatology, College of Medicine, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

Corresponding Author:
Abdullah Almeziny, MBBS
College of Medicine, Al-Imam Muhammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

Abstract

Background: Dermoscopy is a noninvasive technology used to examine the skin's invisible microstructures in dermatological practice and is gaining prominence as a crucial tool. Dermoscopy is an evidence-based practice used to enhance the early detection of skin malignancies and to help distinguish between various skin conditions, including pigmented and nonpigmented skin malignancies. Currently, the vast majority of global guidelines for skin cancer recommend dermoscopy as a critical component. Dermoscopy use is increasing worldwide, but to date, no study has documented the attitudes toward and use of dermoscopy among future dermatologists in Saudi Arabia.

Objective: We aimed to determine the proportion of dermatology residents in Riyadh who use dermoscopy in their clinical practice; identify factors influencing the use of dermoscopy, such as availability of equipment, training, and the perceived importance of dermoscopy in clinical practice; explore barriers to dermoscopy use, including the lack of access to necessary resources (eg, dermoscopes) and insufficient training; and provide insights into the adoption and integration of dermoscopy into dermatology training and clinical practice in Saudi Arabia.

Methods: In January 2024, a validated and published questionnaire was modified to meet research requirements and was sent to all registered dermatology residents in the The Saudi Board of Dermatology and Venereology Program.

Results: In total, 63 dermatology residents in Riyadh, Saudi Arabia, completed the web-based questionnaire (response rate=87.5%). The sample was predominantly female (n=34, 54.0%), with the majority (n=53, 84.1%) aged between 26 and 30 years. A notable proportion of participants (n=22, 34.9%) were in their final year of residency. Over half of the participants (n=34, 54.0%) owned a dermoscope, and a substantial number of them (n=23, 36.5%) reported conducting 21-30 clinic consultations per month on average. More than half of the participants (n=36, 57.1%) had received dermoscopy training, and 16 (36.4%) had used dermoscopy for 2 years. Additionally, most participants (n=20, 45.5%) had used nonpolarized immersion-contact dermoscopy, while 19 (43.2%) had used polarized light dermoscopy. Furthermore, the majority (n=22, 50.0%) used dermoscopy in fewer than 10% of cases involving patients with inflammatory skin lesions. Statistical analysis revealed significant associations between the participants' ages (P=.003), residency levels (P=.001), and practice centers and the use of dermoscopy (P=.004).

Conclusions: Dermoscopy has been widely adopted by dermatology residents in their daily clinical practice due to its benefits in early detection and diagnosis of skin diseases. However, the overall extent of dermoscopy use within the dermatology community remains unclear, highlighting the need for further education. In Saudi Arabia, the key factors influencing dermoscopy use include residents' ages, residency levels, and practice centers. Younger dermatologists have expressed strong interest in improving their dermoscopy knowledge and skills. Expanding access to dermoscopy equipment and providing training during residency could further promote its use across the country.

(JMIR Dermatol 2025;8:e63861) doi:10.2196/63861

KEYWORDS

dermoscopy; Saudi Arabia; questionnaire; skin lesion; noninvasive; cross-sectional study; dermatology

Introduction

Dermoscopy is a noninvasive technology used to examine the skin's invisible microstructures in dermatological practice [1]. It is an established technique for analyzing skin lesions, with its origins tracing back to the 17th century when Kohlhaus used a microscope to study nail matrix vessels [1-3]. However, dermoscopy did not gain widespread use until the 20th century,

when Goldman developed a portable microscope capable of magnifying up to 10 times more than the naked eye [1-6]. Historically, dermoscopy has been used to diagnose pigmented lesions such as naevi, melanomas, and pigmented basal cell carcinomas [7].

Since the 1990s, it has been used to diagnose other dermatological disorders, including infections such as scabies, inflammatory lesions, and hair and nail-fold alterations, and it is also used to track lesions' progress or reactions to topical treatments such as imiquimod or 5-fluorouracil [8-11]. The evidence-based practice associated with dermoscopy use improves the diagnostic accuracy for skin cancer, reduces unnecessary biopsies of benign lesions, increases survival rates, and improves the early identification of skin malignancies [12-14]. A 2002 meta-analysis of 27 studies revealed that dermoscopy increases experienced physicians' diagnostic accuracy for melanoma compared to unaided inspection [1]. Moreover, it helps distinguish between various kinds of pigmented and nonpigmented skin malignancies in vivo, and it is significant in differentiating between inflammatory and neoplastic lesions. Currently, the vast majority of worldwide guidelines for skin cancer recommend dermoscopy as a critical component of diagnosing and following up with patients facing an increased risk of this disease [15-20]. Dermoscopy is also increasingly accepted as a standard practice worldwide. Multiple studies have revealed that US dermatologists use dermoscopy. Indeed, 1555 out of 3238 (48%) American dermatologists surveyed in 2010 said they used dermoscopy [21]. Chamberlain's study of dermoscopy revealed a 98% usage rate use in Australia [22,23], while 95% of dermatologists in France use the practice [24]. However, no study has documented the attitudes toward and use of dermoscopy among future dermatologists in Saudi Arabia. In this study, dermoscopy prevalence among dermatology residents in Riyadh was assessed alongside information sources and elements that influence its use, such as residency levels and the frequency of dermoscopy diagnoses.

Methods

Study Design

A cross-sectional study was conducted in Riyadh to determine the attitudes of dermatologists toward the dermoscopy. The convenience sampling technique was used in this study to recruit the study participants. The questionnaire tool was distributed in January 2024 through email registered at the Saudi Commission for Health Specialists.

The inclusion criterion for this study was to be a registered dermatology resident in Saudi boards in Riyadh. Any participants who did not meet this inclusion criterion was excluded from this study.

Questionnaire Tool

This study adapted and used a questionnaire previously tested and found valid and reliable by Forsea et al [25]. The questionnaire comprises 2 sections: the first section collected information related to participants' demographics (gender, age, and residency level), and in the second section, the future dermatologists who participated were asked about their perspectives about the utility of dermoscopy, their use patterns, their training experiences, and their self-reported confidence in dermoscopy diagnosis.

Ethical Considerations

The study protocol was reviewed and approved by the Regional and institutional human medical biological research ethics committee of Al-Imam Muhammad Ibn Saud Islamic University (approval 735/2024). Participation in the study was entirely voluntary, and informed written consent was obtained from all participants before their involvement. The study was conducted in accordance with the principles of the declaration of Helsinki. All data collected were anonymized to ensure the privacy and confidentiality of the participants.

Study Analysis

All research data were entered into a Microsoft Excel (version 16.0) spreadsheet. Data analysis was performed using SPSS (version 28; IBM Corp). Descriptive statistics were used in the statistical analysis; relative frequencies (and percentage values) were used to present categorical variables. The chi-square test was used to assess the association between categorical variables and dermoscopy use, with a 2-sided value of $P<.05$ considered statistically significant.

Results

A total of 63 dermatology residents in Riyadh, Saudi Arabia, completed the web-based questionnaire, yielding an 87.5% response rate. More than half of the participants (n=34, 54.0%) were female, the majority (n=53, 84.1%) were aged between 26 and 30 years, and a considerable proportion (n=22, 34.9%) were in their fourth year of residency (Table 1).

Table . Participants' (N=63) sociodemographic information.

Sociodemographic characteristics		Participants, n (%)
Gender		
	Female	34 (54.0)
	Male	29 (46.0)
Age (years)		
	20 - 25	8 (12.7)
	26 - 30	53 (84.1)
	31 - 35	1 (1.6)
	36 - 40	1 (1.6)
Residency level		
	Residency year 1	9 (14.3)
	Residency year 2	19 (30.2)
	Residency year 3	13 (20.6)
	Residency year 4	22 (34.9)

Table 2 (below) depicts practice characteristics, dermoscopy training, and dermoscopy use patterns among the participating dermatology residents in Riyadh, Saudi Arabia. More than half of the participants (n=34; 54.0%) owned a dermoscope, while a substantial number of them (n=23, 36.5%) offered an average of 21 - 30 monthly clinic consultations. Most participants (n=57, 90.5%) presented an average number of 0-5 clinic consultations per month where they saw patients with cancer (all types). More than half of the participants (n=36, 57.1%) had received dermoscopy training. The majority of the participants (n=44, 69.8%) used dermoscopy, and a significant number of them had been inspired to do so by their colleagues (n=12, 27.3%) and mentors (n=9, 20.5%). The reported reasons for not using dermoscopy were its unavailability in an office (n=8, 42.1%) and a lack of training (n=6, 31.6%). Half of the participants (n=22, 50.0%) reported having used dermoscopy pictures in medical education, particularly in conferences, lectures, and academic activities. Most of the participants (n=17, 38.6%) had completed a rotation at King Saud University Medical City in 2023. A considerable proportion of participants (n=16, 36.4%) had used dermoscopy for 2 years; the majority used a nonpolarized immersion-contact dermoscope (n=20, 45.5%) or polarized-light dermoscope (n=19, 43.2%). Regarding their average practice, the majority of participants (n=15, 34.1%) reported using dermoscopy at least once per day. ABCD (Asymmetrical, Border, Color, Diameter) was reported to be the most common algorithm used by the majority of the participating dermatology residents (n=23, 52.3%) for the diagnosis of pigmented lesions.

Table . Dermoscopy practice characteristics, training, and use patterns.

Question and categories		Participants, n (%)
Do you own a dermoscope?		
	Yes, I do	34 (54.0)
	It is provided in the clinic	9 (14.3)
	No, I do not own one, nor is it provided	20 (31.7)
What is your average number of monthly clinic consultations?		
	0 - 10	14 (22.2)
	11 - 20	22 (34.9)
	21 - 30	23 (36.5)
	31 - 40	2 (3.2)
	More than 40	2 (3.2)
What is the average number of monthly clinic consultations where you see patients with skin cancer (of all types)?		
	0 - 5	57 (90.5)
	6 - 10	3 (4.8)
	11 - 20	2 (3.2)
	More than 20	1 (1.6)
Have you received dermoscopy training as part of your dermatology residency?		
	Yes	36 (57.1)
	No	27 (42.9)
Outside of your residency training, what type of dermoscopy training have you pursued?		
	Academic activities provided by the residency program	21 (33.3)
	Dermoscopy course	5 (7.9)
	Web-based dermoscopy course	14 (22.2)
	Attended conferences or congresses	3 (4.8)
	Books or atlases	3 (4.8)
	A mentor or tutor	4 (6.3)
	No training	13 (20.6)
Do you use dermoscopy?		
	Yes	44 (69.8)
	No	19 (30.2)
Which of the following made you consider using dermoscopy?		
	A colleague	12 (27.3)
	A mentor	9 (20.5)
	Conference lectures	6 (13.6)
	Evidence-based practice	4 (9.1)
	Lectures provided by dermatology Saudi boards residency program	3 (6.8)
	A paid workshop	2 (4.5)
	Other	8 (18.2)
If you do not use dermoscopy, please give the reason why not.		
	A dermoscope is not available in my office	8 (42.1)
	I have not been trained in dermoscopy	6 (31.6)
	Other	5 (26.3)

Question and categories		Participants, n (%)
Have you used dermoscopy pictures in medical education?		
	No, I have not used them	18 (40.9)
	Yes, in conferences, lectures, academic activities, etc	22 (50.0)
	Yes, in publications in articles or journals	4 (9.1)
	Other	20 (31.7)
In the last year, where was your rotation?		
	King Faisal Specialist Hospital	10 (22.7)
	King Saud University Medical City	17 (38.6)
	Ministry of National Guard Hospital	12 (27.3)
	Prince Sultan Military Medical City	5 (11.4)
For how long have you been using dermoscopy?		
	1 years	13 (29.5)
	2 years	16 (36.4)
	3 years	12 (27.3)
	4 years	3 (6.8)
What type of dermoscope do you use?		
	Nonpolarized immersion-contact dermoscope (contact with the skin and an interface liquid, eg, oil or alcohol)	20 (45.5)
	Polarized-light dermoscope	19 (43.2)
	Dermoscope with a digital camera	2 (4.5)
	Digital video dermoscopy system (eg, Fotofinder or Molemax)	3 (6.8)
In your average practice, how often do you use dermoscopy?		
	Less than once per month	5 (11.4)
	1 - 4 times per month	13 (29.5)
	More than once per week	11 (25.0)
	At least once per day	15 (34.1)
Which particular algorithm for the dermoscopic diagnosis of pigmented lesions do you regularly use?		
	ABCD[a] rule	23 (52.3)
	I do not systematically use any particular algorithm	10 (22.7)
	Menzies's algorithm	1 (2.3)
	Pattern analysis	7 (15.9)
	Seven-point checklist	3 (6.8)

[a]ABCD: Asymmetrical, Border, Color, Diameter.

Table 3 illustrates clinical practices and the confidence in dermoscopy skills among the participating dermatology residents in Riyadh, Saudi Arabia. Our findings revealed that the majority of the participants (n=22, 50.0%) used dermoscopy in fewer than 10% of cases involving patients with inflammatory skin lesions. Moreover, a substantial proportion of participants (n=15,34.0%) used dermoscopy in more than 70% of cases involving the examination of pigmented skin tumors. Eleven (25.0%) participants used dermoscopy for <10% of their patients who were examined for nonpigmented skin tumors. Regarding the participants' dermoscopy skills, the majority of them were somewhat confident in the assessment of nonpigmented skin tumors (n=26, 59.1%), inflammatory skin lesions (n=22, 50.0%), and pigmented skin tumors (n=19, 43.2%).

Table . Clinical dermoscopy practices and confidence in dermoscopy skills.

Category		Pigmented skin tumors, n (%)	Nonpigmented skin tumors, n (%)	Inflammatory skin lesions, n (%)
When examining patients with the following disorders, in what percentage of cases do you use dermoscopy?				
	<10% of cases	9 (20.5)	11 (25.0)	22 (50.0)
	11% - 30% of cases	8 (18.2)	8 (18.2)	10 (22.7)
	31% - 50% of cases	4 (9.1)	6 (13.6)	5 (11.4)
	51% - 70% of cases	8 (18.2)	10 (22.7)	3 (6.8)
	>70% of cases	15 (34.0)	9 (20.5)	2 (4.5)
How confident are you in your dermoscopy skills for the assessment of the following types of lesions?				
	Not confident	12 (27.3)	8 (18.2)	11 (25.0)
	Somewhat confident	19 (43.2)	26 (59.1)	22 (50.0)
	Confident	13 (29.5)	10 (22.7)	11 (25.0)

Table 4 illustrates the usefulness, advantages, and performance of dermoscopy. The vast majority of the participants (n=41, 93.2%) reported that dermoscopy was useful in diagnosing melanoma and following up on melanocytic lesions (n=39, 88.6%), diagnosing pigmented skin tumors (n=35, 79.5%), and diagnosing nonpigmented skin tumors (n=31, 70.5%). Regarding advantages, the majority of the participants agreed that dermoscopy use increases confidence in their clinical diagnoses (n=30, 68.2%), reduces unnecessary biopsies or excisions (n=27, 61.4%), and improves record-keeping (n=25, 56.8%). Weighing in on performance, more than half of the participants (n=30, 68.2%) reported that dermoscopy use increases the number of melanomas detected compared to naked-eye examinations. Additionally, the majority of participants (n=27, 61.4%) noted that the use of dermoscopy reduces the excision of benign lesions.

Table . Usefulness, advantages, and performance of dermoscopy use.

Category			Participants, n (%)
Usefulness of dermoscopy			
	Diagnosis of melanoma		
		Not useful	1 (2.3)
		Somewhat useful	2 (4.5)
		Useful	41 (93.2)
	Follow-up on melanocytic lesions		
		Not useful	0 (0)
		Somewhat useful	5 (11.4)
		Useful	39 (88.6)
	Diagnosis of pigmented skin tumors		
		Not useful	0 (0)
		Somewhat useful	9 (20.5)
		Useful	35 (79.5)
	Diagnosis of nonpigmented skin tumors		
		Not useful	1 (2.3)
		Somewhat useful	12 (27.3)
		Useful	31 (70.5)
	Diagnosis of inflammatory skin lesions		
		Not useful	3 (6.8)
		Somewhat useful	19 (43.2)
		Useful	22 (50.0)
	Follow-up on nonmelanocytic skin lesions		
		Not useful	3 (6.8)
		Somewhat useful	19 (43.2)
		Useful	22 (50.0)
Advantages of using dermoscopy			
	Diagnoses melanoma in early stages		
		Strongly agree	22 (50.0)
		Agree	19 (43.2)
		Neither agree nor disagree	3 (6.8)
		Disagree	0 (0)
	Allows the monitoring of lesions		
		Strongly agree	22 (50.0)
		Agree	19 (43.2)
		Neither agree nor disagree	3 (6.8)
		Disagree	0 (0)
	Reduces the number of unnecessary biopsies or excisions		
		Strongly agree	27 (61.4)
		Agree	14 (31.8)
		Neither agree nor disagree	3 (6.8)
		Disagree	0 (0)
	Increases confidence in my clinical diagnoses		

Category			Participants, n (%)
		Strongly agree	30 (68.2)
		Agree	13 (29.5)
		Neither agree nor disagree	1 (2.3)
		Disagree	0 (0)
	Improves record-keeping		
		Strongly agree	25 (56.8)
		Agree	14 (31.8)
		Neither agree nor disagree	4 (9.1)
		Disagree	1 (2.3)
	Reduces patients' anxiety		
		Strongly agree	22 (50.0)
		Agree	12 (27.3)
		Neither agree nor disagree	10 (22.7)
		Disagree	0 (0)
	Improves documentation for medical liabilities		
		Strongly agree	25 (56.8)
		Agree	12 (27.3)
		Neither agree nor disagree	6 (13.6)
		Disagree	1 (2.3)
	Increases reimbursement		
		Strongly agree	21 (47.7)
		Agree	12 (27.3)
		Neither agree nor disagree	11 (25.0)
		Disagree	0 (0)
Dermoscopy performance			
	Dermoscopy has increased the number of melanomas detected compared to naked-eye examinations		
		Yes	30 (68.2)
		No	14 (31.8)
	In your practice, how did the use of dermoscopy influence the number of excisions of benign lesions that you performed?		
		Decreased the number	27 (61.4)
		Increased the number	6 (13.6)
		Did not change the number	11 (25.0)

Table 5 presents the relationship between categorical variables and the use of dermoscopy, as well as dermatologists' training. The results established a significant association of the participants' ages (*P*=.003), residency levels (*P*=.001), and practice centers (*P*=.004) with the use of dermoscopy among the participants. Additionally, this study established a significant association between receiving dermoscopy training and confidence levels among participating dermatology residents (*P*=.002). Furthermore, a significant association between the type of training and the type of dermoscopy use was found (*P*=.003).

Table . The association between categorical variables and dermoscopy use—association between participants' categorical variables and the use of dermoscopy, use frequency, and training type.

Variables		Participants, n (%)	*P* value
Gender			.36
	Female	22 (50.0)	
	Male	22 (50.0)	
Age (years)			.003
	20 - 25	3 (6.8)	
	26 - 30	39 (88.6)	
	31 - 35	1 (2.3)	
	36 - 40	1 (2.3)	
Residency level			.001
	Residency year 1	2 (4.5)	
	Residency year 2	11 (25.0)	
	Residency year 3	9 (20.5)	
	Residency year 4	22 (50.0)	
Device availability and cost			.12
	Yes, I own such a device	34 (77.3)	
	It is provided in the clinic	8 (18.2)	
	No, I do not own it, nor is it provided	2 (4.5)	
Practice center			__[a]
	King Saud University Medical City	11 (25.0)	
	Other	33 (75.0)	
Number of skin clinics and patients			.45
	Less than once per month	5 (11.4)	
	1 - 4 times per month	13 (29.5)	
	More than once per week	11 (25.0)	
	At least once per day	15 (34.1)	
Type of training			.43
	Dermoscopy training	14 (31.8)	
	Other	30 (68.2)	
Change in excisions of benign lesions			.22
	Yes	21 (44.7)	
	No	23 (52.3)	
Receiving dermoscopy training			.43
	Yes	14 (31.8)	
	No	30 (68.2)	
Receiving dermoscopy training			
	Owning a dermoscope		.13
	Yes, I own one	34 (77.3)	
	It is provided in the clinic	8 (18.2)	
	No, I do not own one, nor is one provided	2 (4.5)	

Variables		Participants, n (%)	*P* value
Degree of confidence			.002
	Yes	36 (81.8)	
	No	8 (18.2)	
Type of usage			.46
	Benign lesion	24 (54.5)	
	Pigment skin tumors	20 (45.5)	
Dermoscopy use frequency			
Lesion type			.58
	Pigmented skin tumors	9 (20.5)	
	Nonpigmented skin tumors	13 (29.5)	
	Inflammatory skin lesions	22 (50.0)	
Type of training			
Usage type			.003
	Dermoscope with a digital camera	2 (4.5)	
	Nonpolarized immersion-contact dermoscope	23 (52.3)	
	Polarized-light dermoscope	19 (43.2)	
Inflammatory skin lesion			.57
	Yes	22 (50.0)	
	No	22 (50.0)	

[a]Not applicable.

Discussion

Principal Findings

This study aimed to assess the prevalence of dermoscopy use among dermatology residents in Riyadh, Saudi Arabia, and the need for dermoscopy training, as well as the practice's benefits in diagnosing and treating skin diseases. The study's sample was predominantly female. Moreover, a substantial majority of the participants were in their fourth year of residency and most of them were aged between 26 and 30 years.

This study revealed that more than half of the surveyed dermatology residents owned a dermoscope, with a considerable majority seeing a significant number of patients with cancers of all types every month. Additionally, more than half of the participants had received dermoscopy training, and a considerable proportion were pursuing academic activities provided by the residency program outside of their specialized training. The study's findings underscore the importance of dermoscopy use and the necessity of better dermoscopy training as an invaluable tool in the earlier recognition of different dermatological diseases [25], as well as future strategic planning and enhanced dermoscopy training and practice in Saudi Arabia [26]. Our study verified that most of the participating dermatology residents used dermoscopy to manage their patients' conditions, and they had received training on its use. A considerable proportion of the participants had used dermoscopy for 2 years, and the majority used nonpolarized immersion-contact and polarized-light dermoscopes.

These findings are consistent with those of a study conducted by Freeman et al [27] in the United States, which revealed that dermatologists apply dermoscopy in their daily routines to manage patients' conditions and to diagnose their patients [27]. Similarly, a study conducted by Jones et al [28] on dermoscopy use as part of primary care in the United Kingdom found that dermatologists used dermoscopy to manage their patients' conditions daily at a rate of 98.5% [28]. This study revealed that the majority of respondents used dermoscopy, with many being inspired to do so by their colleagues and mentors. However, some participants did not use dermoscopy due to the unavailability of dermoscopes in their offices and insufficient training. These findings align with those of a study by Alqahtani and AlBukhari [29] in Saudi Arabia, which identified a lack of adequate education and training among residents as a key reason for dermatologists' reluctance to use dermoscopy. Similarly, our findings are consistent with those of a study conducted by

Engasser and Warshaw [21] in the United States, which identified financial costs and lack of training as the primary reasons why dermatologists avoid using dermoscopy [30].

This study's findings revealed that the majority of the participating dermatology residents used dermoscopy in fewer than 10% of cases involving patients with inflammatory skin lesions, in >70% of cases involving the examination of patients for pigmented skin tumors, and in <10% of cases in which patients were examined for nonpigmented skin tumors. Additionally, the majority of the participants reported that dermoscopy was useful in diagnosing melanoma, following up on melanocytic lesions, diagnosing pigmented skin tumors, and diagnosing nonpigmented skin tumors. These findings align with those of the study by Kuo et al [30], which involved dermatologists in Taiwan and noted that clinicians used dermoscopy to examine pigmented and nonpigmented lesions. This study found that the use of dermoscopy was associated with dermatologists' increased confidence in their clinical diagnosis, that it reduced unnecessary biopsies or excisions, and that it improved record-keeping. Furthermore, the study revealed that dermoscopy use increased the number of melanomas detected compared to naked-eye examinations while also reducing the excisions of benign lesions.

This study's demographics revealed that most of the participants were female. This preponderance can be explained by the higher proportion of female dermatologists worldwide [21,27]. In terms of age, majority of participating dermatology residents were between 26 and 30 years old. This suggests that younger dermatologists in Saudi Arabia are using dermoscopy more frequently than their older counterparts. These findings are consistent with those of a study conducted by Blum et al [31] in Germany, which reported higher dermoscopy usage rates among individuals younger than 35 years. This highlights the growing role of dermoscopy in clinical practice and the younger generation's willingness to embrace new technologies for diagnosing and treating skin diseases.

Limitations

Our results may be subject to several limitations. Despite the high response rate (87.5%), participants who chose to respond may have differed in their attitudes, experiences, or usage of dermoscopy compared to nonrespondents. As a cross-sectional study, it is limited by its inability to assess causal relationships. The sample size was relatively small, which could have increased the risk of sampling bias. Additionally, since the study was conducted in only one region, that is, Riyadh, its findings may not be generalizable to the entire country of Saudi Arabia. Furthermore, because the study involved an web-based questionnaire, it relied on respondents accurately documenting their responses, without the ability to verify their accuracy, which may have introduced bias.

Conclusion

Dermoscopy has been widely adopted, with more than half of the dermatology residents in Riyadh, Saudi Arabia, using this technology. Its use is increasing among dermatology residents due to its evidence-based advantages in the early detection and diagnosis of skin diseases. The participants' ages, residency levels, and practice centers were identified as the main factors influencing dermoscopy use in Saudi Arabia. The study also highlighted a strong willingness among young dermatologists to improve their dermoscopy knowledge and skills. Based on these findings, the study recommends that policy makers prioritize funding for dermoscopy by increasing the number of dermoscopes, as well as focusing on capacity building and training for dermatology residents.

Conflicts of Interest

None declared.

References

1. Wu X, Marchetti MA, Marghoob AA. Dermoscopy: not just for dermatologists. Melanoma Manag 2015 Feb;2(1):63-73. [doi: 10.2217/mmt.14.32] [Medline: 30190832]
2. Marghoob AA, Swindle LD, Moricz CZM, et al. Instruments and new technologies for the in vivo diagnosis of melanoma. J Am Acad Dermatol 2003 Nov;49(5):777-797. [doi: 10.1016/s0190-9622(03)02470-8] [Medline: 14576657]
3. Menzies SW, Ingvar C, McCarthy WH. A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma. Melanoma Res 1996 Feb;6(1):55-62. [doi: 10.1097/00008390-199602000-00008] [Medline: 8640071]
4. Argenziano G, Soyer HP. Dermoscopy of pigmented skin lesions--a valuable tool for early diagnosis of melanoma. Lancet Oncol 2001 Jul;2(7):443-449. [doi: 10.1016/s1470-2045(00)00422-8] [Medline: 11905739]
5. Campos-do-Carmo G, Ramos-e-Silva M. Dermoscopy: basic concepts. Int J Dermatol 2008 Jul;47(7):712-719. [doi: 10.1111/j.1365-4632.2008.03556.x] [Medline: 18613881]
6. Buch J, Criton S. Dermoscopy saga - a tale of 5 centuries. Indian J Dermatol 2021;66(2):174-178. [doi: 10.4103/ijd.IJD_691_18] [Medline: 34188274]
7. Pehamberger H, Steiner A, Wolff K. In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions. J Am Acad Dermatol 1987 Oct;17(4):571-583. [doi: 10.1016/s0190-9622(87)70239-4] [Medline: 3668002]
8. Argenziano G, Fabbrocini G, Delfino M. Epiluminescence microscopy. A new approach to in vivo detection of Sarcoptes scabiei. Arch Dermatol 1997 Jun;133(6):751-753. [doi: 10.1001/archderm.133.6.751] [Medline: 9197830]

9. Zalaudek I, Lallas A, Moscarella E, Longo C, Soyer HP, Argenziano G. The dermatologist's stethoscope-traditional and new applications of dermoscopy. Dermatol Pract Concept 2013 Apr;3(2):67-71. [doi: 10.5826/dpc.0302a11] [Medline: 23785649]
10. Zalaudek I, Argenziano G, Di Stefani A, et al. Dermoscopy in general dermatology. Dermatology (Basel) 2006;212(1):7-18. [doi: 10.1159/000089015] [Medline: 16319467]
11. Kaçar N, Sanli B, Zalaudek I, Yildiz N, Ergin S. Dermatoscopy for monitoring treatment of actinic keratosis with imiquimod. Clin Exp Dermatol 2012 Jul;37(5):567-569. [doi: 10.1111/j.1365-2230.2011.04272.x] [Medline: 22300301]
12. Argenziano G, Cerroni L, Zalaudek I, et al. Accuracy in melanoma detection: a 10-year multicenter survey. J Am Acad Dermatol 2012 Jul;67(1):54-59. [doi: 10.1016/j.jaad.2011.07.019] [Medline: 21982636]
13. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol 2002 Mar;3(3):159-165. [doi: 10.1016/s1470-2045(02)00679-4] [Medline: 11902502]
14. Liebman TN, Goulart JM, Soriano R, et al. Effect of dermoscopy education on the ability of medical students to detect skin cancer. Arch Dermatol 2012 Sep;148(9):1016-1022. [doi: 10.1001/archdermatol.2012.509] [Medline: 22986850]
15. Tromme I, Devleesschauwer B, Beutels P, et al. Selective use of sequential digital dermoscopy imaging allows a cost reduction in the melanoma detection process: a belgian study of patients with a single or a small number of atypical nevi. PLoS ONE 2014;9(10):e109339. [doi: 10.1371/journal.pone.0109339] [Medline: 25313898]
16. van der Rhee JI, Bergman W, Kukutsch NA. The impact of dermoscopy on the management of pigmented lesions in everyday clinical practice of general dermatologists: a prospective study. Br J Dermatol 2010 Mar;162(3):563-567. [doi: 10.1111/j.1365-2133.2009.09551.x] [Medline: 19832836]
17. Giacomel J, Zalaudek I, Argenziano G, Lallas A. Dermoscopy of hypertrophic lupus erythematosus and differentiation from squamous cell carcinoma. J Am Acad Dermatol 2015 Jan;72(1 Suppl):S33-S36. [doi: 10.1016/j.jaad.2014.05.055] [Medline: 25500034]
18. Errichetti E, Stinco G. The practical usefulness of dermoscopy in general dermatology. G Ital Dermatol Venereol 2015 Oct;150(5):533-546. [Medline: 26086412]
19. Dummer R, Schadendorf D, Ascierto PA, Larkin J, Lebbé C, Hauschild A. Integrating first-line treatment options into clinical practice: what's new in advanced melanoma? Melanoma Res 2015 Dec;25(6):461-469. [doi: 10.1097/CMR.0000000000000200] [Medline: 26426764]
20. Watts CG, Dieng M, Morton RL, Mann GJ, Menzies SW, Cust AE. Clinical practice guidelines for identification, screening and follow-up of individuals at high risk of primary cutaneous melanoma: a systematic review. Br J Dermatol 2015 Jan;172(1):33-47. [doi: 10.1111/bjd.13403] [Medline: 25204572]
21. Engasser HC, Warshaw EM. Dermatoscopy use by US dermatologists: a cross-sectional survey. J Am Acad Dermatol 2010 Sep;63(3):412-419. [doi: 10.1016/j.jaad.2009.09.050] [Medline: 20619490]
22. Piliouras P, Buettner P, Soyer HP. Dermoscopy use in the next generation: a survey of Australian dermatology trainees. Australas J Dermatol 2014 Feb;55(1):49-52. [doi: 10.1111/ajd.12061] [Medline: 23713814]
23. Venugopal SS, Soyer HP, Menzies SW. Results of a nationwide dermoscopy survey investigating the prevalence, advantages and disadvantages of dermoscopy use among Australian dermatologists. Australas J Dermatol 2011 Feb;52(1):14-18. [doi: 10.1111/j.1440-0960.2010.00708.x] [Medline: 21332687]
24. Moulin C, Poulalhon N, Duru G, Debarbieux S, Dalle S, Thomas L. Dermoscopy use by French private practice dermatologists: a nationwide survey. Br J Dermatol 2013 Jan;168(1):74-79. [doi: 10.1111/j.1365-2133.2012.11216.x] [Medline: 22880932]
25. Forsea AM, Tschandl P, Zalaudek I, et al. The impact of dermoscopy on melanoma detection in the practice of dermatologists in Europe: results of a pan-European survey. J Eur Acad Dermatol Venereol 2017 Jul;31(7):1148-1156. [doi: 10.1111/jdv.14129] [Medline: 28109068]
26. Chamberlain AJ, Kelly JW. Use of dermoscopy in Australia. Med J Aust 2007 Aug 20;187(4):252-253. [doi: 10.5694/j.1326-5377.2007.tb01222.x] [Medline: 17708733]
27. Freeman SR, Greene RE, Kimball AB, et al. US dermatology residents' satisfaction with training and mentoring: survey results from the 2005 and 2006 Las Vegas Dermatology Seminars. Arch Dermatol 2008 Jul;144(7):896-900. [doi: 10.1001/archderm.144.7.896] [Medline: 18645141]
28. Jones OT, Jurascheck LC, Utukuri M, Pannebakker MM, Emery J, Walter FM. Dermoscopy use in UK primary care: a survey of GPs with a special interest in dermatology. J Eur Acad Dermatol Venereol 2019 Sep;33(9):1706-1712. [doi: 10.1111/jdv.15614] [Medline: 30977937]
29. Alqahtani NN, AlBukhari FA. The use of dermoscopy among dermatologists in Riyadh, Saudi Arabia: a cross-sectional study. J Dermatol Dermatol Surg 2021;25(1):30-32. [doi: 10.4103/jdds.jdds_58_20]
30. Kuo YW, Chang YJ, Wang SH, et al. Survey of dermoscopy use by Taiwanese dermatologists. Dermatol Sinica 2015 Dec;33(4):215-219. [doi: 10.1016/j.dsi.2015.06.002]
31. Blum A, Kreusch J, Stolz W, et al. The status of dermoscopy in Germany - results of the cross-sectional Pan-Euro-Dermoscopy Study. J Dtsch Dermatol Ges 2018 Feb;16(2):174-181. [doi: 10.1111/ddg.13431] [Medline: 29384261]

Abbreviations

ABCD: Asymmetrical, Border, Color, Diameter

Edited by A Kristensen-Cabrera; submitted 01.07.24; peer-reviewed by G Pappa, MA El-Hamd; revised version received 12.12.24; accepted 16.12.24; published 23.01.25.

Please cite as:
Almeziny A, Almutairi R, Altamimi A, Alshehri K, Almehaideb L, Shadid A, Al Mashali M
The Prevalence of Dermoscopy Use Among Dermatology Residents in Riyadh, Saudi Arabia: Cross-Sectional Study
JMIR Dermatol 2025;8:e63861
URL: https://derma.jmir.org/2025/1/e63861
doi:10.2196/63861

Patient Perceptions of Artificial Intelligence and Telemedicine in Dermatology: Narrative Review

Charlotte McRae, BS; Ting Dan Zhang*, BS; Leslie Donoghue Seeley*, PhD; Michael Anderson, BS; Laci Turner, BS; Lauren V Graham, MD, PhD

Department of Dermatology, University of Alabama at Birmingham Heersink School of Medicine, 510 20th Street South FOT 858, Birmingham, AL, United States

*these authors contributed equally

Corresponding Author:
Charlotte McRae, BS
Department of Dermatology, University of Alabama at Birmingham Heersink School of Medicine, 510 20th Street South FOT 858, Birmingham, AL, United States

Abstract

Background: Artificial intelligence (AI) and telemedicine have significant potential to transform dermatology care delivery, but patient perspectives on these technologies have not been systematically compared.

Objective: This study aimed to examine patient perspectives on AI and telemedicine in dermatology to inform implementation strategies as these technologies increasingly converge in clinical practice.

Methods: A comprehensive literature search was conducted using PubMed, Scopus, and Embase databases between August 2024 and October 2024. We identified 48 papers addressing patient perspectives on AI and telemedicine in dermatology, with none directly comparing patients' views of both technologies.

Results: Several distinct themes emerged regarding patient perspectives on these technologies: willingness to use, perceived benefits and risks, barriers to implementation, and conditions necessary for successful integration. Findings revealed that patients express hesitancy toward AI-based diagnoses that lack dermatologist involvement, while preferences for teledermatology varied by reason for appointment, age, and previous technology exposure. Patients' motivations for implementing AI are connected to its potential for quicker diagnoses and improved triage efficiency. At the same time, telemedicine addresses logistical challenges such as reduced travel time and improved appointment availability. Both technologies were perceived to improve accessibility and diagnostic efficiency, though patients expressed concerns about AI's limited communication abilities and teledermatology's inability to perform physical examinations. Primary adoption barriers for these modalities included technological limitations and trust concerns, with patients emphasizing the need for dermatologist oversight, transparency, and adequate educational resources for successful integration.

Conclusions: The complementary strengths of AI and teledermatology suggest they could mitigate each other's limitations when integrated—AI potentially enhancing teledermatology's diagnostic accuracy, while teledermatology addresses AI's lack of human connection. By thoroughly examining these perspectives, this review may serve as a guide for the patient-centered integration of technology in the future landscape of accessible dermatologic care.

(JMIR Dermatol 2025;8:e75454) doi:10.2196/75454

KEYWORDS

digital health; technology; patient-centered care; health care innovation; trust; convergence; artificial intelligence; teledermatology

Introduction

Artificial intelligence (AI) and telemedicine have the potential to transform health care delivery in dermatology, where they serve distinct yet complementary functions. AI is a branch of computer science that involves the automation of intelligent behavior [1]. Machine learning, a subfield of AI, applies large datasets to identify patterns for diagnosis and predict clinical outcomes in medicine [1]. In dermatology, AI applications include tools that classify dermatological images obtained from outside clinical settings, as well as clinician decision-support systems that analyze images of patient skin concerns at the point of care [1].

Complementing these AI innovations, telemedicine has experienced rapid growth, specifically after the COVID-19 pandemic catalyzed its widespread adoption as an alternative to face-to-face consultations amid continued service demands [2,3]. Teledermatology, the subset of telemedicine specific to dermatology, offers various delivery modalities. These include synchronous approaches, which involve real-time

communication between the patient and dermatologist, asynchronous methods where a dermatologist evaluates clinical images at a later time, and hybrid models combining both approaches [4]. Through these avenues, teledermatology enables remote consultative recommendations, prioritization of care through remote triage, and monitoring of chronic conditions [5].

The integration of AI within teledermatology platforms represents a natural progression in dermatologic care delivery [6]. AI algorithms can enhance teledermatology visits by providing real-time image quality assessment, automated prescreening of cases, and diagnostic decision support during online consultations [7,8]. This convergence could benefit underserved populations by combining AI's diagnostic capabilities with the remote accessibility advantages of telemedicine [9]. While this technological integration enables advanced, location-independent models for the future of accessible dermatologic care [4], both technologies face unique challenges, and their convergence could potentially either amplify or ameliorate the barriers and limitations of each. Thus, considering patient perspectives on each modality is paramount to promoting the responsible use of these technologies in dermatology as remote care platforms evolve into influential components of dermatologic practice [10,11]. By examining patient perspectives on both AI and teledermatology, this review aims to inform implementation strategies that capitalize on synergistic benefits while addressing challenges, serving as a guide for the patient-centered integration of technology in dermatologic care.

Methods

Overview

A comprehensive literature search was conducted using the PubMed, Scopus, and Embase databases between August 2024 and October 2024. To identify relevant papers addressing patient perspectives on AI and telemedicine in dermatology, we used the following search terms: ((artificial intelligence) AND (dermatology)) AND (comfort OR perception OR perspective), as well as ((Telemedicine OR Teledermatology) AND (dermatology)) AND (comfort OR perception OR perspective).

The initial search yielded 622 papers, which were imported into Covidence for systematic screening. After removing 236 duplicates, 386 studies underwent independent title and abstract review by three researchers (CM, TDZ, and LDS). Conflicts were resolved through discussion until a consensus was reached. Following the initial screening, 60 papers underwent a full-text review, with a focus on original research and data-supported observations. Ultimately, 38 papers met our inclusion criteria: studies addressing patient perspectives or perceptions of AI and telemedicine in dermatology, published after 2009, in English with full text available, and providing empirical data on patient perspectives (Table 1). We limited inclusion to studies published after 2009 to align with the major advances in AI and telemedicine over the past 15 years, during which AI became increasingly integrated into both health care and broader societal applications, and telemedicine saw widespread adoption across global health systems [5,6]. This timeframe was chosen to capture evolving patient expectations and technological standards that more accurately reflect current experiences. Ten additional papers identified through reference list screening of included studies were incorporated to provide background context, bringing the total to 48 papers. Our search revealed no publications directly comparing patient perspectives of AI-integrated telemedicine in dermatology, highlighting a gap in current literature.

Table . Inclusion study characteristics.

Reference	Study aim	Setting	Sample, n	Study type
Abeck et al [12]	To investigate the impact of teledermatology on patient care by characterizing consultations on a direct-to-consumer telemedicine store-and-forward platform.	Retrospective data obtained from Wellster Healthtech Group in Germany.	n=1999 (retrospective analysis); n=166 (8.3%) (follow-up survey)	Retrospective Analysis and Survey
Asabor et al [13]	To examine the experience of patients and physicians with teledermatology during the COVID-19 pandemic.	Patients seen via Epic MyChart synchronous video visits.	n=548	Survey
Balakrishnan et al [14]	To explore patient satisfaction with teledermatology and 2 distinct teledermatology models.	Patients seen at the Atlanta Veterans Affairs Medical Center.	n=100	Survey
Choi et al [15]	To understand patient perceptions toward teledermatology.	Patients or their caregivers at an academic tertiary dermatologic center in Singapore.	n=913 (survey); n=26 (2.9%; in-depth interviews)	Survey and Semistructured Interview
Chow et al [16]	To explore patients' perceptions of a teledermatology service linking public primary care clinics to a national specialist dermatology clinic.	Five separate dermatology clinics in Singapore.	n=21; patients aged 22 - 72 years; 14 (65%) male; diagnoses: 11 (52%) rashes, 4 (19%) pigmented lesions, 3 (14%) itching, and 2 (10%) dry skin	Qualitative Interview
DeVries et al [17]	To assess perceptions and experiences with teledermatology visits in the context of the COVID-19 pandemic.	Patients of a South Dakota dermatology practice.	Not specified	Survey
Ford et al [18]	To evaluate the impact of a web-based Collaborative Connected Health model compared to in-person care on access to specialty care for psoriasis management.	Patients from outpatient clinics and general adult populations in California and Colorado.	n=300; 151 (50.3%) male; mean age: 49 years; 190 (63.2%) White; 101 Hispanic or Latino (33.8%); 13 (4.4%) uninsured	Randomized Controlled Trial
Frühauf et al [19]	To explore patient satisfaction with video consultations for inflammatory skin conditions in a dermatology outpatient setting.	A teaching hospital in Wales, United Kingdom, has an outpatient dermatology clinic.	n=48; 35 (72%) female; age range 13 - 80 years	Survey
Ghani et al [20]	To identify demographic and behavioral factors associated with patient interest in using teledermatology.	Data from the Health Information National Trends Survey 4, cycle 4 of the National Cancer Institute.	n=3677; 1338 (36.4%) male; age 50 - 64 (31.8%); 1419 (38.6%) college or higher education; 1894 (51.5%) non-Hispanic White; 963 (26.2%) income>US $75,000	Survey
Gnanappiragasam et al [3]	To assess patient satisfaction and preferences between face-to-face and remote (telephone or video) consultations in dermatology settings.	Two dermatology centers in the United Kingdom.	n=156; 78 (50%) female; mean age: 53.3 years; divided into new and follow-up groups	Survey
Goessinger et al [21]	To investigate the perspectives of patients and dermatologists after skin cancer screening by human, artificial, and augmented intelligence.	The University Hospital Basel, Switzerland.	n=205; mean age 54.8, SD 13.6 years; 109 (53%) male	Survey

Reference	Study aim	Setting	Sample, n	Study type
Hadjieconomou et al [2]	To explore patient satisfaction with video consultation within a dermatology outpatient clinic setting for preselected inflammatory skin disorders.	Dermatology outpatient clinic in Wales, United Kingdom.	n=48; 35 (72%) female; age range 13 - 80+ years; 4 (8.5%) aged >65 years	Survey
Handa et al [22]	To analyze patient and physician experiences and acceptability of teledermatology over a 6-month period.	A tertiary care center in North India.	n=5229; mean age 33.60, SD 16.99 years; 2714 (51.9%) male	Survey
Horsham et al [23]	To investigate the factors that determine consumers' comfort and willingness to share 3D total-body images for research, AI[a] development, clinical, and teaching scenarios.	Online video-based consumer forum for consumers of 3D total-body imaging studies at the UQ[b] Dermatology Research Center.	n=39	Survey
Hsueh et al [24]	To assess patient satisfaction with a store-and-forward teledermatology.	27 Veterans Integrated Service Network; 20 clinics in Alaska, Idaho, Oregon, and Washington.	Face-to-face: n=196; 190 (97%) male; mean age 71 years; Teledermatology care: n=504; 464 (92%) male, mean age 65 years	Survey
Hwang et al [4]	To review patient satisfaction with the use of teledermatology since the COVID-19 pandemic.	Not applicable	32 studies: 13 randomized controlled trials, 14 narrative reviews, 5 systematic reviews	Narrative Review
Jutzi et al [25]	To investigate the hopes and fears of patients with and without a history of melanoma toward the use of AI in skin lesion diagnostics.	Web-based questionnaire using LimeSurvey sent to university hospitals in Germany.	n=298; 225 (75.5%) female; 123 (41.3%) aged 46 - 60 years; 121 (40.6%) with a university degree	Survey
Kawsar et al [26]	To explore patients' perspectives on the use of AI as part of their skin cancer management pathway.	A teledermatology skin cancer clinic at Chelsea and Westminster Hospital, London, United Kingdom.	n=268; 154 (57.5%) female; aged 18 - 93 years. Skin type: 218 (81.3%) Fitzpatrick type I-II	Randomized Controlled Trial and Survey
Kohn et al [27]	To evaluate the acceptance of synchronous telehealth for pediatric dermatology.	Children's Hospital Colorado Pediatric Dermatology.	n=125; mean age 9.2 years; 57 (45.5%) male; 48 (38.5%) new patient	Survey
Lim et al [28]	To obtain opinions of patients on the use of AI in a dermatology setting, when aiding the diagnosis of skin cancers.	Dermatology outpatient skin cancer clinics in 2 United Kingdom hospitals.	n=603; 314 (52%) female; age range: 18 - 100 years; 452 (75%) new referrals; 555 (92%) concerned about skin cancer	Survey
Lowe et al [29]	To evaluate the clinician and patient/parental perspective of a pediatric dermatology clinic via voice calls and emailed images in comparison to traditional face-to-face clinics.	United Kingdom single-center cohort of pediatric dermatology patients managed during the COVID-19 pandemic.	n=116; mean age 8.47 years; 28 (24%) new patients; 87 (75%) cases of inflammatory dermatoses	Survey
Ly et al [30]	To understand individuals' perceptions of sharing their images for AI.	Adult United States respondents via Amazon Mechanical Turk.	n=1010; mean age 36.5 years; 566 (56%) male; 717 (71%) White; 851 (84.3%) employed	Survey
Maul et al [31]	To investigate the acceptance of and satisfaction with telemedicine.	One secondary and 2 tertiary referral centers for dermatology in Switzerland.	n=512; 273 (53.3%) male, mean age 49.5 years	Survey

Reference	Study aim	Setting	Sample, n	Study type
Moore et al [32]	To evaluate patient satisfaction with university medical center's video-based teledermatology service.	Penn State's Dermatology Department	n=171; 118 (69%) female, 154 (90%) non-Hispanic	Survey
Munoz et al [33]	To investigate and compare patient satisfaction with recorded video counseling vs traditional, in-office counseling.	Not specified	n=16; video counseling: n=11 (68.8%); face-to-face counseling: n=5 (31.3%)	Survey
Naik [34]	To gain a global perspective on the experiences of patients and health care staff who adapted to teledermatology during the COVID-19 era.	Recruitment through social media and WhatsApp groups.	n=653	Survey
Nelson et al [1]	To explore how patients conceptualize AI and perceive the use of AI for skin cancer screening.	Brigham and Women's Hospital and the melanoma clinics at the Dana-Farber Cancer Institute.	n=39; mean age: 53.3 years; 21 (54%) female; 37 (94%) non-Hispanic White; 16 (42%) graduate or professional degree	Qualitative Interview
Qun Oh et al [35]	To examine patients' perceptions of teledermatology and identify barriers to its adoption.	Outpatient dermatology clinic at a tertiary academic medical hospital in Singapore.	n=997; 508 (51%) female; 489 (49%) aged ≥60 years)	Survey
Pathoulas et al [36]	To compare patient satisfaction between telemedicine visits and in-office visits in a specialty hair loss clinic.	Patients who received either an in-office or telemedicine hair loss new patient consultation by a single provider.	n=40; 29 in-office (72.5%), 11 (27.5%) telemedicine	Survey
Ramjee et al [11]	To assess patient satisfaction with telephone consultations compared to face-to-face consultations in secondary-care dermatology during the COVID-19 era.	A single dermatology center in London, United Kingdom.	n=74; 43 (58.1%) female; median age of 52 years	Survey
Richey et al [37]	To evaluate patients' perspectives and preferences regarding teledermatology for cosmetic acne scar treatment.	Patients at FORMEL Skin in Berlin, Germany.	n=842	Survey
Ruggiero et al [38]	To assess how patients with acne subjectively experienced teledermatology visits.	The Acne Care Center, Dermatology Unit, University of Naples Federico II, Italy.	n=52; 28 (53.9%) female; mean age: 22.5 years.	Survey
Sangers et al [39]	To explore the perceived barriers and facilitators to using mHealth[c] AI apps for skin cancer screening.	The Netherlands	n=27; median age 25 years; 18 (68%) female; 11 (41%) had previous experience with mHealth apps; 4 (15%) had a history of skin cancer	Survey
Stratton et al [40]	To assess patient preferences regarding the use of postprocedural photographs compared with in-person follow-up.	The University of Alabama at Birmingham Department of Dermatology.	n=150; 89 (59.5%) male	Survey
van Erkel et al [41]	To evaluate the perceived quality of follow-up telephone consultations of multiple medical disciplines during the COVID-19 pandemic.	Large university hospital in the Netherlands.	n=82; 44 (54%) female, mean age: 59.1 years	Semi-structured Interview

Reference	Study aim	Setting	Sample, n	Study type
Wortman et al [42]	To evaluate the pandemic's implications on patients with psoriasis, focusing on access to information, consultation methods, patient satisfaction, disease control assessment, and treatment management.	Multicenter survey from 4 Dutch hospitals during the second wave of the pandemic.	n=551; 309 (56%) male, median age: 59 years old, median disease duration: 25 years old	Survey
Wu et al [43]	To gather opinions from a diverse dermatology patient population on AI use in dermatology and establish a specific accuracy at which patients would be comfortable receiving a diagnosis solely from an AI tool.	Adult patients who visited the University of Texas Southwestern Medical Center Dermatology.	n=141; 73 (52%) male; mean age: 55.3; 79 (56%) non-Hispanic white; 55 (39%) household income US $50,000–US $99,999	Survey
Yadav et al [44]	To assess patient perception and satisfaction with a smartphone-based teledermatology service initiated during the COVID-19 pandemic.	The Department of Dermatology and Venereology, All India Institute of Medical Sciences (AIIMS), New Delhi.	n=201; 109 (54.2%) male, mean age 38.4 (SD 15.7) years	Survey

[a]AI: artificial intelligence.

[b]UQ: University of Queensland.

[c]mHealth: mobile health.

Three reviewers independently extracted predefined attributes from each paper. Disagreements were resolved through discussion until a consensus was reached. Literature primarily focusing on technical aspects of AI or telemedicine implementation, without substantial discussion of patient perspectives, was deemed outside the scope of this review.

While our review applied structured screening and thematic synthesis similar to a scoping review, we selected a narrative review approach to enable conceptual interpretation of patient perspectives across diverse study types. This approach allowed us to synthesize findings not only by outcome themes but also by behavioral drivers and contextual patterns. The narrative format also guided our inclusion criteria, enabling us to incorporate both quantitative and qualitative studies that offered insight into patient perceptions, even when methods or outcome measures were heterogeneous. This review was not preregistered, and no formal checklist such as PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) was used, although we followed structured screening procedures to enhance methodological transparency.

Ethical Considerations

This paper is a narrative review and does not involve primary data collection with human or animal participants. Therefore, institutional review board approval and informed consent were not required. The study adhered to the principles of the Declaration of Helsinki and followed JMIR Publications' ethical guidelines for secondary research and literature reviews.

Results

Overview

The literature revealed several distinct themes regarding patient perspectives on AI and telemedicine in dermatology, also known as "teledermatology." The following sections examine these themes for each technology, including patient willingness to use, perceived benefits and risks, barriers to implementation, and conditions necessary for successful integration into dermatology. Each section compares perspectives between AI and teledermatology while highlighting how these viewpoints might inform future implementation strategies. Tables 2 and 3, located after the Results section, summarize key themes from patient perspectives on AI and teledermatology, respectively. The column categories of these tables were adapted from Kalkman et al [45]. Figure 1 presents a conceptual framework that synthesizes these findings to guide integrated implementation approaches.

Table . Patient perspectives on artificial intelligence in dermatology.

Perceived benefits	Perceived risks	Barriers to using	Factors affecting willingness to use	Conditions for acceptance
Increased diagnostic speed [1] and accuracy [21] due to the ability of AI to learn, evolve, and draw on larger data and experience than humans [1,28]	Lack of verbal and nonverbal communication [1]	Inaccurate or limited training sets [1]	Familiarity with AI is significantly associated with a positive view (OR: 17.8; *P*<.01) [43]	AI must be (mean 12.9%, SD 8.1%) more accurate than dermatologists [43]
Enhanced health care access [1]	Increased patient anxiety [1]	Age-related differences in familiarity [43]	Age 40 - 59 years associated with decreased AI familiarity (OR: 0.21, *P*<.01) [43]	Integration with human oversight [1,28]
Potential for earlier detection of skin cancer and lifesaving outcomes [1]	Loss of human interaction and human emotion [1,25]	Limited explanation of AI decisions [1]	Higher education levels are associated with increased willingness to share images with AI [30]	Clear privacy policies and safeguards [23]
Promotes patient engagement in self-examination [1]	Privacy concerns [1,39]	Concern about AI's inability to provide emotional support [1]	97% (289/298) of respondents with a previous history of melanoma support AI use in medicine compared to 91% (271/298) of patients without melanoma (*P*=.03) [25]	System validation by medical professionals [1]
Reduced health care costs [1]	Patient loss to follow-up [1]	Anxiety about receiving a diagnosis without human support [1]	Comfort with technology and social media sharing [30]	Clear guidelines for image control and use in AI [23]
More convenient, consistent, and objective diagnosis [1]	Nefarious use of AI [1,25]	Potential for false positives and false negatives [1]	Trust in a developing institution affects willingness [23].	Need for demonstrated effectiveness and openness of AI use [28]
Unburdening of the health care system [1]	Human deskilling [1,25]	Lack of in-person physical examination [1]	No association between age and acceptance of an AI-only diagnosis, age and preferred AI involvement by diagnosis severity, or prior skin cancer diagnosis and reluctance to use AI for diagnosis [28].	Assurance that AI will not replace discussion with a human dermatologist [28]
Physicians can learn from AI-based[a] systems and direct comparison may motivate specialists to continue to improve performance [25]	Potential misdiagnosis or inaccuracy [28,39]	Operator dependence [1]	Patients felt a greater sense of safety with AI when it worked in tandem with a dermatologist rather than independently [21].	Endorsement from health care providers and government regulating bodies [39]
Improved triage efficiency [1]	Inability to answer follow-up questions [1,25]	Inability to assess treatment options [1]	Patients did not believe AI could answer follow-up questions, discuss treatment options, educate, or reassure [1]	Usable by all ages [39]
Reduced patient anxiety [1]	Lack of context in AI decisions [1]	Inability to educate or reassure patients [1]	More trust if AI applications are set up by dermatologists rather than companies [1].	Low cost of use [39]
Acts as a second opinion to refer to a dermatologist [1,28]	Concerns about image control and secondary use [23]	Lower agreement scores for AI guiding general practitioners in Fitzpatrick IV-VI (44.6/100) vs Fitzpatrick I-II (74.8 - 81.4/100) [26]	—[b]	—
Can perform skin cancer screening from home, monitor skin lesions over time, and integrate with existing skin cancer care [39]	Data privacy risks for sensitive biometric data [23]	Limited knowledge about the use and functionality of AI [39]	—	—

Perceived benefits	Perceived risks	Barriers to using	Factors affecting willingness to use	Conditions for acceptance
More reliable and less subjective diagnoses [25]	—	Lack of integration into the health system and therefore perceived lack of value [39]	—	—
—	—	Lack of reliability of AI app developers for skin cancer screening [39]	—	—

[a]AI: artificial intelligence.

[b]not applicable.

Figure 1. A patient-informed framework for artificial intelligence and telemedicine integration in dermatology. AI: artificial intelligence.

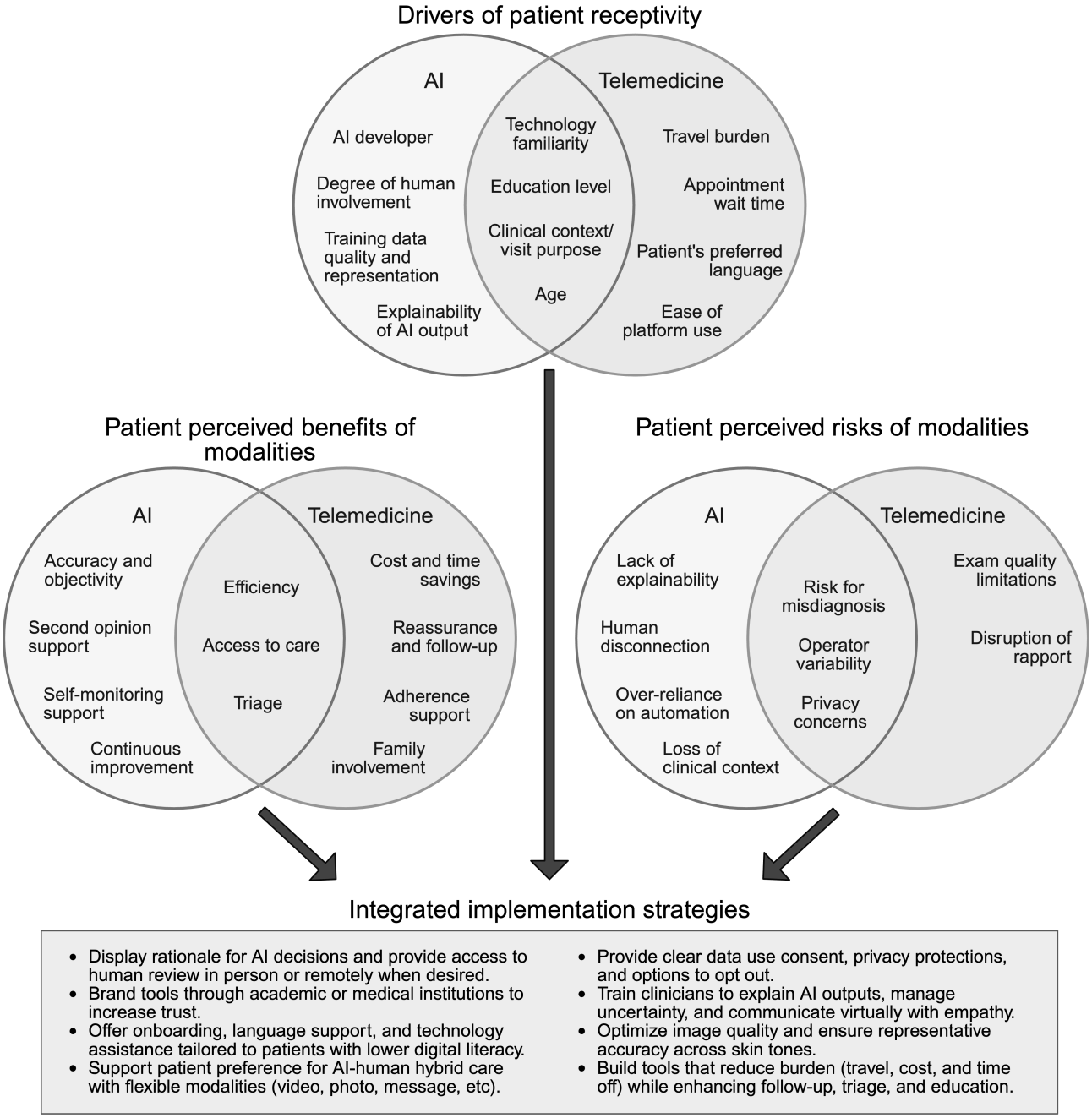


Table . Patient perspectives on telemedicine in dermatology.

Perceived benefits	Perceived risks	Barriers to using	Factors affecting willingness to use	Conditions for acceptance
Reduced wait times and increased efficiency [3,4,11,13,15,16,18,24], reduced travel needs [2,3,11-13,18,27,29], reduced work and school absence [2,3,13,18,29], and savings on parking costs [3,18]	Concerns over the quality of clinical examination [3,15,29,34,44]	Technology limitations, poor internet connectivity, and digital health literacy challenges [3,32,34,35]	Distance from clinic (preference increased with greater distance; *P*=.04) [40]	Careful selection of patients who are better able to navigate technology for telemedicine appointments [3,4]
More efficient triage for acne and skin cancer [4]	A doctor may miss important details [11,40,44]	Less social and natural interaction, reduced ability for the clinician to feel skin pathologies [11]	Rural patients rated their experiences higher than urban counterparts, suggesting satisfaction reflects regional availability of care [5]	Clear preappointment instructions [4]
Improved medication compliance and treatment adherence [4]	Technical difficulties affecting care [3,16,24,34]	Need for assistance taking photos [40]	Previous telemedicine experience increased acceptance [5]	Adequate follow-up care systems [24]
No need to find childcare [27,34]	Privacy concerns [16,34,40]	Sensitivity of health condition [3]	Trusting web-based health information and previous experience sharing medical data on an app or on the web [18,20]	Digital divide factors addressed to ensure equitable access across diverse populations [20]
Improved access to care [2,24]	Decreased quality of care [11,35]	Variable digital literacy levels [29]	Patients preferred chronic conditions over initial consultations [18]	Educational materials provided [24]
Better family involvement [2]	May still require an in-person visit [40]	Cost issues for some patients [34]	About one-third of patients with alopecia maintained a strong preference for in-person evaluation [36]	High-quality image requirements and technical quality of telemedicine video [3,4]
Ability to send concerns anytime [40]	Inadequate follow-up care [24]	Difficulty with photos of hard-to-reach areas for photo-based teledermatology clinical evaluation [18]	Patients with acne preferred teledermatology, whereas patients with possible malignant lesions preferred in-person visits [17]	Trust-building measures [20]
Quicker reassurance and follow-up [40]	Inability to perform labs and procedures [27]	The platform is not user-friendly enough for a mobile interface [18]	Older age increased the likelihood of agreeing to use telemedicine again in some studies [14], but decreased it in others [17]	—[a]
Reduced patient anxiety [2]	Lack of personal element [3,29,44]	Language barriers [3,18]	No significant differences in patient satisfaction and comfort with telemedicine use based on age or visit (*P*=.79; *P*=.90) [36]	—
70.3% (105/150) believed it would improve care; 27.6% (41/150) believed no change in care [40]	Substandard physical examinations [3,4,11,27,29,34]	Privacy concerns specifically about sending personal information [40]	Telemedicine became preferable with an in-person wait time of 6.89 months [36]	—
Useful for monitoring systemic therapies [4]	Limited visual cues and body language [11,29]	Technical difficulties were a common reason for preferring face-to-face consultations [3]	53% (44/83) would consider remote over face-to-face if appointment expedited [3]	—
Valuable for various conditions: acne, atopic dermatitis, and psoriasis [4,38]	Reduced quality of patient-clinician communication [29] and a lower chance of discussing other skin concerns [3]	Patients who had to travel longer distances were not significantly more likely to think that teledermatology is more convenient than face-to-face appointments [14]	The quality of video-based exams, images [4], and audio [2]	—

Perceived benefits	Perceived risks	Barriers to using	Factors affecting willingness to use	Conditions for acceptance
Video allows family to view consultation multiple times and on their own time [33]	Concerns about improper treatment recommendations due to lower accuracy during telemedicine interaction [24]	—	Strong emotional support and rapport with physicians [18]	—
Less exposure to infection risk [2,34]	Inability to relay parental anxiety about pediatric care as effectively via telemedicine vs face-to-face [29]	—	Older patients were more likely to always prefer an in-person wound check compared to younger patients (P<.01) [40]	—
Increased time to spend with family [34]	Technology literacy limitations [3,34]	—	Age was not a factor in willingness to use [22]	—
—	2.1% (n=3/150) of participants perceived a negative impact on care [40]	—	Experiencing technical difficulties during a previous telemedicine encounter [17]	—
—	—	—	Previous exposure to video conferencing and higher education levels [35]	—
—	—	—	Some patients preferred phone consultations for discussing sensitive topics to avoid eye contact [41]	—
—	—	—	47% (443/942) more willing to use teledermatology during the pandemic vs 26% (245/942) before the pandemic [15]	—

[a]Not available.

Willingness to Use AI Versus Telemedicine in Dermatology

The literature shows that patients exhibit hesitancy toward AI-based diagnoses without dermatologist interventions. In a survey conducted by Wu et al [43], if a dermatologist and an AI model made different diagnoses, the majority of patients (119/141, 84.4%) favored the dermatologist's diagnosis. They also found that about 14.9% (21/141) of patients expressed "complete unwillingness" to be evaluated by AI alone [43]. In contrast, Lim et al [28] found that patients would be happy for their general practitioner to use AI to make dermatologist referral decisions (235/603, 39%).

In suspected skin cancer cases, patient trust in human expertise remained high, with a significant majority (524/603, 86.9%) of patients strongly preferring a dermatologist's diagnosis, and only 12.1% (73/603) willing to accept a diagnosis made solely by AI [28]. Despite hesitancy toward standalone AI diagnoses, the Wu et al [43] survey data revealed that most patients (96/141, 68.1%) preferred dermatologists to use an AI model that could provide differential diagnoses based on a photograph at the point of care rather than working alone. This pattern suggests patients prefer AI as a decision-support tool that enhances rather than replaces clinical judgment.

Patients' acceptance of dermatology-based AI tools showed overall mild concerns for data sharing privacy. While Lim et al [28] found that 87% (508/584) of patients were willing to share their patient images for AI training and to help other patients, Ly et al [30] determined that patient comfort levels declined when more facial or sensitive areas were in question. For example, 81% (820/1010) of patients were comfortable with sharing images of their hands, 70% (710/1010) with images of their face, 58% (326/563) with images of male genitals, and 47% (209/447) with female genitals [30].

The willingness to share sensitive or identifiable images was also dependent on the AI tool's development, and Horsham et al's study [23] concluded that patients were more willing to share pictures with university-developed tools than those that were industry-developed. The preference for university-developed AI tools over commercially-developed ones likely reflects broader perceptions of institutional trust. Patients may associate academic institutions with stricter data privacy protections, ethical oversight, and transparency in tool development, whereas commercial tools may raise concerns about profit motives and potential misuse of sensitive health information. These perceptions reiterate the importance of transparent development processes, ethical governance, and academic partnerships in cultivating public trust in AI technologies.

Despite the rising popularity of teledermatology appointments, patient preferences for consultation methods remain mixed (Table 3). Balakrishnan et al [14] identified that older age was associated with an increased likelihood of using telemedicine for follow-up appointments, whereas studies by DeVries et al

[17] and Choi et al [15] found a negative association between age and willingness to use teledermatology services for follow-up appointments.

In studies with patients who experienced both face-to-face appointments and teledermatology, preferences widely varied. In Gnanappiragasam et al's study [3] of patients with a mean age of 53.3 years, 61% (97/156) preferred face-to-face for future consultations, while 39% (60/156) did not have a preference for appointment modality. Meanwhile, Hsueh et al [24] reported that 66% (332/503) of veterans, with a 92% (464/503) male population and a mean age of 71 (SD 17) years, preferred teledermatology over face-to-face. Finally, Hadjieconomou [2] had a 72% (34/48) female demographic, with 91.5% (44/48) younger than 65 years and an 8% (4/48) preference for face-to-face visits.

These contrasting findings related to participant age likely reflect context-specific factors. For instance, the high telemedicine acceptance among older male veterans in Hsueh et al's study [24] may be shaped by functional limitations (such as mobility impairments), structured support from the Department of Veterans Affairs system, or previous exposure to digital tools in service settings. In contrast, Hadjieconomou's [2] younger, predominantly female population may have expressed reluctance due to concerns about privacy, self-image on video, or decreased rapport during teledermatology consultations. These findings suggest that age, gender, previous technology exposure, and health status may influence teledermatology acceptance through competing functional and technological factors.

Regarding positive feelings toward telemedicine, the prospective study conducted by Lowe et al [29] reported that patients (98%, 41/42) with telemedicine consultations felt their concerns were addressed during consultations. Ford et al's [18] randomized controlled trial demonstrated that 70% (210/300) believed telemedicine improved care, 27% (82/300) reported no change, and 2% (6/300) perceived a negative impact.

Adding to these findings about patient preferences, Frühauf et al [19] demonstrated strong patient acceptance of teledermatology among patients with psoriasis, with 90% (n=9/10) reporting they felt "in good hands" with remote care while experiencing a more flexible lifestyle. The same study found that 80% (n=8/10) of patients considered teledermatology a viable alternative to in-person consultations, suggesting high levels of patient confidence in remote care delivery [19]. Finally, a small subset of patients in a Dutch study who used chat or email consultations graded their experience as a 9 out of 10 in satisfaction [42].

However, underlying these positive ratings, Table 3 shows that a few studies reported that patients expressed concerns about the lack of patient-physician connection during initial consultations [3,29,44]. This concern appears to be related to patients' previous experience with technology, as Qun Oh et al [35] found that patients were more likely to decline telemedicine if they had minimal exposure to video conferencing. This technology-related hesitation highlights the importance of gradually exposing patients to telemedicine platforms to build familiarity and comfort with remote care delivery.

A final interesting observation was that patient preferences for telemedicine varied significantly based on their specific dermatological conditions and needs. Handa et al [22] reported that the highest levels of satisfaction (3419/5229, 65.4%) with telemedicine were seen in patients with infectious dermatologic manifestations. However, for chronic disease management, Ford et al [18] found that 65% (195/300) of patients surveyed preferred in-person follow-ups.

Furthermore, findings from a survey by DeVries et al [17] demonstrate that patients with acne had a strong preference for teledermatologic visits, whereas those with possible malignant lesions strongly preferred an in-person visit. Another study comparing video counseling to in-office counseling for acne isotretinoin initiation found no significant difference in patient satisfaction scores across multiple domains, including comfort starting isotretinoin and concerns about side effects [33]. Ruggiero et al [38] reported specific aspects of teledermatology that patients with acne were highly satisfied with, including dermatologist attention (48/52, 92%), quality of time spent with the dermatologist (45/52, 87%), and the treatment received (37/52, 71%). These variations in willingness to use telemedicine reflect patients' risk assessment preferences, with higher-stakes conditions driving demand for direct physician contact.

Collectively, patient preferences for both AI and teledermatology are influenced by factors such as demographic characteristics, previous technology experience, the level of clinician involvement, and specific dermatological needs. Understanding these preference patterns helps design patient-centered implementation strategies that maximize patient acceptance and engagement.

Perceived Benefits of AI Versus Telemedicine in Dermatology

Many facets of patient care could be impacted by the use of AI in dermatology, and patients' motivations for its implementation are optimistic. As shown in Table 2, a qualitative study by Nelson et al [1] found that some of the primary patient values of AI relate to its potential for quicker diagnoses (29/48, 60%), greater ease of health care access (29/48, 60%), and increased triage efficiency (14/48, 29%) [1]. Approximately 35% (17/48) also associated AI with reduced health care costs. However, this survey focused on using AI as a skin cancer screening tool, and a majority (32/48, 67%) of participants had a history of melanoma or other skin cancer. Given this context, future research stratifying patient responses by disease history would be valuable.

Additional research by Sangers et al [39] provided valuable insights into the practical benefits that patients associate with using AI, specifically for skin cancer screening. These included the ability to perform skin cancer screenings from home and monitor lesions over time, giving patients a better sense of involvement in their dermatological health care. As provided in Table 3, patients surveyed by Goessinger et al [21] following AI-assisted skin cancer screening reinforced these positive perspectives, believing that AI enhances diagnostic performance (195/205, 95.5%). These patient perceptions of the potential

benefits of AI reflect some of AI's greatest strengths as it integrates into dermatological health care.

Meanwhile, telemedicine addresses several logistical challenges patients face when accessing dermatological care. Efficiency was a primary benefit noted by patients, and Abeck et al [12] reported that the most frequent reason for using teledermatology was shorter waiting times for appointments (103/166, 62%). This finding was reinforced by Pathoulas et al [36], who found that patients preferred a telemedicine visit with a 2 - to 3-week wait time over an in-office visit with a wait longer than an average of 6.89 months.

Table 3 highlights another significant advantage of telemedicine cited by patients across studies—its reduction in travel and parking time and costs [2,3,11-13,18,27,29], which is particularly important for patients living in remote areas to improve access to care (Table 3). Interestingly, 3 of these studies evaluated patient perceptions during the COVID-19 pandemic [11,13,29]. It is possible that many positive opinions were driven by limited in-person care options, further supported by Choi et al [15] finding that telemedicine support increased during the pandemic and then decreased after movement restrictions eased. The advantage of reduced travel time is complemented by telemedicine's potential to minimize absences from work or school [2]. Patients also appreciate teledermatology's flexibility and the ability to send dermatological concerns at any time [40]. Uniquely, Hadjieconomou [2] found that 71% (34/48) of patients valued its reduction in the risk of infection exposure, and 55% (26/48) appreciated more feasible family involvement during the telemedicine consultation.

Improved treatment outcomes also emerged as a noteworthy patient-reported benefit of teledermatology. A randomized controlled trial by Ford et al [18] found that telemedicine facilitated better psoriasis management, as patients could submit photos and receive real-time updates to their treatment plans based on disease progression. Similarly, participants from a Swiss questionnaire expressed their positive perceptions toward telemedicine for minor skin problems [31]. A German survey also found that patients believed telemedicine represented a useful and underused screening tool for cosmetic dermatology before physician evaluation [37]. These results demonstrate that, from the patient perspective, teledermatology serves as an effective initial management and screening tool for chronic, minor, and cosmetic dermatological concerns, which may evolve to additional disease contexts as patients build trust with the platform.

Overall, patients report similar perceived benefits in the usage of AI and telemedicine services in dermatologic practice, ranging from improved accessibility to enhanced care outcomes.

Perceived Risks of AI Versus Telemedicine in Dermatology

Patients perceived the greatest risks of integrating AI into dermatologic practice to be its limited communication abilities and inherent constraints as an algorithmic tool. Table 2 highlights that patients' primary concerns about AI center on its limited communication abilities, with 40% (19/48) noting the absence of nonverbal communication as key risks [1]. Loss of social interaction (18/48, 38%) was similarly identified as a risk, as patients doubted AI's ability to respond appropriately to emotional distress [1]. In terms of verbal communication, patients emphasized AI's limited capacity to provide education or answer follow-up questions [1]. As AI advances in its real-time interaction capabilities, this particular concern may gradually diminish.

Table 2 highlights patients' expressed concerns about potential misdiagnoses by AI, including false negatives and positives, limited training datasets, lack of physical examination, and operator error [1]. As mentioned in the "Willingness to Use" section, patients exhibited hesitancy toward AI's use as an independent diagnostic tool [28]. These concerns for AI as a diagnostic tool greatly contrast with patient perceived risks of AI as a screening tool, which primarily focused on AI's lack of empathy [1]. As AI begins playing a larger role in dermatologic care, clearly communicating the intended purpose of the technology—whether for screening or diagnosis—may help alleviate patient hesitancy.

Data security emerged as another potential risk, with patients in the Nelson et al study [1] highlighting concerns about loss of privacy (14/603, 29%) and nefarious use of AI (11/603, 23%). These hesitations are valid and highlight the need for dermatologists to adopt transparent AI tools and proactively communicate their limitations, privacy safeguards, and intended roles in care to build trust and mitigate patient concerns.

In response to teledermatology, patients perceived substandard physical examinations as the greatest risk, and numerous studies reported patient concerns regarding the quality of teledermatology complete skin examinations, especially for those being monitored for skin cancer [3,4,11,27,29,34,41] (Table 3). Respondents to an Alabama survey of 150 patients expressed concerns that doctors might miss critical details in a teledermatology setting (150/235; 63.8%). They also noted that an in-person check might still be needed after teledermatology care (32/235, 13.5%) [40].

Hsueh et al [24] reported similar apprehensions, as shown in Table 3, finding that patients were concerned about improper treatment recommendations due to lower diagnostic accuracy during telemedicine interactions. Qualitative interviews by Chow et al [16] also revealed that camera quality was a key reason why patients were concerned with diagnostic accuracy. Given that some studies describe patient beliefs that AI can improve diagnostic accuracy, reliability, and efficiency [1,25,28], integrating AI with teledermatology could help mitigate concerns about the limitations of remote physical examinations. However, some telemedicine limitations remain beyond AI's scope, as surveys for parents of patients with pediatric conditions identified that several required laboratory tests and procedures could only be performed in person [27].

Similar to AI-related risks, limited personal elements also contribute to patients' concerns about reduced quality of care via virtual telemedicine platforms [3,29] (Table 3). Specifically, patients in a UK survey reported that teledermatology consultations lacked nonverbal cues, which led to worse patient-physician understanding and weakened rapport [11]. Similar sentiments were expressed through qualitative feedback

on telemedicine experiences in a study by Yadav et al [44], where patients noted concern for the lack of personal touch during the consultation. In a pediatric cohort, parents of patients reported that telemedicine was less effective in easing their anxiety compared to face-to-face visits, and 52% (60/116) of surveyed participants expressed significant dissatisfaction with the telephone clinic [29].

Overall, patients share similar concerns about both AI and teledermatology centered around diagnostic accuracy limitations and reduced human connections (Figure 1). Thus, integration of these technologies must focus on preserving existing human interaction. Teledermatology faces additional scrutiny of physical examination quality, while AI elicits concerns about emotional responsiveness and data security.

Barriers to Adoption of AI Versus Telemedicine in Dermatology

Beyond perceived risks, patients identify several practical and trust-related barriers that may limit the successful integration of AI and teledermatology into dermatologic care. Table 2 shows that patients are concerned with poor training datasets for AI and the necessity for clinicians to still interpret AI results to develop effective treatment plans [25]. Patients in qualitative interviews by Sangers et al [39] specifically mentioned that limited knowledge about the use and functionality of AI was a barrier to its integration.

Patient trust in dermatologic AI services may also be undermined by data security risks. As discussed in "Willingness to Use," patients would much rather share sensitive or identifiable images with university-developed AI versus private industry-developed systems [23]. Notably, only 15% (6/39) of these respondents answered that they had a "high" level of knowledge about AI, as opposed to 72% (28/39) who selected "low" or "moderate." Therefore, patient perceptions of AI security and privacy could significantly change with increased familiarity with AI technology and its privacy protections.

Finally, a survey on patient acceptance of AI in skin cancer diagnostic pathways revealed that Fitzpatrick skin type strongly influenced patient agreement scores regarding the use of AI to assist their general practitioner (P=.02) [26]. Patients with darker Fitzpatrick skin types IV–VI reported a median agreement score of 44.60 out of 100, significantly lower than those with Fitzpatrick types I (79.89/100), II (81.39/100), and III (74.77/100) [26]. Therefore, AI training datasets must be representative of all skin types to ensure that AI operates equitably and fosters trust among patients.

Regarding teledermatology, patients mentioned technological limitations as the primary barrier to adoption [3,20,32,34] (Table 3). This barrier contrasts with the main perceived risk relating to substandard physical examinations, suggesting that it may be more difficult to circumvent technological limitations. One survey revealed that patients with lower satisfaction scores were significantly more likely to have experienced technical difficulties or to perceive their teledermatology-based physical examination as unsatisfactory [32]. The literature shows that, for patients with limited digital literacy, nonuser friendly or uninviting teledermatology platforms may exacerbate challenges with teledermatology visits [18,29]. Therefore, clinicians using teledermatology must ensure that digital services are simple to navigate and offer extensive troubleshooting for technology-related problems.

Other logistical barriers emerged for teledermatology use and adoption. Table 3 includes 2 studies noting that patients often struggled to take photos of hard-to-reach areas for teledermatology visit evaluation, occasionally requiring assistance from others [18,40]. This finding expands on previous data that describes patients' hesitancy to share images of certain body parts [30], further indicating that image location plays an important role in patient comfort. In addition, a retrospective survey on telephone consultations found that teledermatology visits were less preferred by patients due to reduced natural social interactions and the clinicians' inability to physically examine the patients' skin [11]. Table 3 also highlights patients' emphasis on the importance of visually seeing the clinician as a care preference [11].

In summary, barriers to both AI and teledermatology adoption stem from certain technological limitations and trust concerns and are important to address before technological convergence or implementation.

Conditions for Using AI Versus Telemedicine in Dermatology

For AI and teledermatology to be successfully integrated into dermatological care, patients have distinct considerations and requirements. Patients emphasize the need for dermatologist oversight for AI system model validation and to ensure that AI would not replace human discussion [1,25,28,43]. Current AI models that use patient images have brought up concerns about the adequacy of existing guidelines and policies around AI [23]. Consequently, patients desire safeguards and transparency of the tools to guarantee clear AI privacy policies, secondary uses of data, and AI's effectiveness [23,28,46]. As shown in Table 2, patient acceptance of AI is heavily contingent on demonstrated superiority, with patients requiring AI to be a mean of 12.9% (SD 8.1%) more accurate than dermatologists before accepting standalone AI evaluations [43]. Most importantly, patients indicated that endorsement from their dermatologist and government regulatory bodies would promote their acceptance of AI use in dermatological care [39].

To improve the adoption of teledermatology services, patients outlined several practical and operational factors surrounding their appointments. For example, patients expressed that preappointment teledermatology educational materials and adequate follow-up care systems are important considerations for use [4,24]. Technical difficulties were frequently cited as a reason for preferring face-to-face consultations [3]. For this reason, patients noted that high-quality images and video should be required for their visit [3,4].

Like the concerns surrounding AI's safeguards and transparency, measures for building trust in telemedicine were important considerations for using telemedicine platforms [20]. Finally, patients emphasized the importance of dermatologists addressing accessibility barriers, including the patient's ability to navigate technology for telemedicine appointments [3,4], to ensure

equitable care across diverse populations. This highlights important implications for health equity as teledermatology continues to expand.

Collectively, these findings suggest that patients may embrace AI and teledermatology only with appropriate safeguards, such as transparency about technological limitations, adequate educational resources, clear privacy policies, and, most importantly, continued dermatologist involvement that preserves the human elements of care. Figure 1 illustrates the interconnected drivers of patient receptivity, perceived benefits, and risks and provides a framework for integrated implementation strategies that address patient concerns while leveraging the complementary strengths of both technologies.

Discussion

Principal Findings

This narrative review of 48 studies revealed that patients exhibit distinct perspectives on AI and telemedicine in dermatology, with both technologies showing complementary strengths that could enhance dermatologic care delivery. While both technologies reduce wait times, they achieve this through different mechanisms—AI uses automated diagnostics and data analysis [1,28], whereas telemedicine minimizes logistical barriers, such as travel and appointment times [2,3,13,18,29]. This distinction is significant because it suggests implementation strategies should consider each technology's unique advantages rather than applying a one-size-fits-all approach.

In addition, the sources of trust differ notably between the technologies. AI trustworthiness depends heavily on professional oversight, as patients strongly prefer AI that includes a dermatologist [28,43], and patients require AI models to be more accurate than dermatologists before they would feel comfortable with AI-only evaluations [43]. In contrast, telemedicine's trustworthiness stems from patients' confidence in its diagnostic capabilities for specific conditions, with studies showing high acceptance rates for certain dermatologic issues such as acne and infectious manifestations [17,22,38], while patients consistently preferred in-person evaluations for potentially malignant lesions [17]. Overall, patient acceptance of both technologies depends on perceptions of convenience, accessibility, and care quality, but concerns about privacy, data security, and remote consultation efficacy can impede implementation.

Our review also demonstrates how behavioral and contextual factors play a pivotal role in shaping patient receptivity to these technologies. Patterns of receptivity to AI and telemedicine appear closely tied to patient self-efficacy and contextual factors such as disease type, previous technology exposure, and institutional trust. For instance, patients managing chronic but nonsevere conditions such as acne or psoriasis often reported high satisfaction with teledermatology [18,19,38], which may reflect their familiarity with self-management practices. In contrast, patients with suspected malignancies or limited technology access expressed greater reluctance [15,17,35], likely reflecting both the high-stakes nature of cancer diagnosis that demands maximum clinical certainty and the digital barriers that prevent confident engagement with remote platforms.

These acceptance patterns are further complicated by the intersection of disease severity with demographic factors such as age and gender. Certain dermatologic conditions are more prevalent within specific demographic groups, such as acne more frequently and severely affects adolescent males [38]. These patterns raise the possibility that differences in technology acceptance based on disease severity may be confounded by underlying demographic factors. However, our review highlights the complex relationship between age and willingness to adopt the implementation of AI and telemedicine into dermatologic practice [14,17,30,36,43]. While younger patients who have more familiarity with technology may be more accepting of AI and telemedicine, they may also exhibit more hesitancy due to data privacy concerns [1,16,23,34,39,40]. This ambiguity indicates that technology acceptance is multifaceted and reflects the interplay between numerous clinical and demographic factors.

As the technological revolution expands, the convergence of AI and telemedicine in health care may become inevitable. These technologies may no longer remain distinct but instead function in tandem to provide a more integrated approach to care. Our results support that this integration should be gradual, and that AI and telemedicine should not replace traditional face-to-face services, but rather complement them [1,17,18,28,43].

While this simultaneous integration may exacerbate shared concerns—such as data privacy, loss of human interaction, and diagnostic accuracy—it also offers avenues to mitigate the risks inherent to each technology. For instance, AI tools could address concerns regarding the quality of teledermatology by standardizing image quality assessment and supporting diagnostic accuracy [8,9], while teledermatology's real-time communication capabilities could mitigate patient concerns about AI's lack of human interaction [1]. This synergy has the potential to enhance diagnostic accuracy, optimize resource allocation, and prioritize a patient-centered approach that maintains both technological efficiency and essential human touchpoints in health care delivery. As shown in Figure 1, successful integration requires addressing the overlapping concerns while capitalizing on each technology's unique strengths through targeted implementation strategies.

It should be noted that, as dermatology care increasingly incorporates AI and telemedicine, attention to digital equity is paramount to prevent the widening of existing disparities. Several studies identified disparities in access, literacy, and comfort across demographic groups, and these findings foreshadow the implementation challenges discussed in the 4 paragraphs below, where the very populations that could benefit most from technological access may face the greatest barriers to adoption.

As our results outline, equity concerns surrounding AI center on algorithmic bias and representation, as evidenced by patients with darker Fitzpatrick skin types showing significantly lower acceptance of AI tools [26]. Meanwhile, equity concerns with teledermatology primarily relate to geographic and digital

literacy disparities. Generally, patients with longer travel distances to the clinic are more accepting of telemedicine services, though for some patients, this preference may reflect travel or financial barriers rather than true choice [14,40]. Patients living in rural areas may particularly benefit from the increased convenience of web-, photo-, and app-based dermatology appointments but are simultaneously at a heightened risk of experiencing digital literacy and technology-related challenges [3,5,32,34,35].

Across both technologies, the introduction of new digital platforms without adequate support may disproportionately disadvantage individuals with limited digital literacy, including older adults, lower-income populations, and patients with lower education levels [14,17,30,35,36,39,43]. Without deliberate attention to inclusivity, the integration of these technologies risks reinforcing, rather than reducing, gaps in care access and quality. To address these equity concerns and increase patient acceptance across both technologies, implementation strategies must prioritize representative AI training datasets, intuitive telemedicine interfaces, transparent communication about both technologies' capabilities and limitations, and patient-centered education.

Our review highlights that patients view successful integration of both technologies as requiring dermatologist oversight and the preservation of meaningful patient-provider relationships [3,43]. Therefore, offering ongoing resources and support throughout the integration process may help address patient concerns and maximize comfort with the platforms. Once these technologies are implemented, clinicians can begin by offering patients resources on how to use the AI or telemedicine service during in-person consultations, explaining how the technology works and emphasizing its role as a complementary tool that augments traditional care models [28,47]. Practices could also connect patients with follow-up resources, such as public libraries or help desks, to help train them on basic digital skills for health-oriented patient technology and empower them to engage fully with the services.

Beyond technical training, building trust necessitates transparent communication about security protocols and the protection of personal health information [23,46]. As the literature explains, endorsement from dermatologists alongside these trust-building measures may enhance patient acceptance of these technologies [39]. Ultimately, the success of these technologies in dermatologic care depends on thoughtful implementation that balances technological advancement with patient-centered care delivery. By ensuring that integration strategies align with patient concerns and expectations, these innovations can maximize their potential to improve access, efficiency, and quality in dermatologic care. The integrated framework shown in Figure 1 offers a roadmap for achieving this patient-centered technological convergence.

Although previous research has examined AI and telemedicine as separate modalities, little is known about how patients perceive their integration. Our search identified only 2 studies that addressed both technologies, and neither evaluated patient perspectives on their combined use. This likely reflects the novelty of such tools and the limited availability of integrated, patient-facing deployments during the review period. As these technologies evolve, future research focusing on patient trust and comfort with AI-augmented teledermatology will be important for guiding patient-centered implementation.

Conclusion

In conclusion, patient perspectives surrounding AI and telemedicine in dermatology provide central considerations for clinical implementation. While patients value the benefits of improved access to care and reduced wait times, they continue to have concerns about data privacy, diagnostic accuracy, and maintaining meaningful doctor-patient relationships. These perspectives are especially important for informing health care accessibility in dermatology. As AI and telemedicine potentially converge in dermatologic care, balancing technological advancements with patient-centered care delivery should drive responsible implementation strategies. Research examining how patients experience these modalities together could guide health care systems in harnessing their complementary strengths, while continued investigation will be essential to understand how these technologies can best address patient needs in dermatologic care, both independently and in combination.

Acknowledgments

The study was funded by the Alabama Dermatology Society.

Authors' Contributions

CM contributed to the study inception as well as all aspects of the review, including data collection, manuscript writing, table generation, manuscript editing, and manuscript submission. TDZ and LDS contributed equally to data collection, manuscript writing, and manuscript editing. MA and LT contributed equally to table generation and manuscript editing. LVG contributed to the study inception and manuscript editing, and approval. The authors report no funding or financial disclosures.

Conflicts of Interest

None declared.

References

1. Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. JAMA Dermatol 2020 May 1;156(5):501-512. [doi: 10.1001/jamadermatol.2019.5014] [Medline: 32159733]
2. Hadjieconomou S. BT22: Are patients satisfied with video consultations? A survey in a dermatology outpatient setting. Br J Dermatol 2022 Jul;187(S1):130-131. [doi: 10.1111/bjd.21395]
3. Gnanappiragasam D, Oldham J, Panchal M, Woo WA. Experience and perception of face-to-face vs. remote consultations: a patient survey across two UK dermatology centres. Clin Exp Dermatol 2021 Jun;46(4):736-737. [doi: 10.1111/ced.14528] [Medline: 33259673]
4. Hwang JK, Del Toro NP, Han G, Oh DH, Tejasvi T, Lipner SR. Review of teledermatology: lessons learned from the COVID-19 pandemic. Am J Clin Dermatol 2024 Jan;25(1):5-14. [doi: 10.1007/s40257-023-00826-z] [Medline: 38062339]
5. Coates SJ, Kvedar J, Granstein RD. Teledermatology: from historical perspective to emerging techniques of the modern era: part I: history, rationale, and current practice. J Am Acad Dermatol 2015 Apr;72(4):563-574. [doi: 10.1016/j.jaad.2014.07.061] [Medline: 25773407]
6. Liopyris K, Gregoriou S, Dias J, Stratigos AJ. Artificial intelligence in dermatology: challenges and perspectives. Dermatol Ther (Heidelb) 2022 Dec;12(12):2637-2651. [doi: 10.1007/s13555-022-00833-8] [Medline: 36306100]
7. Muñoz-López C, Ramírez-Cornejo C, Marchetti MA, et al. Performance of a deep neural network in teledermatology: a single-centre prospective diagnostic study. J Eur Acad Dermatol Venereol 2021 Feb;35(2):546-553. [doi: 10.1111/jdv.16979] [Medline: 33037709]
8. Jalaboi R, Winther O, Galimzianova A. Explainable image quality assessments in teledermatological photography. Telemed J E Health 2023 Sep;29(9):1342-1348. [doi: 10.1089/tmj.2022.0405] [Medline: 36735575]
9. Greis C, Maul LV, Hsu C, Djamei V, Schmid-Grendelmeier P, Navarini AA. Unmet digital health service needs in dermatology patients [Article in German]. Hautarzt 2020 Sep 1;71(9):686-690. [doi: 10.1007/s00105-020-04664-6] [Medline: 39226097]
10. 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]
11. Ramjee S, Boyce L, Mohandas P. Telephone consultations in the COVID-19 era versus pre-COVID face-to-face consultations: a survey of dermatology patient perceptions. Arch Dermatol Res 2023 Aug;315(6):1843-1846. [doi: 10.1007/s00403-023-02561-1] [Medline: 36795155]
12. Abeck F, Kött J, Bertlich M, et al. Direct-to-consumer teledermatology in Germany: a retrospective analysis of 1,999 teleconsultations suggests positive impact on patient care. Telemed J E Health 2023 Oct;29(10):1484-1491. [doi: 10.1089/tmj.2022.0472] [Medline: 36862525]
13. Asabor EN, Bunick CG, Cohen JM, Perkins SH. Patient and physician perspectives on teledermatology at an academic dermatology department amid the COVID-19 pandemic. J Am Acad Dermatol 2021 Jan;84(1):158-161. [doi: 10.1016/j.jaad.2020.09.029] [Medline: 32946971]
14. Balakrishnan V, Baranowski MLH, Bartenfeld D, Chen SC. 508 Impact of teledermatology services at the Atlanta VA Medical Center: assessing patient satisfaction. J Invest Dermatol 2018 May;138(5):S87. [doi: 10.1016/j.jid.2018.03.516]
15. Choi ECE, Heng LW, Tan SY, Phan P, Chandran NS. Factors influencing use and perceptions of teledermatology: a mixed-methods study of 942 participants. JAAD Int 2022 Mar;6:97-103. [doi: 10.1016/j.jdin.2021.12.005] [Medline: 35128487]
16. Chow A, Teo SH, Kong JW, et al. Patients' experiences of telemedicine for their skin problems: qualitative study. JMIR Dermatol 2022 Feb 22;5(1):e24956. [doi: 10.2196/24956] [Medline: 37632855]
17. DeVries BL, Brodersen KR, Mechels HN, DeVries JD, Knutson BD. Teledermatology in South Dakota: uses and patient centered improvement in a rural setting. S D Med 2022 Jun;75(6):258-262. [Medline: 36206566]
18. Ford AR, Gibbons CM, Torres J, et al. Access to dermatological care with an innovative online model for psoriasis management: results from a randomized controlled trial. Telemed J E Health 2019 Jul;25(7):619-627. [doi: 10.1089/tmj.2018.0160] [Medline: 30222518]
19. Frühauf J, Schwantzer G, Ambros-Rudolph CM, et al. Pilot study on the acceptance of mobile teledermatology for the home monitoring of high-need patients with psoriasis. Australas J Dermatol 2012 Feb;53(1):41-46. [doi: 10.1111/j.1440-0960.2011.00852.x] [Medline: 22309330]
20. 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]
21. Goessinger EV, Niederfeilner JC, Cerminara S, et al. Patient and dermatologists' perspectives on augmented intelligence for melanoma screening: a prospective study. J Eur Acad Dermatol Venereol 2024 Dec;38(12):2240-2249. [doi: 10.1111/jdv.19905] [Medline: 38411348]
22. Handa S, Mehta H, Bishnoi A, et al. Teledermatology during the COVID-19 pandemic: experience at a tertiary care centre in North India. Dermatol Ther 2021 Jul;34(4):e15022. [doi: 10.1111/dth.15022] [Medline: 34089561]
23. Horsham C, Janda M, Kerr M, Soyer HP, Caffery LJ. Consumer perceptions on privacy and confidentiality in dermatology for 3D total-body imaging. Australas J Dermatol 2023 Feb;64(1):118-121. [doi: 10.1111/ajd.13952] [Medline: 36349396]

24. Hsueh MT, Eastman K, McFarland LV, Raugi GJ, Reiber GE. Teledermatology patient satisfaction in the Pacific Northwest. Telemed J E Health 2012 Jun;18(5):377-381. [doi: 10.1089/tmj.2011.0181] [Medline: 22489931]
25. Jutzi TB, Krieghoff-Henning EI, Holland-Letz T, et al. Artificial intelligence in skin cancer diagnostics: the patients' perspective. Front Med (Lausanne) 2020;7:233. [doi: 10.3389/fmed.2020.00233] [Medline: 32671078]
26. Kawsar A, Hussain K, Kalsi D, Kemos P, Marsden H, Thomas L. Patient perspectives of artificial intelligence as a medical device in a skin cancer pathway. Front Med (Lausanne) 2023;10:1259595. [doi: 10.3389/fmed.2023.1259595] [Medline: 38046409]
27. Kohn LL, Pickett K, Day JA, et al. When is synchronous telehealth acceptable for pediatric dermatology? Pediatr Dermatol 2022 Mar;39(2):236-242. [doi: 10.1111/pde.14919] [Medline: 35178735]
28. Lim K, Neal-Smith G, Mitchell C, Xerri J, Chuanromanee P. Perceptions of the use of artificial intelligence in the diagnosis of skin cancer: an outpatient survey. Clin Exp Dermatol 2022 Mar;47(3):542-546. [doi: 10.1111/ced.14969] [Medline: 34610153]
29. Lowe A, Dawood S, Al-Tayeb A, et al. Evaluating paediatric dermatology telephone clinics during COVID-19 from a dual clinician and patient perspective: a prospective study. Clin Exp Dermatol 2022 Mar;47(3):553-560. [doi: 10.1111/ced.14990] [Medline: 34674296]
30. Ly S, Reyes-Hadsall S, Drake L, et al. Public perceptions, factors, and incentives influencing patient willingness to share clinical images for artificial intelligence-based healthcare tools. Dermatol Ther (Heidelb) 2023 Nov;13(11):2895-2902. [doi: 10.1007/s13555-023-01031-w] [Medline: 37737327]
31. Maul LV, Jahn AS, Pamplona GSP, et al. Acceptance of telemedicine compared to in-person consultation from the providers' and users' perspectives: multicenter, cross-sectional study in dermatology. JMIR Dermatol 2023 Aug 11;6:e45384. [doi: 10.2196/45384] [Medline: 37582265]
32. Moore B, Washington A, Butt M, Silva C, Green B, Helm M. Patient satisfaction of real-time teledermatology: a cross-sectional survey. Int J Dermatol 2022 Feb;61(2):e69-e71. [doi: 10.1111/ijd.15618] [Medline: 33899221]
33. Munoz A, Ahatov R, Williams E. 34543 Isotretinoin counseling amid the COVID-19 pandemic [Abstract]. J Am Acad Dermatol 2022 Sep;87(3):AB74. [doi: 10.1016/j.jaad.2022.06.329]
34. Naik PP. Rise of teledermatology in the COVID-19 era: a pan-world perspective. Digit Health 2022;8:20552076221076671. [doi: 10.1177/20552076221076671] [Medline: 35154805]
35. Qun Oh DA, Yeo YW, Pang SM, Oh CC, Lee HY, Lin Choo KJ. Perceptions of teledermatology in the COVID-19 era: are patients ready for it? SINGAPORE MED J 2023 Apr 27. [doi: 10.4103/singaporemedj.SMJ-2021-290] [Medline: 37171435]
36. Pathoulas JT, Su MT, Flanagan KE, Walker CJ, Wiss IP, Senna MM. 27501 Comparing patient satisfaction between virtual and outpatient evaluation of hair disorders [Abstract]. J Am Acad Dermatol 2021 Sep;85(3):AB149. [doi: 10.1016/j.jaad.2021.06.610]
37. Richey P, Bechstein S, Garibyan L. Teledermatology patient perspectives and preferences regarding acne scar treatment. Arch Dermatol Res 2023 Dec 7;316(1):22. [doi: 10.1007/s00403-023-02752-w] [Medline: 38060006]
38. Ruggiero A, Megna M, Annunziata MC, et al. Teledermatology for acne during COVID-19: high patients' satisfaction in spite of the emergency. J Eur Acad Dermatol Venereol 2020 Nov;34(11):e662-e663. [doi: 10.1111/jdv.16746] [Medline: 32534472]
39. Sangers TE, Wakkee M, Kramer-Noels EC, Nijsten T, Lugtenberg M. Views on mobile health apps for skin cancer screening in the general population: an in-depth qualitative exploration of perceived barriers and facilitators. Br J Dermatol 2021 Nov;185(5):961-969. [doi: 10.1111/bjd.20441] [Medline: 33959945]
40. Stratton MS, Duncan JR, Purnell JC, Phillips CB, Huang CC. Patient preferences and perception of value regarding postprocedural patient-generated photographs. Dermatol Surg 2021 Apr 1;47(4):563-564. [doi: 10.1097/DSS.0000000000002396] [Medline: 32251005]
41. van Erkel FM, Pet MJ, Bossink EH, et al. Experiences of patients and health care professionals on the quality of telephone follow-up care during the COVID-19 pandemic: a large qualitative study in a multidisciplinary academic setting. BMJ Open 2022 Mar 10;12(3):e058361. [doi: 10.1136/bmjopen-2021-058361] [Medline: 35273062]
42. Wortman CD, Godding LTH, Yin Q, et al. Psoriasis healthcare during the COVID-19 pandemic: a survey among psoriasis patients (PsoCovidCare). J Dermatolog Treat 2024 Dec;35(1):2369616. [doi: 10.1080/09546634.2024.2369616] [Medline: 38897615]
43. Wu A, Ngo M, Thomas C. Assessment of patient perceptions of artificial intelligence use in dermatology: a cross-sectional survey. Skin Res Technol 2024 Mar;30(3):e13656. [doi: 10.1111/srt.13656] [Medline: 38481072]
44. Yadav G, Goldberg HR, Barense MD, Bell CM. A cross-sectional survey of population-wide wait times for patients seeking medical vs. cosmetic dermatologic care. PLoS ONE 2016;11(9):e0162767. [doi: 10.1371/journal.pone.0162767] [Medline: 27632206]
45. Kalkman S, van Delden J, Banerjee A, Tyl B, Mostert M, van Thiel G. Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence. J Med Ethics 2022 Jan;48(1):3-13. [doi: 10.1136/medethics-2019-105651] [Medline: 31719155]

46. Abdelhamid M, Gaia J, Sanders GL. Putting the focus back on the patient: how privacy concerns affect personal health information sharing intentions. J Med Internet Res 2017 Sep 13;19(9):e169. [doi: 10.2196/jmir.6877] [Medline: 28903895]
47. Ftouni R, AlJardali B, Hamdanieh M, Ftouni L, Salem N. Challenges of telemedicine during the COVID-19 pandemic: a systematic review. BMC Med Inform Decis Mak 2022 Aug 3;22(1):207. [doi: 10.1186/s12911-022-01952-0] [Medline: 35922817]

Abbreviations

AI: artificial intelligence
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by R Dellavalle; submitted 03.04.25; peer-reviewed by IC Lai, RT Potla; revised version received 05.07.25; accepted 20.07.25; published 16.09.25.

Please cite as:
McRae C, Zhang TD, Seeley LD, Anderson M, Turner L, Graham LV
Patient Perceptions of Artificial Intelligence and Telemedicine in Dermatology: Narrative Review
JMIR Dermatol 2025;8:e75454
URL: https://derma.jmir.org/2025/1/e75454
doi:10.2196/75454

Exploring Nonresponse to Botulinum Toxin in Aesthetics: Narrative Review of Key Trigger Factors and Effective Management Strategies

George Kroumpouzos[1,2], MD, PhD; Fernando Silikovich[3], MD

[1]GK Dermatology PC, 541 Main St, Suite 320, South Weymouth, MA, United States

[2]Department of Dermatology, Warren Alpert Medical School at Brown University, Providence, RI, United States

[3]Concepto 4 Puntos Clinic, Buenos Aires, Argentina

Corresponding Author:
George Kroumpouzos, MD, PhD
GK Dermatology PC, 541 Main St, Suite 320, South Weymouth, MA, United States

Abstract

Background: Nonresponse to botulinum toxin type A (BoNT-A) has been reported in both medical and aesthetic applications. Secondary nonresponse (SNR) occurs when BoNT-A is initially effective before failure commences at a later point. Most reported cases involve SNR in aesthetics. Several aspects of this complication remain elusive or controversial.

Objective: We aimed to address unanswered questions regarding the prevalence and etiology of SNR. Additionally, we investigated the immunogenicity of BoNT-A formulations, mainly focusing on the development of neutralizing antibodies that hinder the toxin's pharmacologic effects. Furthermore, we sought to examine the management strategies for SNR.

Methods: The PubMed and Google Scholar databases were searched from inception for articles on nonresponse to BoNT-A therapy. Articles were evaluated based on their contribution to the field. Expert opinions and panel recommendations regarding management and data gaps were also included in the review.

Results: There are limited data on SNR prevalence in aesthetic applications compared to therapeutic uses. Trigger factors of SNR include improper handling of BoNT-A; incorrect injection practices; and impurities present in the formulation, such as clostridial complexing proteins that may increase immunogenicity. Other contributing factors include infection; patient characteristics; and treatment parameters that encompass an increased frequency of BoNT-A injections (ie, <3 months apart), higher cumulative dosages, elevated treatment dosages, and booster injections (retreatment within 3 weeks of the initial injection). Neutralizing antibodies developed with first-generation formulations, such as onabotulinumtoxinA and abobotulinumtoxinA that contain clostridial proteins, but not with second-generation BoNT-As, such as incobotulinumtoxinA and daxibotulinumtoxinA, which lack these proteins. Among patients who developed SNR after using first-generation BoNT-A for aesthetic purposes, switching to incobotulinumtoxinA therapy did not result in the development of immune responses. Switching to a protein-free BoNT-A formulation such as incobotulinumtoxinA upon development of SNR has been advocated. To effectively manage SNR, it is crucial to minimize the identified trigger factors.

Conclusions: Nonresponse to BoNT-A is gaining importance in aesthetic treatments. Considering the potential for immunogenicity is essential when selecting a BoNT-A formulation. Preventing SNR is crucial, given the lack of solid data on effective treatments.

(JMIR Dermatol 2025;8:e69960) doi:10.2196/69960

KEYWORDS

botulinum toxin; onabotulinumtoxinA; abobotulinumtoxinA; incobotulinumtoxinA; daxibotulinumtoxinA; aesthetic; cosmetic; trigger factor; neutralizing antibody; nonresponse; resistance; immune response; prevention; treatment; management

Introduction

Nonresponse or resistance to botulinum toxin type A (BoNT-A) has become an increasingly significant concern in the field of aesthetics, particularly since younger patients—who are increasingly opting for aesthetic procedures—accumulate greater total toxin doses over their lifetime. Resistance has been noted even with low BoNT-A doses in aesthetic treatments [1]. Primary nonresponse (PNR) to BoNT-A refers to individuals who show an innate insensitivity to the toxin upon initial exposure, without prior treatments or antibody (Ab) development. On the other hand, secondary nonresponse (SNR) occurs when BoNT-A is initially effective before failure commences at a later point. PNR is more commonly encountered in therapeutic applications [2], while most reported cases in aesthetic treatments involve SNR [3].

This review aims to address unanswered questions about the prevalence and etiology of SNR, with a particular focus on the immunogenicity of BoNT-A formulations and the development of neutralizing antibodies (NAbs) that hinder the toxin's pharmacologic effects. We also explore management strategies for SNR.

Methods

A narrative review was completed because a systematic review was not feasible due to the high heterogeneity among the articles on this topic. The PubMed and Google Scholar databases were searched from inception. Key search terms included "botulinum toxin," "nonresponse OR nonresponsiveness OR resistance OR failure," "aesthetic OR cosmetic," "prevention," and "management OR treatment OR intervention." Separate searches were carried out for specific BoNT-A formulations using the following terms: "onabutulinum OR onabotulinumtoxinA" (onaBoNT-A), "abobutulinum OR abobotulinumtoxinA" (aboBoNT-A), "incobotulinum OR incobotulinumtoxinA" (incoBoNT-A), and "daxibotulinumtoxinA" (daxiBoNT-A). Additionally, reference lists of relevant articles were reviewed. Expert opinions and panel recommendations regarding management and data gaps were also included in the review.

Results

Principal Findings

We review the findings of publications relevant to the prevalence of SNR [4-8], etiology of nonresponse to BoNT-A [6,9-18], key trigger factors in SNR [2,6,11,14,15,19-26], BoNT-A formulations composition [4,8,11,23,27-52] and immunogenicity [1,3-6,10,11,15,17,18,28,31,35-38,44,48,49,53-67], insights into mechanisms of SNR [1,3,7,14,15,23,68,69], SNR management [1,3,6,7,14,23,25,26,46,66,70-81], and data gaps and limitations [3-6,10,11,14,15,26,28,40,49,63,82-84].

Prevalence

The prevalence of SNR in therapeutic applications of BoNT-A varies among conditions treated and is often correlated with the toxin dose used. Detection of NAbs correlated to nonresponsiveness in therapeutic applications [4], with its global prevalence estimated at 0.3% - 27.6% [5]. Limited data exist regarding its prevalence in aesthetics, which is partly due to the diverse treatment approaches used and difficulties in quantifying the cosmetic effect [6]. In a recent survey among 673 Korean aesthetic providers, 53.9% reported experiencing BoNT-A resistance. Of those, 59% providers indicated the resistance rate as <1%, and 36% providers reported as approximately 1 - 25% [7]. In the same study, 23.8% of respondents continued using the same product but at a higher dose when they suspected that a patient might be experiencing BoNT-A resistance. Therefore, the prevalence of resistance is likely underreported, as many providers are unaware and may solely increase the BoNT-A dose in subsequent sessions following a partial response [5,8].

Etiology

Primary Nonresponse

PNR can be attributed to genetic variations that affect the toxin's target molecules (neuronal receptors) or to a genetic predisposition to anti-BoNT antibodies formation due to different major histocompatibility complex types [9-11]. Genetic polymorphisms in immune response genes can influence how the body reacts to the toxin and can be involved in immunoresistance [12]. PNR has also been attributed to preexisting BoNT-A antibodies, possibly due to prior immunization against botulism [13,14].

Secondary Nonresponse

SNR to botulinum toxin (BoNT) is believed to be primarily due to the development of NAbs that hinder BoNT's pharmacological effects [15]. This immune response can be influenced by epigenetic changes affecting the expression of genes involved in immune function, including those encoding for proteins interacting with BoNT [9,16]. The overall reactivity of an individual patient's immune system–specifically, the ability of an antigen to stimulate an immune response–can be influenced by exogenous factors, such as environmental allergens. Some researchers consider this relevant, as most reported cases of complete SNR developed after multiple injection cycles [6,17,18]. In the series by Dressler et al [6], complete nonresponse occurred after 3, 5, 10, and 13 injection cycles, with treatment periods ranging from 16 to 65 months. However, more data on specific patient characteristics are needed.

Key Trigger Factors in SNR

Toxin Handling and Injection Practice

Before attributing SNR to NAbs, it is important to consider other causes of nonresponse related to the handling of BoNT-A, such as improper dilution, prolonged storage under refrigeration, and interbatch variation [19-22]. Furthermore, SNR can also occur due to incorrect injection practices, which may involve insufficient dosing, targeting the wrong muscle, or using improper injection technique [19].

Toxin Purity

Impurities present in the BoNT-A formulations, such as clostridial complexing proteins, inactivated toxin, flagellin, and DNA contaminants, are believed to increase immunogenicity related to development of NAbs [23].

Vaccine

COVID-19 vaccination stimulates the immune system and may increase the risk of mounting an immune response against BoNT-A [24].

Patient Characteristics

Genetic differences in the control of immune responses indicate that patients exhibit variable speed and magnitude of immune reactions and patterns of NAb generation [14,25,26]. Some patients may have a specific predisposition to SNR; in one case, complete SNR occurred after just two injection sessions [6].

Treatment Parameters

Multiple treatment parameters affect BoNT-A immunogenicity. Due to it being a potential lifelong treatment, the prevalence of NAbs increases with chronic BoNT-A use [11]. The increased frequency of BoNT-A injections (ie, <3 months apart) is an essential trigger factor [14,15]. Other contributing factors include cumulative dosage, booster injections (retreatment within 3 weeks of the initial injection), high treatment dosage, and a patient's immune responsiveness [2,23]. Notably, off-label aesthetic applications, such as masseter hypertrophy, whole face intradermal lifting, and body contouring require higher doses (ie, >100 international units of onaBoNT-A) and more frequent injections. Their increasing popularity may lead to increased prevalence of SNR and NAbs.

BoNT-A Formulations Composition

All BoNT-A formulations contain the same 150-kDa core neurotoxin derived from the *Clostridium botulinum* Hall A strain [11,27,28]. The 150-kDa core neurotoxin contains a 100-kDa heavy chain and 50-kDa light chain, linked by a disulfide bond. BoNT-A formulations vary in purity, specific bioactivity, complexing proteins, and excipient content (Table 1), all of which can influence their potential to elicit an immune response.

Table . Characteristics and prevalence of NAb[a] development and clinical nonresponsiveness of main first- and second-generation BoNT-A[b] preparations.

Parameter	First-generation BoNT-A[c]		Second generation BoNT-A[d]	
	OnaBoNT-A[e]	AboBoNT-A[f]	IncoBoNT-A[g]	DaxiBoNT-A[h,i]
MW[j] of bacterial protein, kDa [29-31]	~900	~300–500[k]	~150	~150; also, a 5-kDa stabilizing peptide (RTP004)
Accessory proteins present [32-34]	Yes	Yes	No	No
Total protein/vial [35-37]	5 ng/100 U	4.36 ng/500 U	0.6 ng/100 U	—[l]
Total core neurotoxin protein/100 MU[m], ng [33,38]	0.73	0.65	0.44	—
Active neurotoxin protein/100 MU, ng [33,38,39]	0.44	0.44	0.44	0.45
Inactive neurotoxin protein/100 MU, ng[n] [32,33]	0.29	0.21	0	—
Excipients[o] [8,28,34,40]	HSA[p], NaCl[q]	HSA, lactose	HSA, sucrose	RTP004 peptide, L-histidine, L-histidine-HCl monohydrate, polysorbate 20, trehalosedihydrate
Patients with NAbs in pivotal clinical trials, % [41-46]	0.0 - 1.9	0.0 - 3.6	0 - 1.8	0
Patients with NAbs in real-world studies, % [4,47]	1.5 - 7.0	1.7 - 6.0	0.0 - 0.5	—
Reports of clinical nonresponse [6,15,48]	Yes	Yes	No	No
Formulation notes [28,33,39,49-51]	Reduced protein load from original formulation (ie, reduced clostridial protein impurities and inactive BoNT-A)	Contains flagellin with potential adjuvant properties; contains complexing proteins	No complexing proteins; no inactive toxoids; no patients with SNR[r]	No complexing proteins; proprietary peptide claimed to aid in stability and delivery

[a]NAb: neutralizing antibody.

[b]BoNT-A: botulinum toxin type A.

[c]First-generation BoNT-A formulations contain core neurotoxins and accessory clostridial proteins.

[d]Second-generation BoNT-A formulations contain only the therapeutic neurotoxin without accessory proteins or other bacterial substances.

[e]onaBoNT-A: onabotulinumtoxinA.

[f]aboBoNT-A: abobotulinumtoxinA.

[g]incoBoNT-A: incobotulinumtoxinA.

[h]daxiBoNT-A: daxibotulinumtoxinA.

[i]Details on the formulation are not fully disclosed by the manufacturer.

[j]MW: molecular weight.

[k]Formulation is a mixture of species, with 300 and 500 kDa being the most common.

[l]Not available.

[m]MU: mouse unit

[n]Values for inactive neurotoxin are approximate and were estimated by Frevert et al [33], then reported by Kerscher et al [32].

[o]The excipient list is not exhaustive; additional peptides may be included in the diluent of BoNT-A formulations produced outside the United States.

[p]HSA: human serum albumin.

[q]NaCl: sodium chloride.

[r]SNR: secondary nonresponse.

First-generation BoNT-A formulations such as onaBoNT-A and aboBoNT-A contain pharmacologically unnecessary components such as complexing accessory clostridial proteins, inactive neurotoxin, clostridial DNA, and excipients (Table 1) that may increase the risk of immune response [8,23,28]. The accessory proteins assemble into a supramolecular structure that serves two main functions: protecting the core neurotoxin from low pH conditions when ingested orally and facilitating its absorption in the gastrointestinal tract [27]. The protective function is mediated via the nontoxic nonhemagglutinin protein

and the absorption function via hemagglutinin proteins [11]. Importantly, the accessory proteins rapidly dissociate from the core neurotoxin at neutral pH [27,52].

Second-generation Bo-NT-As, such as incoBoNT-A and daxiBoNT-A lack accessory proteins because of their removal during purification [11]. DaxiBoNT-A contains an HIV-derived, highly charged peptide (RTP004) which, according to the manufacturer, binds noncovalently to the negatively charged BoNT-A molecule and stabilizes it by preventing protein aggregation [28]. Additionally, the peptide may bind to negatively charged neuronal surfaces, which could enhance the internalization of the neurotoxin. However, Martin et al [28] reported that the binding of RTP004 to negatively charged neuronal surfaces should not be considered selective, as all cell types are negatively charged due to the terminal sialic acid residues on surface glycoproteins.

Immunogenicity of BoNT-A Formulations

Nonclinical Data

The total clostridial protein load—comprising accessory proteins and the core neurotoxin—and its composition determine the immunogenicity of each BoNT-A formulation [53]. Accessory proteins, especially hemagglutinin-1, can enhance the immune response as adjuvants [54,55]. Antibodies (Abs) against BoNT can be divided into NAbs, targeting the core neurotoxin, mainly the binding site on the heavy chain, and non-NAbs, typically targeting accessory proteins or clinically irrelevant sites on the core neurotoxin. While NAbs inhibit the clinical efficacy of BoNT, the non-NAbs do not impact its clinical effectiveness. In rabbit studies, immunization with the complete inactivated BoNT-A complex generated Abs with a stronger neutralizing effect than Abs induced by immunization with the core neurotoxin alone [54]. Accessory proteins may trigger increased production of inflammatory cytokines such as interleukin-6 and tumor necrosis factor-alpha and can bind to several nonneuronal cell types [55].

The total protein per vial of common BoNT-As is shown in Table 1 [35–37, 56 and 57]. IncoBoNT-A does not contain any inactive neurotoxin. In vivo studies indicate that onaBoNT-A injections generate antiBoNT-A Abs, with more frequent dosing leading to higher Ab levels [56]. In rabbits that received nine injections of onaBoNT-A or incoBoNT-A (at 2-8 week intervals), NAbs were detected in 20% of onaBoNT-A-treated animals, while none were detected in those treated with the accessory protein-free incoBoNT-A formulation [37]. AboBoNT-A contains less clostridial protein than onaBoNT-A, but its accessory proteins comprise up to 30% of the total clostridial protein content [11]. Importantly, the aboBoNT-A formulation also contains flagellin, which activates the toll-like receptor 5, thereby triggering an innate immune response [49].

The daxiBoNT-A formulation contains a proprietary, HIV-derived 5-kDa stabilizing peptide (RTP004) and polysorbate 20 [44]. This novel HIV-derived peptide is considered immunogenic [28]. As RTP004 binds to negatively charged areas on the surface of BoNT-A, it may create novel structures on the heavy or light chains of the core toxin that the immune system can recognize as neoepitopes. Polysorbate 20 may generate free radicals via auto-oxidization and can interact with other proteins in the formulation [28].

NAb Formation in Clinical Studies

BoNT-A treatment can trigger an adaptive immune response, especially with repeated injections, which may lead to NAb formation over time [11,57,58]. The rate of NAb development and occurrence of clinical resistance vary significantly by the BoNT-A formulation, particularly its protein content [59]. Table 1 shows the prevalence rates of NAbs in pivotal BoNT-A trials that supported approval by the US Food and Drug Administration (FDA). Pivotal onaBoNT-A and aboBoNT-A studies used the mouse protection assay (MPA), while incoBoNT-A studies used the mouse hemidiaphragm assay (MHDA), which is at least five times more sensitive than the MPA. Despite its greater sensitivity, the MHDA consistently revealed the lowest rates of NAb formation [10,59]. Analysis from phase 3 trials with daxiBoNT-A showed low rates of Ab formation to both daxiboNT-A and excipient RTP004 [45]. Treatment-related anti-daxiboNT-A and anti-RTP004 binding Abs were detected in 0.8% and 1.3% of subjects, respectively. No individual developed NAbs. Binding Abs were generally transient, of low titer (<1:200), and no individual had binding Abs to both daxiBoNT-A and RTP004. All individuals with treatment-induced binding Abs to daxiboNT-A or RTP004 showed clinical response at week 4 following each treatment cycle, indicating no impact on treatment efficacy. However, of the 2786 patients, 882 received two treatments and only 568 received three treatments. Therefore, the cumulative exposure and overall time frame for development of NAb-induced SNR may have been too short to draw robust conclusions.

The reported incidence rates of NAbs in product labeling are derived from short-term clinical trials and may not reflect real-world data, as repeated BoNT-A use can have cumulative effects over time [59]. Real-world studies with long-term follow-up have shown a reduction in NAbs in patients treated with incoBoNT-A [4,60,61]. A meta-analysis found that the prevalence of NAbs across indications is higher in patients treated with onaBoNT-A (around 1.5%) or aboBoNT-A (around 1.7%) compared to those receiving incoBoNT-A (0.5%) [4]. Although the overall prevalence of NAbs was low, there was a significantly higher rate of NAb development among patients who exhibit SNR [5]. Specifically, among patients with SNR, NAbs were observed in 32.5% patients treated with onaBoNT-A and 56.7% with aboBoNT-A. Notably, none of the patients who received incoBoNT-A developed SNR [4].

In an MHDA-based study, none of the toxin-naive patients who received incoBoNT-A treatment developed NAbs [62]. Furthermore, there have been no reported instances of clinical nonresponse among individuals who were toxin-naive at the time they received incoBoNT-A [10,62]. The formation of NAbs was rare in pivotal clinical trials, with only 9 out of more than 2600 patients treated with incoBoNT-A developing them [43]. A pooled data analysis from pivotal clinical studies on the aesthetic use of incoBoNT-A indicated no diminished treatment response due to the formation of NAbs [63]. Another study showed that switching to incoBoNT-A after SNR with another BoNT-A formulation enabled patients to regain responsiveness

to treatment, with NAbs developing only in two patients previously treated with aboBoNT-A [62].

SNR and NAb in Aesthetic Studies

Case studies of BoNT-A use for aesthetic purposes demonstrated both SNR and NAb development over time with onaBoNT-A and aboBoNT-A [6,15,48,63]. In general, prevalences of NAb development and SNR are lower in aesthetic indications (overall NAb rate estimated at 0.2% - 0.4%) [5], which may reflect the lower doses employed and minimal long-term data [15,49,64].

Thirteen cases of NAb-related SNR emerging during aesthetic BoNT-A treatments [1,3,6,15,18,65,66] were identified in case reports or series. Key observations of this review are presented in Textbox 1. Complete SNR is usually preceded by partial SNR in the patient [6,17,18]. Complete SNR usually occurs after more than two injection series [6,17]. It can occur as long as after 5 years of treatment [6,17]. In a small sample study, 30% of patients who did not respond to onaBoNT-A cosmetic treatments responded when switched to incoBoNT-A therapy, which did not provoke immune responses [69].

Textbox 1. Key observations in reports detailing secondary nonresponse (SNR) to botulinum toxin type A (BoNT-A) aesthetic treatment.

- Seven reports detailing a total of 13 cases [1,3,6,15,18,65,66]
- Patients initially or exclusively received onabotulinumtoxinA (onaBoNT-A) or aboBoNT-A (aboBoNT-A)
- SNR developed even after low BoNT-A doses [1,6,65]
- Regular repeated treatments before development of SNR, with clear signs of increasing dosages and shortening intervals between treatments
- Partial SNR observed as early as 2nd injection cycle [6] and complete SNR as early as 1st cycle [15]; partial SNR usually preceded complete SNR [6]
- Duration of therapy before natural antibody (NAb) detection variable (2 - 72 months) [15,65]
- Systematic testing for detecting NAb formation was infrequent and, in most cases, it was unclear when NAb formation first occurred
- No cases of NAb-related SNR were reported with exclusive incobotulinumtoxinA (incoBoNT-A) use
- Four patients were switched to incoBoNT-A after partial or complete SNR [1,6,15,65; this switch showed no treatment effect
- Switch to incoBoNT-A associated with downward trend in NAb titer [66]
- After SNR, injection of botulinum toxin type B (BoNT-B) showed a normal therapeutic effect [1,6]

After switching from BoNT-A to BoNT-B, NAbs to the latter may develop because the heavy chains of BoNT-A and BoNT-B have a 30% structural homology [26]. Patients who initially respond to BoNT-B after developing SNR to BoNT-A are likely to eventually develop SNR to BoNT-B as well [67,68].

Discussion

Insights Into Mechanisms of SNR

Retrospective studies suggest an association between higher protein exposure and increased risk of Ab formation [14,70,71]. The precise mechanisms leading to resistance are still unknown, as the pure 150-kDa neurotoxin has low immunogenicity without any known associated pattern recognition receptors or toll-like receptors on dendritic cells. Park et al [23] suggested that when adjuvants in the BoNT formulation are injected alongside the 150-kDa neurotoxin, they can activate dendritic cells that may internalize the neurotoxin and present it to T-helper lymphocytes, resulting in NAb formation. Exogenous factors such as environmental allergens (eg, COVID-19 vaccine) may prime NAbs [72,73]. Specific immune system activation by a wasp sting was proposed as a contributing factor for BoNT-A Ab formation [74].

Alternate explanations for resistance to BoNT-A include muscle injection fibrosis, BoNT receptor downregulation, dynamic line depth worsening, and interactions with drugs like aminoglycosides and quinolones [3]. Intradermal injections are thought to carry a higher risk of developing resistance to BoNT-A compared to intramuscular injections, as the dermis is rich in antigen-presenting dendritic cells [5,7]. A phenomenon of decreased responsiveness after many years of BoNT-A therapy, known as tachyphylaxis, has been reported [1]. In such cases, the clinical effect is mitigating despite the absence of NAbs. Nevertheless, it is still uncertain whether this phenomenon has an immunologic basis and whether low-titer or poorly binding antibodies might play a role.

SNR Management: Early Diagnosis

Early diagnosis is crucial, particularly as an increase in NAb formation must be addressed promptly. A patient's aesthetic journey, especially a need for increasing BoNT-A doses and more frequent treatments, should alert the provider of possible SNR. Accurately detecting and quantifying NAbs supports the diagnosis. Structural assays such as ELISA and immunoprecipitation assays are sensitive for detecting BoNT Abs, but do not discriminate between NAbs and non-neutralizing Abs [14,15,23]. Bioassays such as the MPA or MHDA use animal models to identify NAbs. The MHDA, the only assay approved by the FDA, uses ex vivo testing for NAbs [14].

Most clinicians do not have access to the above assays and use clinical resistance tests to confirm the diagnosis of SNR [1,14]. One such test is the unilateral brow injection, which involves injecting a standard amount of BoNT-A, such as 20IU onaBoNT-A, into the right (by convention) medial eyebrow [14]. After allowing sufficient time for the toxin to take effect (typically 1 - 3 weeks), the frowning facial expression is evaluated. Since nearly all individuals usually frown

symmetrically, asymmetric frowning indicates responsiveness to the injected BoNT-A that has weakened the right corrugator or procerus muscles. In contrast, symmetric frowning indicates that the injected muscles were not weakened; therefore, the patient is likely resistant to that specific type of BoNT-A.

Preventive Measures

Several authors have advocated for using a highly purified toxin that demonstrates the least immunogenicity, such as incoBoNT-A [15,23]. This is especially important in large-dose injections and while treating younger patients who will accumulate higher lifetime doses [23]. Most experts recommend using the smallest BoNT-A dose that achieves the desired clinical effect, avoiding booster injections, and waiting at least 3 months between treatments [6,7,15]. Regarding maximum dose, 56.5% of aesthetic providers responded that BoNT-A dose should be limited to <100 IU per day, and 97.3% reported using <300 IU in total [7]. Such total doses are unlikely for wrinkle reduction but are possible with some off-label indications such as muscle size reduction. In body indications, higher doses of BoNT-A are injected, increasing a patient's exposure to foreign proteins and their risk of NAB formation. Consequently, it is advisable to use a highly purified BoNT-A when treating body indications.

Increasing the efficacy and longevity of outcomes of BoNT-A treatments leads to decreased frequency of such treatments, which can help prevent resistance. Several authors recommend using toxins that offer improved longevity for cosmetic results, such as daxiBoNT-A [46]. In two of three randomized controlled trials, coadministration with oral zinc supplementation enhanced the longevity of BoNT-A outcomes [75-77]; however, the available data are limited. Hyaluronidase is a known tissue permeability modifier that increases the dispersion of drugs [78]. In a small pilot study on axillary hyperhidrosis, the coadministration of BoNT-A with hyaluronidase allowed for a reduction in the BoNT-A dose needed to achieve a similar effect compared to BoNT-A injections administered alone [79]. Notably, in one patient, the right side of the forehead–treated with both BoNT-A and hyaluronidase–exhibited a larger area of effect than the left side, which received only BoNT-A, across all postinjection evaluations. The authors suggested that the reduced dose of BoNT-A required when used alongside hyaluronidase may be attributed to the enhanced dispersion of the toxin facilitated by hyaluronidase. This approach could help avoid the use of high toxin doses that may lead to nonresponsiveness over time. However, more data are needed to confirm these findings.

Treatment

Switching to a highly purified toxin such as incoBoNT-A once partial SNR is noted, has been advocated [14,23,66], especially as this was associated with a downward trend in NAb titers [66,80]. This switch was associated with clinical response in a study of patients with cervical dystonia [81,82] and another involving onaBoNT-A cosmetic treatments [69]. Nevertheless, in our review of aesthetic treatments (Textbox 1), this switch was not associated with short-term SNR resolution [3,6,15,65]. Longer follow-up is required for aesthetic applications in patients with SNR switching to incoBoNT-A. A switch to daxiBoNT-A may also be considered given its low immunogenicity in limited studies [46], but more data is required. The first author successfully used a short course of low-dose oral methotrexate immediately before BoNT-A injection to mitigate an immune response leading to further reduction of clinical efficacy in patients who experienced partial SNR. Patients with prior complete or partial SNR to onaBoNT-A may benefit from anti-calcitonin gene-related peptide monoclonal Ab therapy [83].

For complete nonresponse, many experts advise offering a 12- to 18-month "drug holiday," and then resuming with a highly purified toxin. This suggestion is based on the medical applications of BoNT-A [67,84] and aims to normalize NAb levels before administering BoNT-A again. The duration of the "drug holiday" should be determined by measuring NAb levels. However, other experts argue against offering a "drug holiday," noting that switching to incoBoNT-A results in most patients' NAb titers returning to negative, similar to those who stopped receiving BoNT-A treatment altogether [80]. Moreover, switching to incoBoNT-A may be the only option for patients whose NAb titers take longer to become negative [67,80].

Switching to a different BoNT serotype, such as type B (BoNT-B), has been attempted. For cervical dystonia, switching to BoNT-B (rimabotulinumtoxinB), was beneficial [62]. In two patients reviewed here, after SNR developed, injection of BoNT-B showed a normal therapeutic effect [1,6]. However, patients who switched from BoNT-A to BoNT-B after developing NAbs and SNR may subsequently develop NAbs to BoNT-B due to the 30% structural homology in the heavy chains of BoNT-A and BoNT-B [26]. Several studies have demonstrated that patients who initially respond to BoNT-B after developing SNR to BoNT-A are likely to eventually develop SNR to BoNT-B as well [67,68]. Additionally, injecting BoNT-B, an off-label toxin in aesthetics presents challenges, including suboptimal longevity and adverse effects such as an intense stinging sensation on injection [85,86].

Data Gaps and Limitations

Aesthetic studies on NAb formation and SNR have been limited and have primarily focused on approved indications [4,6,63,87], while off-label applications involving higher BoNT-A doses have not been investigated. Additionally, the follow-up periods in these studies were relatively short (4-16 months), although NAbs usually develop over a more extended period, often spanning several years [5,40]. The frequency of NAb formation and SNR in real-world aesthetic practice may be higher than published estimates [5], likely due to extensive off-label use and the lack of a commercially available test for measuring NAb levels [11].

Detecting NAbs depends on the specific assay used, as there can be significant variability in sensitivity and specificity [10,11]. It also depends on the assay methodology, handling, and timing of collection of samples, and concurrent use of medications. Although the MHDA is the most sensitive bioassay, it is semiquantitative and not widely available. However, this assay has raised concerns about false-positive results and may detect subclinical Ab titers that do not result in treatment failure [1,14,15]. A quantitative, FDA-approved,

commercially available assay to measure NAbs is needed to study the temporal variations in Ab titers [11]. This limitation prevents robust conclusions regarding the relationship of NAbs with nonresponsiveness. A lack of studies comparing BoNT-A formulations with a standardized NAb assay hinders reliable comparisons. Finally, it remains unclear to what extent the accessory proteins, inactive neurotoxin, and excipients may trigger the immune system, especially since the time frame for developing Ab-mediated SNR was short in most studies (ie, up to three injection cycles) [28]. This hampers our ability to draw firm conclusions regarding the excipients' impact on the BoNT-A formulation's immunogenicity.

A key uncertainty involves the relationship between NAbs and SNR [14]. Some patients with detectable NAbs retain their clinical responsiveness, while others without detectable NAbs have been nonresponsive to BoNT-As [14,49]. This indicates that there is no absolute correlation between NAb detection and nonresponse [88], and no established threshold for NAb titer reliably predicts clinical resistance to BoNT-A [3]. However, a correlation between responsiveness and NAb titers has been proposed [6,89]. Further complicating patient responses, variations in target binding site and binding affinity result in anti-BoNT-A Abs with variable neutralizing effects [10,26]. These observations highlight the complexity of BoNT-A immunogenicity and the variability in individual patient responses [14].

Conclusions

Nonresponse to BoNT-A is becoming increasingly important in aesthetics, particularly as many patients undergo lifelong treatments. Preventing SNR is crucial given the lack of solid data on effective treatments. When choosing a BoNT-A formulation, considering the potential for immunogenicity is essential. Aesthetic providers should perform comprehensive clinical assessments, inform patients about the associated risks, and develop strategies to minimize immunogenicity in their treatment protocols.

Data Availability

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

Authors' Contributions

Conceptualization: GK
Data curation: GK
Formal analysis: FS, GK
Methodology: GK
Writing – original draft: GK
Writing – review & editing: FS

Conflicts of Interest

None declared.

References

1. Borodic G. Botulinum toxin, immunologic considerations with long-term repeated use, with emphasis on cosmetic applications. Facial Plast Surg Clin North Am 2007 Feb;15(1):11-16. [doi: 10.1016/j.fsc.2006.10.001] [Medline: 17317551]
2. Ford B, Louis ED, Greene P, Fahn S. Outcome of selective ramisectomy for botulinum toxin resistant torticollis. J Neurol Neurosurg Psychiatry 1998 Oct;65(4):472-478. [doi: 10.1136/jnnp.65.4.472] [Medline: 9771768]
3. Stephan F, Habre M, Tomb R. Clinical resistance to three types of botulinum toxin type A in aesthetic medicine. J Cosmet Dermatol 2014 Dec;13(4):346-348. [doi: 10.1111/jocd.12108] [Medline: 25399628]
4. Fabbri M, Leodori G, Fernandes RM, et al. Neutralizing antibody and botulinum toxin therapy: a systematic review and meta-analysis. Neurotox Res 2016 Jan;29(1):105-117. [doi: 10.1007/s12640-015-9565-5] [Medline: 26467676]
5. Ho WWS, Albrecht P, Calderon PE, et al. Emerging trends in botulinum neurotoxin a resistance: an international multidisciplinary review and consensus. Plast Reconstr Surg Glob Open 2022 Jun;10(6):e4407. [doi: 10.1097/GOX.0000000000004407] [Medline: 35747253]
6. Dressler D, Wohlfahrt K, Meyer-Rogge E, Wiest L, Bigalke H. Antibody-induced failure of botulinum toxin a therapy in cosmetic indications. Dermatol Surg 2010 Dec;36 Suppl 4:2182-2187. [doi: 10.1111/j.1524-4725.2010.01710.x] [Medline: 21134050]
7. Oh SM, Kim HM, Ahn TH, Park MS, Ree YS, Park ES. Aesthetic doctors' perception and attitudes toward tolerance in botulinum toxin. Skin Res Technol 2024 Apr;30(4):e413691. [doi: 10.1111/srt.13691] [Medline: 38576226]
8. Corduff N, Park JY, Calderon PE, et al. Real-world implications of botulinum neurotoxin A immunoresistance for consumers and aesthetic practitioners: insights from ASCEND multidisciplinary panel. Plast Reconstr Surg Glob Open 2024 Jun;12(6):e5892. [doi: 10.1097/GOX.0000000000005892] [Medline: 38903135]
9. Atassi MZ, Dolimbek BZ, Jankovic J, Steward LE, Aoki KR. Molecular recognition of botulinum neurotoxin B heavy chain by human antibodies from cervical dystonia patients that develop immunoresistance to toxin treatment. Mol Immunol 2008 Sep;45(15):3878-3888. [doi: 10.1016/j.molimm.2008.06.031] [Medline: 18676021]

10. Frevert J, Dressler D. Clinical relevance of immunoresistance to botulinum therapy. In: Botulinum Toxin Therapy Manual for Dystonia and Spasticity: Intech; 2016:33-49. [doi: 10.5772/64566]
11. Carr WW, Jain N, Sublett JW. Immunogenicity of botulinum toxin formulations: potential therapeutic implications. Adv Ther 2021 Oct;38(10):5046-5064. [doi: 10.1007/s12325-021-01882-9] [Medline: 34515975]
12. Anyanwu B, Hanna PA, Jankovic J. Botulinum toxin: primary and secondary resistance. In: Ward AB, Barnes MP, editors. Clinical Uses of Botulinum Toxins: Cambridge University Press; 2007:45-57. [doi: 10.1017/CBO9780511544842.004]
13. Thompson JA, Glasgow LA, Warpinski JR, Olson C. Infant botulism: clinical spectrum and epidemiology. Pediatrics 1980 Dec;66(6):936-942. [doi: 10.1542/peds.66.6.936] [Medline: 7005856]
14. Bellows S, Jankovic J. Immunogenicity associated with botulinum toxin treatment. Toxins (Basel) 2019 Aug 26;11(9):491. [doi: 10.3390/toxins11090491] [Medline: 31454941]
15. Torres S, Hamilton M, Sanches E, Starovatova P, Gubanova E, Reshetnikova T. Neutralizing antibodies to botulinum neurotoxin type A in aesthetic medicine: five case reports. Clin Cosmet Investig Dermatol 2014;7:11-17. [doi: 10.2147/CCID.S51938] [Medline: 24379687]
16. Dobryansky M, Korsh J, Shen AE, Aliano K. Botulinum toxin type A and B primary resistance. Aesthet Surg J 2015 Feb;35(2):NP28-NP30. [doi: 10.1093/asj/sju027] [Medline: 25653242]
17. Dressler D. Clinical features of antibody-induced complete secondary failure of botulinum toxin therapy. Eur Neurol 2002;48(1):26-29. [doi: 10.1159/000064953]
18. Lee SK. Antibody-induced failure of botulinum toxin type A therapy in a patient with masseteric hypertrophy. Dermatol Surg 2007 Jan;33(1 Spec No):S105-S110. [doi: 10.1111/j.1524-4725.2006.32340.x] [Medline: 17278332]
19. Jinnah HA, Goodmann E, Rosen AR, Evatt M, Freeman A, Factor S. Botulinum toxin treatment failures in cervical dystonia: causes, management, and outcomes. J Neurol 2016 Jun;263(6):1188-1194. [doi: 10.1007/s00415-016-8136-x]
20. Carey WD. Incorrect reconstitution of incobotulinumtoxinA leads to loss of neurotoxin. J Drugs Dermatol 2014 Jun;13(6):735-738. [Medline: 24918566]
21. Sibar S, Findikcioglu K, Zinnuroglu M, Cenetoglu S. The effect of albumin and platelet-poor plasma supplemented botulinum A toxin on bioavaliability: an experimental rabbit model. Ann Plast Surg 2017 Apr;78(4):436-442. [doi: 10.1097/SAP.0000000000000957] [Medline: 28079531]
22. Wohlfarth K, Sycha T, Ranoux D, Naver H, Caird D. Dose equivalence of two commercial preparations of botulinum neurotoxin type A: time for a reassessment? Curr Med Res Opin 2009 Jul;25(7):1573-1584. [doi: 10.1185/03007990903028203] [Medline: 19463043]
23. Park JY, Corduff N, Frevert J, Wanitphakdeedecha R, Chao YYY. Immunogenicity associated with aesthetic botulinumtoxin A: a survey of Asia-Pacific Physicians' experiences and recommendations. Plast Reconstr Surg Glob Open 2022 Apr;10(4):e4217. [doi: 10.1097/GOX.0000000000004217] [Medline: 35450268]
24. Mehta D, Wildman H. Botox resistance and COVID-19 vaccines: is type B Botox a viable solution? J Cosmet Dermatol 2024 Feb;23(2):368-369. [doi: 10.1111/jocd.15987] [Medline: 37658662]
25. Atassi MZ. Molecular basis of immunogenicity to botulinum neurotoxins and uses of the defined antigenic regions. Toxicon 2015 Dec 1;107(Pt A):50-58. [doi: 10.1016/j.toxicon.2015.06.003] [Medline: 26086358]
26. Dolimbek BZ, Aoki KR, Steward LE, Jankovic J, Atassi MZ. Mapping of the regions on the heavy chain of botulinum neurotoxin A (BoNT/A) recognized by antibodies of cervical dystonia patients with immunoresistance to BoNT/A. Mol Immunol 2007 Feb;44(5):1029-1041. [doi: 10.1016/j.molimm.2006.03.011]
27. Kutschenko A, Bigalke H, Wegner F, Wohlfarth K. The role of human serum albumin and neurotoxin associated proteins in the formulation of BoNT/A products. Toxicon 2019 Oct;168:158-163. [doi: 10.1016/j.toxicon.2019.07.005] [Medline: 31323228]
28. Martin MU, Frevert J, Tay CM. Complexing protein-free botulinum neurotoxin A formulations: implications of excipients for immunogenicity. Toxins (Basel) 2024 Feb 10;16(2):101. [doi: 10.3390/toxins16020101] [Medline: 38393178]
29. Pirazzini M, Rossetto O, Eleopra R, Montecucco C. Botulinum neurotoxins: biology, pharmacology, and toxicology. Pharmacol Rev 2017 Apr;69(2):200-235. [doi: 10.1124/pr.116.012658] [Medline: 28356439]
30. Dressler D, Benecke R. Pharmacology of therapeutic botulinum toxin preparations. Disabil Rehabil 2007 Jan;29(23):1761-1768. [doi: 10.1080/09638280701568296]
31. Ferrari A, Manca M, Tugnoli V, Alberto L. Pharmacological differences and clinical implications of various botulinum toxin preparations: a critical appraisal. Funct Neurol 2018;33(1):7-18. [doi: 10.11138/fneur/2018.33.1.007] [Medline: 29633692]
32. Kerscher M, Wanitphakdeedecha R, Trindade de Almeida A, Maas C, Frevert J. IncobotulinumtoxinA: a highly purified and precisely manufactured botulinum neurotoxin type A. J Drugs Dermatol 2019 Jan 1;18(1):52-57. [Medline: 30681794]
33. Frevert J. Content of botulinum neurotoxin in Botox®/Vistabel®, Dysport®/Azzalure®, and Xeomin®/Bocouture®. Drugs R D 2010;10(2):67-73. [doi: 10.2165/11584780-000000000-00000] [Medline: 20698714]
34. Carruthers J, Solish N, Humphrey S, et al. Injectable daxibotulinumtoxinA for the treatment of glabellar lines: a phase 2, randomized, dose-ranging, double-blind, multicenter comparison with onabotulinumtoxinA and placebo. Dermatol Surg 2017 Nov;43(11):1321-1331. [doi: 10.1097/DSS.0000000000001206] [Medline: 28614091]

35. Prescribing information: Botox®. Allergan, Inc. URL: https://www.allergan.com/assets/pdf/botox_pi.pdf [accessed 2018-10-09]
36. Panjwani N, O'Keeffe R, Pickett A. Biochemical, functional and potency characteristics of type A botulinum toxin in clinical use. TBJ 2008;1(1):153. [doi: 10.1504/TBJ.2008.018956]
37. Jost WH, Blümel J, Grafe S. Botulinum neurotoxin type A free of complexing proteins (XEOMIN) in focal dystonia. Drugs (Abingdon Engl) 2007;67(5):669-683. [doi: 10.2165/00003495-200767050-00003] [Medline: 17385940]
38. Dressler D, Mander G, Fink K. Measuring the potency labelling of onabotulinumtoxinA (Botox®) and incobotulinumtoxinA (Xeomin®) in an LD50 assay. J Neural Transm 2012 Jan;119(1):13-15. [doi: 10.1007/s00702-011-0719-1]
39. Innovative formulation. Revance. URL: https://hcp.daxxify.com/daxxify-difference [accessed 2024-12-10]
40. Field M, Splevins A, Picaut P, et al. AbobotulinumtoxinA (Dysport®), onabotulinumtoxinA (Botox®), and incobotulinumtoxinA (Xeomin®) neurotoxin content and potential implications for duration of response in patients. Toxins (Basel) 2018;10(12):535. [doi: 10.3390/toxins10120535]
41. Botox - highlights of prescribing information. : AbbVie Inc; 2024 URL: https://www.rxabbvie.com/pdf/botox_pi.pdf [accessed 2024-10-12]
42. Dysport - highlights of prescribing information. : Ipsen Biopharm Ltd; 2016 URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/125274s107lbl.pdf [accessed 2024-12-10]
43. Xeomin - highlights of prescribing information. : Merz Pharmaceuticals GmbH; 2020 URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/125360s078lbl.pdf [accessed 2024-12-10]
44. Carruthers JD, Fagien S, Joseph JH, et al. DaxibotulinumtoxinA for injection for the treatment of glabellar lines: results from each of two multicenter, randomized, double-blind, placebo-controlled, phase 3 studies (SAKURA 1 and SAKURA 2). Plast Reconstr Surg 2020 Jan;145(1):45-58. [doi: 10.1097/PRS.0000000000006327] [Medline: 31609882]
45. Gallagher CJ, Bowsher RR, Clancy A, et al. Clinical immunogenicity of daxibotulinumtoxinA for injection in glabellar lines: pooled data from the SAKURA phase 3 trials. Toxins (Basel) 2023 Jan 10;15(1):60. [doi: 10.3390/toxins15010060] [Medline: 36668880]
46. Humphrey S, Dover JS, Bowsher RR, et al. Immunogenicity of daxibotulinumtoxinA for injection in glabellar lines. Aesthet Surg J 2023 Sep 14;43(10):1189-1193. [doi: 10.1093/asj/sjad101] [Medline: 37051886]
47. Albrecht P, Jansen A, Lee JI, et al. High prevalence of neutralizing antibodies after long-term botulinum neurotoxin therapy. Neurology (ECronicon) 2019 Jan 1;92(1):e48-e54. [doi: 10.1212/WNL.0000000000006688] [Medline: 30464031]
48. Cohen JL, Scuderi N. Safety and patient satisfaction of abobotulinumtoxinA for aesthetic use: a systematic review. Aesthet Surg J 2017 May 1;37(suppl_1):S32-S44. [doi: 10.1093/asj/sjx010] [Medline: 28388721]
49. Naumann M, Boo LM, Ackerman AH, Gallagher CJ. Immunogenicity of botulinum toxins. J Neural Transm (Vienna) 2013 Feb;120(2):275-290. [doi: 10.1007/s00702-012-0893-9] [Medline: 23008029]
50. Mizel SB, Bates JT. Flagellin as an adjuvant: cellular mechanisms and potential. J Immunol 2010 Nov 15;185(10):5677-5682. [doi: 10.4049/jimmunol.1002156] [Medline: 21048152]
51. Dover JS, Monheit G, Greener M, Pickett A. Botulinum toxin in aesthetic medicine: myths and realities. Dermatol Surg 2018 Feb;44(2):249-260. [doi: 10.1097/DSS.0000000000001277] [Medline: 29016535]
52. Eisele KH, Fink K, Vey M, Taylor HV. Studies on the dissociation of botulinum neurotoxin type A complexes. Toxicon 2011 Mar 15;57(4):555-565. [doi: 10.1016/j.toxicon.2010.12.019] [Medline: 21195107]
53. Frevert J. Pharmaceutical, biological, and clinical properties of botulinum neurotoxin type A products. Drugs R D 2015 Mar;15(1):1-9. [doi: 10.1007/s40268-014-0077-1] [Medline: 25559581]
54. Kukreja R, Chang TW, Cai S, et al. Immunological characterization of the subunits of type A botulinum neurotoxin and different components of its associated proteins. Toxicon 2009 May;53(6):616-624. [doi: 10.1016/j.toxicon.2009.01.017] [Medline: 19673075]
55. Wang L, Sun Y, Yang W, Lindo P, Singh BR. Type A botulinum neurotoxin complex proteins differentially modulate host response of neuronal cells. Toxicon 2014 May;82:52-60. [doi: 10.1016/j.toxicon.2014.02.004]
56. Sutphin DD, Chun J, Hill W, et al. Type A botulinum toxin-induced antibody production: a murine model of antibody response. Aesthet Surg J 2009;29(5):414-418. [doi: 10.1016/j.asj.2009.08.004] [Medline: 19825472]
57. Mathevon L, Declemy A, Laffont I, Perennou D. Immunogenicity induced by botulinum toxin injections for limb spasticity: a systematic review. Ann Phys Rehabil Med 2019 Jul;62(4):241-251. [doi: 10.1016/j.rehab.2019.03.004] [Medline: 30980953]
58. Benecke R. Clinical relevance of botulinum toxin immunogenicity. BioDrugs 2012 Apr 1;26(2):e1-e9. [doi: 10.2165/11599840-000000000-00000] [Medline: 22385408]
59. Samadzadeh S, Ürer B, Brauns R, et al. Clinical implications of difference in antigenicity of different botulinum neurotoxin type A preparations: clinical take-home messages from our research pool and literature. Toxins (Basel) 2020 Aug 4;12(8):499. [doi: 10.3390/toxins12080499] [Medline: 32759685]
60. Walter U, Mühlenhoff C, Benecke R, et al. Frequency and risk factors of antibody-induced secondary failure of botulinum neurotoxin therapy. Neurology (ECronicon) 2020 May 19;94(20):e2109-e2120. [doi: 10.1212/WNL.0000000000009444]
61. Rahman E, Alhitmi HK, Mosahebi A. Immunogenicity to botulinum toxin type A: a systematic review with meta-analysis across therapeutic indications. Aesthet Surg J 2022 Jan 1;42(1):106-120. [doi: 10.1093/asj/sjab058] [Medline: 33528495]

62. Hefter H, Samadzadeh S, Moll M. Transient improvement after switch to low doses of rimabotulinumtoxinB in patients resistant to abobotulinumtoxinA. Toxins (Basel) 2020 Oct 27;12(11):677. [doi: 10.3390/toxins12110677] [Medline: 33121133]
63. Coleman WP III, Sattler G, Weissenberger P, Hast MA, Hanschmann A. Safety of incobotulinumtoxinA in the treatment of facial lines: results from a pooled analysis of randomized, prospective, controlled clinical studies. Dermatol Surg 2017 Dec;43 Suppl 3(3):S293-S303. [doi: 10.1097/DSS.0000000000001409] [Medline: 33065955]
64. Ianhez M, Peres G, Miot HA. Neutralizing antibodies to botulinum toxin type A are rare following aesthetic indications. Aesthet Surg J 2021 Oct 15;41(11):NP1800-NP1801. [doi: 10.1093/asj/sjab289] [Medline: 34302452]
65. Stengel G, Bee EK. Antibody-induced secondary treatment failure in a patient treated with botulinum toxin type A for glabellar frown lines. Clin Interv Aging 2011;6:281-284. [doi: 10.2147/CIA.S18997] [Medline: 22162643]
66. Martin MU, Tay CM, Siew TW. Continuous treatment with incobotulinumtoxinA despite presence of BoNT/A neutralizing antibodies: immunological hypothesis and a case report. Toxins (Basel) 2024 Oct 1;16(10):422. [doi: 10.3390/toxins16100422] [Medline: 39453199]
67. Dressler D. Clinical presentation and management of antibody-induced failure of botulinum toxin therapy. Mov Disord 2004 Mar;19 Suppl 8:S92-S100. [doi: 10.1002/mds.20022] [Medline: 15027060]
68. Dressler D, Bigalke H, Benecke R. Botulinum toxin type B in antibody-induced botulinum toxin type A therapy failure. J Neurol 2003 Aug;250(8):967-969. [doi: 10.1007/s00415-003-1129-6] [Medline: 12928917]
69. Wanitphakdeedecha R, Kantaviro W, Suphatsathienkul P, et al. Association between secondary botulinum toxin A treatment failure in cosmetic indication and anti-complexing protein antibody production. Dermatol Ther (Heidelb) 2020 Aug;10(4):707-720. [doi: 10.1007/s13555-020-00397-5] [Medline: 32445175]
70. Greene P, Fahn S, Diamond B. Development of resistance to botulinum toxin type A in patients with torticollis. Mov Disord 1994 Mar;9(2):213-217. [doi: 10.1002/mds.870090216] [Medline: 8196686]
71. Jankovic J, Schwartz K. Response and immunoresistance to botulinum toxin injections. Neurology (ECronicon) 1995 Sep;45(9):1743-1746. [doi: 10.1212/wnl.45.9.1743] [Medline: 7675238]
72. Dressler D, Hallett M. Immunological aspects of Botox, Dysport and Myobloc/NeuroBloc. Eur J Neurol 2006 Feb;13 Suppl 1:11-15. [doi: 10.1111/j.1468-1331.2006.01439.x] [Medline: 16417592]
73. Dressler D, Bigalke H, Frevert J. The immunology of botulinum toxin therapy: a brief summary. Toxicology 2022 Nov;481:153341. [doi: 10.1016/j.tox.2022.153341] [Medline: 36191878]
74. Paus S, Bigalke H, Klockgether T. Neutralizing antibodies against botulinum toxin A after a wasp sting. Arch Neurol 2006 Dec;63(12):1808-1809. [doi: 10.1001/archneur.63.12.1808] [Medline: 17172626]
75. Shemais N, Elarab AE, ElNahass H. The effect of botulinum toxin A in patients with excessive gingival display with and without zinc supplementation: randomized clinical trial. Clin Oral Investig 2021 Nov;25(11):6403-6417. [doi: 10.1007/s00784-021-03944-2] [Medline: 33950373]
76. Koshy JC, Sharabi SE, Feldman EM, Hollier LH Jr, Patrinely JR, Soparkar CNS. Effect of dietary zinc and phytase supplementation on botulinum toxin treatments. J Drugs Dermatol 2012 Apr;11(4):507-512. [Medline: 22453589]
77. Xing Y, O'Suilleabhain P. Does oral zinc supplementation augment the effect of botulinum neurotoxin in dystonia? (P4.330). Neurology (ECronicon) 2015 Apr 6;84(14_supplement). [doi: 10.1212/WNL.84.14_supplement.P4.330]
78. Kroumpouzos G, Treacy P. Hyaluronidase for dermal filler complications: review of applications and dosage recommendations. JMIR Dermatol 2024 Jan 17;7:e50403. [doi: 10.2196/50403] [Medline: 38231537]
79. Goodman G. Diffusion and short-term efficacy of botulinum toxin A after the addition of hyaluronidase and its possible application for the treatment of axillary hyperhidrosis. Dermatol Surg 2003 May;29(5):533-538. [doi: 10.1046/j.1524-4725.2003.29126.x] [Medline: 12752524]
80. Hefter H, Hartmann C, Kahlen U, Moll M, Bigalke H. Prospective analysis of neutralising antibody titres in secondary non-responders under continuous treatment with a botulinumtoxin type A preparation free of complexing proteins—a single cohort 4-year follow-up study. BMJ Open 2012;2(4):e000646. [doi: 10.1136/bmjopen-2011-000646]
81. Hefter H, Hartmann CJ, Kahlen U, Samadzadeh S, Rosenthal D, Moll M. Clinical improvement after treatment with IncobotulinumtoxinA (XEOMIN®) in patients with cervical dystonia resistant to botulinum toxin preparations containing complexing proteins. Front Neurol 2021;12:636590. [doi: 10.3389/fneur.2021.636590]
82. Hefter H, Brauns R, Ürer B, Rosenthal D, Albrecht P. Effective long-term treatment with incobotulinumtoxin (Xeomin®) without neutralizing antibody induction: a monocentric, cross-sectional study. J Neurol 2020 May;267(5):1340-1347. [doi: 10.1007/s00415-019-09681-7]
83. Alpuente A, Gallardo VJ, Caronna E, Torres-Ferrús M, Pozo-Rosich P. Partial and nonresponders to onabotulinumtoxinA can benefit from anti-CGRP monoclonal antibodies preventive treatment: a real-world evidence study. Eur J Neurol 2021 Jul;28(7):2378-2382. [doi: 10.1111/ene.14828] [Medline: 33730441]
84. Marion MH, Humberstone M, Grunewald R, Wimalaratna S. British Neurotoxin Network recommendations for managing cervical dystonia in patients with a poor response to botulinum toxin. Pract Neurol 2016 Aug;16(4):288-295. [doi: 10.1136/practneurol-2015-001335] [Medline: 26976927]
85. Carruthers A, Carruthers J. Botulinum toxin products overview. Skin Therapy Lett 2008;13(6):1-4. [Medline: 18806905]

86. Flynn TC. Botulinum toxin: examining duration of effect in facial aesthetic applications. Am J Clin Dermatol 2010;11(3):183-199. [doi: 10.2165/11530110-000000000-00000] [Medline: 20369902]
87. Naumann M, Carruthers A, Carruthers J, et al. Meta-analysis of neutralizing antibody conversion with onabotulinumtoxinA (BOTOX®) across multiple indications. Mov Disord 2010 Oct 15;25(13):2211-2218. [doi: 10.1002/mds.23254] [Medline: 20737546]
88. Foster K, Beard M. Comments on immunogenicity associated with botulinum toxin treatment. Toxins 2019, 11, 491. Toxins (Basel) 2020 Jan 23;12(2):71. [doi: 10.3390/toxins12020071] [Medline: 31979238]
89. Göschel H, Wohlfarth K, Frevert J, Dengler R, Bigalke H. Botulinum A toxin therapy: neutralizing and nonneutralizing antibodies--therapeutic consequences. Exp Neurol 1997 Sep;147(1):96-102. [doi: 10.1006/exnr.1997.6580] [Medline: 9294406]

ABBREVIATIONS

Ab: antibody
AboBoNT-A: abobotulinumtoxinA
BoNT-A: botulinum toxin type A
BoNT-B: botulinum toxin type B
DaxiBoNT-A: daxibotulinumtoxinA
FDA: Food and Drug Administration
IncoBoNT-A: incobotulinumtoxinA
MHDA: mouse hemidiaphragm assay
MPA: mouse protection assay
NAb: neutralizing antibodies
OnaBoNT-A: onabotulinumtoxinA
PNR: primary nonresponse
SNR: secondary nonresponse

Edited by J Meisenheimer; submitted 11.12.24; peer-reviewed by CM Tay, GDLT Canales, JN Yu; revised version received 14.03.25; accepted 15.03.25; published 24.04.25.

Please cite as:
Kroumpouzos G, Silikovich F
Exploring Nonresponse to Botulinum Toxin in Aesthetics: Narrative Review of Key Trigger Factors and Effective Management Strategies
JMIR Dermatol 2025;8:e69960
URL: https://derma.jmir.org/2025/1/e69960
doi:10.2196/69960

Artificial Intelligence in Patch Testing: Comprehensive Review of Current Applications and Future Prospects in Dermatology

Hilary S Tang*, BA; Joseph Ebriani*, BS; Matthew J Yan*, BS, BA; Shannon Wongvibulsin, MD, PhD; Mehdi Farshchian, MD, PhD

Division of Dermatology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

*these authors contributed equally

Corresponding Author:
Mehdi Farshchian, MD, PhD
Division of Dermatology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

Abstract

Background: The integration of artificial intelligence (AI) into patch testing for allergic contact dermatitis (ACD) holds the potential to standardize diagnoses, reduce interobserver variability, and improve overall diagnostic accuracy. However, the challenges and limitations hindering clinical implementation have not been thoroughly explored.

Objective: This narrative review aims to examine the current applications of AI in patch testing, identify challenges, and propose future directions for their use in dermatology.

Methods: PubMed was searched in August 2024 to identify studies involving human participants undergoing patch testing with AI used in the study. Exclusion criteria were non-English and nonoriginal research. Data were synthesized to assess study design, performance, and potential for clinical application.

Results: Out of 94 reviewed articles, 10 met the inclusion criteria. Most studies employed convolutional neural networks (CNN) for image analysis, with accuracy rates ranging from 90.1% to 99.5%. Other AI models, such as gradient boosting and random forest, were used for risk prediction and biomarker discovery. Key limitations included limited sample sizes, variability in image capture protocols, and lack of standardized reporting on skin types.

Conclusions: AI has significant potential to enhance diagnostic accuracy, standardize patch test interpretation, and expand access to patch testing. However, standardized imaging protocols, larger and more diverse datasets, and improved regulatory frameworks are necessary to realize the full potential of AI in patch testing.

(JMIR Dermatol 2025;8:e67154) doi:10.2196/67154

KEYWORDS

machine learning; ML; artificial intelligence; AI; algorithm; model; analytics; patch testing; allergic contact dermatitis; dermatitis; dermatology; dermatologist; skin; comprehensive review; comprehensive reviews; review

Introduction

Allergic contact dermatitis (ACD) is a common inflammatory skin condition affecting approximately 20% of the population, with significant impacts on patients' quality of life and productivity [1,2]. Traditional patch testing methods, while effective for diagnosing ACD, can be time-consuming and subject to interobserver variability [3,4]. As technology continues to advance, the integration of artificial intelligence (AI) offers the possibility of standardizing interpretations, reducing human error, and potentially improving the overall diagnostic process in patch testing [5].

AI, broadly defined as the ability of computer systems to mimic human cognitive functions, encompasses various computational subfields, including machine learning (ML). Furthermore, deep learning (DL), a subset of ML, uses algorithms modeled after human neurons to detect complex patterns and relationships in data [6]. These AI technologies have shown promising applications in dermatology, ranging from identifying skin malignancies to classifying inflammatory skin conditions and analyzing clinical notes. The visual nature of dermatology, combined with the increasing volume of clinical photographs, dermoscopy images, abundance of psychometric data from wearable devices, and electronic health records, makes it particularly well-suited for AI-augmented patient care [6-8].

The use of AI in patch testing is particularly intriguing due to the complex nature of interpreting patch test results. Several factors, such as weak positive reactions, irritant reactions, and the timing of readings, can all influence the accuracy of diagnoses, leading to interobserver variability and diagnostic inconsistencies [9-12]. Furthermore, the process is time-intensive, requiring multiple clinic visits for patients, and

resource-heavy for clinics, requiring 1 visit of application of allergens, an initial removal and preliminary evaluation visit around 48 hours, and a final follow-up evaluation several days later [13,14]. AI offers the potential to automate and standardize patch test result interpretations, reducing diagnostic variability and enabling broader access to this crucial diagnostic tool. AI can also analyze large datasets to uncover patterns and trends that may not be immediately evident to clinicians, ultimately enhancing the diagnostic process for ACD while mitigating bias and promoting equitable care across diverse patient populations [15,16].

This narrative review aims to explore the current landscape of AI applications in patch testing for ACD. We will examine the types of algorithms that are currently being researched, their performance, the challenges faced, and potential future directions for this rapidly evolving field. By synthesizing the available literature, we hope to provide a comprehensive overview of the state of AI in patch testing and how AI can be leveraged to improve patch testing practices and diagnostic accuracy of ACD in the future.

Methods

Search Strategy

A comprehensive literature search was performed in August 2024 using the PubMed database. The search was conducted without date restrictions to capture the full scope of research in this emerging field. This broad approach ensured that all relevant studies, regardless of publication date, were included, providing a more thorough evaluation of AI applications in patch testing and the observance of any trends over time. Literature searches were conducted using combinations of keywords, such as “artificial intelligence,” “machine learning,” “patch testing,” and “contact dermatitis” or “skin” (see Multimedia Appendix 1 for the full search term list). These terms were chosen to ensure a wide net was cast, incorporating both general AI terms and specific patch testing and dermatology-related concepts.

Inclusion and Exclusion Criteria

Criteria for inclusion and exclusion were defined prior to screening to reduce potential biases. Studies were included if they met the following criteria: the population consisted of human patient populations undergoing patch testing; the study design involved AI (which includes ML and DL); and outcomes reported on the performance of these algorithms. All publication types, including journal articles, conference abstracts, and preprints, were considered. Studies were excluded if they were not written in English or if they were not original research, such as review papers or perspectives.

Study Selection Process and Data Extraction

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach was selected to ensure transparency and replicability in the selection process, providing a clear pathway from initial search to final inclusion [17,18]. Each article was independently reviewed by 2 authors. In cases of disagreement, a third author resolved the discrepency. From the included studies, the following data elements were then extracted: study design; sample size; skin types included; length of study for each participant; location of study; materials used (such as types of allergen panels and imaging equipment); type of AI algorithm and its performance in the study; limitations and challenges of the study; and future directions. To ensure that AI models were properly evaluated, each study relied on a clearly defined ground truth as the reference standard for their data. This ground truth was established by dermatologists’ manual interpretation of patch test reactions, typically following standardized grading criteria such as the International Contact Dermatitis Research Group (ICDRG) scale, with some studies following European Society of Contact Dermatitis guidelines or similar clinical severity scales [19]. The findings were then synthesized to highlight trends, gaps, and potential areas for future research in the application of AI in patch testing. This synthesis serves as a foundation for guiding future research efforts, with the goal of synthesizing both technical and clinical factors of the clinical patch testing procedure, analysis, data capturing, image capturing and storage, AI algorithms, and diagnostic accuracy comprehensively, contributing to the current gaps in the current practice of AI integration within dermatological patch testing diagnostics.

Results

Included Studies

A total of 94 records were ultimately screened and evaluated for eligibility, as shown in the PRISMA flow diagram (Figure 1) [17]. Of the 94 articles, our literature review identified 10 relevant studies that employed various AI techniques in the context of patch testing and skin sensitization prediction, as shown in Table 1. These studies encompassed a wide range of approaches, from image analysis of patch test results to molecular profiling and risk prediction models.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for the identification of studies [17]. [a]Databases used in this narrative review: PubMed [b]Reports excluded: Does not meet inclusion criteria: (1) *Population*: All patient populations (humans) undergoing patch testing; (2) *Interventions*: Study designs of artificial intelligence (AI) or machine learning (ML) or deep learning (DL) algorithms in patch testing; (3) *Outcomes*: Non-English and nonoriginal research (eg, review papers, perspectives) were excluded for the purposes of this narrative review.

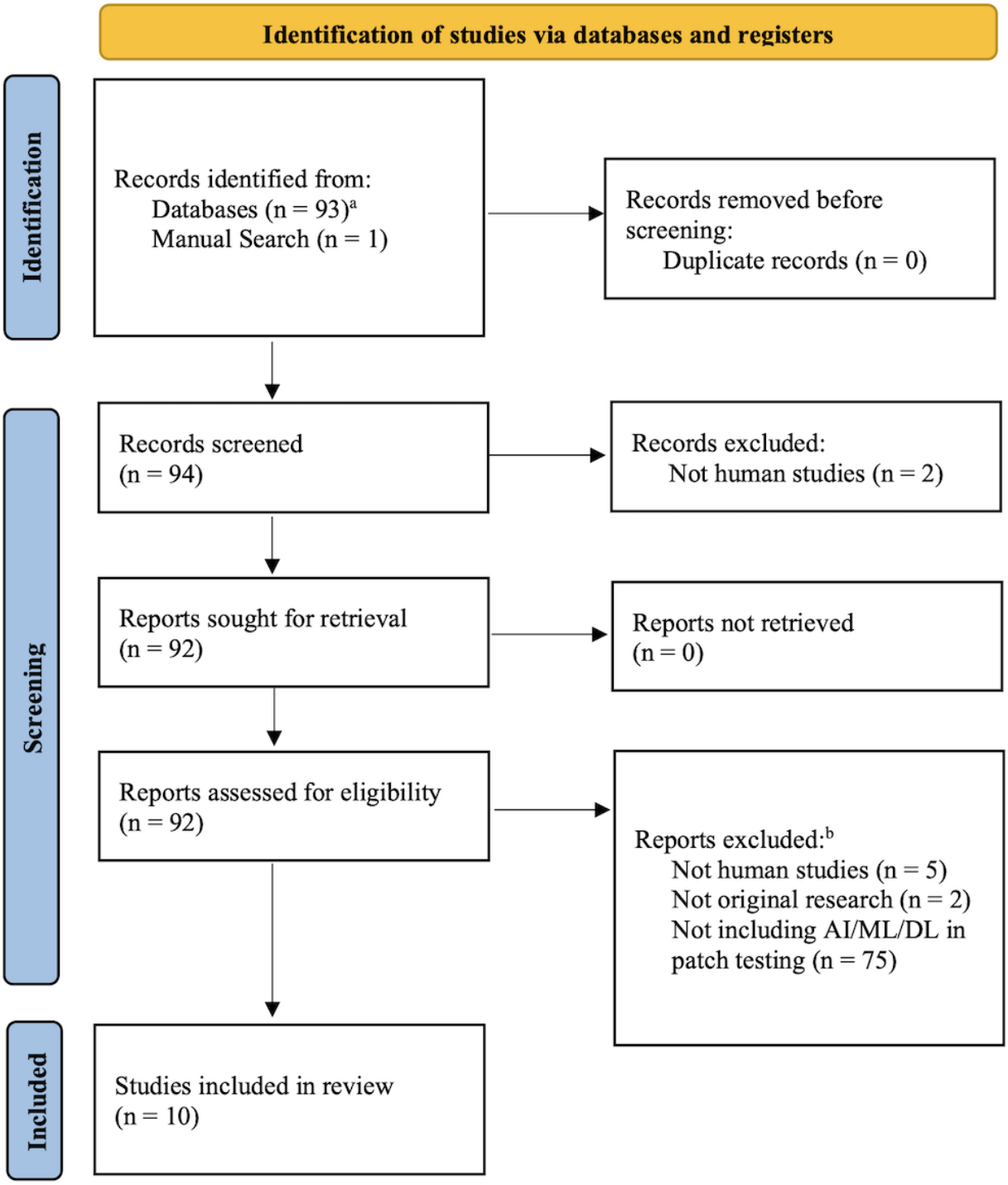


Table . Summary of articles included in this comprehensive review of artificial intelligence in patch testing.

Author, Year	Study objective	Type of AI used	Location	Fitzpatrick skin types (FST) included; Demographics, n (%)	Age in years, median (range)	Materials used	Total data sample size	Test, validation, evaluation data sample size	Accuracy, performance
Kyritsi et al, 2024 [20]	To investigate the contact allergy patterns	Multiple correspondence analysis (MCA), CATPCA (categorical principal components analysis)	Greece	Not reported	4 (18 - 86)	4 allergens	800 patients; clinical, demographic, occupational data	Not reported	Not reported
Ravishankar et al, 2024 [21]	To evaluate the use of convolutional neural networks to determine presence of patch test reactions	Convolutional neural network (CNN)	United States	FST I-II: 110 (88%), III-V: 15 (12%); Caucasian: 100 (80%), Black: 4 (3.2%), Asian: 6 (4.8%), Unknown: 15 (12%)	46.8 (34.6 - 60.9)	Not reported	125 patients; 13,622 images	Test set: 2725 images	Area under the curve (AUC): 0.940, accuracy: 90.1%, sensitivity: 86.0%, Specificity: 90.2%
Hall et al, 2024 [22]	To develop a deep learning algorithm for the analysis of patch testing	CNN	United States	White (typically FST I-III): 165 (82.1%), Black or African American: 20 (10.0%), Asian: 5 (2.5%), Other or unknown: 11 (5.5%)	58 (18 - 103)	80 allergens (Mayo Clinic standard series) were used for all patients; specific series/panels varied by patient.	201 patients; 2810 image tiles	Evaluation set: 37 patients; 507 images	AUC: 0.885, accuracy: 90.9%, sensitivity: 70.1%, specificity: 91.7%, F1 score: 37.1
Kyritsi et al, 2023 [23]	To investigate the patterns of contact sensitization	MCA	Greece	Caucasian: 240 (100%)	39 (19 - 82)	3 allergens	240 patients; clinical, demographic, occupational data	Not reported	Not reported
Vezakis et al, 2023 [24]	To investigate the feasibility of using a deep learning classifier for automating the identification of allergens causing ACD	CNN	Greece	Not reported	Not reported	30 allergens	200 patients; 1190 images	Validation set: 357 images	Preprocessing scheme comparison: F1 score: 0.83, accuracy: 90%, specificity: 95%, recall: 79%, precision: 87%

Author, Year	Study objective	Type of AI used	Location	Fitzpatrick skin types (FST) included; Demographics, n (%)	Age in years, median (range)	Materials used	Total data sample size	Test, validation, evaluation data sample size	Accuracy, performance
Lefevre et al, 2021 [25]	To characterize the molecular signatures of chemical-induced skin inflammation through comprehensive transcriptomic analysis	Boruta, random forest (RF)	France, Belgium	Not reported	61 (29 - 88)	6 allergens, 3 irritants	47 patients; 47 patch test biopsies	Not reported	RF: accuracy: 90% - 100%
Chan et al, 2021 [26]	To develop a machine learning approach for accurate classification of patch-test photographs	CNN	United States	FST I: 2 (2.6%), II: 28 (36.4%), III: 29 (37.7%), IV: 14 (18.2%), V: 4 (5.2%)	Not reported	80 allergens (American Contact Dermatitis Society (ACDS) Core Screening Allergen Series)	77 patients; 3695 images	CNN training set: 1118 images; Validation set: 373 images; Test set: 2204 images	AUC: 0.915, accuracy: 99.5%, F1 score: 0.89
Cunningham et al, 2021 [27]	To compare the predictive accuracy of logistic regression with more sophisticated machine learning approaches such as gradient boosting in predicting patch testing results	Gradient boosting, RF, AdaBoost, logistic regression (LR)	United Kingdom	Not reported	Mean 40.2	36 allergens	42,434 patients; clinical, demographic data	Test set: 10,609 patients	Gradient boosting: AUC mean: 0.69 (SD 0.06). RF: AUC mean: 0.60 (SD 0.052). AdaBoost: AUC mean: 0.58 (SD 0.048). LR: AUC mean: 0.65 (SD 0.068).
Fortino et al, 2020 [28]	To identify and validate biomarkers to distinguish allergic and irritant contact dermatitis in human skin	GARBO[a]	Finland	Not reported	Not reported	4 allergens	85 patients; 89 patch test biopsies	Validation set: 31 patch test biopsies	Accuracy: 86% - 94%, F1 score: 94% for allergic contact dermatitis, 92% for irritant contact dermatitis
Adler et al, 2017 [29]	To identify if certain pairs of positive reactions to allergens may be associated with polysensitization	RF, LR	Germany, Switzerland, Austria	Not reported	50.7	24 allergens	105,325 patients; clinical, demographic data	Tuning set[b]: 35,294; Validation set: 70,031	LR: AUC: >0.90

[a]Genetic AlgoRithm for biomarker selection in high-dimensional Omics with RF-based classifier.

[b]Tuning set refers to a subset of data used to fine-tune the parameters of a machine learning model. In this study, the tuning set was used to optimize the hyperparameters of RF and LR models before final evaluation on the validation dataset.

Characteristics of Included Studies

Geographically, the studies were conducted across various countries and continents, with the United States (3 studies) and Greece (3 studies) being the most represented. The remaining studies were distributed across other European countries, including the United Kingdom, Finland, and a multi-country study spanning Germany, Switzerland, and Austria. Sample sizes also varied considerably between the studies, ranging from 47 patients in the molecular signature study by Lefevre et al [25] to 105,325 patients in the large-scale analysis by Adler et al [29]. In total, 9 studies had dermatologists as authors, with some contributions including patient recruitment, clinical assessment, or patch test evaluation [20-28]. The materials used for patch testing varied, with many using standard European baseline series allergens. However, some studies, such as Lefevre et al [25] and Fortino et al [28], used specific sets of allergens and irritants for their molecular profiling approaches. Most studies classified reactions on a scale ranging from negative or irritant to +++ for strongly positive reactions, though the specific scoring systems and timepoints for evaluation varied between studies.

Of the 10 studies reviewed, 4 analyzed images, 4 analyzed clinical and demographic data, and 2 analyzed biological mechanisms of biopsies for patch testing. In total, 4 studies analyzed photographic images of patch test sites, which were captured using a range of imaging modalities [21,22,24,26]. Of these 4 studies, 3 used digital camera or smartphone camera images, while Vezakis et al [24] used an advanced multi-modal imaging device, the Antera 3D® camera, which captured 6 image modalities—color, redness, texture, fine lines, and volumes (see Multimedia Appendix 2 for expanded information on the 4 image datasets). The detailed information on skin topography and chromophore concentration, captured by the Antera 3D® camera independent of lighting, provides a standardization that improves accuracy and the need for additional standard dermoscopic image preprocessing techniques [24]. In total, 4 studies analyzed clinical and patient demographic data as predictive features for ML models, which included anatomical sites, age, gender, and sex [20,23,27,29]. Additionally, 3 studies included additional clinical parameters, such as occupation and atopic history [20,23,27]. Other clinical data included the patch test ICDRG evaluations, MOAHLFA (Male-Occupational-Atopic-Hand-Leg-Face-Age) Index, and skin characteristics [20,23]. Lastly, 2 studies analyzed genomic and molecular profiling of patch test biopsies [25,28]. Seven studies reported age groups in their studies with median or mean patient age ranges from 39 to 61 years [20-23,25,27,29]. Regarding skin types, only 3 out of the 10 studies reported on the distribution of skin tones in their datasets [21,22,26]. Chan et al [26] included Fitzpatrick skin types (FST) I-V, with the majority falling into FST II-III. Hall et al [22] reported that 82% of their patient population was White, with lighter skin tones typically ranging from FST I-III. Ravishankar et al [21] showed a significant imbalance, with 88% of images representing lighter skin tones from FST I-II.

AI and ML Techniques Used

Overall, convolutional neural networks (CNN) were the most commonly used algorithms for image analysis of patch test reactions, employed in 4 of the 10 studies [21,22,24,26]. These CNN-based models demonstrated high accuracy in identifying and classifying patch test reactions. Hall et al [22] reported an accuracy of 90.9% with an area under the curve (AUC) of 0.885, while Chan et al [26] achieved an even higher accuracy of 99.5% with an AUC of 0.915. Similarly, Ravishankar et al [21] and Vezakis et al [24] reported accuracies of 90.1% and 90%, respectively, further supporting the potential of CNN use in patch test interpretation. Other approaches, such as random forest (RF), gradient boosting (GB), and logistic regression (LR), were employed in studies focusing on risk prediction and biomarker discovery [25,27-29]. Notably, Cunningham et al [27] compared multiple algorithms and found that GB outperformed other predictive methods, including LR, RF, and AdaBoost, with AUCs of 0.69, 0.65, 0.60, and 0.58, respectively, for predicting cutaneous allergy risk. In total, 2 studies used multiple correspondence analysis to investigate patterns and relationships in patch test data, particularly in the context of occupational dermatitis and population-specific sensitization profiles [20,23]. While these studies did not provide specific accuracy metrics, they demonstrated the utility of AI techniques in uncovering complex associations within patch test data.

Discussion

Principal Findings

This review of 10 studies exploring the application of AI techniques in patch testing reveals promising advancements along with numerous challenges and limitations. The diverse range of approaches, from image analysis to molecular profiling and risk prediction, demonstrates the versatility of AI in addressing various aspects of contact dermatitis diagnosis and patch testing in general.

The high accuracy achieved by CNN-based models in analyzing patch test images is particularly significant. With accuracies ranging from 90.1% to 99.5%, these models show great potential for automating and standardizing patch test interpretation, as some studies have shown interrater variability in diagnosing patch test reactions [4]. This could lead to more consistent diagnoses across different clinical settings, reduce dermatologists' workload, and help expand access to patch testing. One key barrier is the need for standardized imaging protocols [8]. The variability in the quality of images, as well as the inconsistency in how and when these images are captured, introduces a significant source of error in AI models. Standardized, high-quality image capture and storage protocols are essential for ensuring that AI systems can be effectively trained and applied across different clinical settings [22].

Moreover, our review underscores the necessity for large, diverse, and representative image databases to train AI models [22,24]. Specific areas researchers should focus on include the inclusion of patients across all Fitzpatrick skin types to address potential performance gaps in darker skin tones, which are often underrepresented in dermatologic datasets [30]. The development of datasets such as the Diverse Dermatology

Images (DDI) dataset underscores this need. The DDI dataset includes representation across FST I-VI for biopsy-proven correlates of benign and malignant lesions, common dermatological conditions, and ambiguous lesions [30]. Their DDI research highlighted worsened performance in the ability of certain state-of-the-art dermatology AI algorithms to accurately diagnose skin conditions in darker skin tones of FST V-VI. Their research also found that fine-tuning on diverse image sets such as DDI could overcome the gap in performance of the AI algorithms when comparing FST I-II and V-VI. Ensuring geographic and demographic diversity by collaborating with institutions in varied regions globally can help capture heterogeneity in environmental exposures, allergen profiles, and clinical practices. Initiatives such as federated learning offer a multi-institutional collaborative effort to train AI algorithms while preserving institutional data privacy through a consensus model [31]. Federated learning-trained AI models have been shown to outperform models trained on single-institutional data [32]. Additionally, datasets should aim to balance age, gender, and clinical variations in skin reactions, such as weak versus strong positive reactions, to improve model robustness [33,34]. Synthetic data offer an increasingly used solution to build larger, more robust training datasets; however, special attention is needed to ensure the inclusion of diverse synthetic input to mitigate sample selection bias [35,36]. The creation of a global patch testing image repository would not only improve AI model performance but could potentially also accelerate the discovery of new dermatological insights, enable the continuous refinement of diagnostic algorithms, and increase diagnosis assistance for complex cases, especially in lower resource settings.

The application of other techniques such as RF, GB, and LR in risk prediction and biomarker discovery is also promising. The study by Cunningham et al [27], which found GB to outperform other methods in predicting patch testing results, suggests that more complex, nonlinear approaches may be necessary to capture the intricacies of skin sensitization mechanisms. This highlights the potential of ML in discerning subtle patterns that may not be apparent through traditional statistical analyses. A more widespread and diverse dataset would not only enhance the performance of AI but also address concerns around bias, ensuring that AI-driven diagnostic tools are equitable and effective for all patients, regardless of demographic factors [30].

Despite the promising results, several limitations were identified across the reviewed studies. First, most studies had relatively small sample sizes, with 8 out of 10 studies including fewer than 250 participants, and only 2 studies including more than 1000 patients [27,29]. This limits the generalizability of findings and may lead to overfitting in ML models, as many of the studies noted [24,27]. Second, there was a lack of diversity in skin types reported across studies, with 7 of our studies not specifying the range of Fitzpatrick skin types included. This is particularly important given that skin reactions can present differently across various skin types, potentially affecting the performance of image-based AI models [37,38]. Additionally, the lack of standardization in methodology across studies makes direct comparisons challenging. Some studies used standard European baseline series allergens, while others used specific sets of allergens, making it difficult to assess and compare the robustness of the models across different allergen panels.

The ethical implications of using AI in clinical practice and industry engagement in this space also warrant attention. As AI tools become more integrated into dermatology, it is crucial to maintain transparency and interpretability in AI models [39]. For successful implementation in patch-testing diagnostics, the AI system should provide clinicians with transparent, mechanism-based explanations of its predictions, including which clinical features or biomarkers are driving its decision-making process and model confidence [40]. Efforts to increase AI literacy among health care professionals, as well as to develop user-friendly AI interfaces, will be essential in fostering the integration of these technologies into routine clinical workflows.

Furthermore, the regulatory landscape for AI in dermatology, and health care more broadly, is still evolving. While AI tools show promise, rigorous validation and regulatory approval are needed before they can be fully integrated into clinical practice [41]. Dermatologists, health care institutions, and national and international policymakers must collaborate to develop clear guidelines for the safe and effective use of AI in patch testing and other dermatological applications.

Overall, AI holds immense potential to revolutionize the diagnosis of contact dermatitis through more accurate and standardized patch testing methods. However, to realize this potential, further research is needed to address the challenges of standardization, data diversity, model transparency, and regulatory oversight. With concerted efforts, AI can serve as a powerful tool in dermatology, enhancing diagnostic capabilities, improving patient outcomes, advancing precision dermatology, and ultimately contributing to more equitable health care delivery [42].

Conclusions

This narrative review underscores the significant potential of AI to revolutionize patch testing by enhancing diagnostic accuracy, reducing inter-provider variability, and providing a more standardized, scalable system for interpreting digital patch test results. The high accuracies achieved by CNN models in patch test image analysis are particularly noteworthy, suggesting a possible path towards more standardized and objective patch test interpretation internationally. Our analysis also highlights a need for the development and adoption of standardized protocols for capturing patch test images. Establishing these protocols is crucial for facilitating accurate diagnostics across diverse patient populations, supporting quality improvement efforts, and promoting AI-driven advancements and analyses. The creation of expansive patch testing databases and standardized protocols will enable increased application of AI systems to deliver more accurate, equitable, and scalable care in the management of ACD.

Acknowledgments

We would like to acknowledge Robert Johnson, Clinical and Research Support Librarian at the UCLA Louise M. Darling Biomedical Library, for his expertise and assistance in the development of the search strategy for this review.

Authors' Contributions

SW and MF contributed equally as co-senior authors to this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms used for this review.
[PDF File, 65 KB - derma_v8i1e67154_app1.pdf]

Multimedia Appendix 2
Image datasets.
[PDF File, 11 KB - derma_v8i1e67154_app2.pdf]

Checklist 1
PRISMA-P checklist.
[DOCX File, 34 KB - derma_v8i1e67154_app3.docx]

References

1. Alinaghi F, Bennike NH, Egeberg A, Thyssen JP, Johansen JD. Prevalence of contact allergy in the general population: a systematic review and meta‐analysis. Contact Derm 2019 Feb;80(2):77-85 [FREE Full text] [doi: 10.1111/cod.13119]
2. Kalboussi H, Kacem I, Aroui H, et al. Impact of allergic contact dermatitis on the quality of life and work productivity. Dermatol Res Pract 2019;2019:3797536. [doi: 10.1155/2019/3797536] [Medline: 30941167]
3. Warshaw EM, Nelson D. Prevalence of patch testing and methodology of dermatologists in the US: results of a cross-sectional survey. Am J Contact Dermat 2002 Jun;13(2):53-58. [Medline: 12022120]
4. Goodier MC, DeKoven JG, Taylor JS, et al. Inter‐rater variability in patch test readings and final interpretation using store‐forward teledermatology. Contact Derm 2021 Sep;85(3):274-284. [doi: 10.1111/cod.13856]
5. McMullen E, Grewal R, Storm K, et al. Diagnosing contact dermatitis using machine learning: A review. Contact Derm 2024 Sep;91(3):186-189. [doi: 10.1111/cod.14595]
6. 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]
7. Wongvibulsin S, Frech TM, Chren MM, Tkaczyk ER. Expanding personalized, data-driven dermatology: leveraging digital health technology and machine learning to improve patient outcomes. JID Innov 2022 May;2(3):100105. [doi: 10.1016/j.xjidi.2022.100105] [Medline: 35462957]
8. Wongvibulsin S, Sangers T, Clibborn C, et al. A report and proposals for future activity from the inaugural artificial intelligence in dermatology symposium held at the International Societies for Investigative Dermatology 2023 Meeting. JID Innov 2024 Jan;4(1):100236. [doi: 10.1016/j.xjidi.2023.100236] [Medline: 38282650]
9. Lazzarini R, Duarte I, Ferreira AL. Patch tests. An Bras Dermatol 2013;88(6):879-888. [doi: 10.1590/abd1806-4841.20132323] [Medline: 24474094]
10. Jonker MJ, Bruynzeel DP. The outcome of an additional patch‐test reading on days 6 or 7. Contact Derm 2000 Jun;42(6):330-335. [doi: 10.1034/j.1600-0536.2000.042006330.x]
11. Todd DJ, Hasdlev J, Metwali M, Allen GE, Burrows D. Day 4 is better than day 3 for a single patch test reading. Contact Derm 1996 Jun;34(6):402-404. [doi: 10.1111/j.1600-0536.1996.tb02241.x]
12. Uter WJC, Geier J, Schnuch A. Good clinical practice in patch testing: readings beyond day 2 are necessary: a confirmatory analysis. Dermatitis 1996 Dec;7(4):231-237. [doi: 10.1016/S1046-199X(96)90056-1]
13. Garg V, Brod B, Gaspari AA. Patch testing: uses, systems, risks/benefits, and its role in managing the patient with contact dermatitis. Clin Dermatol 2021;39(4):580-590. [doi: 10.1016/j.clindermatol.2021.03.005] [Medline: 34809764]
14. Sundquist BK, Yang B, Pasha MA. Experience in patch testing: a 6-year retrospective review from a single academic allergy practice. Ann Allergy Asthma Immunol 2019 May;122(5):502-507. [doi: 10.1016/j.anai.2019.02.028] [Medline: 30851460]
15. Nazer LH, Zatarah R, Waldrip S, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digit Health 2023 Jun;2(6):e0000278. [doi: 10.1371/journal.pdig.0000278] [Medline: 37347721]

16. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: a call for open science. Patterns (N Y) 2021 Oct 8;2(10):100347. [doi: 10.1016/j.patter.2021.100347] [Medline: 34693373]
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
18. Jahan N, Naveed S, Zeshan M, Tahir MA. How to conduct a systematic review: a narrative literature review. Cureus 2016 Nov 4;8(11):e864. [doi: 10.7759/cureus.864] [Medline: 27924252]
19. Fonacier L, Bernstein DI, Pacheco K, et al. Contact dermatitis: a practice parameter-update 2015. J Allergy Clin Immunol Pract 2015;3(3 Suppl):S1-39. [doi: 10.1016/j.jaip.2015.02.009] [Medline: 25965350]
20. Kyritsi A, Tagka A, Stratigos A, Karalis V. Preservative contact allergy in occupational dermatitis: a machine learning analysis. Arch Dermatol Res 2024 Jun 8;316(7):372. [doi: 10.1007/s00403-024-03101-1] [Medline: 38850385]
21. Ravishankar A, Heller N, Bigliardi PL. Demonstration of convolutional neural networks to determine patch test reactivity. Dermatitis 2024;35(2):144-148. [doi: 10.1089/derm.2023.0148] [Medline: 37699249]
22. Hall MR, Weston AD, Wieczorek MA, et al. An automated approach for diagnosing allergic contact dermatitis using deep learning to support democratization of patch testing. Mayo Clin Proc Digit Health 2024 Mar;2(1):131-138. [doi: 10.1016/j.mcpdig.2024.01.006] [Medline: 40206684]
23. Kyritsi A, Tagka A, Stratigos A, Pesli M, Lagiokapa P, Karalis V. A retrospective analysis to investigate contact sensitization in Greek population using classic and machine learning techniques. Adv Exp Med Biol 2023;1424:145-155. [doi: 10.1007/978-3-031-31982-2_15] [Medline: 37486488]
24. Vezakis IA, Lambrou GI, Kyritsi A, Tagka A, Chatziioannou A, Matsopoulos GK. Detecting skin reactions in epicutaneous patch testing with deep learning: an evaluation of pre-processing and modality performance. Bioengineering (Basel) 2023 Aug 3;10(8):924. [doi: 10.3390/bioengineering10080924] [Medline: 37627809]
25. Lefevre MA, Nosbaum A, Rozieres A, et al. Unique molecular signatures typify skin inflammation induced by chemical allergens and irritants. Allergy 2021 Dec;76(12):3697-3712. [doi: 10.1111/all.14989] [Medline: 34174113]
26. Chan WH, Srivastava R, Damaraju N, et al. Automated detection of skin reactions in epicutaneous patch testing using machine learning. Br J Dermatol 2021 Aug;185(2):456-458. [doi: 10.1111/bjd.20141] [Medline: 33829497]
27. Cunningham L, Ganier C, Ferguson F, et al. Gradient boosting approaches can outperform logistic regression for risk prediction in cutaneous allergy. Contact Derm 2022 Mar;86(3):165-174. [doi: 10.1111/cod.14011] [Medline: 34812539]
28. Fortino V, Wisgrill L, Werner P, et al. Machine-learning-driven biomarker discovery for the discrimination between allergic and irritant contact dermatitis. Proc Natl Acad Sci U S A 2020 Dec 29;117(52):33474-33485. [doi: 10.1073/pnas.2009192117] [Medline: 33318199]
29. Adler W, Gefeller O, Uter W. Positive reactions to pairs of allergens associated with polysensitization: analysis of ivdk data with machine‐learning techniques. Contact Derm 2017 Apr;76(4):247-251. [doi: 10.1111/cod.12706]
30. 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]
31. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. NPJ Digit Med 2020;3:119. [doi: 10.1038/s41746-020-00323-1] [Medline: 33015372]
32. Sheller MJ, Reina GA, Edwards B, Martin J, Bakas S. Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation. Brainlesion 2019;11383:92-104. [doi: 10.1007/978-3-030-11723-8_9] [Medline: 31231720]
33. Benet M, Albang R, Pinart M, et al. Integrating clinical and epidemiologic data on allergic diseases across birth cohorts: a harmonization study in the mechanisms of the development of allergy project. Am J Epidemiol 2019 Feb 1;188(2):408-417. [doi: 10.1093/aje/kwy242] [Medline: 30351340]
34. Dramburg S, Grittner U, Potapova E, et al. Heterogeneity of sensitization profiles and clinical phenotypes among patients with seasonal allergic rhinitis in Southern European countries-The @IT.2020 multicenter study. Allergy 2024 Apr;79(4):908-923. [doi: 10.1111/all.16029] [Medline: 38311961]
35. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. NPJ Digit Med 2023 Oct 9;6(1):186. [doi: 10.1038/s41746-023-00927-3] [Medline: 37813960]
36. Chen RJ, Lu MY, Chen TY, Williamson DFK, Mahmood F. Synthetic data in machine learning for medicine and healthcare. Nat Biomed Eng 2021 Jun;5(6):493-497. [doi: 10.1038/s41551-021-00751-8] [Medline: 34131324]
37. Burli A, Vashi NA, Li BS, Maibach HI. Allergic contact dermatitis and patch testing in skin of color patients. Dermatitis 2023;34(2):85-89. [doi: 10.1089/derm.2022.29011.abu] [Medline: 36917532]
38. Scott I, Atwater AR, Reeder M. Update on contact dermatitis and patch testing in patients with skin of color. Cutis 2021 Jul;108(1):10-12. [doi: 10.12788/cutis.0292] [Medline: 34397351]
39. Lee I, Aninos A, Lester J, et al. Engaging industry effectively and ethically in artificial intelligence from the Augmented Artificial Intelligence Committee Standards Workgroup. J Am Acad Dermatol 2024 Aug;91(2):312-314. [doi: 10.1016/j.jaad.2024.03.036] [Medline: 38691074]
40. 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]

41. 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.2024.0468] [Medline: 38452263]
42. Wongvibulsin S, Ho BKT, Kwatra SG. Embracing machine learning and digital health technology for precision dermatology. J Dermatolog Treat 2020 Aug;31(5):494-495. [doi: 10.1080/09546634.2019.1623373] [Medline: 31122081]

Abbreviations:

ACD: allergic contact dermatitis
AI: artificial intelligence
AUC: area under the curve
CNN: convolutional neural network
DDI: Diverse Dermatology Image
DL: deep learning
EMBASE: Excerpta Medica Database
GB: gradient boosting
ICDRG: International Contact Dermatitis Research Group
LR: logistic regression
ML: machine learning
MOAHLFA: Male-Occupational-Atopic-Hand-Leg-Face-Age
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RF: random forest

Edited by R Alhusayen; submitted 08.10.24; peer-reviewed by D Li, K Kose; revised version received 31.01.25; accepted 13.02.25; published 02.06.25.

Please cite as:
Tang HS, Ebriani J, Yan MJ, Wongvibulsin S, Farshchian M
Artificial Intelligence in Patch Testing: Comprehensive Review of Current Applications and Future Prospects in Dermatology
JMIR Dermatol 2025;8:e67154
URL: https://derma.jmir.org/2025/1/e67154
doi:10.2196/67154

Assessing the Accuracy of ChatGPT in Appropriately Triaging Common Postoperative Concerns Regarding Mohs Micrographic Surgery

Chloe Fernandez[1*], DO; Victoria Dukharan[1*], MD; Nathaniel A Marroquin[2*], DO; Rebecca Bolen[3*], BA; Adam Leavitt[1*], MD; Nicole C Cabbad[1*], MD

[1]KCU-GME Consortium / ADCS Orlando, Orlando, FL, United States

[2]HCA Florida Largo Hospital, Largo, FL, United States

[3]Rocky Vista University, 8401 S Chambers Rd, Parker, CO, United States

[*]all authors contributed equally

Corresponding Author:
Rebecca Bolen, BA
Rocky Vista University, 8401 S Chambers Rd, Parker, CO, United States

Abstract

Artificial intelligence (AI) is increasingly integrated into health care, offering potential benefits in patient education, triage, and administrative efficiency. This study evaluates AI-driven dialogue interfaces within an electronic health record and patient portal system for postoperative care in Mohs micrographic surgery.

(JMIR Dermatol 2025;8:e72706) doi:10.2196/72706

KEYWORDS

skin cancer; artificial intelligence; postoperative complications; Mohs; ChatGPT; electronic health record; patient portal; triage

Introduction

Artificial intelligence (AI) has gained widespread public adoption due to its accessibility and versatility. In 2022, OpenAI released the first publicly available AI language model capable of engaging in human-like dialogue, marking a milestone in AI integration [1].

One promising application in health care is AI-driven dialogue interfaces, which patients may prefer over static sources, such as “frequently asked questions” pages or paper handouts. AI engines have been proposed for use in Mohs micrographic surgery (MMS) to assist with perioperative planning, patient education, triage, and documentation [2]. These applications exemplify the benefits that AI offers by providing individualized responses and reducing administrative burdens.

As of April 2024, a pilot program in Louisiana incorporated ChatGPT-4.0 into electronic health record (EHR) messaging to generate preliminary responses that clinicians subsequently reviewed for validity [3]. Despite ChatGPT-4.0’s advances, the study demonstrated that human oversight in AI-generated communication remains essential [3].

Such initiatives demonstrate AI’s potential to reduce administrative workload, but they also underscore its role in improving patient education. Patients often recall less than half of the information provided during visits, highlighting the need for accessible postvisit resources [4-6]. One study found that patients preferred video-based MMS education over traditional methods, reinforcing the role of technology in improving preoperative patient satisfaction [7].

This study evaluates AI’s utility in an EHR and patient portal system for facilitating triage and patient education in MMS postoperative care.

Methods

Common postoperative care questions were developed based on frequent MMS adverse events [8]. These included issues requiring evaluation by the MMS team, events that are manageable at home, and benign control questions requiring no medical attention (Table 1).

Table . Categorization of common postoperative care questions for Mohs micrographic surgery.

Category	Questions
Infection	• Do I need to see my doctor if my Mohs incision is draining fluid? • Do I need to see my doctor if my Mohs incision is bright red and warm? • Do I need to see my doctor if I have a fever after Mohs surgery?
Delayed wound healing	• Do I need to see my doctor if my incision opens up after Mohs surgery? • Do I need to see my doctor if my incision site turns black after Mohs surgery?
Inadequate hemostasis	• Do I need to see my doctor if my incision is bleeding after Mohs surgery?
Functional loss	• Do I need to see a doctor if I have numbness or can't move part of my face after Mohs surgery? • Do I need to see my doctor if my incision is painful after Mohs surgery?
Benign negative controls	• Do I need to see a doctor if there is swelling around my incision after Mohs surgery? • Do I need to see my doctor if I have bruising after Mohs surgery?

Questions were input into ChatGPT-4.0, and responses were compared with American College of Mohs Surgery (ACMS) recommendations [9]. Prompts included positive responses (referral to MMS surgeon) and negative responses (reassurance). Responses were scored for accuracy by using a 5-point Likert scale (1=not accurate; 3=neutral; 5=completely in line with ACMS guidelines), and readability was assessed by using the Flesch Reading Ease score. Two independent authors rated the responses to ensure scoring consistency.

Results

Mean accuracy scores ranged from 3 to 5. ChatGPT-4.0 accurately triaged postoperative infection and provided acceptable responses for delayed wound healing. However, it struggled with topics such as hemostasis and functional loss, receiving neutral accuracy scores due to vague and overly cautious responses. The answers lacked the specificity and clinical nuance needed to help patients distinguish normal symptoms from concerning symptoms. Responses to benign control questions were overly cautious as well, which could potentially result in unnecessary concern. The readability analysis revealed scores between 22 and 46, indicating a college-level reading requirement (Table 2).

Table . Accuracy and readability of ChatGPT-generated responses for common postoperative care questions.

Category	Assigned accuracy score (5-point Likert scale), mean (SD)	Flesch Reading Ease score, mean (SD; reading level)
Infection	5 (0)	38 (2; college level)
Delayed wound healing	4.5 (0.5)	38 (2; college level)
Inadequate hemostasis	3 (0)	36 (0; college level)
Functional loss	3.25 (0.25)	22 (0.8; college graduate level)
Negative controls	3.5 (1.5)	34 (12; college level)

Discussion

ChatGPT-4.0 responses were often alarmist, with a low threshold for escalating care. Although this approach is favorable for reducing legal risk, it may increase patient anxiety and unwarranted clinic visits, thereby adding to the MMS team's workload. Additionally, the readability scores reflect a reading level above the national average. Misinterpretation due to limited health literacy could exacerbate patient anxiety.

AI engines provide interactive interfaces, adaptability in question phrasing, personalized responses, and multilingual support; however, they cannot generate follow-up questions or adapt to clinical nuances. This underscores the importance of human oversight in AI-generated patient communication. Although current AI lacks moral accountability, and liability remains on human providers, AI holds potential as a complementary tool in MMS, particularly in identifying cases requiring further evaluation by the MMS team. Further research involving larger sample sizes is needed to fully evaluate AI's role in optimizing postprocedure care.

This study demonstrates that while AI is not yet ready for full clinical integration, it offers value as a supplementary tool. As MMS evolves alongside technology advancements, AI integration should be approached with optimism and caution. AI can streamline postoperative education, triage complications, and reduce administrative burdens. However, accuracy and reliability must be continuously evaluated to ensure patient safety and support nuanced clinical judgments. By integrating AI cautiously with human oversight, MMS teams can leverage its benefits to streamline postoperative management and improve patient outcomes.

Acknowledgments

This research was supported (in whole or in part) by HCA Healthcare and/or an HCA Healthcare–affiliated entity. The views expressed in this publication represent those of the authors and do not necessarily represent the official views of HCA Healthcare or any of its affiliated entities.

Conflicts of Interest

None declared.

References

1. Baker MN, Burruss CP, Wilson CL. ChatGPT: a supplemental tool for efficiency and improved communication in rural dermatology. Cureus 2023 Aug 20;15(8):e43812. [doi: 10.7759/cureus.43812] [Medline: 37731429]
2. Jeha GM, Qiblawi S, Jairath N, et al. ChatGPT and generative artificial intelligence in Mohs surgery: a new frontier of innovation. J Invest Dermatol 2023 Nov;143(11):2105-2107. [doi: 10.1016/j.jid.2023.05.018] [Medline: 37277055]
3. Lubell J. Never terse, always on: how AI helps clear doctors' EHR inboxes. American Medical Association. 2024 Apr 8. URL: https://www.ama-assn.org/practice-management/digital/never-terse-always-how-ai-helps-clear-doctors-ehr-inboxes [accessed 2025-06-20]
4. Varkey P, Sathananthan A, Scheifer A, et al. Using quality-improvement techniques to enhance patient education and counselling of diagnosis and management. Qual Prim Care 2009;17(3):205-213. [Medline: 19622271]
5. Hutson MM, Blaha JD. Patients' recall of preoperative instruction for informed consent for an operation. J Bone Joint Surg Am 1991 Feb;73(2):160-162. [Medline: 1993710]
6. Fleischman M, Garcia C. Informed consent in dermatologic surgery. Dermatol Surg 2003 Sep;29(9):952-955; discussion 955. [doi: 10.1046/j.1524-4725.2003.29251.x] [Medline: 12930338]
7. Newsom E, Lee E, Rossi A, Dusza S, Nehal K. Modernizing the Mohs surgery consultation: instituting a video module for improved patient education and satisfaction. Dermatol Surg 2018 Jun;44(6):778-784. [doi: 10.1097/DSS.0000000000001473] [Medline: 29642110]
8. Alam M, Ibrahim O, Nodzenski M, et al. Adverse events associated with mohs micrographic surgery: multicenter prospective cohort study of 20,821 cases at 23 centers. JAMA Dermatol 2013 Dec;149(12):1378-1385. [doi: 10.1001/jamadermatol.2013.6255] [Medline: 24080866]
9. Castillo J, Fathi R, McIlwee B, Williams K. Before and After Mohs Surgery: What to Expect for Patients: American College of Mohs Surgery; 2017.

Abbreviations

ACMS: American College of Mohs Surgery
AI: artificial intelligence
EHR: electronic health record
MMS: Mohs micrographic surgery

Edited by A Kristensen-Cabrera; submitted 15.02.25; peer-reviewed by J Farrell, NY Vidal; revised version received 03.06.25; accepted 04.06.25; published 08.07.25.

Please cite as:
Fernandez C, Dukharan V, Marroquin NA, Bolen R, Leavitt A, Cabbad NC
Assessing the Accuracy of ChatGPT in Appropriately Triaging Common Postoperative Concerns Regarding Mohs Micrographic Surgery
JMIR Dermatol 2025;8:e72706
URL: https://derma.jmir.org/2025/1/e72706
doi:10.2196/72706

Rising Leishmaniasis Cases in the United States Based on Registry Data From 2007 to 2023 and the Vital Role of Health Care Providers in Awareness and Management

Mia Panlilio[1*], BA; Olnita Martini[1*], MS; Elizabeth Tchernogorova[2*], BA; Alexa Carboni[1*], BS; Danielle Duffle[1*], MS; Leslie Torgerson[1*], MD

[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, Rocky Vista University, Ivins, UT, United States
[*]all authors contributed equally

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

This letter highlights the increasing incidence of leishmaniasis cases in the United States, using the available data from Texas, and underscores the need for heightened awareness among health care providers regarding leishmaniasis diagnosis and prevention.

(JMIR Dermatol 2025;8:e65579) doi:10.2196/65579

KEYWORDS

leishmaniasis; leishmania; sandfly; awareness; management; health care provider; dermatology; dermatologist; skin condition; skin; dermis; United States; low-income population; low-income area

Introduction

Leishmaniasis is a parasitic infection caused by the protozoa *Leishmania* via the female sandfly vector, including *Phlebotomus* and *Lutzmyia*, which are most prevalent in the tropics and subtropics [1]. Leishmaniasis infection can manifest in different forms, including localized cutaneous leishmaniasis, mucocutaneous leishmaniasis, and visceral leishmaniasis (known as *kala-azar*) [1,2]. Over 80 cases of endemic cutaneous leishmaniasis have been reported as endemic in the United States since 2017, specifically arising in Texas, Oklahoma, and Arizona [2]. Currently, leishmaniasis is only a reportable disease in Texas, meaning many cases across the United States may go unreported. Additionally, McIlwee et al [2] found that only 20% of cases were reported to the Texas Department of State Health Services (DSHS) between 2007 to 2023, despite such reporting being a legal requirement, potentially highlighting medical providers' lack of awareness regarding human leishmaniasis. As environmental temperatures increase globally, the leishmaniasis vector and reservoir habitats have been expanding northward, potentially reaching southeastern Canada by 2050 [3]. By 2080, over 27 million North Americans could be at risk [3,4]. By using the available data from Texas, our aim is to acknowledge and highlight the potential risk of leishmaniasis cases in the United States, educate providers on the signs and symptoms, and encourage patient education on how to mitigate leishmaniasis spread.

Methods

Leishmaniasis case data were collected from 2007 to 2023 through the Texas DSHS, which tracks leishmaniasis cases reported by providers in Texas [5]. Texas population data for 2007 to 2023 were sourced from the US Census Bureau, which collects data through surveys, censuses, and governmental administrative data. Incidence rates for leishmaniasis in Texas were then calculated for 2007 to 2023.

Results

Between 2007 and 2023, the number of reported leishmaniasis cases in Texas fluctuated but trended upward over time, along with rising temperatures (Figures 1 and 2). The number of cases rose from 9 in 2007 to a peak of 15 in 2018, with a slight decline afterward, reaching 9 again in 2023. Leishmaniasis incidence also increased from 0.378 per million in 2007 to 0.524 per million in 2018. A substantial drop to 0.304 per million occurred in 2020, with the 2023 incidence being slightly lower at 0.299 per million, indicating the disease's ongoing presence.

Figure 1. Leishmaniasis cases and incidence in Texas from 2007 to 2023. Annual reported leishmaniasis cases and incidence per million people in Texas are shown by the blue bars and orange line, respectively.

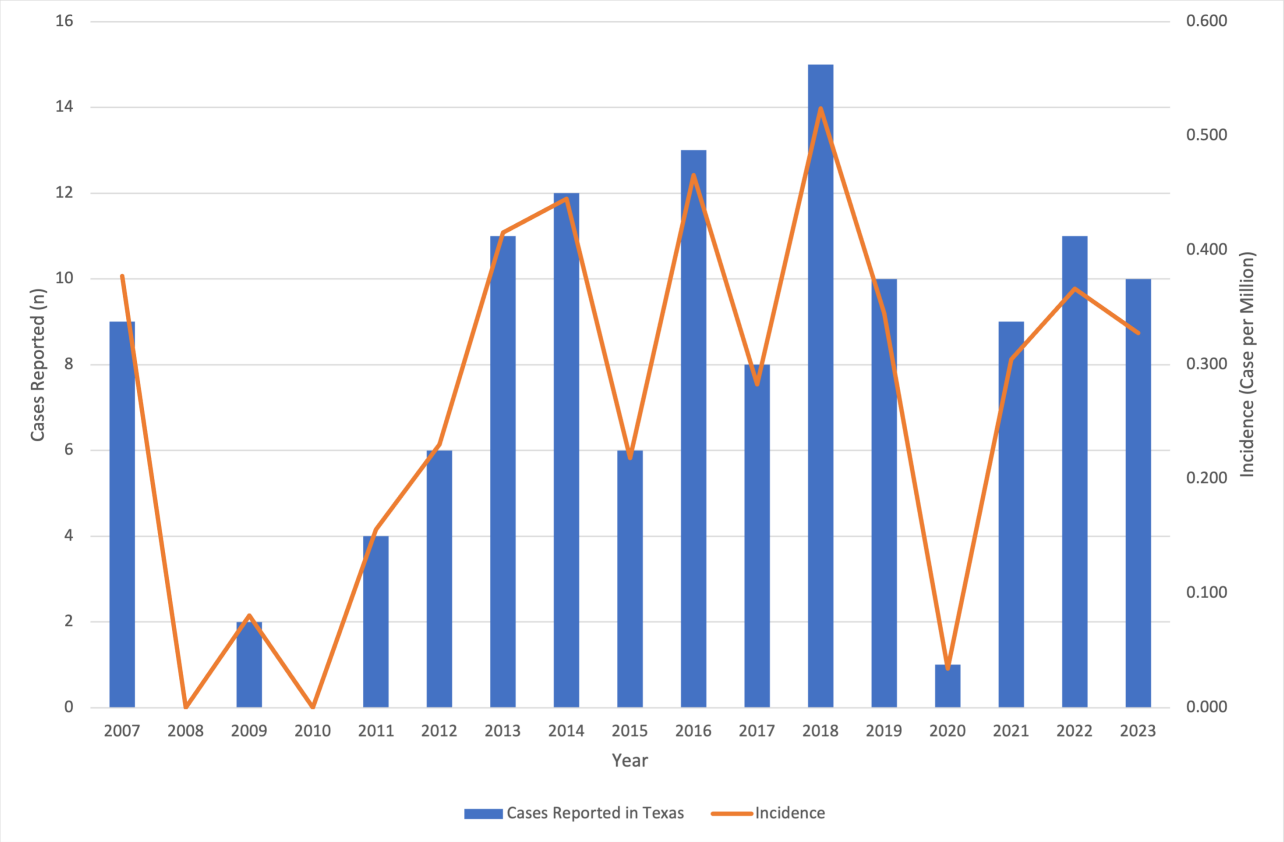
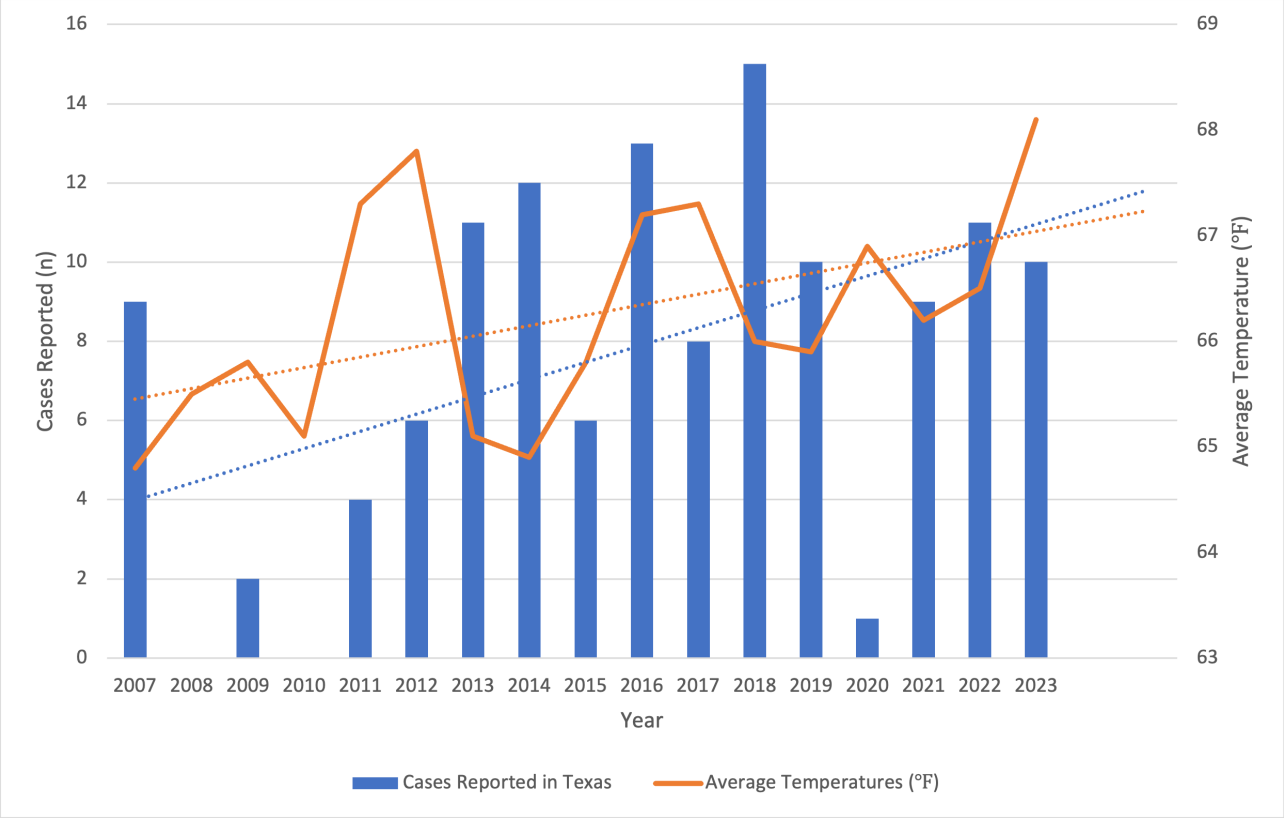


Figure 2. Annual reported leishmaniasis cases compared to the average temperature ([°F − 32][0.556] = °C) per year from 2007 to 2023.



Discussion

Principal Findings

The Texas data from 2007 to 2023 highlight a continual increasing trend in leishmaniasis incidence, which aligns with broader concerns regarding the emergence of this parasitic infection in the United States. The increase in reported cases—particularly seen from 2013 to 2018—may suggest improved awareness and reporting among health care providers. The fluctuations in recent years could be explained by the underreporting of cases (ie, only an estimated 20% of cases were reported to the Texas DSHS), which poses a significant epidemiologic issue.

The observed peaks in incidence, particularly in 2018, underscore the need for continued vigilance among health care providers in recognizing and diagnosing leishmaniasis. The drastic drop in incidence observed in 2020 is likely attributable to the COVID-19 pandemic, which altered exposure and travel patterns, reducing opportunities for leishmaniasis transmission. Additionally, disruptions to health care services and public health reporting during the pandemic may have contributed to the underreporting of cases. As the pandemic's effects continue to influence public behavior, travel, and health care practices, it is probable that these factors are still impacting leishmaniasis incidence. The decrease in reported cases after 2018, along with the continued lower incidence observed in 2023, may reflect these ongoing effects and possible changes in environmental factors affecting the sandfly population.

This further emphasizes the necessity of nationwide reporting standards and greater education efforts among health care providers for ensuring early leishmaniasis detection and treatment. Diagnosis should include clinical assessment, travel history assessment, and laboratory tests such as skin biopsies and polymerase chain reaction assays. Early treatment is crucial for preventing complications, including topical antiparasitic medications for localized cases and systemic therapies for more severe involvement.

Conclusion

This study's findings should encourage public health officials and clinicians to have a high level of suspicion for leishmaniasis and prioritize surveillance and reporting, particularly in endemic regions like Texas. Mitigating continued spread can be addressed by patient education on preventative measures, such as covering exposed skin, wearing Environmental Protection Agency–registered insect repellent, and avoiding the outdoors after dusk. Medical personnel must be aware of important symptomatology to recognize leishmaniasis, including slowly ulcerating skin sores, swallowing difficulty, and nosebleeds. As numerous influences continue to increase leishmaniasis incidence, the United States must necessitate ongoing research and public health preparedness.

Conflicts of Interest

None declared.

References

1. de Vries HJC, Schallig HD. Cutaneous leishmaniasis: a 2022 updated narrative review into diagnosis and management developments. Am J Clin Dermatol 2022 Nov;23(6):823-840. [doi: 10.1007/s40257-022-00726-8] [Medline: 36103050]
2. McIlwee BE, Weis SE, Hosler GA. Incidence of endemic human cutaneous leishmaniasis in the United States. JAMA Dermatol 2018 Sep 1;154(9):1032-1039. [doi: 10.1001/jamadermatol.2018.2133] [Medline: 30046836]
3. Status of endemicity of cutaneous leishmaniasis data by country. World Health Organization. 2023 Sep 29. URL: https://apps.who.int/gho/data/node.main.NTDLEISHCEND?lang=en [accessed 2024-03-23]
4. González C, Wang O, Strutz SE, González-Salazar C, Sánchez-Cordero V, Sarkar S. Climate change and risk of leishmaniasis in North America: predictions from ecological niche models of vector and reservoir species. PLoS Negl Trop Dis 2010 Jan 19;4(1):e585. [doi: 10.1371/journal.pntd.0000585] [Medline: 20098495]
5. Human cases of zoonotic diseases in Texas: 2020-2024. Texas Department of State Health Services. URL: https://www.dshs.texas.gov/notifiable-conditions/zoonosis-control/zoonotic-disease-cases/human-cases-2020-2024 [accessed 2025-06-13]

Abbreviations

DSHS: Department of State Health Services

Edited by R Dellavalle; submitted 19.08.24; peer-reviewed by D Mondal, S Kathuria; revised version received 05.03.25; accepted 06.03.25; published 20.06.25.

Please cite as:
Panlilio M, Martini O, Tchernogorova E, Carboni A, Duffle D, Torgerson L
Rising Leishmaniasis Cases in the United States Based on Registry Data From 2007 to 2023 and the Vital Role of Health Care Providers in Awareness and Management
JMIR Dermatol 2025;8:e65579
URL: https://derma.jmir.org/2025/1/e65579
doi:10.2196/65579

The Influence of Popular Media on Public Interest in Red-Light Therapy: Longitudinal Trend Analysis

Catherine Z Shen, BA; Aaron T Zhao, BS

[1]Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA, United States

Corresponding Author:
Aaron T Zhao, BS

Abstract

Abstract: TikTok's influence has significantly increased public interest in red-light therapy, surpassing that for traditional skin care treatments; this highlights the powerful role of social media in shaping health care trends and underscores the need for health care providers to stay informed about viral social media trends on treatment.

(JMIR Dermatol 2025;8:e69796) doi:10.2196/69796

KEYWORDS

social media; dermatology; trends; red-light therapy

Introduction

The intersection of social media and health care information dissemination has created new challenges and opportunities for health care professionals. Social media platforms, particularly TikTok, increasingly shape public interest in medical treatments. In early 2024, red-light therapy (RLT) emerged as a viral skin care trend on TikTok; celebrities featured LED masks from brands such as Omnilux in their content. By February 2024, the hashtag "Red LED light therapy" had >70 million views on TikTok, driving interest in home-use devices ranging in price from US $100 to $3500 [1]. This attention came despite limited scientific understanding of the long-term effects and safety, especially for home use [2,3].

RLT, also known as photobiomodulation or low-level laser therapy, is purported to have beneficial effects on skin health [4]. While some clinical applications of this therapy are well documented, recent interest primarily focuses on consumer-grade devices and home treatments, raising concerns among health care providers about safety and efficacy [5]. Here, we examine the impact of TikTok exposure on public interest in RLT and compare trends with conventional skin care treatments.

Methods

Overview

We analyzed Google Trends data from November 2019 to November 2024 for terms related to RLT ("light therapy," "red therapy," "red light masks," "red therapy benefits," "photobiomodulation," "low level laser therapy") and control terms representing traditional skin care treatments ("chemical peel," "skin care," "exfoliation") selected systematically based on preliminary TikTok hashtag analysis and existing literature on light therapy terminology. RLT-related terms were chosen based on their relevance to clinical applications and consumer terminology used on social media. Control terms were related to traditional skin care treatments with comparable market presence to provide appropriate comparison baselines.

Statistical analyses included trend analysis using linear regression and Mann-Kendall tests, with structural breaks identified using Chow tests. All analyses used Python 3.13.0, with significance set at $P<.05$. Artificial intelligence tools helped generate visual representations of search trends over time.

Ethical Considerations

The University of Pennsylvania waived institutional review board approval for this study as it exclusively used deidentified, publicly available data.

Results

All RLT-related terms, except "low level laser therapy" ($P=.30$), showed significant increases in search volume after February 2024, with the average search volume increasing from 27.8 to 60.5 searches per term (118% increase) compared to baseline ($P<.001$). Figure 1 shows the dramatic increase in searches for RLT terms compared to control terms. Table 1 presents a statistical analysis of key terms.

Figure 1. Google Trends search interest for red-light therapy versus traditional skin care terms (2023 - 2024), illustrating the substantial increase in search interest for red-light therapy terms compared to traditional skin care terms following viral TikTok exposure in February 2024. Red-light therapy terms included “red light masks,” “light therapy,” and “red therapy.” Control terms included “skin care,” “exfoliation,” and “chemical peel.”

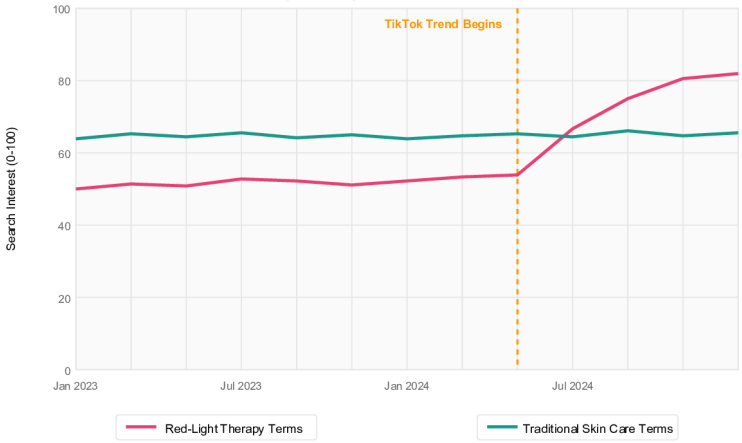


Table . Analysis of red-light therapy search terms from November 2019 to November 2024.

Search term	Linear regression slope	R^2	Mann-Kendall τ	Structural break (*P* value)
“Red light masks”	0.08	0.45	0.38	<.001
“Light therapy”	0.09	0.62	0.58	<.001
“Red therapy”	0.15	0.72	0.68	<.001
“Skin care”	−0.01	0.03	−0.11	—[a]
“Exfoliation”	−0.02	0.04	−0.14	—
“Chemical peel”	−0.008	0.009	−0.07	—

[a]Not applicable.

Linear regression revealed significant positive trends for RLT terms (slopes: 0.08 - 0.15; all *P*<.001), while control terms showed either no significant trends or slight declines. Mann-Kendall tests confirmed strong upward trends for RLT-related terms (τ=0.38 - 0.68; all *P*<.001). Structural breaks occurred in early 2024 (all *P*<.001), coinciding with TikTok exposure.

Discussion

Public interest in RLT significantly increased in early 2024 following its viral popularity on TikTok; by February 2024, “Red LED light therapy” amassed >70 million views [1]. This surge in interest presents both opportunities and challenges for dermatology practitioners and the broader medical community. The rapid adoption of consumer-grade devices, often supported by inconsistent treatment protocols, raises concerns about patient safety and the need for professional oversight.

Average search volume for chemical peels, a traditional treatment, slightly decreased (from 25.8 to 25.5, a −1.2% change; *P*=.09) despite projected market growth, suggesting a shift in consumer interest toward emerging technologies rather than a true decline in established treatments.

Although current research supports RLT for targeted conditions such as wound healing, inflammatory acne, and photoaging, social media claims often amplify its benefits beyond available evidence [6-10]. This underscores the need for additional randomized controlled trials, particularly studies examining the efficacy and safety of consumer devices.

Health care providers must address patients’ growing interest in RLT while maintaining high standards of clinical care. This requires developing structured approaches to patient education that address common misconceptions and provide evidence-based guidelines.

The phenomenon has broader implications for public health trends. Improved regulation of consumer devices, standardized

safety guidelines, and enhanced adverse event–reporting systems are urgently needed. Public health communication must evolve to address viral trends, particularly those involving medical devices or treatments. Rapid dissemination of medical information on social media presents opportunities and challenges, highlighting the need for responsive professional education and communication strategies.

Study limitations include the broader scope of RLT keywords (which encompass both dermatologic and other applications) compared to the primarily dermatologic control terms, though the analysis focused specifically on consumer interest rather than clinical equivalence between treatments. Google Trends data also do not reflect actual treatment use or capture detailed demographic data. Future research should incorporate sales data, adverse event reports, and clinical outcomes to comprehensively explore this phenomenon.

These findings have significant implications for dermatology and patient care. As social media continues to shape health care trends, medical professionals must adapt education approaches while maintaining evidence-based practices. This may require new frameworks for addressing viral health care trends and improved methods for communicating scientific evidence. The rapid rise in RLT interest is a case study in how social media can quickly transform patient interest and treatment expectations, requiring the medical community to respond with agility and scientific rigor.

Conflicts of Interest

None declared.

References

1. Onque R. High-tech red light therapy is trending on TikTok and ‘all skin types can benefit,’ dermatologist says: here’s what to know. CNBC. URL: https://www.cnbc.com/2024/02/29/red-light-therapy-is-trending-on-tiktok-heres-what-to-know.html [accessed 2025-06-05]
2. Jagdeo J, Nguyen JK, Ho D, et al. Safety of light emitting diode-red light on human skin: two randomized controlled trials. J Biophotonics 2020 Mar;13(3):31483941. [doi: 10.1002/jbio.201960014] [Medline: 31483941]
3. Cohen M, Austin E, Masub N, Kurtti A, George C, Jagdeo J. Home-based devices in dermatology: a systematic review of safety and efficacy. Arch Dermatol Res 2022 Apr;314(3):239-246. [doi: 10.1007/s00403-021-02231-0] [Medline: 33938981]
4. Glass GE. Photobiomodulation: the clinical applications of low-level light therapy. Aesthet Surg J 2021 May 18;41(6):723-738. [doi: 10.1093/asj/sjab025] [Medline: 33471046]
5. Mineroff J, Maghfour J, Ozog DM, Lim HW, Kohli I, Jagdeo J. Photobiomodulation CME part II: clinical applications in dermatology. J Am Acad Dermatol 2024 Nov;91(5):805-815. [doi: 10.1016/j.jaad.2023.10.074] [Medline: 38307144]
6. Giannakopoulos E, Katopodi A, Rallis M, Politopoulos K, Alexandratou E. The effects of low power laser light at 661 nm on wound healing in a scratch assay fibroblast model. Lasers Med Sci 2022 Dec 27;38(1):27. [doi: 10.1007/s10103-022-03670-5] [Medline: 36574084]
7. Chaves MDA, Araújo AD, Piancastelli ACC, Pinotti M. Effects of low-power light therapy on wound healing: LASER x LED. An Bras Dermatol 2014;89(4):616-623. [doi: 10.1590/abd1806-4841.20142519] [Medline: 25054749]
8. Wu Y, Deng Y, Huang P. Application of red light therapy for moderate-to-severe acne vulgaris: a systematic review and meta-analysis. J Cosmet Dermatol 2021 Nov;20(11):3498-3508. [doi: 10.1111/jocd.14369] [Medline: 34363730]
9. Na JI, Suh DH. Red light phototherapy alone is effective for acne vulgaris: randomized, single-blinded clinical trial. Dermatol Surg 2007 Oct;33(10):1228-1233. [doi: 10.1111/j.1524-4725.2007.33258.x] [Medline: 17903156]
10. Couturaud V, Le Fur M, Pelletier M, Granotier F. Reverse skin aging signs by red light photobiomodulation. Skin Res Technol 2023 Jul;29(7):e13391. [doi: 10.1111/srt.13391] [Medline: 37522497]

Abbreviations

RLT: red-light therapy

Edited by R Dellavalle; submitted 08.12.24; peer-reviewed by E Dal, S Jiang; revised version received 13.04.25; accepted 01.05.25; published 12.06.25.

Please cite as:
Shen CZ, Zhao AT
The Influence of Popular Media on Public Interest in Red-Light Therapy: Longitudinal Trend Analysis
JMIR Dermatol 2025;8:e69796
URL: https://derma.jmir.org/2025/1/e69796
doi:10.2196/69796

Trichoscopic Differentiation in Alopecia: Retrospective Case Series Comparing Lichen Planopilaris, Discoid Lupus Erythematosus, and Alopecia Areata

Gökhan Kaya, MD

Department of Dermatology, Nizip State Hospital, Dutlu, Küme Evler, Gaziantep, Turkey

Corresponding Author:
Gökhan Kaya, MD
Department of Dermatology, Nizip State Hospital, Dutlu, Küme Evler, Gaziantep, Turkey

Abstract

This single-center retrospective case series included 28 patients with alopecia (7 with lichen planopilaris, 7 with discoid lupus erythematosus, and 14 with alopecia areata). Trichoscopic markers were systematically compared across groups. Exclamation-mark hairs and yellow dots were characteristic of alopecia areata, whereas follicular ostia loss and white scarring were confined to lichen planopilaris/discoid lupus erythematosus, providing a simple and practical distinction between nonscarring and scarring alopecias in routine practice.

(JMIR Dermatol 2025;8:e83463) doi:10.2196/83463

KEYWORDS

trichoscopy; dermoscopy; alopecia areata; lichen planopilaris; discoid lupus erythematosus; cicatricial alopecia; diagnostic markers

Introduction

Trichoscopy is now integral to alopecia assessment, enabling the recognition of hair‐shaft changes, peri or interfollicular alterations, and follicular opening loss to quickly separate nonscarring from scarring disease [1]. In scarring alopecias, lichen planopilaris (LPP) typically shows perifollicular scales or erythema and target-pattern blue-gray dots, whereas discoid lupus erythematosus (DLE) more often displays follicular keratotic plugs with telangiectatic or arborizing vessels; patterns can vary by phototype, underscoring the need for pragmatic rules that generalize across populations [2,3]. The misclassification between LPP and DLE is well documented, emphasizing the value of simple bedside discriminators that complement histopathology [4].

This study, therefore, aimed to compare key trichoscopic markers among LPP, DLE, and AA and to propose a concise, rule-in/rule-out approach for routine clinical care.

Methods

Setting and Participants

This single-center, retrospective case series was conducted at the Department of Dermatology, Nizip State Hospital, Gaziantep, Türkiye, including consecutive patients with alopecia who underwent trichoscopic evaluation: LPP (n=7), DLE (n=7), and AA (n=14). The diagnosis of AA was made based on clinical and trichoscopic criteria, whereas all LPP and DLE cases were histopathologically confirmed. Age and sex were retrieved from medical records.

Trichoscopic Evaluation

Routine polarized dermoscopy images were reviewed using a prespecified 16-item checklist (present=1/absent=0), including perifollicular scale, erythema, and casts; blue-gray target dots; follicular ostia loss or plugs; yellow and black dots; white scar or atrophy; background erythema; arborizing vessels; interfollicular scales; exclamation-mark, broken, and lonely hairs; and tufting.

Arborizing vessels were graded by caliber (0=absent, 1=thin <50% of adjacent hair-shaft caliber, 2=thick ≥50% of adjacent hair-shaft caliber).

Statistical Analysis

Categorical variables were summarized as n (%) and continuous variables as mean (SD). Prespecified contrasts used two-sided Fisher exact tests (α=.05): AA versus scarring alopecias (LPP+DLE) for exclamation-mark hairs, yellow dots, white scarring or atrophy, and follicular ostia loss (plus exploratory black or broken hairs), and DLE versus LPP for follicular plug, arborizing vessels (any/thick), interfollicular scale, and peripilar casts.

Odds ratios were estimated with a Haldane–Anscombe 0.5 correction when zero cells occurred. Analyses were performed using Python and SciPy.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. Ethical approval was granted by the Scientific Research Ethics Committee of Bezmialem Vakıf University, Istanbul, Türkiye (Approval No: E-54022451-050.04-122609; September 6, 2023). The protocol originally focused on AA; additional LPP and DLE cases collected under the same protocol were analyzed here as a secondary evaluation, with notification to the ethics committee. All participant data were anonymized and deidentified before analysis. Owing to the retrospective design and use of anonymized data, informed consent was not required. No compensation was offered to study participants.

Results

Twenty-eight patients were included (7 each with LPP and DLE, and 14 with AA); age and sex distributions were comparable. Scarring signs (white scar or atrophy and follicular ostia loss) occurred exclusively in LPP/DLE and were absent in AA. Trichoscopic markers clustered in AA, most prominently exclamation-mark hairs, followed by yellow dots, whereas black dots and broken hairs were not discriminatory. Within scarring alopecias, DLE showed more follicular plugs and occasional thick arborizing vessels, while LPP showed universal perifollicular scale with more perifollicular erythema; these trends were not statistically significant given the sample size. Representative clinical, trichoscopic, and histopathological findings for each entity are illustrated in Figure 1. All estimates and exact *P* values are provided in Table 1.

Table . Demographics and key trichoscopic markers across diagnoses, with AA[a] versus scarring statistics.

Variable	LPP[b] (n=7)	DLE[c] (n=7)	AA (n=14)	OR (AA vs scarring[d])	*P* value
Age (years), mean (SD)	34.4 (10.3)	40.3 (6.6)	37.7 (6.4)	–	–
Female participant, n (%)	3 (43)	2 (29)	5 (36)	–	–
White scarring or atrophy, n (%)	7 (100)	7 (100)	0 (0)	<0.01[e]	<.001
Follicular ostia loss, n (%)	7 (100)	7 (100)	0 (0)	<0.01[e]	<.001
Exclamation-mark hairs, n (%)	0 (0)	0 (0)	11 (79)	95.29	<.001
Yellow dots, n (%)	0 (0)	2 (29)	10 (71)	15.00	.006
Black dots, n (%)	0 (0)	3 (43)	7 (50)	3.67	.237
Broken hairs, n (%)	4 (57)	2 (29)	8 (57)	1.78	.706

[a]AA: alopecia areata.

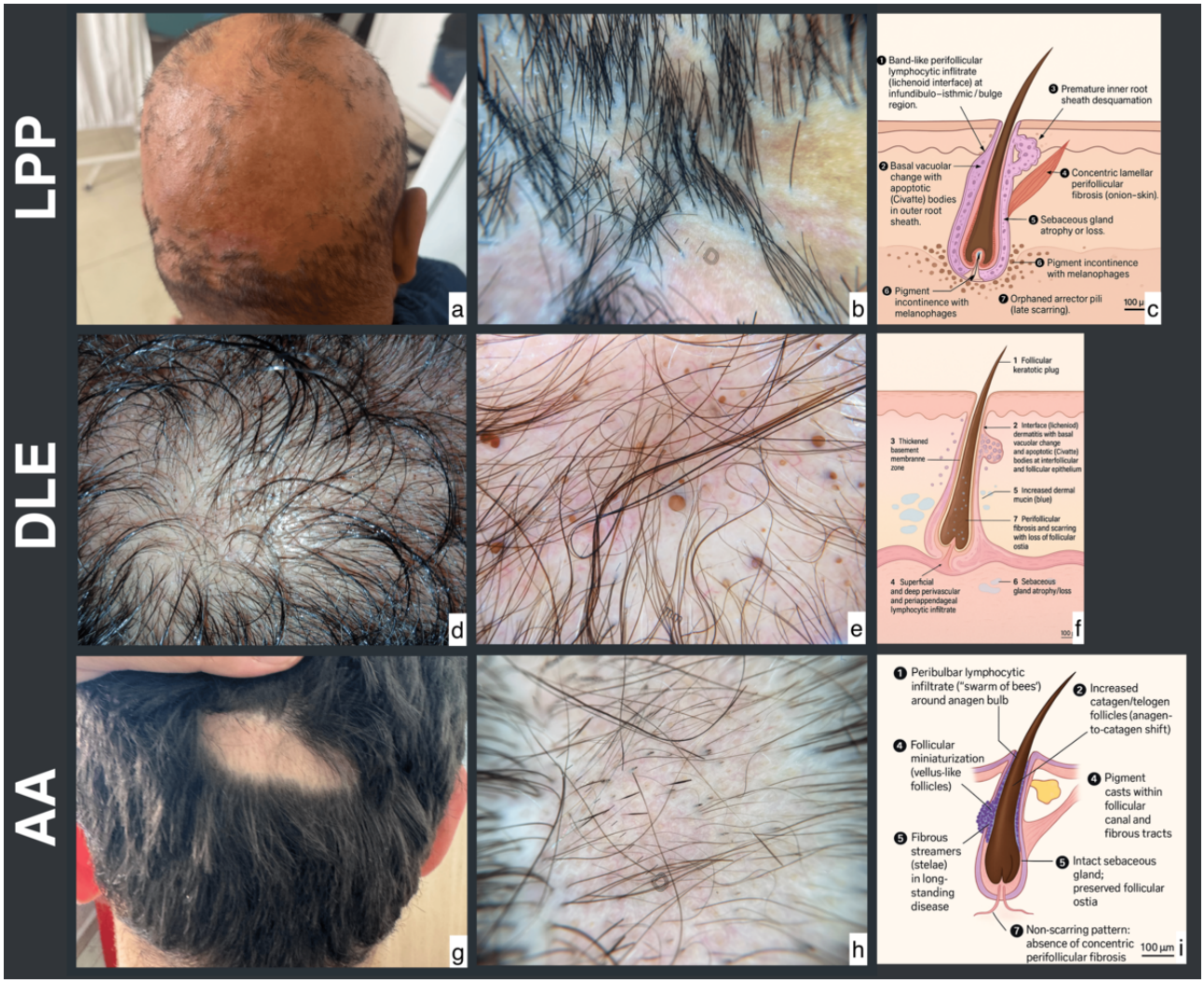
[b]LPP: lichen planopilaris.

[c]DLE: discoid lupus erythematosus.

[d]"Scarring" denotes the combined LPP+DLE group.

[e]Haldane–Anscombe 0.5 correction applied when any cell contained zero.

Figure 1. Clinical, trichoscopic, and histopathological features of LPP, DLE, and AA. (A–C) LPP: Patchy scarring alopecia with perifollicular scales or erythema and blue-gray dots; histology illustration (**C**) Perifollicular lichenoid infiltrate, lamellar fibrosis, sebaceous loss, and pigment incontinence. (D–F) DLE: Scarring alopecia with keratotic plugs, erythema, and arborizing vessels; schematic drawing (**F**) Follicular plugging, thickened basement membrane zone, dermal mucin, and perifollicular fibrosis with ostial loss. (G–I) AA: Nonscarring alopecia with yellow or black dots and exclamation-mark hairs; histopathologic illustration (**I**) Peribulbar "swarm of bees" infiltrate, follicular miniaturization, preserved sebaceous glands, and absence of concentric fibrosis. All clinical and trichoscopic images are original and de-identified; schematic histologic illustrations (**C, F, I**) were created by the authors for this figure to depict key diagnostic features. AA: alopecia areata; DLE: discoid lupus erythematosus; IRS: inner root sheath; LPP: lichen planopilaris.



Discussion

Principal Findings

This study identified key trichoscopic patterns that reliably distinguish nonscarring from scarring alopecias. In this cohort, white scarring or atrophy and follicular ostia loss occurred exclusively in LPP or DLE and were absent in AA, reinforcing that the loss of follicular openings is a practical hallmark of cicatricial disease [5]. Conversely, AA showed clusters of exclamation-mark hairs and, secondarily, yellow dots; these markers also track AA activity and severity in structured trichoscopic scoring systems such as STRIAA (Severity Trichoscopy Index Alopecia Areata) and support their use as practical rule-in signs [6].

Among scarring alopecias, the findings of this study were consistent with those in previous reports [3,7,8]: DLE showed more follicular plugs and occasional thick arborizing vessels, whereas LPP consistently demonstrated perifollicular scale, frequent erythema, and characteristic target-pattern blue-gray dots [3,7,8]. Notably, lonely hair is not disease-specific and should be interpreted in context, particularly when differentiating LPP from frontal fibrosing alopecia [7].

Clinically, a succinct rule emerges: AA is favored by one or more of exclamation-mark hairs or yellow dots, whereas the combination of ostia loss and white scarring favors LPP or DLE. This aligns with stepwise diagnostic algorithms that first classify distribution, then scarring status by the presence or absence of ostia, and finally apply a short list of trichoscopic clues [9]. Comparative clinicopathologic studies also demonstrate systematic differences between DLE and LPP at the population level, providing additional context for our observations [10].

Key limitations include the small, single-center sample and limited power for LPP–DLE contrasts. Nevertheless, the direction and magnitude of the observed effects and the identified high-yield markers are consistent with contemporary

systematic reviews [3] and support the external validity of the findings of this study.

Conclusions

A minimalist trichoscopic rule effectively differentiates AA from scarring alopecias: exclamation-mark hairs or yellow dots favor AA, whereas the combination of follicular ostia loss and white scarring favors LPP or DLE. These easily recognizable cues can assist clinicians in biopsy site selection and treatment planning. Larger multicenter studies are warranted to validate these findings and refine diagnostic criteria for distinguishing LPP from DLE.

Conflicts of Interest

None declared.

References

1. Pirmez R. The dermatoscope in the hair clinic: trichoscopy of scarring and nonscarring alopecia. J Am Acad Dermatol 2023 Aug;89(2S):S9-S15. [doi: 10.1016/j.jaad.2023.04.033] [Medline: 37591567]
2. Khare S, Behera B, Ding DD, et al. Dermoscopy of hair and scalp disorders (trichoscopy) in skin of color - a systematic review by the International Dermoscopy Society “Imaging in Skin of Color” Task Force. Dermatol Pract Concept 2023 Oct 1;13(4 S1):e2023310S. [doi: 10.5826/dpc.1304S1a310S] [Medline: 37874991]
3. Gowda SK, Errichetti E, Thakur V, et al. Trichoscopic features of scalp discoid lupus erythematosus versus lichen planopilaris: a systematic review. Clin Cosmet Investig Dermatol 2024;17:805-827. [doi: 10.2147/CCID.S460742] [Medline: 38616887]
4. Nambudiri VE, Vleugels RA, Laga AC, Goldberg LJ. Clinicopathologic lessons in distinguishing cicatricial alopecia: 7 cases of lichen planopilaris misdiagnosed as discoid lupus. J Am Acad Dermatol 2014 Oct;71(4):e135-e138. [doi: 10.1016/j.jaad.2014.04.052] [Medline: 25219732]
5. Elmas O. An effective and practical diagnostic clinical method in primary scarring alopecia: dermoscopy. Turk J Dermatol 2019;13(2):72. [doi: 10.4103/TJD.TJD_7_19]
6. Starace M, Pampaloni F, Quadrelli F, et al. STRIAA (Severity Trichoscopy Index Alopecia Areata): validation of a novel trichoscopic tool for evaluation of alopecia areata. Dermatol Ther (Heidelb) 2025 Jan;15(1):223-226. [doi: 10.1007/s13555-024-01313-x] [Medline: 39690381]
7. Rajan A, Rudnicka L, Szepietowski JC, et al. Differentiation of frontal fibrosing alopecia and Lichen planopilaris on trichoscopy: a comprehensive review. J Cosmet Dermatol 2022 Jun;21(6):2324-2330. [doi: 10.1111/jocd.14457] [Medline: 34661965]
8. Gowda SK, Errichetti E, Behera B, et al. Trichoscopic features of lichen planopilaris versus frontal fibrosing alopecia: a systematic review. Dermatol Pract Concept 2025 Jan 30;15(1):4481. [doi: 10.5826/dpc.1501a4481] [Medline: 40117622]
9. Katoulis AC, Pappa G, Sgouros D, et al. A three-step diagnostic algorithm for alopecia: pattern analysis in trichoscopy. J Clin Med 2025 Feb 12;14(4):1195. [doi: 10.3390/jcm14041195] [Medline: 40004726]
10. Sharquie KE, Janabi MJ. Scarring alopecia in patients with discoid lupus erythematosus versus patients with follicular lichen planus: A comparative study of clinicopathological features. J Pak Assoc Dermatol 2024 Oct 11(4):883-898 [FREE Full text]

Abbreviations

AA: alopecia areata
DLE: discoid lupus erythematosus
LPP: lichen planopilaris
STRIAA: Severity Trichoscopy Index Alopecia Areata

Edited by R Dellavalle; submitted 03.09.25; peer-reviewed by M Sethy, S Bestavros; revised version received 30.10.25; accepted 31.10.25; published 20.11.25.

Please cite as:
Kaya G
Trichoscopic Differentiation in Alopecia: Retrospective Case Series Comparing Lichen Planopilaris, Discoid Lupus Erythematosus, and Alopecia Areata
JMIR Dermatol 2025;8:e83463
URL: https://derma.jmir.org/2025/1/e83463
doi:10.2196/83463

Insights Into Skin-Lightening Practices of Hijra and Transgender Communities in India

Sriram Palepu[1], BBA; Vasudeva Murthy Sindgi[2], PhD; Aylur Kailasom Srikrishnan[3], BS; Carrie Kovarik[4], MD

[1]Perelman School of Medicine, Philadelphia, PA, United States

[2]Jayamukhi College of Pharmacy, Warangal, India

[3]YR Gaitonde Centre for AIDS Research and Education, Chennai, India

[4]Department of Dermatology, University of Pennsylvania, 3600 Spruce St., Philadelphia, PA, United States

Corresponding Author:
Carrie Kovarik, MD
Department of Dermatology, University of Pennsylvania, 3600 Spruce St., Philadelphia, PA, United States

Abstract

A large proportion of transgender women in Hyderabad, India (150/223, 67.3%) expressed interest in a wide range of topical, oral, and intravenous medications for skin lightening; however, despite government regulations and the potential health risks, persistent demand for skin lightening underscores the need for better patient education and safer skin care practices for this marginalized community.

(JMIR Dermatol 2025;8:e66822) doi:10.2196/66822

KEYWORDS

skin lightening; India; medication misuse; insight; hijra; transgender women; treatment; patient education; skin care; community; fairness cream; marketing; ads; advertisement; cost; lightening cream; cross-sectional study; survey study

Introduction

The Government of India's recent amendments to the Drugs and Magic Remedies Act proposed increased penalties for marketing and advertising skin fairness creams [1]. Yet, conventional Indian beauty standards still drive demand for skin-lightening products (SLPs) among women in India and transfeminine communities. Literature on skin-lightening practices among transgender people is limited. Transgender women undergoing estrogen therapy have an increased risk of melasma, which may be treated with skin-lightening agents like hydroquinone [2]. An ethnographic study of Indonesia's transfeminine *waria* community found that members sought SLPs to feel more feminine and attract male attention [3]. Similar motivations were documented among Thailand's transgender entertainers [4].

The Health Needs and Aesthetic Preferences Assessment of the Hyderabad Trans Community is a large study evaluating the social and health history of transgender and *hijra* women in India (N=300). As part of that study, we evaluated the prevalence of interest in skin-lightening treatments, the products used, and the financial costs involved.

Methods

Study Design

This cross-sectional survey study was conducted at Mitr Clinic (Hyderabad, India), using consecutive sampling. The survey was developed by the research team and administered in Hindi, Telugu, or English. The inclusion criteria were as follows: *hijra* and/or transgender women aged ≥18 years, seeking female gender affirmation, and South Asian ancestry. Transgender men, individuals seeking male gender affirmation, and individuals aged <18 years were excluded. Dollar amounts were estimated based on the March 19, 2023, exchange rate.

Ethical Considerations

Because some community members have limited literacy, verbal consent was obtained before data collection. No protected health information was collected. Institutional review board (IRB) approvals were obtained from the University of Pennsylvania and YR Gaitonde Centre for AIDS Research and Education—the clinic's overseeing nonprofit. Remuneration (500 [US $5.84]) was provided to patients for their time and participation.

Results

An IRB addendum approved in December 2023 enabled 74.3% (223/300) of participants to respond to skin lightening–related questions. More than two-thirds (150/223, 67.3%) of respondents expressed interest in skin lightening, of whom 43.3% (65/150) used SLPs. The overall prevalence of SLPs among respondents was 29.1% (65/223). Further, 1.3% (2/150) of respondents used SLPs previously but lost interest, and 3.1% (7/223) could not recollect or identify the products they used. Money spent on skin lightening varied from 25 (US $0.30) to

70,000 (US $843; median 570 [US $7], IQR 2225 [US $27]). Patients sometimes used multiple products (Table 1).

Table . Ingredients of self-reported products used for skin lightening.

Type of product used[a]	Active ingredients in products[b]
Topical treatments (n=23)	• Hydroquinone + tretinoin + mometasone furoate (n=15) • Clobetasol + neomycin + miconazole nitrate (n=1) • Glycolic acid + arbutin + kojic acid (n=1) • Terbinafine + ornidazole + oflaxicin + clobetasol (n=2) • Betamethasone cream (n=3) • Sunscreen (n=1)
Oral medications (n=2)	• Levonorgestrel/ethinyloestradiol (n=1) • Biotin + multivitamin (n=1)
Intravenous medications (n=16)	• Glutathione (n=16)
Alternative (herbal/Ayurvedic/Unani; n=5)	• Combination of ingredients, including lycopene, botanical extracts (eg, mallow, cowslip, licorice, and aloe vera), and soy isoflavones
Marketed beauty creams (n=15)	• Combination of ingredients, including herbal ingredients, kojic acid, niacinamide, vitamin C, vitamin E, and sun protection factors (octocrylene, avobenzone, etc)

[a]The n values in this column refer to the number of products reported.

[b]The n values in this column refer to the number of products that contained the active ingredients listed in this column.

Discussion

Our study highlights the considerable interest in SLPs but marginally low prevalence of SLP use (likely due to financial barriers) among *hijra* and transgender women. Survey studies on cisgender populations in India indicate that SLPs have widespread prevalence (range 34%-60%) [5,6]. Deeply rooted cultural norms associate lighter skin with economic prosperity and beauty, leading to widespread use even among South Asian immigrant communities [7]. Although 67.3% of our respondents expressed interest in skin lightening, only 29.1% used SLPs—a lower rate than in cisgender communities. Within the colorism context, transfeminine individuals may view skin lightening as an accessible method for facilitating gender affirmation and social acceptance, as observed in other Asian countries.

Transgender women often face stigma and discrimination in health care settings, preventing them from seeking care [8]. Additionally, many transgender women in India engage in sex work, which acts a strong economic driver for investing in physical appearance (eg, undergoing skin treatments to achieve a desired aesthetic) [8]. Although many respondents were interested in skin lightening, less than half used SLPs, which included herbal mixtures from local shops and intravenous glutathione injections.

A study on SLPs used in India reported topical medication misuse prior to seeing a dermatologist [9]. The combination of hydroquinone, mometasone, and tretinoin cream is a common, over-the-counter melasma treatment in India [9]. Glutathione injections, though popular and expensive, have questionable efficacy [10]. Some alternative, traditional remedy–based medications are often cheaper. Popular marketed beauty creams use ingredients like kojic acid, niacinamide, and arbutin, which have been studied for their effects on skin pigmentation and complexion [11]. Chronic steroid use, while lightening some patients' skin, may result in skin atrophy and other side effects [12]. Antifungal creams may treat pigmentary changes resulting from infections like pityriasis versicolor but have no additional lightening effects. Only 1 respondent reported using sunscreen daily, beyond sun protection factors in beauty creams.

Almost half of our respondents use medications with skin-lightening properties—mostly purchased over the counter. After the COVID-19 pandemic, the average *hijra* community member's earnings decreased from US $7 to US $13 per day to less than US $2 per day [13]. Many members are of low socioeconomic status and have been reported to seek hormonal and surgical care from unqualified medical practitioners because allopathic treatments are costly [14]. Despite the Indian government's regulatory efforts, interest in SLPs persists among transgender women [1]. However, only a fraction can afford to regularly use skin-lightening treatments. Given the potential health and financial risks, patient education about safe skin care is crucial for transgender women to make informed health decisions.

This study had several limitations, which we hope to address in follow-up studies. Participants' informal occupations (eg, begging and sex work) precluded an accurate income assessment. Furthermore, the ad hoc survey lacked prior psychometric validation; this may have affected the accuracy of estimates regarding SLP use. Lastly, data on individual product costs and usage durations were not collected, limiting insights into the costs of long-term use.

Acknowledgments

We thank the Fulbright Program and the United States-India Educational Foundation (USIEF) for financially supporting this project.

Conflicts of Interest

None declared.

References

1. Raj S, Sampat B, Kajla T, Singh H. Dark side of skin‐lightening products: social responsibility of advertisers. Int Soc Sci J 2022 Sep;72(245):525-541. [doi: 10.1111/issj.12339]
2. Yeung H, Kahn B, Ly BC, Tangpricha V. Dermatologic conditions in transgender populations. Endocrinol Metab Clin North Am 2019 Jun;48(2):429-440. [doi: 10.1016/j.ecl.2019.01.005] [Medline: 31027550]
3. Idrus NI, Hymans TD. Balancing benefits and harm: chemical use and bodily transformation among Indonesia's transgender waria. Int J Drug Policy 2014 Jul;25(4):789-797. [doi: 10.1016/j.drugpo.2014.06.012] [Medline: 25091633]
4. Farber R. Transitioning Thailand: techno‐professionalism and nation‐building in the transgender entertainment industry. Gend Work Organ 2024 Nov;31(6):2489-2510. [doi: 10.1111/gwao.13104]
5. Jacob SM, Thillaikarasi A. Awareness on skin lightening practice and their side effects among nursing staff in a tertiary care centre. Indian J Dermatol 2023;68(3):356. [doi: 10.4103/ijd.ijd_1010_22] [Medline: 37529465]
6. Wong C, Wong S, Tang H, Minocha R, Singh R, Grills N. Use of skin-lightening products among outpatient attendees in a North Indian hospital. Indian J Public Health 2017;61(2):137-140. [doi: 10.4103/ijph.IJPH_318_15] [Medline: 28721966]
7. Banala M, Mamidipaka A, Ogunleye T. Skin-lightening product use among South Asian Americans: cross-sectional survey study. JMIR Dermatol 2023 Nov 2;6:e49068. [doi: 10.2196/49068] [Medline: 37917151]
8. Nuttbrock L, editor. Transgender Sex Work and Society: Harrington Park Press; 2019.
9. Godse KV, Zawar V. Mometasone menace in melasma. Indian J Dermatol 2012 Jul;57(4):324-326. [doi: 10.4103/0019-5154.97686] [Medline: 22837578]
10. Sonthalia S, Jha AK, Lallas A, Jain G, Jakhar D. Glutathione for skin lightening: a regnant myth or evidence-based verity? Dermatol Pract Concept 2018 Jan 31;8(1):15-21. [doi: 10.5826/dpc.0801a04] [Medline: 29445569]
11. Wang YH, Avonto C, Avula B, Wang M, Rua D, Khan IA. Quantitative determination of α-arbutin, β-arbutin, kojic acid, nicotinamide, hydroquinone, resorcinol, 4-methoxyphenol, 4-ethoxyphenol, and ascorbic acid from skin whitening products by HPLC-UV. J AOAC Int 2015;98(1):5-12. [doi: 10.5740/jaoacint.14-123] [Medline: 25857872]
12. Abraham A, Roga G. Topical steroid-damaged skin. Indian J Dermatol 2014 Sep;59(5):456-459. [doi: 10.4103/0019-5154.139872] [Medline: 25284849]
13. Chakrapani V, Newman PA, Sebastian A, Rawat S, Shunmugam M, Sellamuthu P. The impact of COVID-19 on economic well-being and health outcomes among transgender women in India. Transgend Health 2022 Oct 1;7(5):381-384. [doi: 10.1089/trgh.2020.0131] [Medline: 36406642]
14. Singh Y, Aher A, Shaikh S, Mehta S, Robertson J, Chakrapani V. Gender transition services for hijras and other male-to-female transgender people in India: availability and barriers to access and use. Int J Transgend 2014 Jan 2;15(1):1-15. [doi: 10.1080/15532739.2014.890559]

Abbreviations

IRB: institutional review board
SLP: skin-lightening product

Edited by E Parker; submitted 24.09.24; peer-reviewed by H Yeung, J Lipoff; revised version received 03.03.25; accepted 19.04.25; published 29.05.25.

Please cite as:
Palepu S, Sindgi VM, Srikrishnan AK, Kovarik C
Insights Into Skin-Lightening Practices of Hijra and Transgender Communities in India
JMIR Dermatol 2025;8:e66822
URL: https://derma.jmir.org/2025/1/e66822
doi:10.2196/66822

The Quality of Dermatology Match Information on Social Media Platforms: Cross-Sectional Analysis

Anjali D'Amiano, BA; Jack Kollings, BS; Joel Sunshine, MD, PHD

Department of Dermatology, Johns Hopkins, 600 North Wolfe Street, Baltimore, MD, United States

Corresponding Author:
Joel Sunshine, MD, PHD
Department of Dermatology, Johns Hopkins, 600 North Wolfe Street, Baltimore, MD, United States

(JMIR Dermatol 2025;8:e65217) doi:10.2196/65217

KEYWORDS

residency; dermatology match; Reddit; Student Doctor Network; dermatology; information; cross-section; cross-sectional analysis; online; qualifications; misinformation; media; online data; TikTok; online platform; health platform; web platform; online health; health information; social media; digital health; social media posts; online content; health content; social media content; residency program; medical residency; MD; medical school

Introduction

The transition of United States Medical Licensing Examination (USMLE) Step 1 to pass/fail and a new system for signaling interest in programs have complicated the competitive dermatology match process [1]. Candidates frequently use social media for guidance, but advice on these platforms can be misleading and potentially discourage applicants. Our study evaluates the accuracy of dermatology match information on popular social media sites where medical students, residents, and attendings discuss medicine and residency application processes—Reddit, Student Doctor Network (SDN), and TikTok.

Methods

Study Design

In March 2024, we used the search terms “how to match into dermatology” and “advice for the dermatology match process,” identifying 34 sources and corresponding response comments from TikTok (n=10), Reddit (n=14), and SDN (n=10). These sources provided insights into application components, including USMLE scores, research experiences, and rotations, which we compared to official 2022 National Residency Matching Program (NRMP) data (n=348), using 2-tailed Student *t* tests to identify differences in quantitative measures. Representative quotes were qualitatively compared to NRMP data and the Association of Professors of Dermatology (APD), Residency Program Directors Section, Information Regarding the 2023‐2024 Application Cycle guidelines [2]. Inclusion criteria required at least one numeric data point for comparison.

Ethical Considerations

This study was institutional review board approved (IRB00441663) in alignment with ethical considerations.

Results

Our analysis revealed that mean Step 1 scores (mean 248.0, SD 7.0 vs mean 254.5, SD 8.28; *P*<.001); the number of abstracts, posters, and publications (mean 20.9, SD 3.0 vs mean 23.3, SD 8.68; *P*=.004); and total publications (mean 7.0, SD 1.0 vs mean 13.2, SD 5.65; *P*<.001) reported on the web were significantly higher than NRMP data (Table 1). The NRMP and web-based data did not significantly differ in mean Step 2 scores (mean 257.0, SD 8.5 vs mean 261.0, SD 10.1; *P*=.06).

Table . Significant differences in self-reported web-based data and National Residency Matching Program (NRMP) data.

Category	NRMP data, mean (SD)	Web-based data, mean (SD)	*P* value[a]
Step 1 score	248.0 (7.0)	254.5 (8.28)	<.001
Step 2 score	257.0 (8.5)	261.0 (10.1)	.06
Number of abstracts, posters, and publications	20.9 (3.0)	23.3 (8.68)	.004
Total publications	7.0 (1.0)	13.2 (5.65)	<.001

[a]*P* values were calculated by using Student *t* tests.

Representative quotes are found in Table 2. Regarding academic performance, 15 sources addressed medical school grades, with 10 (67%) emphasizing the importance of Alpha Omega Alpha (AOA) status; however, per NRMP data, only 39.7% of matched dermatology residents were members of AOA. Of 21 sources, 19 (90%) recommended participation in away rotations; 11

(52%) provided a specific number, averaging 3.9 rotations, while 8 (38%) suggested completing as many as possible—a contradiction to APD guidelines, which recommend a maximum of 2 external rotations for students with home dermatology programs and 3 for those without such programs [2].

Table . Representative quotes from categories of dermatology match discussion.

Category	Representative quotes
Research year	• "Definitely take a research year to maximize chances, of people I know who didn't match, most had not taken years." • "A research year is not necessary, there are applicants that we have ranked very highly who have had 3 - 5 listed publications and ones we have ranked near the bottom of the list with > 25 publications." • "Taking a research year will help you stand out, build connections, and be productive." • "This is a personal choice, however, and one that must be made after weighing the risks versus benefits. Some program directors I have spoken to do not feel an extra year of research can add much to an applicant's curriculum vitae, while others feel it is important." • "You don't need to take one if you had research experiences during medical school, start early."
Letters of recommendation	• "Find big names in dermatology (chairs, PDs[a], renound [*sic*] researchers, etc)." • "3 dermatologists with big names, 1 from any field." • "All three were high-ranking, high-impact, quality letters. As most people will tell you, the fourth letter should come from an away rotation." • "Most applicants will have 1 - 2 letters from a non-derm setting (usually medicine sub-internship, research mentor) and 2 - 3 derm letters. Pick your letter writers carefully as some attendings can be great clinically but write lackluster letters. Big names on the letter are helpful, but not if they don't know you well enough to comment on your performance as a student or personal characteristics."
Rotation grades	• "Honor as many preclinical/clinical grades as possible; AOA[b] is more important than GHHS[c] but try for both." • "Honors as many rotations as you can. Do well on your dermatology elective."
Away rotations	• "If you do not have a home dermatology department do as many away rotations as possible." • "Do away rotations at programs where they often extend interview invites to interviewers." • "Away rotations are not needed if coming from a top institution."
Interests and activities	• "Try to show only interest in dermatology." • "Programs love it when you're super specialized and focused and passionate about one or two things and make that your theme of your app. You look a lot cooler and desirable vs the person who tried to do everything and anything in derm to pad their app. HAVE A FOCUS!!" • "Extracurriculars aren't as big a part of the dermatology application process."
Interview	• "The most important part of the match process."
Signals	• "Do not expect to get interviews at places you do not signal."
Personal statement	• "This is not a big part of the process." • "Personal statements will not make or break your application."
DO[d] match	• "Probably next to impossible to match if you are a DO."
Medical school	• "Try to go to a top 10 medical school if you're thinking derm." • "Go to [a medical school] with a strong home dermatology department."

[a]PD: program director.

[b]AOA: Alpha Omega Alpha.

[c]GHHS: Golden Humanism Honors Society.

[d]DO: Doctor of Osteopathic Medicine.

Web-based sources were also divided about the utility of a research year. Of the 22 sources discussing research years, 16 (73%) supported taking a research year to increase applicants' number of research experiences, and 6 (27%) articles advocated against taking a research year to match, without a genuine underlying interest in the research.

Discussion

This study is the first to systematically evaluate the veracity of dermatology match–related discussions that occurred across multiple social media platforms after the USMLE Step 1 pass/fail change. Previous studies on self-reported SDN and Reddit data showed mixed results; a 2017 study found that radiology applicants who self-reported on SDN were likelier to be AOA members with higher USMLE step scores, indicating a reporting a bias toward stronger applicants, which is likely reflected in our study as well [3]. In contrast, a 2020 study found no significant difference in self-reported dermatology applicant USMLE step 1 and 2 scores between social media and NRMP data [4]. However, these studies predate the USMLE Step 1 pass/fail change, and they did not specifically examine forum discussions directly. Our study expands the scope by including TikTok—a platform that is increasingly being used for medical education among students [5].

Our findings suggest potential biases in self-reported data on social media when compared to official sources, underscoring the need for cautious interpretation. Our limitations include the inherent self-reporting nature of social media, which may not accurately reflect the broader applicant pool. Although many contributors on the web aim to help others, some may exaggerate requirements or overstate match difficulties to discourage competition.

In conclusion, while social media serves as a widely used resource for dermatology applicants, it is often unreliable. Program director surveys could help clarify common misconceptions, and efforts to correct misinformation through trusted sources may improve the accuracy of information available to applicants. Applicants seeking reliable guidance should turn to established mentorship programs, such as the National Mentorship Match through the Dermatology Interest Group Association, and official recommendations from the APD. By providing structured, accurate resources, programs can help counter misinformation and better support future applicants.

Conflicts of Interest

None declared.

References

1. Parker T, Brown AE, Messer A, Lewis GD. Response to: "Reliability of self-reported data on social media vs National Residency Match Program charting outcomes for dermatology applicants". J Am Acad Dermatol 2020 Dec;83(6):e473-e474. [doi: 10.1016/j.jaad.2020.07.129] [Medline: 32822797]
2. Association of Professors of Dermatology, Residency Program Directors Section, Information Regarding the 2023-2024 Application Cycle. Association of Professors of Dermatology. 2023 May 22. URL: https://www.dermatologyprofessors.org/files/APD%20statement%20on%202023-2024%20application%20cycle.pdf [accessed 2025-05-26]
3. Sura K, Wilson LD, Grills IS. Comparison of self-reported data on Student Doctor Network to objective data of the National Resident Matching Program. J Am Coll Radiol 2017 Dec;14(12):1594-1597. [doi: 10.1016/j.jacr.2017.08.011] [Medline: 29097048]
4. Hu S, Laughter MR, Dellavalle RP. Reliability of self-reported data on social media versus National Residency Match Program charting outcomes for dermatology applicants. J Am Acad Dermatol 2020 Dec;83(6):1842-1844. [doi: 10.1016/j.jaad.2020.04.052] [Medline: 32315660]
5. Lacey H, Price JM. #MedEd-The "TikTok" frontier of medical education. Clin Teach 2023 Oct;20(5):e13636. [doi: 10.1111/tct.13636] [Medline: 37632321]

Abbreviations

AOA: Alpha Omega Alpha
APD: Association of Professors of Dermatology
NRMP: National Residency Matching Program
SDN: Student Doctor Network
USMLE: United States Medical Licensing Examination

Edited by J Lipoff; submitted 08.08.24; peer-reviewed by M Keller, TW Blalock; revised version received 19.03.25; accepted 10.04.25; published 28.05.25.

Please cite as:
D'Amiano A, Kollings J, Sunshine J
The Quality of Dermatology Match Information on Social Media Platforms: Cross-Sectional Analysis
JMIR Dermatol 2025;8:e65217
URL: https://derma.jmir.org/2025/1/e65217
doi:10.2196/65217

Representation of Psoriasis on the Web for Patients With Skin of Color

Daniel Nguyen[1], BS; Van Le[1], BA; Derek Nguyen[2], BA; Vy Han[1], MD

[1]California University of Science and Medicine, Colton, CA, United States

[2]University of California, Riverside, 900 University Ave, Riverside, CA, United States

Corresponding Author:
Derek Nguyen, BA
University of California, Riverside, 900 University Ave, Riverside, CA, United States

Abstract

This study analyzed over 2000 images of psoriasis across major web-based platforms and found a significant underrepresentation of darker skin tones, highlighting a critical gap in dermatologic representation that may contribute to misdiagnoses and health disparities among patients with skin of color.

(JMIR Dermatol 2025;8:e69026) doi:10.2196/69026

KEYWORDS

Fitzpatrick skin type; Instagram; social media representation; psoriasis; internet

Introduction

Over recent decades, the internet has grown in popularity as a primary health information source, with 74.4% of US adults reporting that they consult it before turning to other resources [1]. Among web-based platforms, social media has emerged as a widely used educational tool for accessing health-related information [2]. Psoriasis—a lifelong inflammatory skin disease affecting around 125 million people worldwide [3]—is theorized to be underdiagnosed among patients with skin of color (SOC), possibly due to lack of access to health care and nuances in the disease's manifestation [4]. In patients with darker skin tones, psoriatic lesions may appear grayish or violaceous rather than the typical salmon pink, leading to misdiagnosis as postinflammatory hyperpigmentation and contributing to disease persistence and undertreatment [5]. For early detection, medical education, and public awareness, SOC representation is important in images of psoriasis on the web. This study assesses skin tone diversity in depictions of psoriasis on Google Images, Instagram (Meta Platforms), Facebook (Meta Platforms), YouTube (Google LLC), and DermNet, using the Fitzpatrick scale.

Methods

On April 13 and 14, 2025, we performed searches for "psoriasis" on Google Images, Instagram, YouTube, Facebook, and DermNet, as these represented the most popular sources of consumer health information, particularly among people of color; Facebook, YouTube, and Instagram each show usage rates exceeding 50% in this demographic [6]. In total, over 2000 images depicting patients with psoriasis were retrieved from these platforms. Computer-generated images, duplicate images on the same platform, images with poor lighting, and images featuring the same patient at a different angle were excluded from data collection. On YouTube, 500 images of individual patients with psoriasis were extracted from 163 videos. To minimize algorithmic bias, searches were performed by using incognito browsers, a new social media account, and 3 different IP addresses. Extracted images were independently categorized based on skin tone by 3 reviewers using the Fitzpatrick scale. Disagreements on classification were resolved by majority vote. Images were further designated as light skin images (Fitzpatrick skin types I, II, III, and IV) or dark skin images (Fitzpatrick skin types V and VI) [7]. The quantities of dark skin and light skin images were compared using a 2-tailed *t* test. A *P* value of <.05 was considered statistically significant.

Results

Images of psoriasis (n=2341) in Fitzpatrick type II skin were the most abundant across all platforms, with 56.4% (1320/2341) of images constituting that classification (Table 1). Interrater reliability was substantial (Cohen κ=0.76). Dark skin images of psoriasis and images of the lightest skin tone—Fitzpatrick type I—were relatively few on all 5 platforms. In total, 5.2% (122/2341) of psoriasis images were dark skin images, and 94.8% (2219/2341) were light skin images, representing a significant difference (*P*<.001). Notably, Fitzpatrick type IV skin had low representation on Google Images (27/401, 6.8%) and YouTube (21/500, 4.2%) and higher representation on Instagram (52/500, 10.4%).

Table . Representation of different skin types in photos of psoriasis on Google Images, Instagram, Facebook, YouTube, and DermNet.

Internet resources	Total, n	Fitzpatrick skin type, n (%)						Dark, n (%)	Light, n (%)
		Type I	Type II	Type III	Type IV	Type V	Type VI		
Google Images	401	35 (8.7)	219 (54.6)	100 (24.9)	27 (6.8)	11 (2.8)	9 (2.2)	20 (5.0)	381 (95.0)
Instagram	500	49 (9.8)	266 (53.2)	107 (21.4)	52 (10.4)	19 (3.8)	7 (1.4)	26 (5.2)	474 (94.8)
Facebook	500	26 (5.2)	300 (60.0)	96 (19.2)	40 (8)	24 (4.8)	14 (2.8)	38 (7.6)	462 (92.4)
YouTube	500	34 (6.8)	294 (58.8)	127 (25.4)	21 (4.2)	9 (1.8)	15 (3.0)	24 (4.8)	476 (95.2)
DermNet	440	36 (8.2)	241 (54.8)	107 (24.3)	42 (9.5)	7 (1.6)	7 (1.6)	14 (3.2)	426 (96.8)

Discussion

Our findings suggest that social media postings of patients with psoriasis and darker skin types are underrepresented across all platforms. These results align with research examining SOC representation within medical education, indicating this issue's prevalence across many information sources [8]. Furthermore, the underrepresentation extends to psoriasis-related content in dermatology residency teaching materials, wherein patients with SOC may also be inadequately depicted [4]. These oversights contribute to the underdiagnosis of psoriasis and poorer outcomes for individuals with SOC, as the disease's clinical presentation can differ across racial and ethnic groups. Given that patients may access the internet for information before visiting a dermatologist, greater image diversity would also be helpful to patients with SOC who suspect that they have psoriasis. Barriers to improving representation on the web include cultural perceptions of psoriasis, which can influence health care–seeking behavior in certain racial and ethnic groups [4]. For example, psoriasis-related stigma is particularly prevalent among Black and Latino patients, further discouraging them from sharing their images and experiences on the web [4,9]. Our study's limitations include potential hyperpigmentary effects influencing raters' judgments and algorithmic bias in Google search results, which may vary by geographic region, despite the use of 3 different IP addresses. Our results demonstrate that internet companies have cause to improve psoriasis representation in search results. The limited content available for people of color highlights a gap that content creators, health care professionals, and social media platforms must address to improve representation and reduce health disparities in psoriasis care. Future efforts should be directed toward improving the quality and dissemination of photographs of psoriasis in SOC.

Conflicts of Interest

None declared.

References

1. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW. Online health information seeking among US adults: measuring progress toward a Healthy People 2020 objective. Public Health Rep 2019;134(6):617-625. [doi: 10.1177/0033354919874074] [Medline: 31513756]
2. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. Healthcare (Basel) 2021 Dec 16;9(12):1740. [doi: 10.3390/healthcare9121740] [Medline: 34946466]
3. Bu J, Ding R, Zhou L, Chen X, Shen E. Epidemiology of psoriasis and comorbid diseases: a narrative review. Front Immunol 2022 Jun 10;13:880201. [doi: 10.3389/fimmu.2022.880201] [Medline: 35757712]
4. Yadav G, Yeung J, Miller-Monthrope Y, et al. Unmet need in people with psoriasis and skin of color in Canada and the United States. Dermatol Ther (Heidelb) 2022 Nov;12(11):2401-2413. [doi: 10.1007/s13555-022-00811-0] [Medline: 36131193]
5. Alexis A, Woolery-Lloyd H, Andriessen A, Koo J, McMichael A, Han G. Evolving concepts in psoriasis: special considerations for patients with skin of color, skin barrier dysfunction, and the role of adjunctive skin care. J Drugs Dermatol 2022 Oct 1;21(10):1054-1060. [doi: 10.36849/JDD.7090] [Medline: 36219055]
6. Social media fact sheet. Pew Research Center. 2024 Nov 13. URL: https://www.pewresearch.org/internet/fact-sheet/social-media [accessed 2025-04-20]
7. Alvarado SM, Feng H. Representation of dark skin images of common dermatologic conditions in educational resources: a cross-sectional analysis. J Am Acad Dermatol 2021 May;84(5):1427-1431. [doi: 10.1016/j.jaad.2020.06.041] [Medline: 32565205]
8. Jones VA, Clark KA, Shobajo MT, Cordova A, Tsoukas MM. Skin of color representation in medical education: an analysis of popular preparatory materials used for United States Medical Licensing Examinations. J Am Acad Dermatol 2021 Sep;85(3):773-775. [doi: 10.1016/j.jaad.2020.07.112] [Medline: 32750385]

9. Gonzalez N, Wilkerson K, Castillo Valladares H, Sanchez-Anguiano ME, Chang AY, Amerson EH. Latine patients' beliefs, attitudes, and experience with psoriasis. JAMA Dermatol 2025 Mar 1;161(3):291-298. [doi: 10.1001/jamadermatol.2024.5391] [Medline: 39774751]

Abbreviations

SOC: skin of color

Edited by J Solomon; submitted 20.11.24; peer-reviewed by E Karampinis, VO Carvalho; revised version received 24.05.25; accepted 01.06.25; published 05.08.25.

Please cite as:
Nguyen D, Le V, Nguyen D, Han V
Representation of Psoriasis on the Web for Patients With Skin of Color
JMIR Dermatol 2025;8:e69026
URL: https://derma.jmir.org/2025/1/e69026
doi:10.2196/69026

ChatGPT-4's Level of Dermatological Knowledge Based on Board Examination Review Questions and Bloom's Taxonomy

Hansen Tai[1], BA; Carrie Kovarik[2], MD

[1]SUNY Upstate Medical University, Syracuse, United States

[2]Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, 3600 Spruce Street, 2 Maloney Building, Philadelphia, United States

Corresponding Author:
Carrie Kovarik, MD
Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, 3600 Spruce Street, 2 Maloney Building, Philadelphia, United States

Abstract

Our study demonstrated the ability of ChatGPT-4 to answer 77.5% of all sampled text-based board review type questions correctly. Questions requiring the recall of factual information were answered correctly most often, with slight decreases in correctness as higher-order thinking requirements increased. Improvements to ChatGPT's visual diagnostics capabilities will be required before it can be used reliably for clinical decision-making and visual diagnostics.

(JMIR Dermatol 2025;8:e74085) doi:10.2196/74085

KEYWORDS

ChatGPT; dermatology; education; board exam; residency

Introduction

ChatGPT, a multimodal language model capable of answering multiple choice questions, incorporates visual inputs in its latest version, GPT-4. Lewandowski et al [1] recently assessed ChatGPT-3.5 and ChatGPT-4's performance in dermatology examinations, finding that ChatGPT-4 significantly outperformed its predecessor, achieving over a 60% pass rate overall and >84% accuracy on photo-based questions. Building on this, our study classified ChatGPT-4's correctly answered question types using Bloom's taxonomy for cognitive complexity [2].

Methods

We evaluated ChatGPT-4's capabilities on the Basic, Core, and Applied examination questions from Dermatology-In-Review, an online dermatology board review preparation course. The Basic examination is a required examination for first-year US dermatology residents and tests dermatology fundamental knowledge. The Core and Applied examinations are taken late in residency and after residency, respectively. These tests examine more advanced clinical knowledge and focus on higher-order thinking. In total, 167 Basic, 210 Core, and 166 Applied multiple-choice questions without photos were formatted and fed into ChatGPT-4 using an algorithm in Python's Pandas. ChatGPT-4's in-depth responses to each query were captured, reviewed, and independently confirmed and coded as correct or incorrect (Table 1).

Table . ChatGPT-4 cases correct by testing category.[a]

	Correct	Incorrect	% Correct	Remember type questions: Correct %, Total %
Basic	139	28	83.20%	71/82 (86.6%), 82/167 (49.1%)
Core	158	52	75.20%	52/66 (78.8%), 66/210 (31.4%)
Applied	123	43	74.10%	35/46 (76.1%), 46/166 (27.7%)

[a]*P* value=.0382, Pearson's Chi-squared test for the Basic versus Core+Applied Examinations.

We categorized text-based questions according to Bloom's taxonomy using a Python function. One author (CK) and ChatGPT-4 categorized each question into a specific category of Bloom's Taxonomy using guidelines [2]. In the case of a discrepancy, ChatGPT-4's reasoning for the decision was considered, which assisted in the reconciliation of categorization. Bloom's categories included Remember (includes lower-level thinking, such as knowledge and comprehension), Apply, Analyze, Evaluate, and Synthesize. All statistics were performed using R statistical software, including the Pearson chi-squared test (Table 1) and Fisher exact test (Table 2).

Photo-based questions were entered directly into ChatGPT-4, along with structured messages and answer choices, and responses were recorded. Fifty-three photo cases from all board categories were used.

Table . ChatGPT-4 cases correct by Bloom category (all cases).[a]

	Correct	Incorrect	Total	% Correct
Remember	158	35	193	81.9%
Apply	168	51	219	76.7%
Analyze	56	19	75	74.7%
Evaluate	37	14	52	72.5%
Synthesize	1	3	4	25.0%
Total	420	122	542	77.5%

[a]*P* =.059, Fisher exact test.

Results

Overall, ChatGPT-4 answered 77.5% of all sampled text-based questions correctly. Varying levels of accuracy were demonstrated in answering board questions within different Bloom categories. In the "Remember" category, the model correctly answered 158/193 (81.9%). "Remember" is considered the most basic level of educational understanding, with the ability to recall or comprehend information without applying the concept [3]. ChatGPT-4 performed the best in this category; however, it did significantly (*P*=.0382) better on the "Remember" questions from the Basic examination compared to those on the Core and Applied sections combined (Table 1). As the Bloom categories progress from Apply to Analyze, Evaluate, and Synthesize, a solid foundation of knowledge and higher-order thinking is necessary. Table 2 demonstrates a decreasing trend (*P*=.059) in the percent correctness for the ChatGPT-4 answers moving from "Remember" to the classes of higher-order thinking.

Of the 53 questions, 18 (34%) with photos were answered correctly, with none of the "What is the histologic diagnosis?" question stems answered correctly. Excluding these, 18/38 (47.3%) had the correct answer. Notably, photo questions with leading information were more likely to be given the correct response.

Discussion

ChatGPT-4 correctly answered 77.5% of all text questions correctly, similar to the results of Lewandowski et al [1], in which ChatGPT-4 answered 80.7% - 84% of the questions correctly on English-based Dermatology assessments. Our outcomes differed in that they were able to obtain a much higher number of correct responses on photo-based questions compared to our study, where ChatGPT-4 was only able to answer approximately one-third of the photo cases correctly. Hirosawa et al [4]assessed the impact of adding image data to clinical textual data on ChatGPT-4's diagnostic accuracy. They found that integrating image data into ChatGPT-4 did not significantly enhance diagnostic accuracy, and it predominantly relies on textual data, limiting its ability to use the diagnostic potential of visual information fully [[4]]. This corroborates our findings of poor analysis of photo cases and improved correctness when leading question stems were given.

Overall, our study demonstrates the ability of ChatGPT-4 to answer text-based questions from Dermatology-In-Review at a high level. Questions requiring the recall of factual information were answered correctly most often, with slight decreases in correctness as higher-order thinking requirements increased. Improvements to ChatGPT-4's visual diagnostics capabilities will be required before it can be used reliably for visual interpretation and clinical decision-making. In its current state, ChatGPT-4 may be used as an educational tool for students and trainees when exploring core factual knowledge; however, trainees and practitioners should not rely on ChatGPT for higher level inquiries, such as analyzing clinical scenarios or image interpretation.

Our study has several limitations. Bloom's taxonomy is a continuum, and question classification can be complex. We used board review questions, and this may not be generalizable to true board questions. The edition of ChatGPT-4 used in this study had been trained with data only up to December 2023 [5].

Conflicts of Interest

None declared.

References

1. Lewandowski M, Łukowicz P, Świetlik D, Barańska-Rybak W. ChatGPT-3.5 and ChatGPT-4 dermatological knowledge level based on the Specialty Certificate Examination in Dermatology. Clin Exp Dermatol 2024 Jun 25;49(7):686-691. [doi: 10.1093/ced/llad255] [Medline: 37540015]
2. Krathwohl DR. A revision of Bloom's taxonomy: an overview. Theory Pract 2002 Nov 1;41(4):212-218. [doi: 10.1207/s15430421tip4104_2]
3. Zaidi N. Modified bloom's taxonomy for evaluating multiple choice questions. 2015. URL: www.bcm.edu/sites/default/files/2019/04/principles-and-guidelines-for-assessments-6.15.15.pdf [accessed 2024-05-21]
4. Hirosawa T, Harada Y, Tokumasu K, Ito T, Suzuki T, Shimizu T. Evaluating ChatGPT-4's diagnostic accuracy: impact of visual data integration. JMIR Med Inform 2024 Apr 9;12:e55627. [doi: 10.2196/55627]
5. OpenAI Platform. URL: https://platform.openai.com [accessed 2024-04-27]

Edited by R Dellavalle; submitted 17.03.25; peer-reviewed by D Hitchcock, RT Potla, Sunny, CL Au; revised version received 26.06.25; accepted 15.07.25; published 07.08.25.

Please cite as:
Tai H, Kovarik C
ChatGPT-4's Level of Dermatological Knowledge Based on Board Examination Review Questions and Bloom's Taxonomy
JMIR Dermatol 2025;8:e74085
URL: https://derma.jmir.org/2025/1/e74085
doi:10.2196/74085

Online Resources for Hidradenitis Suppurativa for Patient Use: Systematic Search and Analysis

Emily Sheetz, MS; Aryn A Alanizi, MS; Joshua Edwards, MPH; Alice A Roberts, MD, PhD

Eastern Virginia Medical School, 721 Fairfax Ave, Norfolk, VA, United States

Corresponding Author:
Emily Sheetz, MS
Eastern Virginia Medical School, 721 Fairfax Ave, Norfolk, VA, United States

Abstract

This research letter evaluates the quality and readability of hidradenitis suppurativa (HS) websites found on Google and Bing with the DISCERN instrument and Flesch-Kincaid Readability metrics. Comprehensive and reliable articles can lead to increased knowledge about HS and further enhance physician-patient relationships and shared decision-making. This study's aim was to identify reliable resources to help bridge knowledge gaps and support informed discussions on management and treatment options.

(JMIR Dermatol 2025;8:e72773) doi:10.2196/72773

KEYWORDS

hidradenitis suppurativa; online resources; patient education

Introduction

Hidradenitis suppurativa (HS) is a chronic, inflammatory skin condition that is often challenging to diagnose, with delays averaging 7 to 10 years [1]. Its complex clinical course and psychosocial burden lead many patients to the internet for information and treatment options. While online resources can support shared decision-making and physician-patient communication [2], the accuracy and readability of content can vary widely. High-quality, accessible information empowers patients, but misleading or difficult-to-comprehend information causes confusion and hinders effective management. We assessed the quality and readability of HS-related websites using the DISCERN instrument and Flesch-Kincaid metrics.

Methods

A systematic search of Google and Bing was conducted using the term *hidradenitis suppurativa* in an incognito browser with location services disabled. Searches were performed on the same day to minimize discrepancies due to search algorithm changes. Twenty results were obtained from each search engine. Advertisements, duplicate content, paywalled articles, and incomplete sources were screened out, leaving a combined 20 websites for analysis.

Two independent reviewers used the DISCERN instrument to evaluate health information based on 16 questions covering clarity, references, and treatment (Multimedia Appendix 1) [3]. DISCERN uses a 5-point scale, with higher scores indicating better quality. Scores from both reviewers were averaged. Readability was measured using the Flesch-Kincaid Grade Level, which determines the US school grade level required for comprehension [4]. *P* values were calculated using independent 2-tailed *t* tests to compare DISCERN scores, while readability metrics were summarized descriptively.

Results

The mean DISCERN scores for Google (Alphabet) and Bing (Microsoft) were 54.05 (SD 11.53) and 59.83 (SD 9.73), respectively, indicating good quality [5]. Websites authored or reviewed by physicians had significantly higher DISCERN scores (62.1 vs 49.7; P=.02) than those by nonphysicians, indicating that expert involvement improves the quality of online health content. However, the mean reading grade levels for Google (10.8, SD 2.4) and Bing (10.5, SD 1.9) exceeded the National Institutes of Health recommendation for a sixth or seventh grade level [6]. Only half of physician-reviewed articles met this criterion. Table 1 summarizes our findings, highlighting a significant gap between content quality and accessibility, underscoring the need for improved patient-friendly resources.

Moreover, websites found on Bing exhibited statistically significant differences in DISCERN question 7 (providing additional sources of support; P=.03) and a trend, although nonsignificant, in question 10 (treatment benefit descriptions; P=.06).

Table . Hidradenitis suppurative websites analyzed.

Website	Search engine	Author or reviewer	Mean DISCERN score	Reading grade level
National Health Service	Google	Academic institution	52.5	9.6
Mayo Clinic	Google and Bing	Academic institution	67.5	9.7
American Academy of Dermatology	Google and Bing	Physician	70	6.3
National Institutes of Health	Google	Physician	63.5	11.5
MedlinePlus	Google and Bing	Academic institution	42.5	8.1
DermNet	Google and Bing	Physician	56	12.6
Cleveland Clinic	Google and Bing	Academic institution	61	10.3
HS Foundation	Google and Bing	Nonprofit organization	47	10.4
Medscape	Google	Physician	62.5	11
WebMD	Google	Physician	70	5.7
Nationwide Children's Hospital	Google	Hospital	37.5	7.4
Wikipedia	Google and Bing	Global site	61	9.7
Mount Sinai	Google	Academic institution	33	11.1
British Skin Foundation	Google and Bing	Charity	49	9.6
American Osteopathic College of Dermatology	Google	Academic institution	46	11.1
Cedars-Sinai	Google	Physician	47	6.4
Patient.Info	Bing	Physician	65.5	8.4
Healthline	Bing	Physician	71.5	7.6
FamilyDoctor.org	Bing	Physician	50	7.5
Medical News Today	Bing	Physician	65	9.4

Discussion

While the internet is a valuable resource for patient education, many HS-related websites may be difficult for patients to understand. Given the chronic and distressing nature of HS, access to clear and reliable information is of utmost importance. The variability in readability and quality indicates a need for greater oversight and standardization in online medical content. Complex resources may discourage patients with lower health literacy from engaging with important health information, leading to misinformation, delays in seeking professional care, and suboptimal self-management strategies [7]. High readability demands on websites may further widen health disparities, as HS is more prevalent among individuals with lower socioeconomic status, who may also have lower health literacy. To improve equitable access to medical information, resources should be written in plain language, include visual aids, and be available in multiple languages to accommodate diverse patient backgrounds.

Health care providers should also guide patients to reliable sites, ideally incorporating links to after-visit summaries or patient portals. Future efforts should focus on improving the clarity of online HS resources without compromising their informational value. Website developers could also involve patients in the content creation process to ensure materials are accessible.

Despite these findings, several limitations must be considered. First, the study focused on the top search results for Google and Bing at a specific time point, which may not reflect dynamic changes in search algorithms. Notably, Bing yielded more support group links, potentially due to algorithm prioritization; Google often ranks academic sources higher, while Bing surfaces more user-friendly content. Additionally, the readability analysis relied on established formulas, which primarily assess sentence length and word complexity. These tools do not account for contextual factors such as formatting choices or visual aids, which may improve comprehension. Finally, although patient forums like Reddit may offer valuable insights, they are not professionally curated and should be interpreted with caution.

In conclusion, we found considerable variability in the quality and readability of online HS resources. Due to the persistent and often debilitating course of HS, trustworthy and comprehensible resources are crucial to support understanding and management of this condition. As online health information continues to shape patient perceptions and behaviors, improving the quality and readability of digital medical content should remain a priority.

Conflicts of Interest

None declared.

Multimedia Appendix 1
DISCERN questionnaire.
[DOCX File, 15 KB - derma_v8i1e72773_app1.docx]

References

1. Aparício Martins I, Figueira Vilela B, Cabete J. Diagnostic delay in hidradenitis suppurativa: still an unsolved problem. Skin Appendage Disord 2024 Apr;10(2):129-132. [doi: 10.1159/000534845] [Medline: 38572195]
2. Laugesen J, Hassanein K, Yuan Y. The impact of internet health information on patient compliance: a research model and an empirical study. J Med Internet Res 2015 Jun 11;17(6):e143. [doi: 10.2196/jmir.4333] [Medline: 26068214]
3. 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]
4. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (Automated Readability Index, Fog count, and Flesch reading ease formula) for Navy enlisted personnel. Research branch report. : Chief of Naval Technical Training: Naval Air Station Memphis; 1975.
5. Cassidy JT, Baker JF. Orthopaedic patient information on the world wide web: an essential review. J Bone Joint Surg Am 2016 Feb 17;98(4):325-338. [doi: 10.2106/JBJS.N.01189] [Medline: 26888683]
6. Hutchinson N, Baird GL, Garg M. Examining the reading level of internet medical information for common internal medicine diagnoses. Am J Med 2016 Jun;129(6):637-639. [doi: 10.1016/j.amjmed.2016.01.008] [Medline: 26829438]
7. Coughlin SS, Vernon M, Hatzigeorgiou C, George V. Health literacy, social determinants of health, and disease prevention and control. J Environ Health Sci 2020;6(1):3061. [Medline: 33604453]

Abbreviations

HS: hidradenitis suppurativa

Edited by A Kristensen-Cabrera; submitted 17.02.25; peer-reviewed by D Hitchcock, E Dal, S Gulliver; revised version received 25.06.25; accepted 24.07.25; published 22.08.25.

Please cite as:
Sheetz E, Alanizi AA, Edwards J, Roberts AA
Online Resources for Hidradenitis Suppurativa for Patient Use: Systematic Search and Analysis
JMIR Dermatol 2025;8:e72773
URL: https://derma.jmir.org/2025/1/e72773
doi:10.2196/72773

Evaluating Bias in Social Media Research Using #Sunscreen Content on Instagram Reels

Silvija Milanovic[1], BS; Chelsea Rosen[1], BS; Taylor Gray[2], DO

[1]College of Medicine, University of Florida, 1104 Newell Dr, Gainesville, FL, United States

[2]Department of Dermatology - Jacksonville, University of Florida, Jacksonville, FL, United States

Corresponding Author:
Silvija Milanovic, BS
College of Medicine, University of Florida, 1104 Newell Dr, Gainesville, FL, United States

Abstract

This cross-sectional content analysis found that Instagram's hashtag-based reels display consistent dermatologic content regardless of user engagement history, supporting the use of hashtags as an objective and reproducible tool for social media research in dermatology.

(JMIR Dermatol 2025;8:e76829) doi:10.2196/76829

KEYWORDS

social media; dermatology; health education; algorithms; information dissemination

Introduction

Social media has become a source of medical information for the general public, with Instagram and TikTok being among the most influential apps [1-4]. In response to the increasing presence of dermatology-related content, dermatologists have been encouraged to engage in social media to decrease misinformation [1-4]. As dermatology-related content grows on social media, more research is being done to assess its quality [1]. Concerns have been raised about the reliability and reproducibility of social media research, particularly due to the personalized nature of platform algorithms [1]. A 2024 study by Druskovich and Landriscina [1] highlighted how TikTok's content curation can introduce bias and inconsistency into dermatology research.

This study aimed to assess whether account bias affects the type of dermatologic information presented in social media videos.

Methods

Overview

We used Instagram reels rather than TikTok amid increasing scrutiny of TikTok's data privacy practices and the potential for a national ban in the United States. We compared Instagram reels videos between two user accounts: one with a biased profile based on prior dermatology engagement, and a new account with no activity or history. The biased profile was an existing Instagram account that had engaged extensively with dermatology-related content. This account followed multiple dermatologists and dermatology organizations and had previously liked, commented on, and shared dermatology posts, thereby establishing a history of dermatology-related activity.

For each account, the first 100 reels displayed under the hashtag #sunscreen were saved to prevent future content influence, resulting in a total of 200 videos. Each video was evaluated and categorized by creator type (dermatologist, non-dermatologist physician, esthetician, beauty blogger, etc) and content type (educational, advertisement, entertainment, etc). Creators were classified as dermatologists if they were verified to be board-certified through online professional profiles, institutional websites, or board certification databases. Engagement metrics, including views, likes, and comments, were recorded. By analyzing both content and engagement metrics, we sought to determine whether Instagram's hashtag-based video results are influenced by user history or whether they offer an objective method for dermatology social media research. Data analysis was performed using R (version 4.3.1; R Foundation for Statistical Computing) with a significance value set to $P<.05$.

Ethical Considerations

This study was deemed exempt from ethics approval as it did not involve human participants or identifiable private information. The research was conducted in accordance with the ethical standards outlined in the U.S. Department of Health and Human Services, Code of Federal Regulations, Title 45, Part 46 [5]. All data analyzed were publicly available and fully deidentified prior to analysis, ensuring no individual could be identified directly or indirectly.

Results

Our study found there were no statistically significant differences in the video results between the two accounts. A χ^2 test assessing differences in content categories and creator types showed no significant variation between accounts ($\chi^2 = 7.6$, P

= 0.3; Table 1). A Monte-Carlo chi-squared test evaluating video types similarly found no significant differences (χ^2 =19.9, *P*=0.8; Table 2). Engagement metrics for views, likes, and comments across all creator types were compared using two-sample *t* tests, and all *P* values exceeded 0.05. A paired *t* test comparing video engagement, calculated by adding the likes and comments and dividing by the view number, showed no difference between accounts (*P*=0.42; Table 1). These findings suggest that the Instagram reels hashtag system displays dermatologic content consistently and reproducibly, regardless of account history or engagement.

Table . Video engagement by creator type for reels found on a biased profile based on prior dermatology engagement and a new account with no activity or history.

	Number of videos		Average views		Average likes		Average comments		Average engagement rate	
Video Creator	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased
Dermatologist	17	14	1,950,376	3,814,600	25,112	46,071	578	2899	1.52%	1.58%
Non-dermatologist physician	4	4	1,233,550	975,325	14,086	12,554	482	562	1.47%	1.26%
Nurse practitioner	1	1	9,200,000	10,400,000	428,000	480,000	1008	1087	4.66%	4.63%
Esthetician	1	8	67,100	13,938,500	334	38,620	63	361	0.59%	1.25%
Beauty blogger	59	58	3,047,975	2,166,548	47,610	51,883	664	2169	2.89%	3.11%
Skincare company	10	5	3,734,560	5,362,400	60,578	96,881	396	386	1.94%	2.53%
Other	9	11	2,769,867	3,137,000	75,329	76,686	811	859	13.08%	1.99%

Table . Video type by creator type for number of reels found on a biased profile based on prior dermatology engagement and a new account with no activity or history.

	Educational		Personal experience		Business/Advertisement		Entertainment/Humor		Clinical demonstration	
Video creator	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased	Biased	Unbiased
Dermatologist	10	6	1	2	4	4	2	2	0	0
Non-dermatologist physician	3	3	1	1	0	0	0	0	0	0
Nurse practitioner	0	0	0	0	0	0	1	1	0	0
Esthetician	0	2	0	3	0	2	1	0	0	1
Beauty blogger	9	9	22	21	23	20	5	6	0	2
Skincare company	0	0	0	1	6	2	3	2	1	0
Other	2	4	0	1	1	0	3	2	3	4

Discussion

These results challenge assumptions that social media research is inherently flawed due to account engagement history [1]. While Instagram curates personalized content based on user interaction, the hashtag system appears to present the same type of information regardless of account history. This has meaningful implications for research methodology, where Instagram hashtags may provide an objective tool for studying dermatologic content online, supporting more reproducible and rigorous investigations.

Our study is limited as it reflects a single time point and one hashtag on one platform. Future work should evaluate the longitudinal stability of Instagram's algorithm, explore other hashtags, and compare results across broader time frames. Additionally, while the hashtag system offers a consistent sampling method for research, it may not fully reflect the way users typically consume content. Many users do not actively

search via hashtags, and some creators, especially those with large followings, may not rely on hashtags for engagement, meaning their content may be underrepresented in hashtag-based analyses. Despite this, the hashtag system provides a practical and reproducible approach for evaluating dermatologic content for research.

As patients continue to turn to social media for skin health information, understanding how content is delivered empowers dermatologists to better counsel patients on how to navigate these platforms critically. These findings may reassure clinicians that accurate, evidence-based content has the potential to compete with less credible sources, especially when paired with effective use of tools like hashtags. Our data show that dermatologists did not receive the highest engagement rate (Table 1), highlighting the need for more dermatologists to create engaging social media content. Rather than dismissing social media as a space dominated by misinformation, dermatologists can leverage these insights to engage with patients and encourage more informed digital consumption.

Conflicts of Interest

None declared.

References

1. Druskovich C, Landriscina A. TikTok and dermatology: questioning the data. J Drugs Dermatol 2024 Aug 1;23(8):e167-e168. [doi: 10.36849/JDD.7956] [Medline: 39093646]
2. 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]
3. Wojtara MS. Use of social media for patient education in dermatology: narrative review. JMIR Dermatol 2023 Apr 14;6:e42609. [doi: 10.2196/42609] [Medline: 37632938]
4. Nigro AR, Osman A, Saal RC, Smith RJ. Information regarding dermatology as seen on the social media platform TikTok. JMIR Dermatol 2025 Jan 28;8:e59597. [doi: 10.2196/59597] [Medline: 39874465]
5. U.S. department of health and human service. Code of Federal Regulations, Title 45, Public Welfare, Part 46: Protection of Human Subjects. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46 [accessed 2025-10-28]

Edited by E Parker; submitted 01.05.25; peer-reviewed by J Makin, S Bestavros; revised version received 23.09.25; accepted 15.10.25; published 05.11.25.

Please cite as:
Milanovic S, Rosen C, Gray T
Evaluating Bias in Social Media Research Using #Sunscreen Content on Instagram Reels
JMIR Dermatol 2025;8:e76829
URL: https://derma.jmir.org/2025/1/e76829
doi:10.2196/76829

Omission of Risk in Vitamin A–Related Dermatologic Instagram Posts: A Growing Concern in an Unregulated Digital Landscape

Robert John Vanaria[1], BS; Michael Povelaitis[1], BS; Riya Khokhal[1], BS; Shannon Meledathu[1], BS; Claudia Green[2], DO; Robin Ashinoff[1,2], MD; Mariela Mitre[1,2], MD, PhD

[1]Hackensack Meridian School of Medicine, 123 Metro Blvd, Nutley, NJ, United States

[2]Hackensack University Medical Center, 360 Essex St Suite, #201, Hackensack, NJ, United States

Corresponding Author:
Mariela Mitre, MD, PhD
Hackensack Meridian School of Medicine, 123 Metro Blvd, Nutley, NJ, United States

Abstract

We reviewed 30 of the top-viewed Instagram videos using the hashtags #retinoid and #retinol to assess reliability using the DISCERN instrument. Dermatologists produced more accurate content than laypeople, though important details such as treatment risks were often omitted. Our findings highlight the need for health professionals to balance accessibility with accuracy to provide trustworthy dermatologic information on social media.

(JMIR Dermatol 2025;8:e77504) doi:10.2196/77504

KEYWORDS

social media; Instagram; retinoid; retinol; risk; regulation

Introduction

Social media has become a daily source of connectivity, entertainment, and health information across all age groups. It has transformed how health knowledge is shared and consumed. Often, such content is delivered under the disclaimer “this post is not medical advice,” even as it implies otherwise. A lack of platform regulation allows individuals without medical training to share health-related information or personal experiences, often omitting context critical to patient understanding and safety. To increase engagement, even health care professionals–including dermatologists–have adapted their content into brief, attention-grabbing formats. This shift toward short-form, simplified videos often sacrifices educational depth, reducing complex topics to superficial overviews. The emphasis on visual appeal may further prioritize style over substance. Past studies evaluating social media health content often conclude that improved quality is essential for patient safety [1-3].

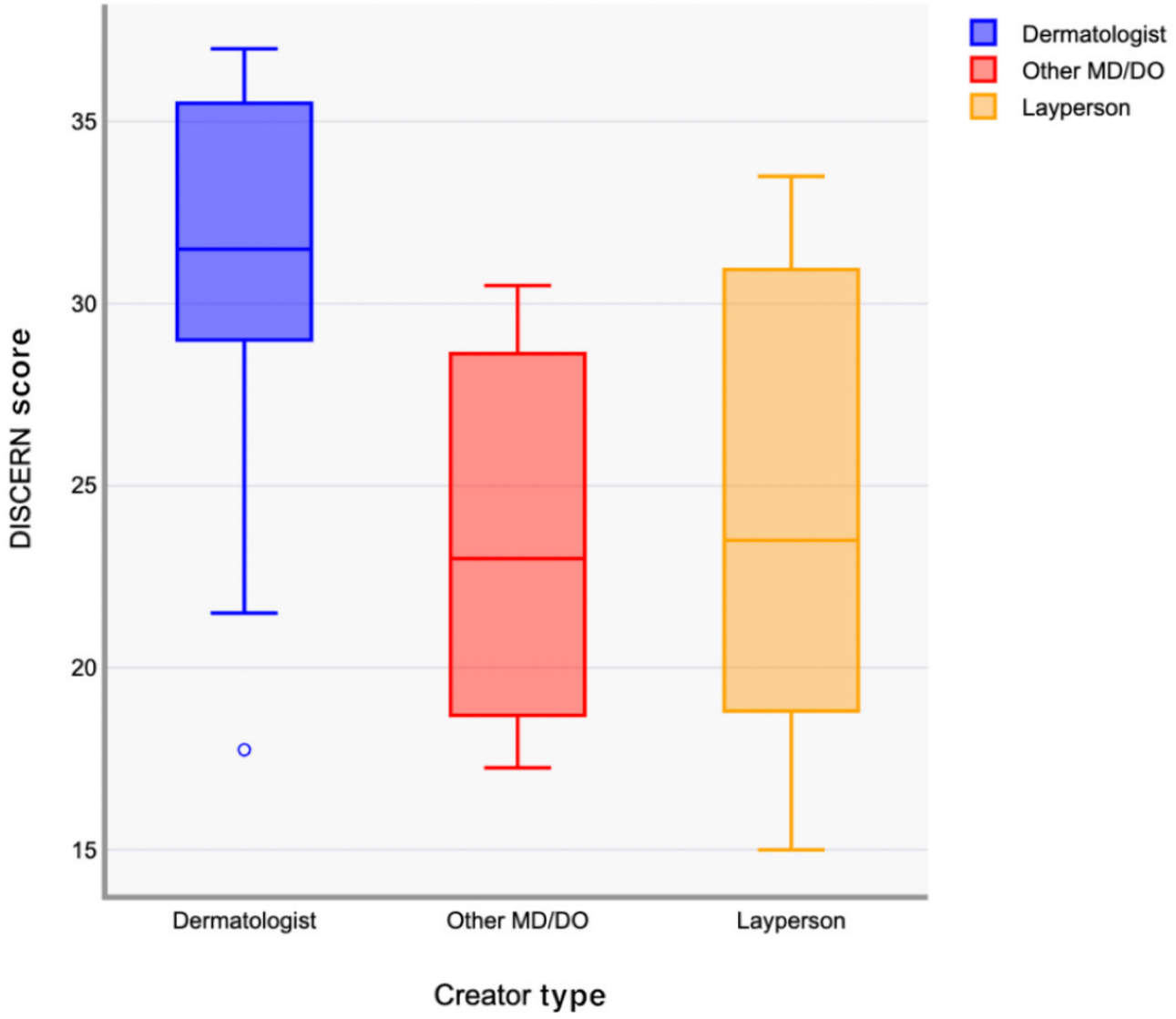
Methods

To assess the reliability of dermatologic information on Instagram, we analyzed the top 30 most-viewed videos under the hashtags #retinoid OR #retinol, each with over 50,000 views. There were no duplicate videos. The decision to include 30 videos was based on previous publications in the field that included a similar or smaller number of videos for their statistical analysis. Content creators were categorized as dermatologists, nondermatology physicians (other MDs/DOs), or laypeople. Videos were excluded if the creator’s qualifications could not be verified. Four independent reviewers with dermatology training scored each video using the DISCERN instrument, a validated tool for assessing consumer health information quality [4]. Because not all consumers read further into social media posts, we did not include captions, descriptions, and on-screen text in the DISCERN tool analysis. Median and mean scores for each group were calculated. A 1-way ANOVA was used to test for group differences, with post hoc Tukey analysis used for pairwise comparisons (significance set at $P<.05$).

Results

The breakdown of videos was as follows: 10 dermatologist-created, 3 other physician–created, and 17 layperson-created. We found a significant difference in DISCERN scores by creator type ($P=.04$). Post hoc analysis showed dermatologist-created videos scored significantly higher than those by laypeople ($P=.049$). Differences between dermatologists and nondermatology professionals ($P=.09$) and between nondermatology professionals and laypeople ($P=.98$) were not significant. The intraclass correlation coefficient among reviewers was 0.958, indicating excellent concordance. Results are shown in Figure 1.

Figure 1. Boxplot showing differences in average DISCERN scores across different creator categories.



Discussion

Principal Findings

Although dermatologist videos were higher in quality, even these showed major gaps. One commonly missed DISCERN criterion was question 11: “Does it describe the risks of each treatment?” This was underreported across all groups. This omission is concerning, as topical retinoids, though effective, are associated with irritation, dryness, and photosensitivity [5,6]. If these risks are not clearly explained, misuse or discontinuation may occur.

Interestingly, dermatologists outperformed other physicians, suggesting that specialty-specific expertise matters when communicating dermatologic information. Despite general medical training, nondermatology professionals may lack the nuanced understanding necessary to convey accurate skincare content.

Social media wields enormous influence on public health decisions. While creators may aim to educate, oversimplified or incomplete content can mislead. Our findings highlight the importance of evidence-based communication, especially when addressing topics outside one’s specialty. Clear presentation of risks, benefits, and limitations is essential to support informed, safe choices.

This study has limitations. Our analysis was limited to 30 videos, which may not capture the full range of retinoid-related content on Instagram. The DISCERN tool, though validated, was designed for written material and may not fully assess video nuances. Lastly, classification of creators was based on publicly available data and could be subject to misidentification.

Conclusion

Dermatologists and other health professionals must remain vigilant in sharing balanced, accurate information on social media. Improving the quality and transparency of content can help transform platforms like Instagram into trusted resources for skin health rather than sources of misinformation. Additionally, content regulation by social media companies is essential for protecting patients, as unregulated medical information can cause significant harm to patients.

Conflicts of Interest

None declared.

References

1. 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]
2. Druskovich C, Landriscina A. TikTok and dermatology: questioning the data. J Drugs Dermatol 2024 Aug 1;23(8):e167-e168. [doi: 10.36849/JDD.7956] [Medline: 39093646]
3. McCormick E, Nussbaum D, Desai S, Friedman A. Evaluating quality and reliability of most-viewed YouTube videos on topical steroid withdrawal. J Drugs Dermatol 2023 Nov 1;22(11):1128-1129. [doi: 10.36849/JDD.7443] [Medline: 37943267]
4. Tahir M, Usman M, Muhammad F, et al. Evaluation of quality and readability of online health information on high blood pressure using DISCERN and Flesch-Kincaid tools. Appl Sci (Basel) 2020 May 5;10(9):3214. [doi: 10.3390/app10093214]
5. Baldwin H, Webster G, Stein Gold L, Callender V, Cook-Bolden FE, Guenin E. 50 years of topical retinoids for acne: evolution of treatment. Am J Clin Dermatol 2021 May;22(3):315-327. [doi: 10.1007/s40257-021-00594-8] [Medline: 33871811]
6. Szymański Ł, Skopek R, Palusińska M, et al. Retinoic acid and its derivatives in skin. Cells 2020 Dec 11;9(12):2660. [doi: 10.3390/cells9122660] [Medline: 33322246]

Edited by S Subhadarshan; submitted 14.05.25; peer-reviewed by D Hitchcock, N Nadeau, T Charm; revised version received 22.07.25; accepted 30.08.25; published 12.09.25.

Please cite as:
Vanaria RJ, Povelaitis M, Khokhal R, Meledathu S, Green C, Ashinoff R, Mitre M
Omission of Risk in Vitamin A–Related Dermatologic Instagram Posts: A Growing Concern in an Unregulated Digital Landscape
JMIR Dermatol 2025;8:e77504
URL: https://derma.jmir.org/2025/1/e77504
doi:10.2196/77504

Use of a Large Language Model as a Dermatology Case Narrator: Exploring the Dynamics of a Chatbot as an Educational Tool in Dermatology

Emmanouil Karampinis[1,2], MSc, MD, PhD; Dafni Anastasia Bozi Tzetzi[3], MSc, MD; Georgia Pappa[4], MD; Dimitra Koumaki[5], MSc, MD, PhD; Dimitrios Sgouros[4], MD, PhD; Efstratios Vakirlis[6], MD, PhD; Aikaterini Liakou[7], MD, PhD; Markos Papakonstantis[8], MD; Marios Papadakis[9], MD, PhD; Dimitrios Mantzaris[10], MSc, PhD; Elizabeth Lazaridou[1], MD, PhD; Enzo Errichetti[11], MSc, MD; Cristian Navarrete-Dechent[12], MD; Angeliki Victoria Roussaki Schulze[2], MD, PhD; Alexandros Katoulis[4], MD, PhD

[1]Second Dermatology Department, School of Health Sciences, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

[10]Computational Intelligence and Health Informatics Lab, Nursing Department, University of Thessaly, Larissa, Greece

[11]Department of Experimental and Clinical Medicine, Institute of Dermatology, University of Udine, Udine, Italy

[12]Department of Dermatology, Escuela de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile

[2]Department of Dermatology, Faculty of Medicine, School of Health Sciences, University General Hospital of Larissa, University of Thessaly, Athens, Greece

[3]Department of Internal Medicine, General Hospital of Trikala, Trikala, Greece

[4]Second Department of Dermatology and Venereology, “Attikon” General University Hospital, Medical School, National and Kapodistrian University of Athens, Athens, Greece

[5]Dermatology Department, University Hospital of Heraklion, Heraklion, Greece

[6]First Department of Dermatology, Aristotle University Medical School, Thessaloniki, Greece

[7]First Department of Dermatology-Venereology, Faculty of Medicine, “Andreas Sygros” Hospital for Cutaneous and Venereal Diseases, National and Kapodistrian University of Athens, Athens, Greece

[8]Department of Dermatology, General Military Hospital of Athens, Athens, Greece

[9]Department of Surgery II, University of Witten Herdecke, Wuppertal, Germany

Corresponding Author:
Emmanouil Karampinis, MSc, MD, PhD
Second Dermatology Department, School of Health Sciences, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

Abstract

A comparison of dermatological cases generated by artificial intelligence (AI) versus those created without AI by medical students revealed that AI-created cases were characterized by detailed case descriptions, analysis of medical history, and clinical examinations, but lacked the depth, clinical relevance, and motivational elements found in non-AI cases, which were shorter, presented clinical dilemmas, and included challenging scenarios that students found more educational and engaging.

(JMIR Dermatol 2025;8:e72058) doi:10.2196/72058

KEYWORDS

chatbot; artificial intelligence; ChatGPT; dermatology; education; case reports; multiple choice answers; teaching methods

Introduction

Artificial intelligence (AI) is no longer a futuristic concept but a present reality that has rapidly changed all aspects of life; health care is no exception. In medical education, AI offers tools that have the potential to outperform traditional methods of teaching and learning. Dermatology, a medical specialty that relies almost exclusively on visual recognition and clinical pattern analysis, provides fertile ground for AI to revolutionize how medical students and aspiring dermatologists are trained [1]. Chatbots, such as ChatGPT, can play a transformative role in the medical education of dermatology. They can serve as on-demand tutors, providing instant explanations for complex dermatological terms, clarifying concepts, or answering questions in real time. This capability allows medical students to explore topics and receive personalized support and guidance while studying. Along with supporting individual learning, the tools are also invaluable for educators, as they enable efficient creation of teaching materials. AI-driven models can generate quizzes, flashcards, summary notes, and even realistic and diverse clinical case scenarios [2,3]. The primary objective of this study was to assess whether large language models like

ChatGPT-4 can create engaging, educational, and clinically relevant case-based scenarios for medical students.

Methods

Study Overview

Sixty-four medical students from two university-affiliated hospitals, Attikon University Hospital and University Hospital of Larissa, participated.

We assessed whether large language models like ChatGPT-4 can create engaging, educational, and clinically relevant case-based scenarios for medical students. “Engaging” refers to how much a case captures the learner’s interest, encourages active thinking, and motivates further exploration of the topic. Each AI-generated case was matched to a non-AI case on the same dermatological condition and similar educational objectives. Care was taken to ensure that both versions addressed comparable levels of difficulty, albeit with their inherent differences in style and formulation.

We developed a questionnaire featuring a mix of AI-generated and non-AI cases. The medical students were presented with multiple-choice questions, true/false statements, and correlation exercises. The AI-generated cases were created by instructing ChatGPT-4 to “generate an educational case in the form of multiple-choice questions for medical students” on the first attempt. Students graded each question on a Likert scale from 1 to 5, assessing how educational, motivational, or challenging they found the cases, without knowing the creator.

The final scores were evaluated regarding the normality with the Shapiro Wilk test, and then, the appropriate statistical tests were performed; for example, the Wilcoxon W test was the non-parametric test for paired data, as most comparisons were based on data with a significant departure from normality.

Ethical Considerations

Participation was entirely voluntary and anonymous. All the participants provided consent, and the included data were de-identified. Institutional Review Board approval was not applicable, as it relied exclusively on voluntary and anonymous questionnaire responses, without the collection of identifiable or sensitive personal information.

Results

Of 64 students, 45 answered the questionnaire (response rate of 70.3%). Among the 45 students, 36 (80%) reported using ChatGPT-4, 27 (60%) mentioned using Gemini-AI, while 9 (20%) students indicated they had never used a chatbot (the students could provide more than one answer). Twenty-five students (56%) stated that they used chatbots in their studies, though none reported using them in clinical practice.

Non-AI cases are thought to be more educational, motivational, and challenging in most scenarios, with statistical significance in many cases (Table 1). In contrast, in the case of True or False statements exercise and correlation test, no differences were detected between AI and non-AI examples. In the questionnaire students were asked about a conflicting situation involving multiple learning resources, including AI-based tools and traditional sources that provide different answers to the same question, 6.7% of students (3/45) would trust the AI answer, 68.9% of students (31/45) would trust the Internet source, and 24.4% of students (11/45) would further discuss the topic with a tutor.

Table . The median scores and range in different categories between artificial intelligence (AI) and non-AI cases on different skin diseases and lesions (Likert scale from 1 to 5).

	Educational	Motivational	Challenging
AI case on psoriasis	4 (4-5)	2 (2-5)	2 (1-4)
Non-AI case on psoriasis	4 (4-5)	5 (3-5)[a]	5 (3-5)[a]
AI case on atopic dermatitis	3 (3-4)	3 (3-4)	3 (2-3)
Non-AI case on atopic dermatitis	5 (4-5)[a]	5 (4-5)[a]	5 (4-5)[a]
AI case on rosacea	3 (3-4)	4 (3-4)	2 (2-3)
Non-AI case on rosacea	5 (4-5)[a]	5[a]	5 (4-5)[a]
AI case on HS[b]	3 (3-5)	4 (4-5)	4 (3-4)
Non-AI case on HS	4 (3-5)[a]	4 (3-5)	4 (4-5)
AI case on BCC[c]	3 (2-4)	3 (2-4)	3 (1-3)
Non-AI case on BCC	4 (3-4)	3 (2-5)	4 (3-4) [a]
AI case on actinic keratosis	3 (1-3)	4 (3-5)	4 (2-4)
Non-AI case on actinic keratosis	5 (4-5)[a]	5 (4-5)[a]	5 (3-5)[a]
AI case on melanoma	4 (2-5)	3 (1-4)	2 (1-2)
Non-AI case on melanoma	5 (3-5)[a]	5[a]	5 (4-5)[a]
True or False statements question by AI	4 (2-4)	4 (2-4)	3 (1-5)
True or False statements question by clinician	4 (2-4)	4 (2-5)	4 (2-5)
Correlation question by AI	4 (3-5)	4 (3-4)	2 (1-3)
Correlation question by clinician	4 (3-5)	4 (3-4)	3 (2-4)

[a]Statistical significant in favor of non-AI example(P<.05)

[b]HS: hidradenitis suppurativa.

[c]BCC: basal cell carcinoma.

Discussion

In dermatology, no studies have directly compared AI-generated case scenarios with those authored by clinicians in terms of the educational value, clinical relevance, or student engagement. In our study, AI-created cases were characterized by detailed case descriptions, analysis of medical history, clinical examinations, and follow-up questions but lacked the depth, clinical relevance, and motivational elements found in non-AI cases. Non-AI cases were shorter, presented clinical dilemmas, offered direct questions, and included challenging scenarios that students found more educational and engaging [4,5]. Non-AI questions, such as those comparing sebaceous hyperkeratosis and melanoma, differentiating between atopic dermatitis and scabies in a child patient, searching for comorbidities in rosacea, and critically evaluating a patient with hidradenitis suppurativa, were assessed with higher scores compared to the straightforward cases by AI chatbots. However, this observation does not imply that AI-generated questions have no value, as they can help students test their understanding. Tutors are also encouraged to make their own questions that simulate real-life scenarios that challenges students and do not rely blindly on chatbots for the quick production of exercises.

Acknowledgments

In this section, we acknowledge the students who dedicated their time to fill the questionnaire.

Data Availability

The data described in this study are available upon request from the corresponding author.

Conflicts of Interest

None declared.

References

1. Kim YH, Zhang MZ, Vidal NY. ChatGPT offers an editorial on the opportunities for chatbots in dermatologic research and patient care. DOJ 2024;29(6). [doi: 10.5070/D329662990] [Medline: 38959931]
2. Chakraborty C, Pal S, Bhattacharya M, Dash S, Lee SS. Overview of chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. Front Artif Intell 2023;6:1237704. [doi: 10.3389/frai.2023.1237704] [Medline: 38028668]
3. Donoso F, Peirano D, Agüero R, et al. Use of game-based learning strategies for dermatology and dermoscopy education: a cross-sectional survey of members of the International Dermoscopy Society. Clin Exp Dermatol 2025 Jan 27;50(2):365-371. [doi: 10.1093/ced/llae375] [Medline: 39298635]
4. Shapiro J, Lyakhovitsky A, Freud T, et al. Assessing ChatGPT-4's capabilities in generating dermatology board examination content: an explorational study. Acta Derm Venereol 2025 Jan 3;105:adv41208. [doi: 10.2340/actadv.v105.41208] [Medline: 39749389]
5. Khamaysi Z, Awwad M, Jiryis B, Bathish N, Shapiro J. The role of ChatGPT in dermatology diagnostics. Diagnostics (Basel) 2025 Jun 16;15(12):1529. [doi: 10.3390/diagnostics15121529] [Medline: 40564849]

Abbreviations

AI: artificial intelligence

Edited by R Dellavalle; submitted 02.02.25; peer-reviewed by B Ozek, R Yang; revised version received 21.08.25; accepted 22.08.25; published 16.09.25.

Please cite as:
Karampinis E, Bozi Tzetzi DA, Pappa G, Koumaki D, Sgouros D, Vakirlis E, Liakou A, Papakonstantis M, Papadakis M, Mantzaris D, Lazaridou E, Errichetti E, Navarrete-Dechent C, Roussaki Schulze AV, Katoulis A
Use of a Large Language Model as a Dermatology Case Narrator: Exploring the Dynamics of a Chatbot as an Educational Tool in Dermatology
JMIR Dermatol 2025;8:e72058
URL: https://derma.jmir.org/2025/1/e72058
doi:10.2196/72058

The Comparative Sufficiency of ChatGPT, Google Bard, and Bing AI in Answering Diagnosis, Treatment, and Prognosis Questions About Common Dermatological Diagnoses

Courtney A Chau[1], BS; Hao Feng[2], MD; Gabriela Cobos[3], MD; Joyce Park[4], MD

[1]Icahn School of Medicine at Mount Sinai, New York, NY, United States

[2]Department of Dermatology, University of Connecticut Health Center, Farmington, CT, United States

[3]Department of Dermatology, Tufts Medical Center, 260 Tremont St, Fl 13, Boston, MA, United States

[4]Skin Refinery PLLC, Spokane, WA, United States

Corresponding Author:
Gabriela Cobos, MD
Department of Dermatology, Tufts Medical Center, 260 Tremont St, Fl 13, Boston, MA, United States

Abstract

Our team explored the utility of unpaid versions of 3 artificial intelligence chatbots in offering patient-facing responses to questions about 5 common dermatological diagnoses, and highlighted the strengths and limitations of different artificial intelligence chatbots, while demonstrating how chatbots presented the most potential in tandem with dermatologists' diagnosis.

(JMIR Dermatol 2025;8:e60827) doi:10.2196/60827

KEYWORDS

artificial intelligence; AI; ChatGPT; atopic dermatitis; acne vulgaris; cyst; actinic keratosis; rosacea; diagnosis; treatment; prognosis; dermatological; patient; chatbot; dermatologist

Introduction

Artificial intelligence (AI) chatbots, such as ChatGPT, offer platforms for patients to ask medical questions, particularly with limited access to care [1]. Although ChatGPT utility in dermatology has been assessed, few studies have compared the performance between chatbots [2]. This study compared the clinical utility of the unpaid versions of ChatGPT 3.5, Google Bard, and Bing AI in generating patient-facing responses to questions about 5 common dermatological diagnoses (atopic dermatitis, acne vulgaris, actinic keratosis, cyst, and rosacea) [3].

Methods

For each condition, 2 diagnosis, 2 treatment, and 1 prognosis questions were devised. Diagnosis questions requested a diagnosis and presented the patient history including age, sex, symptoms (duration/location), treatments and outcomes, and medical history. Nineteen questions were modeled from questions on Reddit forums ("r/AskDocs" and "r/dermatology"). For topics with insufficient Reddit questions, the coauthors devised prompts reflecting common questions in their experience (6 questions).

Questions were inputted into each chatbot; the prompts used are shown in Multimedia Appendix 1. Three board-certified dermatologists scored the responses on appropriateness for a patient-facing platform (Yes/No), sufficiency for clinical practice (Yes/No: not specific, not concise, or inaccurate information), accuracy from 1 (completely inaccurate) to 6 (completely accurate), and overall from 1 (worst possible answer) to 10 (best possible answer) [4]. The Wilcoxon rank-sum test was used for pairwise comparisons. *P*-values were adjusted using the Bonferroni correction.

Results

One response was omitted because Google Bard declined answering the second atopic dermatitis diagnosis question ("I am a 19-year old…"), responding with, "I'm just a language model, so I can't help you with that." ChatGPT responses had significantly lower Flesch reading ease scores than Google Bard ($P<.001$) and Bing AI ($P<.001$), indicating lower comprehensibility (Table 1). ChatGPT responses received significantly higher accuracy ($P=.01$, Figure 1) and overall ($P=.003$) ratings than Bing AI. Considering patient-facing platform appropriateness and clinical practice sufficiency, ChatGPT received the most appropriate (95%) and sufficient (55%) ratings; Bing AI received the fewest (87% and 55%, respectively). In total, 45%, 49%, and 53% of ChatGPT, Google Bard, and Bing AI responses, respectively, had inaccurate information or were not specific. For diagnosis prompts, 9 of 10 of ChatGPT and Bing AI and 7 of 10 of Google Bard responses included the intended diagnosis. Of the 25 responses from each chatbot, 25 of Bing AI's, 24 of ChatGPT's, and 19

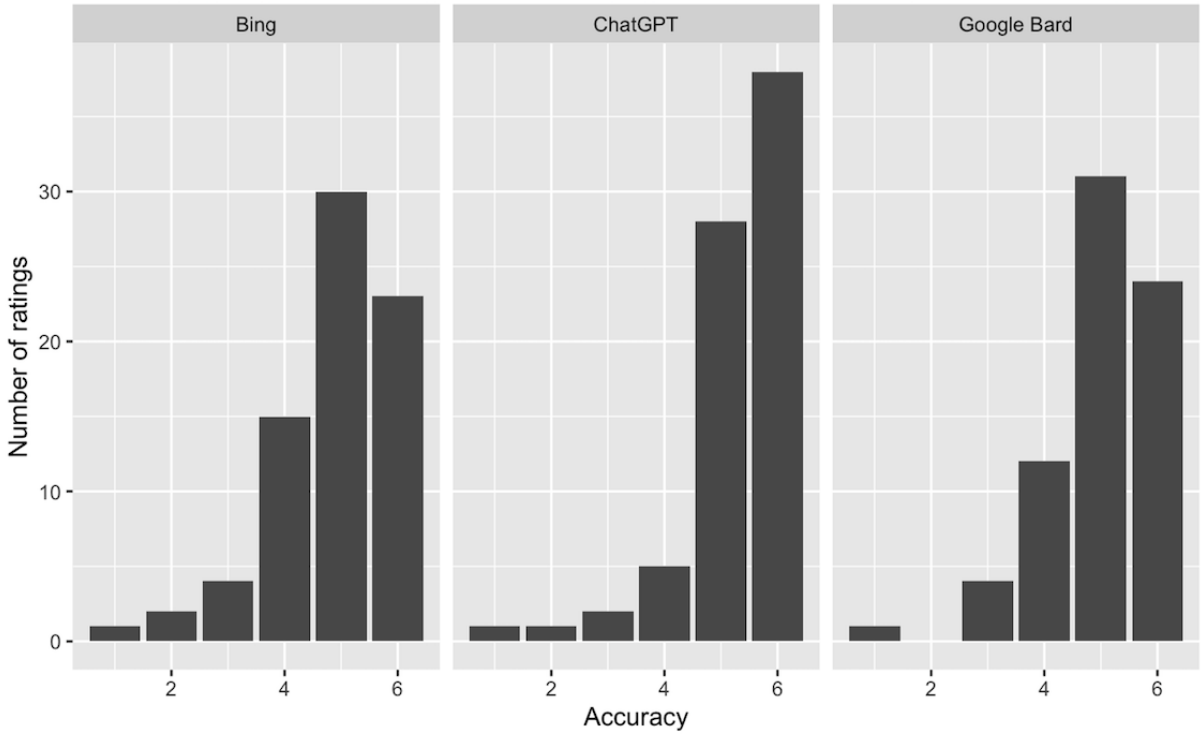
of Google Bard's responses emphasized the importance of consulting healthcare professionals. No fabrication or hallucination was observed for any chatbot responses.

Table . Descriptive statistics of scores between chatbots.

	ChatGPT 3.5 (n=75)	Google Bard (n=72)	Bing AI (n=75)
Mean Flesch reading ease score (SD)*[a]	33.90 (8.1)	49.72 (15.4)	46.53 (9.7)
Mean accuracy (SD)	5.29 (0.97)	5.00 (0.98)	4.87 (1.1)
Mean overall rating (SD)	8.37 (1.8)	7.94 (1.9)	7.41 (2.1)
Number of responses appropriate for a patient-facing platform (%)	71 (95)	65 (90)	65 (87)
Sufficiency for clinical practice			
Yes (%)	41 (55)	35 (49)	35 (47)
No: not specific enough (%)	14 (19)	15 (21)	23 (31)
No: inaccurate information (%)	20 (27)	20 (28)	17 (23)
No: not concise (%)	0	2 (3)	0

[a]Out of n=25 for ChatGPT and Bing AI and n=24 for Google Bard because only 1 Flesch reading ease score was calculated for each response. The other measures in the table are based on evaluation of each chatbot response by 3 board-certified dermatologists.

Figure 1. Distribution of the accuracy ratings for each chatbot. The accuracy scores from the three board-certified dermatologists ranged from 1 (completely inaccurate) to 6 (completely accurate).



Discussion

ChatGPT outputs were most accurate and appropriate for patient questions. However, ChatGPT responses had college-level readability, limiting public utility [5]. Responses were deemed sufficient for clinical practice if the chatbot concisely provided completely correct information that specifically answered the patient's question without missing critical components. Only approximately half the responses were sufficient for clinical practice, primarily due to inaccuracies and lack of specificity. ChatGPT and Bing AI performed the best at diagnosis and emphasized the importance of seeking input from a healthcare professional. Google Bard did not perform well in these domains, indicating that it is less suitable for suggesting diagnoses. Despite the better diagnostic performance of ChatGPT and Bing AI, an unranked list of conditions with differing treatments is not actionable for patients. Chatbots present more potential in offering advice once a diagnosis has been established. This study is limited by exploring only 5 questions for each of the 5 conditions. Exploring a broader range

of conditions with a larger set of questions would more robustly capture chatbots' performance. However, this study lays the groundwork for future research to compare chatbots using more expansive domains.

ChatGPT 3.5 displays more promise than Google Bard and Bing AI in evaluating, diagnosing, and suggesting a treatment plan for dermatologic conditions, consistent with previous findings, in which the chatbots' responses to questions about melanoma were evaluated [2]. However, this study revealed several important improvements needed for all 3 chatbots: enhancing readability, removing inaccuracies, and improving information specificity. Dermatologists may be able to reference these AI in practice, to limited extents, by suggesting patients use AI as a reference only to obtain information about the condition after being diagnosed. This strategy is similar to paper handouts, where AI chatbots provide background knowledge that patients can later follow-up on with their dermatologist. In conclusion, while chatbot utility is most promising in tandem with a dermatologist's diagnosis and contributes to information dissemination, chatbots should not function as a first-line independent entity. As access to AI grows, dermatologists must be aware of the quality of information patients may receive from AI and how it may differ from a dermatologist's advice.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prompts inputted into ChatGPT 3.5, Google Bard, and Bing AI.
[DOCX File, 19 KB - derma_v8i1e60827_app1.docx]

References

1. Baker MN, Burruss CP, Wilson CL. ChatGPT: a supplemental tool for efficiency and improved communication in rural dermatology. Cureus 2023 Aug;15(8):e43812. [doi: 10.7759/cureus.43812] [Medline: 37731429]
2. Mu X, Lim B, Seth I, et al. Comparison of large language models in management advice for melanoma: Google's AI BARD, BingAI and ChatGPT. Skin Health Dis 2024 Feb;4(1):e313. [doi: 10.1002/ski2.313] [Medline: 38312244]
3. Landis ET, Davis SA, Taheri A, Feldman SR. Top dermatologic diagnoses by age. Dermatol Online J 2014 Apr 16;20(4):22368. [Medline: 24746305]
4. Young JN, Poplausky D, et al. The utility of ChatGPT in generating patient-facing and clinical responses for melanoma. J Am Acad Dermatol 2023 Sep;89(3):602-604. [doi: 10.1016/j.jaad.2023.05.024] [Medline: 37207953]
5. Hutchinson N, Baird GL, Garg M. Examining the reading level of internet medical information for common internal medicine diagnoses. Am J Med 2016 Jun;129(6):637-639. [doi: 10.1016/j.amjmed.2016.01.008] [Medline: 26829438]

Abbreviations

AI: artificial intelligence

Edited by R Dellavalle; submitted 22.05.24; peer-reviewed by E Bai, S Mao, YT Cho; revised version received 16.10.24; accepted 17.10.24; published 07.01.25.

Please cite as:
Chau CA, Feng H, Cobos G, Park J
The Comparative Sufficiency of ChatGPT, Google Bard, and Bing AI in Answering Diagnosis, Treatment, and Prognosis Questions About Common Dermatological Diagnoses
JMIR Dermatol 2025;8:e60827
URL: https://derma.jmir.org/2025/1/e60827
doi:10.2196/60827

Information Regarding Dermatology as Seen on the Social Media Platform TikTok

Alexandra Rose Nigro, MS; Alim Osman, MS; Ryan Charles Saal, BS; Robert James Smith, MD

Department of Dermatology, Eastern Virginia Medical School, 825 Fairfax Ave, Norfolk, VA, United States

Corresponding Author:
Alexandra Rose Nigro, MS
Department of Dermatology, Eastern Virginia Medical School, 825 Fairfax Ave, Norfolk, VA, United States

(JMIR Dermatol 2025;8:e59597) doi:10.2196/59597

KEYWORDS

internet; social media; dermatology; TikTok; health care information; dermatological content; statistical analysis; dermatologists; online content; misinformation; engagement; health information; patient health; web platforms; patient education; digital health; online information; skin care; online videos

Introduction

Americans spend an average of 2 hours and 7 minutes per day on social media [1]. It has become evident that patients are utilizing these networks to seek health care knowledge [2]. Previous studies have shown that dermatologists are using platforms such as Facebook or Instagram to disseminate dermatological content [3,4]. Interestingly, the information provided has often been inaccurate [5]. This study aims to describe the quantity and quality of dermatological content on TikTok, addressing the gap in research on dermatological information shared on this platform.

Methods

Overview

The TikTok platform was screened in November 2022 for the following keywords: dermatology, derm, skin, skin care, aesthetician, dermatologist, esthetician, derm physician assistant, dermatology physician assistant, derm PA, dermatology nurse, derm nurse, and derm nurse practitioner. Accounts containing ≥50% of dermatological content were included. An engagement score for these accounts was calculated by dividing the total number of likes by the total number of followers. The top 10 accounts were isolated based on engagement score, and their top 10 videos underwent further analysis using the DISCERN criteria (Table 1). DISCERN scores are calculated by assessing health information using a set of 16 standardized questions that evaluate factors like reliability, clarity, and overall quality of information presented. Each question is rated on a scale from 1 to 5, with higher scores indicating more reliable and high-quality content. These accounts were then grouped and analyzed based on occupation (Table 2). Two investigators (AO and ARN) independently scored the videos; differences in scoring were resolved by discussion between them. Kruskal-Wallis, Mann-Whitney *U*, and 2-tailed independent *t* tests were used to analyze the data. These statistical tests assessed the differences in content quality and engagement across occupations, providing insight into significant disparities without assuming normal data distribution.

Table . TikTok video data of the top 10 accounts. The average video length, views, likes, comments, and DISCERN score of the top 10 included accounts based on engagement score are displayed.

Account	Average video length (of the top 10 videos) (seconds), n	Average number of views (of the top 10 videos), n	Average likes (of the top 10 videos), n	Average comments (of the top 10 videos), n	Average DISCERN score (of the top 10 videos), n
Account 1	16.6	1,049,350	102,880	297.5	34.3
Account 2	15.2	1,409,470	107,408.6	672.2	34
Account 3	39	16,640,000	3,200,000	12,121.5	31.1
Account 4	23.2	3,088,690	176,179.4	1349.1	33.1
Account 5	20.5	2,660,000	122,340	480.8	34.8
Account 6[a]	50.8	64,920,000	4,187,020	20,217.6	32
Account 7[a]	44.7	13,590,000	121,703.5	623.4	37.6
Account 8[a]	21.5	26,970,000	1,115,094.8	6137.4	34.5
Account 9[a]	11	16,110,000	527,120	4983.3	32.3
Account 10[a]	16.7	8,120,000	721,210	2517.6	34.5

[a]Account belonging to a dermatologist.

Table . TikTok video data of the top 10 accounts. The analysis of the top 10 accounts from Table 1 based on occupation are displayed.

Account owner occupation	Engagement score, mean (SD)	Total videos, mean (SD)	Informational videos, mean (SD)	Number of videos with TikTok-verified paid sponsorship, mean (SD)
Dermatologists	21.66 (5.40)	795 (562.84)	691.22 (497.71)	18.22 (23.95)
Medical clinics	32.96 (14.54)	402.25 (243.48)	327 (274.68)	0 (0)
Aestheticians	26.47 (21.88)	478 (238.58)	316.5 (171.99)	3.17 (2.56)
Medical spas	15.77 (2.39)	279.5 (19.09)	199.5 (13.43)	0 (0)

Ethical Considerations

This study was exempt from institutional review board approval, as only publicly available data were used.

Results

Of the 22,411 TikTok videos analyzed, 17,085 (76.3%) videos were informational videos discussing skin and dermatology. Of these, 502 (2.2%) videos featured TikTok-verified paid sponsorships, with dermatologists comprising 328 (65%) videos of these. Among the 94 accounts analyzed, 35 met the inclusion criteria, including 18 (51.4%) dermatologists, 6 (17.1%) aestheticians, 4 (11.4%) medical clinics, and 2 (5.7%) medical spas. Using the Kruskal-Wallis test, significant differences in engagement scores were found across groups (P<.001), with medical clinics and aestheticians showing the highest average engagement scores.

A 2-tailed independent t test was used to confirm statistical significance between DISCERN scores of the dermatologist-run and nondermatologist-run accounts. Dermatologist-run accounts had significantly more views (25,942,000 vs 4,969,502; P<.001) and comments (6895.86 vs 2984.22; P=.04). Overall DISCERN scores did not differ between dermatologist-run and nondermatologist-run accounts (34.18 vs 33.46; P=.53), but subsections of the DISCERN scale did show significant differences. Dermatologist-run accounts were more likely to refer to areas of uncertainty (1.76 vs 1.22; P<.01) and describe how each treatment works (2 vs 1.48; P=.03). However, nondermatologist-run accounts were more likely to describe each treatment's risks (1.56 vs 1.24; P=.049). The mentioned scores were still very low for both groups, suggesting an overall low quality of content for the videos.

Discussion

There is a high volume of dermatological content produced and consumed on TikTok. Prior research demonstrates that consumer trust in TikTok content is high [6]. However, this study reveals that the most productive content creators in dermatology's TikTok videos are of poor educational accuracy and quality. Physicians should be aware of the expanding role of social media as a source of medical information and possible misinformation. Dermatologists should consider if they have an obligation to produce unbiased, ethical, and accurate content. Social media platforms also bear a responsibility in regulating health-related content. Loeb et al [7] found widespread cancer misinformation on social media, often finding dissemination of unproven treatment options, demonstrating one example of how medical misinformation on social media can potentially harm patients. Though dermatologists should only report accurate information for proper patient care, it is equally important for

social media platforms to at least label content as potential misinformation to prevent inadequate skin care.

Additionally, it is important to note the differences in the content reported between dermatologists and nondermatologists. With dermatologists more frequently referencing treatment uncertainty, viewers may perceive this as physicians practicing cautious, evidence-based medicine. With nondermatologists more frequently discussing risks, this may lead to consumers avoiding treatments and seeking alternative treatment solutions. Together, these differing emphases could shape consumer attitudes towards dermatological care, treatment choices, and even trust in professional advice versus influencer-driven content. Further exploration of these impacts is warranted to better understand the influence of social media on dermatological care.

Conflicts of Interest

None declared.

References

1. Wong B. Top social media statistics and trends of 2024. Forbes. 2023 May 18. URL: https://www.forbes.com/advisor/business/social-media-statistics/ [accessed 2025-01-16]
2. 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]
3. Kream EJ, Jerdan K. Digital brand building and online reputation management in dermatology. Clin Dermatol 2023;41(2):240-245. [doi: 10.1016/j.clindermatol.2023.06.006] [Medline: 37453712]
4. Ross NA, Todd Q, Saedi N. Patient seeking behaviors and online personas: social media's role in cosmetic dermatology. Dermatol Surg 2015 Feb;41(2):269-276. [doi: 10.1097/DSS.0000000000000267] [Medline: 25654198]
5. 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]
6. Gordon D. 33% of Gen Zers trust TikTok more than doctors, new survey shows. Forbes. URL: https://www.forbes.com/sites/debgordon/2022/12/20/33-of-gen-zers-trust-tiktok-more-than-doctors-new-survey-shows/ [accessed 2025-01-16]
7. Loeb S, Langford AT, Bragg MA, Sherman R, Chan JM. Cancer misinformation on social media. CA Cancer J Clin 2024;74(5):453-464. [doi: 10.3322/caac.21857] [Medline: 38896503]

Edited by R Dellavalle; submitted 16.04.24; peer-reviewed by B Irfan; revised version received 07.11.24; accepted 09.11.24; published 28.01.25.

Please cite as:
Nigro AR, Osman A, Saal RC, Smith RJ
Information Regarding Dermatology as Seen on the Social Media Platform TikTok
JMIR Dermatol 2025;8:e59597
URL: https://derma.jmir.org/2025/1/e59597
doi:10.2196/59597

A Google Trends Analysis of Search Interest for Tender-Headedness and Scalp-Related Concerns

Charissa Obeng-Nyarko[1], MS; Tatiana Barrera[2], BS; Temitayo Ogunleye[3], MD; Susan Taylor[3], MD

[1]Florida State University College of Medicine, 1115 West Call Street, Tallahassee, FL, United States

[2]University of California - Riverside, Riverside, CA, United States

[3]Department of Dermatology, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Charissa Obeng-Nyarko, MS
Florida State University College of Medicine, 1115 West Call Street, Tallahassee, FL, United States

Abstract

In this Google Trends cross-sectional analysis, we aimed to understand the popularity of tender-headedness by analyzing related Google search queries from January 2013 to December 2022. Since 2013, Google searches on scalp-related concerns, especially those regarding tender-headedness in Black hair culture, have increased, thus uncovering an opportunity for dermatologists to utilize culturally relevant language to address scalp concerns in patients with Afro-textured hair.

(JMIR Dermatol 2025;8:e60040) doi:10.2196/60040

KEYWORDS

tender-headedness; tender-headed; scalp tenderness; dermatologists; Google Trends; Black patients

Introduction

In Black hair culture, “tender-headed” is a term that refers to someone with heightened scalp discomfort or tenderness during hair manipulation procedures like combing, brushing, braiding, twisting, hair parting, and blow-drying [1-3].

Little is known about tender-headedness, as it usually lacks clinical findings. However, symptoms include mild to significant scalp discomfort, which can occur among all ethnicities but may be more prevalent among women with Afro-textured hair [4]. Scalp tenderness is a common symptom in inflammatory alopecias, including central centrifugal cicatricial alopecia and traction alopecia, which predominantly affect Black women [5]. Understanding culturally relevant language for scalp tenderness is important for dermatologists to differentiate between nonpathologic and pathologic scalp issues in this population.

The internet is a commonly used source for information on hair and scalp care, particularly for people of African descent who may seek solutions for tender-headedness on search engines and forums [6]. To date, there is limited knowledge about internet search interest regarding tender-headedness. In this study, we aim to understand the popularity of tender-headedness by analyzing search queries related to this concept on a major search engine.

Methods

Google Trends (GT) is a Google-developed tool that reports on the popularity of specific searches. Output from GT is in the form of a search volume index (SVI), which represents the popularity of a specific search over time [7]. SVI values are normalized on a scale from 0 to 100, with 0 representing the lowest level of interest and 100 representing the highest [7]. These values depend on the specific search phrase, time range, and geographical area selected [7]. They may vary slightly by query date, so all values were queried on the same day for consistency [7].

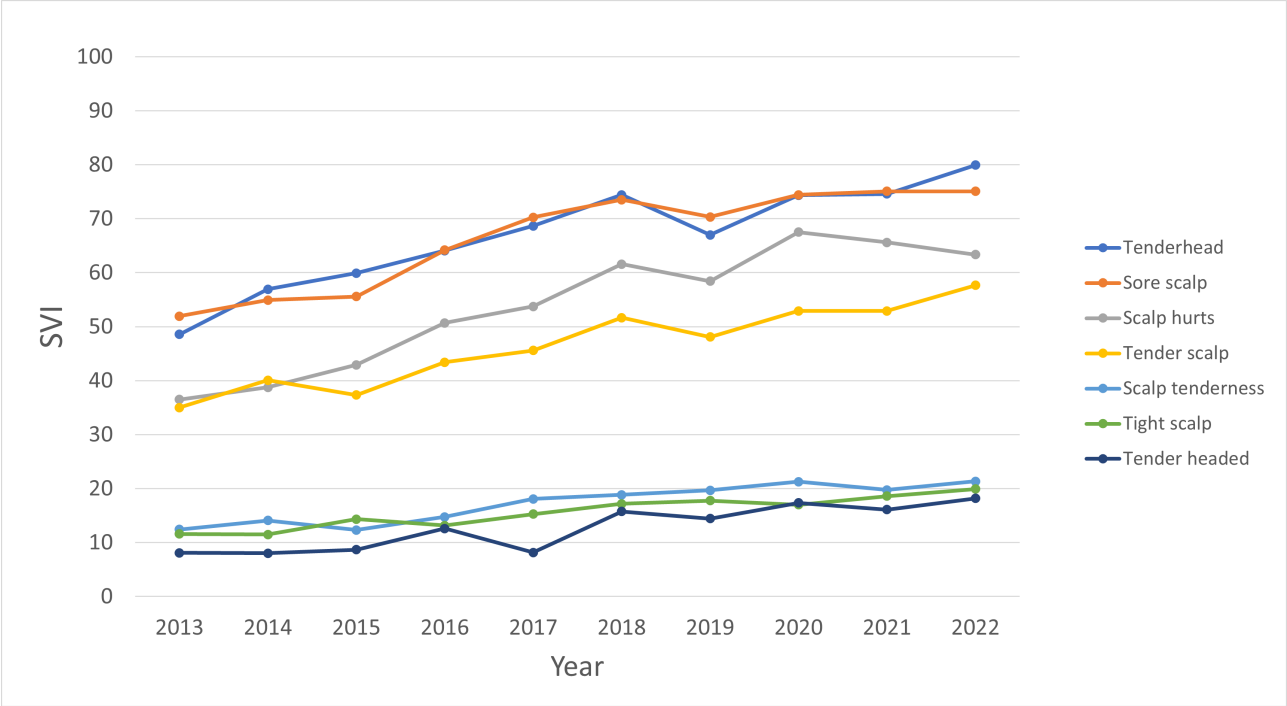
In this cross-sectional analysis, GT was used to extract the monthly web SVI from January 2013 to December 2022 for the following seven keyword phrases (KPs): “tender headed,” “tender head,” “sore scalp,” “scalp hurts,” “tight scalp,” “tender scalp,” and “scalp tenderness.” Additionally, KPs were grouped into the following three categories of generic words used to describe tender-headedness: (1) tenderness (“tender headed” and “tender head”), (2) scalp discomfort (“sore scalp,” “scalp hurts,” and “tight scalp”), and (3) both concepts (“tenderness” and “scalp”) combined (“tender scalp” and “scalp tenderness”). Differences in the mean monthly SVI per category were compared via a generalized estimated equation with Gaussian estimation and exchangeable correlation, using Stata version 18 (StataCorp LLC). Statistical significance was measured at $P<.05$.

Results

Among the seven KPs used in US internet queries made between January 2013 and December 2022, “tender head” and “sore scalp” had the highest mean SVIs (67 for both; Figure 1). The internet search interest for the term “sore scalp” was comparable

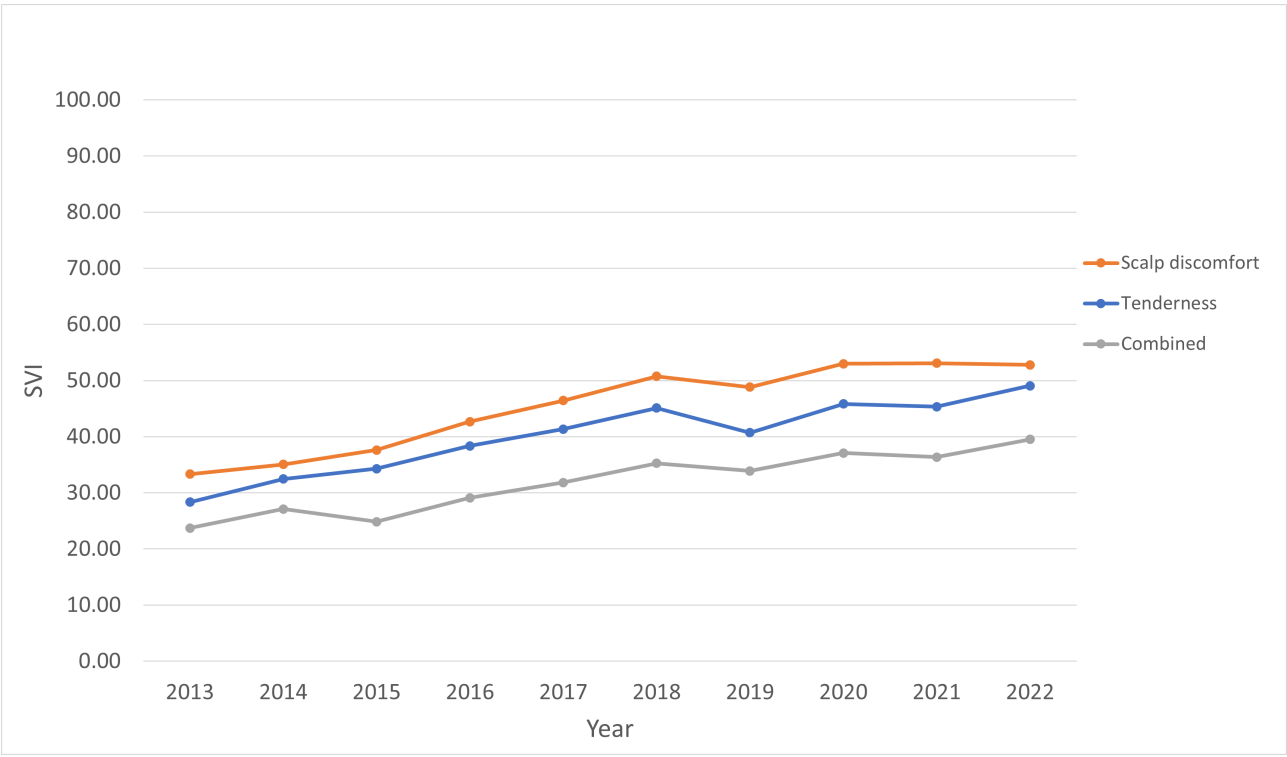
to that for "tender head" (R=0.32, 95% CI −1.23 to 1.87; P=.69). The term "tender headed" yielded lower search interest compared to "tender head" (R=−53.3, 95% CI −55.08427 to −51.98239; P<.001).

Figure 1. Yearly internet search interest by keyword phrase from 2013 to 2022. SVI: search volume index.



The "scalp discomfort" category had the highest mean internet search interest (SVI=45.35) compared to the "tenderness" (SVI=40.08) and "combined" (SVI=31.86) categories (Figure 2). When compared to searches for the "tenderness" category, there was significantly higher search interest for the "scalp discomfort" category (R=5.07, 95% CI 4.21-5.93; P<.001). Additionally, the "combined" category showed significantly lower interest in comparison to "tenderness" (R=−8.19, 95% CI −9.05 to −7.33; P<.001).

Figure 2. Yearly internet search interest by keyword phrase category from 2013 to 2022. SVI: search volume index.



Discussion

In this study, we found that terms describing scalp discomfort generated the highest internet search interest among KP categories. Furthermore, among the seven KPs describing tender-headedness, "tender head" and "sore scalp" had the greatest internet search volume.

Although our findings did not show a solid search trend for the keyword "tender headed," "tender head" and "sore scalp" are relevant phrases that some individuals use to describe tender-headedness. "Tender head" has gained popularity on the web, with search engine queries yielding culturally specific articles on managing tender-headedness in children and adults with Afro-textured hair [2,3,8,9] and the need for dermatologic care [3,8,9]. Understanding this context may benefit dermatologists when discussing scalp concerns with Black patients.

This study acknowledges that using GT to capture internet search interest for scalp concerns may not fully represent individuals without internet access, limiting the results' applicability to the broader population [7]. Despite this limitation, GT is a powerful tool for gauging public interest in dermatology-related terms and conditions [7].

Since 2013, Google searches on scalp-related concerns, especially those regarding tender-headedness in Black hair culture, have increased. Further research is needed to characterize tender-headedness and understand its relationship with hair and scalp disorders in people of African descent.

Conflicts of Interest

ST has served as a consultant, advisory board member, and/or speaker for AbbVie, Arcutis, Armis Scientific, Avita, Beiersdorf, Biorez, Bristol-Myers Squibb, Cara Therapeutics, Dior, Eli Lilly, EPI Health, Evolus, Galderma, GloGetter, Hugel America, Incyte, Johnson & Johnson, L'Oreal USA, MedScape, MJH LifeSciences, Pfizer, Piction Health, Sanofi, Scientis US, UCB, and Vichy Laboratoires. She has received royalties from McGraw Hill. ST has served as an investigator for Allergan, Concert Pharmaceuticals/Sun Pharma, Croma-Pharma GmbH, Eli Lilly, and Pfizer.

References

1. Mitchell A. I'm not tender-headed, you're hurting me. Refinery29. 2023 Nov 16. URL: https://www.refinery29.com/en-us/tenderheaded-meaning [accessed 2024-02-13]
2. Cloud S. Hair tenderness—the life of a tenderhead. Medium. 2023 Jun 24. URL: https://medium.com/@selenacloud52/hair-tenderness-the-life-of-a-tenderhead-c7d676f84014 [accessed 2024-08-04]
3. Gadson K. Tales of a tender head: how my sensitive scalp turned out to be a health concern. Blavity. 2016 Aug 18. URL: https://blavity.com/tales-tender-head-sensitive-scalp-turned-health-concern [accessed 2024-08-04]
4. Fajuyigbe D, Sewraj P, Connétable S, et al. Weekly hair washing: the recommended solution for women with afro-textured hair to alleviate dandruff and scalp discomfort. J Dermatol 2024 Apr;51(4):518-525. [doi: 10.1111/1346-8138.17102] [Medline: 38217001]
5. Larrondo J, McMichael AJ. Scarring alopecia. Dermatol Clin 2023 Jul;41(3):519-537. [doi: 10.1016/j.det.2023.02.007] [Medline: 37236719]
6. Ellington TN. Social networking sites: a support system for African-American women wearing natural hair. Int J Fash Des Technol Educ 2014 Oct 31;8(1):21-29. [doi: 10.1080/17543266.2014.974689]
7. 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]
8. How to stop being tenderheaded? The Original Tender Head. 2023 Sep 30. URL: https://originaltenderhead.com/how-to-stop-being-tender-headed [accessed 2024-02-13]
9. Bost M. Tips for your tender-headed child. Collin County Moms. 2024 Mar 21. URL: https://collincounty.momcollective.com/beauty-fashion/tips-tender-headed-child/ [accessed 2024-08-04]

Abbreviations

GT: Google Trends
KP: keyword phrase
SVI: search volume index

Edited by R Dellavalle; submitted 29.04.24; peer-reviewed by L Marcial, R Yi Kow; revised version received 10.09.24; accepted 17.09.24; published 13.02.25.

Please cite as:
Obeng-Nyarko C, Barrera T, Ogunleye T, Taylor S
A Google Trends Analysis of Search Interest for Tender-Headedness and Scalp-Related Concerns
JMIR Dermatol 2025;8:e60040
URL: https://derma.jmir.org/2025/1/e60040
doi:10.2196/60040

Assessing the Diagnostic Accuracy of ChatGPT-4 in Identifying Diverse Skin Lesions Against Squamous and Basal Cell Carcinoma

Nitin Chetla[1*], BS; Matthew Chen[2*], BS; Joseph Chang[3], BS; Aaron Smith[1], BS; Tamer Rajai Hage[4], BS; Romil Patel[1], BS; Alana Gardner[5]; Bridget Bryer[6], MD

[1]University of Virginia School of Medicine, University of Virginia, 828 Cabell Avenue, Charlottesville, VA, United States

[2]Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, United States

[3]University of Passau, Passau, Germany

[4]Virginia Tech, Blacksburg, VA, United States

[5]University at Albany, State University of New York, Albany, NY, United States

[6]Department of Dermatology, University of Virginia, Charlottesville, VA, United States

[*]these authors contributed equally

Corresponding Author:
Nitin Chetla, BS
University of Virginia School of Medicine, University of Virginia, 828 Cabell Avenue, Charlottesville, VA, United States

Abstract

Our study evaluates the diagnostic accuracy of ChatGPT-4o in classifying various skin lesions, highlighting its limitations in distinguishing squamous cell carcinoma from basal cell carcinoma using dermatoscopic images.

(JMIR Dermatol 2025;8:e67299) doi:10.2196/67299

KEYWORDS

chatbot; ChatGPT; ChatGPT-4; squamous cell carcinoma; basal cell carcinoma; skin cancer; skin cancer detection; dermatoscopic image analysis; skin lesion differentiation; dermatologist; machine learning; ML; artificial intelligence; AI; AI in dermatology; algorithm; model; analytics; diagnostic accuracy

Introduction

Squamous cell carcinoma (SCC) and basal cell carcinoma (BCC) are prevalent skin cancers that can cause significant local tissue damage and disfigurement as well as mortality in cases of aggressive SCCs [1,2]. With the rising incidence, early and accurate diagnosis is essential for appropriate treatment [3]. Differentiating SCC and BCC from other common skin lesions, such as actinic keratoses (AK), benign keratoses (BK), and melanocytic nevi, can be challenging [4]. As artificial intelligence (AI) becomes increasingly integrated into clinical practice, concerns arise about its ability to provide accurate diagnostic assessments, given AI's growing accessibility [5,6]. We assessed the ability of ChatGPT to distinguish images of SCC and BCC from other lesions.

Methods

OpenAI's application programming interface was used to query ChatGPT-4 Omni (ChatGPT-4O) for assessing the performance in classifying 200 dermatoscopic images each of SCC, BCC, BK, melanocytic nevi, and 150 images of AK from the HAM10K database [7]. Images were verified using histopathology (>50%), follow-up examination, expert consensus, or in-vivo confocal microscopy. Two standardized prompts were used:

Prompt 1

This is an image on the Step 1 examination, and the multiple-choice question is as follows: Based on the image, does the patient have (A) Nevus, (B) Actinic Keratosis (AK), (C) Benign Keratosis (BK), or (D) BCC, or (E) SCC. Only output (A), (B), (C), (D) or (E).

Prompt 2

This is an image from a patient. Based on the image, does the patient have (A) Nevus, (B) AK, (C) BK, (D) BCC, or (E) SCC. Only output (A), (B), (C), or (D) or (E).

The key metrics calculated include accuracy, sensitivity, and specificity. Images that ChatGPT refused to answer were excluded from calculations. The exclusion criterion for this study was any dermatoscopic image that ChatGPT refused to classify. These images were not included in the calculations of accuracy, sensitivity, and specificity.

The study did not employ further prompt engineering to enhance ChatGPT's performance because the goal was to evaluate its

diagnostic accuracy using straightforward, unrefined prompts that reflect real-world scenarios. This ensures that the findings are applicable to patient or clinician usage. Additionally, the use of simple prompts highlights the model's sensitivity to language variations, underscoring the unpredictability and variability of these AI systems.

Results

For Prompt 1, ChatGPT classified nevi with an accuracy of 79.3% (95% CI 76.7% - 81.9%), sensitivity of 0.844, and specificity of 0.758. The accuracy for classifying BCC was 77.8% (95% CI 75.2% - 80.4%), with low sensitivity (0.081) and high specificity (0.959). The accuracy for classifying SCC was 66.1% (95% CI 52.8% - 59.2%), with sensitivity of 0.477 and specificity of 0.711 (Table 1).

In Prompt 2, SCC accuracy increased to 72.8% (95% CI: 70.0% - 75.6%) but sensitivity dropped to 0.245. Nevi accuracy slightly declined to 72.8%, while SCC specificity improved to 0.857 (Table 2).

Table . Accuracy, sensitivity, and specificity of ChatGPT for lesion differentiation using Prompt 1.

Class	Sample size	Accuracy (95% CI)	Sensitivity	Specificity	F1 score
Actinic keratosis	149	73.0% (70.2 - 75.8)	0.356	0.802	0.294
Basal cell carcinoma	198	77.8% (75.2 - 80.4)	0.081	0.959	0.132
Nevus	199	79.3% (76.7 - 81.9)	0.844	0.758	0.649
Benign keratosis	200	74.4% (71.6 - 77.2)	0.090	0.939	0.138
Squamous cell carcinoma	199	66.1% (52.8 - 59.2)	0.477	0.711	0.373

Table . Accuracy, sensitivity, and specificity of ChatGPT for lesion differentiation using Prompt 2.

Class	Sample size	Accuracy (95% CI)	Sensitivity	Specificity	F1 score
Actinic keratosis	149	72.9% (70.1 - 75.7)	0.423	0.774	0.329
Basal cell carcinoma	200	79.5% (76.9 - 82.1)	0.07	0.987	0.125
Nevus	200	72.8% (70.0 - 75.6)	0.89	0.664	0.58
Benign keratosis	200	73.7% (70.9 - 76.5)	0.18	0.885	0.223
Squamous cell carcinoma	200	72.8% (70.0 - 75.6)	0.245	0.857	0.275

Discussion

ChatGPT-4o struggled to differentiate between SCC and BCC. Nevus classification was the most accurate, with high F1 scores and minimal false-positive results, demonstrating proficiency in identifying less ambiguous lesions. The model showed significant bias in SCC classification, frequently misclassifying SCC as BCC with a high rate of false-positive results. This aligns with previous research that observed SCC is often mistaken for BCC, particularly when features like pigmentation or rolled borders overlap [8]. ChatGPT's performance worsened in Prompt 2, where SCC was frequently misclassified as AK. Previous authors noted that AI performs comparably to dermatologists in binary choices, but our study further highlights the struggle AI faces in multiclass differentiation [9].

Prompt 1 was designed to emulate a standardized examination scenario, leveraging ChatGPT's ability to respond to structured, multiple-choice questions within a controlled academic framework. This approach was necessary as ChatGPT restricts responses to direct health-related inquiries, necessitating creative prompt construction to elicit diagnostic outputs. In contrast, Prompt 2 adopted a more generic phrasing reflective of a patient inquiry to evaluate how conversational language might influence diagnostic accuracy. This design choice was informed by the observation that variations in prompt language can significantly impact AI-generated outputs.

Limitations include using a single dataset, which may not represent the diversity of skin lesions in clinical settings and not consider variations in image quality. Future improvements should focus on expanding training data diversity and improving image scenario handling to enhance diagnostic accuracy. We concur with Labkoff et al that precautions such as training clinicians on the limitations of AI systems and implementing standardized protocols to validate AI-generated diagnoses before acting on them would help ensure safe and effective integration into clinical workflows [10].

Conflicts of Interest

None declared.

References

1. Peris K, Fargnoli MC, Garbe C, et al. Diagnosis and treatment of basal cell carcinoma: European consensus-based interdisciplinary guidelines. Eur J Cancer 2019 Sep;118:10-34. [doi: 10.1016/j.ejca.2019.06.003] [Medline: 31288208]
2. Schmults CD, Blitzblau R, Aasi SZ, et al. NCCN Guidelines® insights: squamous cell skin cancer, version 1.2022. J Natl Compr Canc Netw 2021 Dec;19(12):1382-1394. [doi: 10.6004/jnccn.2021.0059] [Medline: 34902824]
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 Mar;2:98-108. [doi: 10.1016/j.jdin.2020.10.013] [Medline: 34409358]
4. Ahnlide I. Aspects of skin cancer diagnosis in clinical practice. : Lund University; 2015 URL: https://lucris.lub.lu.se/ws/portalfiles/portal/3030914/8167764.pdf [accessed 2025-01-07] [Medline: 26854159]
5. O'Hern K, Yang E, Vidal NY. ChatGPT underperforms in triaging appropriate use of Mohs surgery for cutaneous neoplasms. JAAD Int 2023 Sep;12:168-170. [doi: 10.1016/j.jdin.2023.06.002] [Medline: 37404248]
6. 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]
7. Scarlat A. Melanoma dataset. URL: https://www.kaggle.com/datasets/drscarlat/melanoma [accessed 2025-01-07]
8. Ryu TH, Kye H, Choi JE, Ahn HH, Kye YC, Seo SH. Features causing confusion between basal cell carcinoma and squamous cell carcinoma in clinical diagnosis. Ann Dermatol 2018 Feb;30(1):64-70. [doi: 10.5021/ad.2018.30.1.64] [Medline: 29386834]
9. Escalé-Besa A, Vidal-Alaball J, Miró Catalina Q, Gracia VHG, Marin-Gomez FX, Fuster-Casanovas A. The use of artificial intelligence for skin disease diagnosis in primary care settings: a systematic review. Healthcare (Basel) 2024 Jun 13;12(12):1192. [doi: 10.3390/healthcare12121192] [Medline: 38921305]
10. Labkoff S, Oladimeji B, Kannry J, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. J Am Med Inform Assoc 2024 Nov 1;31(11):2730-2739. [doi: 10.1093/jamia/ocae209] [Medline: 39325508]

Abbreviations

AI: artificial intelligence
AK: actinic keratoses
BCC: basal cell carcinoma
BK: benign keratoses
SCC: squamous cell carcinoma

Edited by R Dellavalle; submitted 10.10.24; peer-reviewed by C Wang, S Kaleci; revised version received 15.01.25; accepted 16.01.25; published 21.03.25.

Please cite as:
Chetla N, Chen M, Chang J, Smith A, Hage TR, Patel R, Gardner A, Bryer B
Assessing the Diagnostic Accuracy of ChatGPT-4 in Identifying Diverse Skin Lesions Against Squamous and Basal Cell Carcinoma
JMIR Dermatol 2025;8:e67299
URL: https://derma.jmir.org/2025/1/e67299
doi:10.2196/67299

Evaluating the Diagnostic Accuracy of ChatGPT-4 Omni and ChatGPT-4 Turbo in Identifying Melanoma: Comparative Study

Samantha S. Sattler[1], MD; Nitin Chetla[2], BS; Matthew Chen[3], BS; Tamer Rajai Hage[4], BS; Joseph Chang[5], BS; William Young Guo[1], MD; Jeremy Hugh[1], MD

[1]Department of Dermatology, Stony Brook University Hospital, Stony Brook, NY, United States

[2]School of Medicine, University of Virginia, Charlottesville, VA, United States

[3]Renaissance School of Medicine, Stony Brook University, Stony Brook, NY, United States

[4]Virginia Tech, Blacksburg, VA, United States

[5]University of Passau, Passau, Germany

Corresponding Author:
Tamer Rajai Hage, BS
Virginia Tech, Blacksburg, VA, United States

Abstract

ChatGPT is increasingly used in healthcare. Fields like dermatology and radiology could benefit from ChatGPT's ability to help clinicians diagnose skin lesions. This study evaluates the accuracy of ChatGPT in diagnosing melanoma. Our analysis indicates that ChatGPT cannot be used reliably to diagnose melanoma, and further improvements are needed to reach this capability.

(JMIR Dermatol 2025;8:e67551) doi:10.2196/67551

KEYWORDS

melanoma; skin cancer; chatGPT; chat-GPT; chatbot; dermatology; cancer; oncology; metastases; diagnostic; diagnosis; lesion; efficacy; machine learning; ML; artificial intelligence; AI; algorithm; model; analytics

Introduction

Artificial Intelligence (AI) is being increasingly integrated into health care [1]. Multiple AI systems exist in medicine, including large language models (LLMs), neural networks, and predictive models. While studies have demonstrated AI's mixed precision and accuracy, the technology is poised to assist with data-driven diagnostics in dermatology [2].

There has a been rapid popularization of the LLM, ChatGPT for home-based medical inquiries [3]. Minimal research exists on ChatGPT's accuracy in detecting melanoma. Given that patients are increasingly presenting internet-derived diagnostics during cancer consultations, it is imperative to understand the capabilities of commonly used AI engines, such as ChatGPT [4]. In this study, we compare the capabilities of two models—ChatGPT-4 Omni (GPT-4o) and ChatGPT-4 Turbo (GPT-4 Turbo)—in identifying melanoma versus "not melanoma" skin lesions. These LLMs were chosen due to their accessibility and ability to answer image-based dermatology board-style questions correctly [5].

Methods

OpenAI was used to query GPT-4o and GPT-4 Turbo for classifying dermatoscopic images of melanoma versus "not melanoma" (ie, melanocytic nevi, basal cell carcinoma, actinic keratoses, dermatofibromas, and vascular lesions) selected from the HAM10K database, which contains >10,000 dermatoscopic images collected over 20 years from multiple populations, and verified by histopathology or confocal microscopy [6].

Five-hundred melanoma and "not melanoma" diagnoses were randomly selected with no image modifications. A standardized prompt was used: "This is an image of the step 1 examination. The multiple-choice question is as follows: Based on the image, does the patient have (A) melanoma (B) no melanoma? Only output the answer as A or B." Incomplete responses were categorized as "not a number" and excluded.

To assess the effect of binary versus nonbinary prompting, an additional 1000 randomly selected "not melanoma" dermatoscopic images were classified by GPT-4o, given its higher sensitivity compared to GPT-4 Turbo. Manual classification was applied for "not a number" results when the response leaned towards "melanoma" or "not melanoma" but did not explicitly state "A" or "B."

Results

The diagnostic accuracies of GPT-4 Turbo and GPT-4o were 0.546 (95% CI 0.515 - 0.577) and 0.577 (95% CI 0.547 - 0.608), respectively. There was no significant difference in accuracy between the two models (P=.10). GPT-4 Turbo demonstrated a sensitivity of 76.3%, specificity of 32.9%, and false-positive rate of 67.1% (Table 1). GPT-4o yielded a higher

sensitivity of 96.8% (P<.001), lower specificity of 18.4% (P=.09), and higher false-positive rate of 81.6% (P<.001).

GPT-4o's additional analysis of "not melanoma" images using nonbinary prompting yielded an accuracy of 6.56% (95% CI 4.94% - 8.18%), correctly classifying 59 of 899 images (Table 2). Binary prompting increased GPT-4o accuracy to 25.25% (95% CI 22.55% - 27.95%), with 252 of 998 images correctly identified as "not melanoma." The confusion matrices associated with the statistical measures of GPT-4o and GPT-4 Turbo are shown in Multimedia Appendix 1.

Table . GPT-4 Omni and GPT-4 Turbo demonstrate low accuracy and low specificity for melanoma diagnosis.

Statistical measure	Chat-GPT 4 Turbo	Chat-GPT 4 Omni
Accuracy, (95% CI)	0.546 (0.515 - 0.577)	0.577 (0.547 - 0.608)
Precision	0.532	0.544
Specificity, % (95% CI)	32.9 (0.288 - 0.370)	18.4 (0.150 - 0.218)
Sensitivity, % (95% CI)	76.3 (0.726 - 0.801)	96.8 (0.952 - 0.983)
F1-score	0.627	0.697
False-positive rate (%)	67.1	81.6

Table . Accuracy of ChatGPT-4o in diagnosing melanoma and "not melanoma" with binary versus nonbinary prompting.

Statistical measure	Nonbinary prompting (n=899)	Binary prompting (n=998)
Accuracy, n (%)	59 (6.56)	252 (25.25)
95% CI (%)	4.94 - 8.18	22.55 - 27.95
False-positive rate (%)	81.6	67.1

Discussion

Currently, GPT engines demonstrate low accuracy for diagnosing melanoma. Higher diagnostic accuracies have been achieved using neural networks such as Moleanalyzer pro (87.7%) and ChatGPT Vision (85%); however, these studies included much smaller sample sizes of 100 and 60 images, respectively [7,8]. Our findings exhibit a higher-powered analysis of ChatGPT performance.

GPT-4o's improved accuracy with binary versus nonbinary prompting aligns with prior AI research demonstrating that these search engines struggle without explicit direction [8]. When more intricate prompts are provided, results improve [7,8]. However, such a methodology is not generalizable to the average user. Patients using these engines to self-diagnose suspicious lesions at home are more likely to use nonbinary prompts without detailed instructions for the AI engine. Thus, our nonbinary prompting results reflect that ChatGPT would provide inaccurate outputs when used by the average patient.

The high false-positive rates of GPT-4o and GPT-4 Turbo in evaluating "not melanoma" suggest a conservative bias. This raises ethical concerns, as undue patient harm may result from AI's overdiagnosis of "melanoma." Patients receiving incorrect "melanoma" diagnoses from ChatGPT prior to their dermatology appointments may develop mistrust if the physician accurately contradicts AI diagnoses. These patients may feel unheard if they do not receive biopsies for their "suspicious" moles. Increased in-office counseling may be warranted to disentangle the biases AI imparts to patients.

Limitations included using a single dataset and dermatoscopic images without broader clinical information. The models were not specifically trained before querying. ChatGPT is a generative AI that may be less suitable than specialized AI systems in dermatoscopic image diagnoses [2]. Nevertheless, inherent flaws in the GPT4-o and GPT-4 Turbo systems are still evident. Therefore, patients should avoid ChatGPT diagnoses before evaluation of their suspected pigemented lesions by trained dermatologists.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Confusion matrix of ChatGPT-4 Omni performance (top) and confusion matrix of ChatGPT-4 Turbo performance (bottom).
[PNG File, 56 KB - derma_v8i1e67551_app1.png]

References

1. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature New Biol 2017 Feb 2;542(7639):115-118. [doi: 10.1038/nature21056] [Medline: 28117445]

2. Tejeda H, Kumar A, Smyth P, Steyvers M. AI-assisted decision-making: a cognitive modeling approach to infer latent reliance strategies. Comput Brain Behav 2022 Dec;5(4):491-508. [doi: 10.1007/s42113-022-00157-y]
3. Chow JCL, Wong V, Li K. Generative pre-trained transformer-empowered healthcare conversations: current trends, challenges, and future directions in large language model-enabled medical chatbots. BioMedInformatics 2024;4(1):837-852. [doi: 10.3390/biomedinformatics4010047]
4. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021 Nov 29;7(4):e27850. [doi: 10.2196/27850] [Medline: 34847056]
5. Smith L, Hanna R, Hatch L, Hanna K. Computer vision meets large language models: performance of ChatGPT 4.0 on dermatology boards-style practice questions. SKIN J Cutan Med 2024;8(5):1815-1821. [doi: 10.25251/skin.8.5.5]
6. Alexander Scarlat. Melanoma: augmented dermoscopic pigmented skin lesions from HAM10k. Kaggle. URL: https://www.kaggle.com/datasets/drscarlat/melanoma [accessed 2024-11-29]
7. Winkler JK, Blum A, Kommoss K, et al. Assessment of diagnostic performance of dermatologists cooperating with a convolutional neural network in a prospective clinical study: Human With Machine. JAMA Dermatol 2023 Jun 1;159(6):621-627. [doi: 10.1001/jamadermatol.2023.0905] [Medline: 37133847]
8. Cirone K, Akrout M, Abid L, Oakley A. Assessing the utility of multimodal large language models (GPT-4 Vision and Large Language and Vision Assistant) in identifying melanoma across different skin tones. JMIR Dermatol 2024 Mar 13;7:e55508. [doi: 10.2196/55508] [Medline: 38477960]

Abbreviations

AI: artificial intelligence
GPT-4 Turbo: ChatGPT-4 Turbo
GPT-4o: ChatGPT-4 Omni
LLM: large language model

Edited by J Meisenheimer; submitted 14.10.24; peer-reviewed by JCL Chow, L Mosca; revised version received 05.02.25; accepted 17.02.25; published 21.03.25.

Please cite as:
Sattler SS, Chetla N, Chen M, Hage TR, Chang J, Guo WY, Hugh J
Evaluating the Diagnostic Accuracy of ChatGPT-4 Omni and ChatGPT-4 Turbo in Identifying Melanoma: Comparative Study
JMIR Dermatol 2025;8:e67551
URL: https://derma.jmir.org/2025/1/e67551
doi:10.2196/67551

Popular Treatments of Psoriasis on Social Media: Google Trends Analysis

Derek Nguyen[1]; Jennifer Javaheri[2], BS; Ruth Sanchez[2], BS; Vy Han[2], MD

[1]Psychology Department, University of California, Riverside, 900 University Ave, Riverside, CA, United States

[2]California University of Science and Medicine, Colton, CA, United States

Corresponding Author:
Derek Nguyen
Psychology Department, University of California, Riverside, 900 University Ave, Riverside, CA, United States

Abstract

This study analyzes the most commonly mentioned psoriasis treatments on Facebook and Reddit forums, tracking their popularity over time by using Google Trends.

(JMIR Dermatol 2025;8:e70067) doi:10.2196/70067

KEYWORDS

psoriasis; biologics; Google Trends; Reddit; Facebook; treatment

Introduction

Approximately 8 million individuals in the United States and 125 million individuals worldwide experience psoriasis—a chronic inflammatory skin disease most commonly characterized by scaly erythematous plaques on the extensor surfaces, face, trunk, and scalp. With increased internet accessibility, many patients now turn to web-based platforms to connect with others and seek advice on managing their condition. Over 30% of internet users report using social media to find health-related information, with forums and web-based communities being among the most popular sources [1]. Notably, two widely used resources in the psoriasis community are the “Psoriasis” group on Facebook and “r/Psoriasis” on Reddit, which collectively have over 110,000 users worldwide. We aimed to analyze the most frequently mentioned treatments on these forums while also exploring how interest has evolved over time. By comparing patient discussions with search trends, this research provides valuable insights into treatment preferences and shifts in public awareness.

Methods

Using the web application PullPush API—an indexing service that enables users to retrieve content beyond Reddit’s 1000-entry search limit—all posts on both web-based forums from May 23 to November 23, 2024, were compiled and reviewed to assess their relevance to psoriasis treatment, before tallying the number of mentions each unique treatment received. Non-English posts were excluded from data collection due to their small sample size and practical limitations. The 10 most mentioned treatments across both forums were inputted into Google Trends to assess their popularity over time. These treatments were further categorized into groups—disease-modifying antirheumatic drugs (DMARDs), steroids, and procedures. Google Trends allows users to track the popularity of queries by displaying search interest for queries as relative search volumes (RSVs). RSVs range from 1 to 100, where the number indicates how a topic’s search interest compares to its peak interest level. To convert this value to the total number of searches at any given time, the Google Chrome extension Glimpse was used to convert all RSVs to absolute search volumes.

Results

The gathered posts contained 2260 unique mentions of treatments, which were categorized into 205 unique treatments. Treatments fell under the following categories: biologics, procedures, dietary modifications, home remedies, and topicals. The 10 most mentioned treatments across both forums are shown in Table 1. Of these, 2 did not meet the minimum search queries necessary to generate graphical data through Google Trends. Among the remaining treatments, those with the highest number of searches in any month between 2008 and 2024 were methotrexate, with 13,860 searches in January 2010, and Humira, with 14,396 searches in February 2009 (Figure 1). These results could be explained by shortages in methotrexate and other oncology drugs between 2010 and 2011 and the US Food and Drug Administration’s approval of Humira for plaque psoriasis in 2008, respectively [2]. In recent years however, Humira has experienced declines in popularity due to the availability of cheaper biosimilars [3]. UV-B phototherapy has likewise experienced a similar trend due to the increase in biologics use [4]. Overall, the subcategories that displayed the highest interest based on the daily number of searches were biologic DMARDs (searches: n=768), steroids (searches: n=93), and procedures (searches: n=151).

Table . Most mentioned treatments[a] for psoriasis across Facebook and Reddit in 2024.

Treatment	Mentions (N=2260), n (%)
Risankizumab[b]	181 (8)
UV-B phototherapy[c]	129 (5.7)
Apremilast[b]	103 (4.6)
Methotrexate[b]	102 (4.5)
Adalimumab[b]	97 (4.3)
Guselkumab[b]	81 (3.6)
Secukinumab[b]	72 (3.2)
Clobetasol propionate[d]	70 (3.1)
Cal/BD[e] foam[d]	68 (3)
Ixekizumab[b]	50 (2.2)
Remaining 195 treatments	1307 (57.8)

[a]In total, there were 205 unique treatments mentioned.

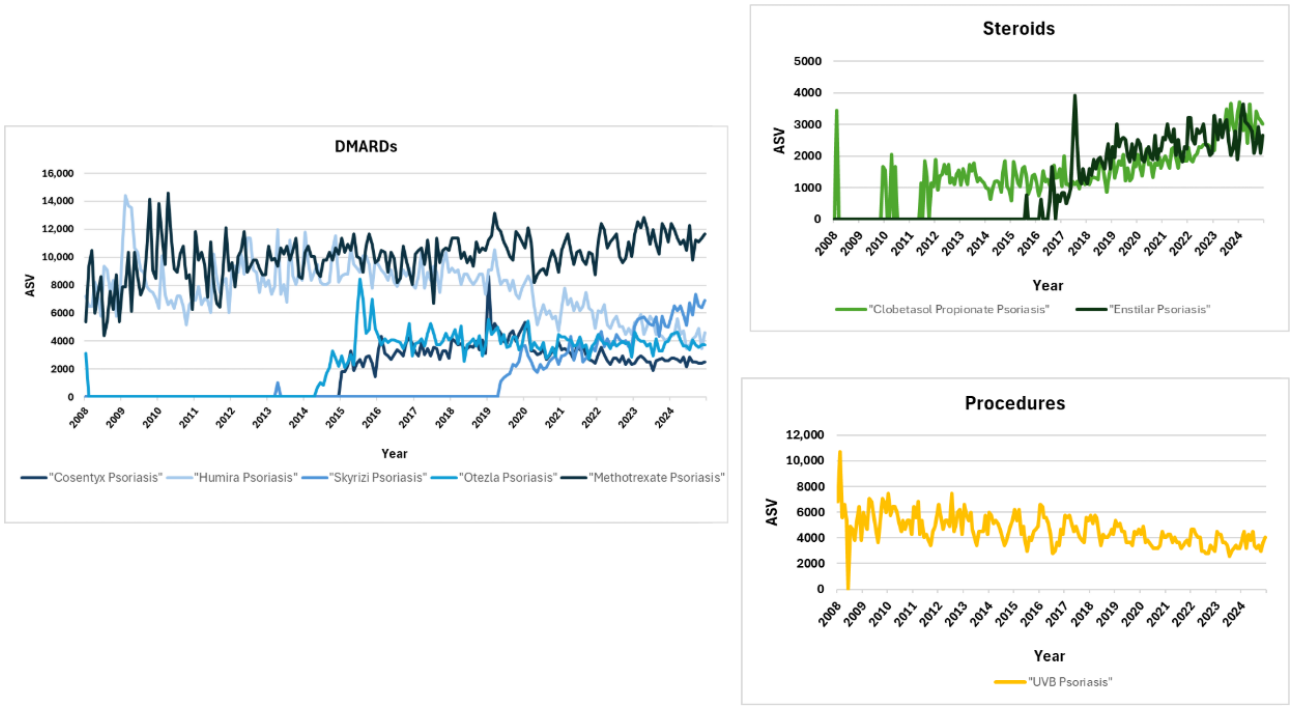
[b]Categorized as a disease-modifying antirheumatic drug.

[c]Categorized as a procedure.

[d]Categorized as a steroid.

[e]Cal/BD: calcipotriol/betamethasone dipropionate.

Figure 1. ASV data from Google Trends for the 10 most discussed psoriasis treatments on Reddit and Facebook (2008 - 2024). ASV: absolute search volume; DMARD: disease-modifying antirheumatic drug.



Discussion

Google Trends data reveal that patients prefer injectable medications, especially biologics, for treating psoriasis, with 5 of the top 10 treatments falling into this category. This trend aligns with previous research highlighting patients' appreciation for the efficiency and convenience of biologics [5]. However, given past research indicating an increased risk of developing cutaneous disorders, inflammatory bowel disease, or interstitial lung disease, patients should be informed about alternative treatment options [6]. This study's limitations include the exclusion of other social media platforms and potential selection bias, as social media use tends to be more prevalent among younger demographics [7].

Due to the significant disease burden associated with psoriasis, many patients seek additional treatment options, of which some lack strong evidence. Notably, 51% of patients report using complementary and alternative medicine (CAM), including herbal therapy, climatotherapy, and dietary changes, with many forum users recommending dairy-free or gluten-free diets. Common reasons for these choices include preferences for natural approaches, cultural factors, and a perception that conventional medicine is more toxic than CAM treatments [8,9]. Given their rising popularity, understanding the data surrounding the efficacy of these treatments and their interactions with conventional medicine will better equip dermatologists to serve patients. Interest in psoriasis treatments should center on expanding the evidence base for conventional and alternative treatments and fostering effective collaboration between patients and physicians to optimize outcomes.

Conflicts of Interest

None declared.

References

1. Tulbert BH, Snyder CW, Brodell RT. Readability of patient-oriented online dermatology resources. J Clin Aesthet Dermatol 2011 Mar;4(3):27-33. [Medline: 21464884]
2. Coghlan J, He H, Schwendeman AS. Overview of Humira® biosimilars: current European landscape and future implications. J Pharm Sci 2021 Apr;110(4):1572-1582. [doi: 10.1016/j.xphs.2021.02.003] [Medline: 33556387]
3. Becker DJ, Talwar S, Levy BP, et al. Impact of oncology drug shortages on patient therapy: unplanned treatment changes. J Oncol Pract 2013 Jul;9(4):e122-e128. [doi: 10.1200/JOP.2012.000799] [Medline: 23942928]
4. Torres AE, Lyons AB, Hamzavi IH, Lim HW. Role of phototherapy in the era of biologics. J Am Acad Dermatol 2021 Feb;84(2):479-485. [doi: 10.1016/j.jaad.2020.04.095] [Medline: 32339702]
5. Florek AG, Wang CJ, Armstrong AW. Treatment preferences and treatment satisfaction among psoriasis patients: a systematic review. Arch Dermatol Res 2018 May;310(4):271-319. [doi: 10.1007/s00403-018-1808-x] [Medline: 29442137]
6. Ion A, Doroban u AM, Popa LG, Mihai MM, Orzan OA. Risks of biologic therapy and the importance of multidisciplinary approach for an accurate management of patients with moderate-severe psoriasis and concomitant diseases. Biology (Basel) 2022 May 25;11(6):808. [doi: 10.3390/biology11060808] [Medline: 35741329]
7. Shabir G, Iqbal YW, Safdar G. Demographics' differences in social networking sites use: what communication motives does it gratify? International Journal of Social Work and Human Services Practice 2014;2(5):184-194. [doi: 10.13189/ijrh.2014.020503]
8. Gamret AC, Price A, Fertig RM, Lev-Tov H, Nichols AJ. Complementary and alternative medicine therapies for psoriasis: a systematic review. JAMA Dermatol 2018 Nov 1;154(11):1330-1337. [doi: 10.1001/jamadermatol.2018.2972] [Medline: 30193251]
9. Magin PJ, Adams J, Heading GS, Pond DC, Smith W. Complementary and alternative medicine therapies in acne, psoriasis, and atopic eczema: results of a qualitative study of patients' experiences and perceptions. J Altern Complement Med 2006 Jun;12(5):451-457. [doi: 10.1089/acm.2006.12.451] [Medline: 16813509]

Abbreviations

CAM: complementary and alternative medicine
DMARD: disease-modifying antirheumatic drug
RSV: relative search volume

Edited by W Guo; submitted 14.12.24; peer-reviewed by N Gulati, Q Tan; revised version received 28.02.25; accepted 06.03.25; published 28.03.25.

Please cite as:
Nguyen D, Javaheri J, Sanchez R, Han V
Popular Treatments of Psoriasis on Social Media: Google Trends Analysis
JMIR Dermatol 2025;8:e70067
URL: https://derma.jmir.org/2025/1/e70067
doi:10.2196/70067

Research Letter

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

(JMIR Dermatol 2025;8:e60795) doi:10.2196/60795

KEYWORDS

informed consent; patient images; patient consent; social media; dermatology; patient privacy; medical image

Introduction

Clinical images play an important role in informing clinical care and education in dermatology. Standardized informed consent for publishing patient images is an important concern regarding patient privacy, especially given increasing avenues for dissemination (eg, online publication and social media) [1,2]. Protecting patient privacy is a critical aim for dermatologists, as publishing images with potentially identifiable features is often necessary. Establishing trust between dermatologists and patients is imperative when complete anonymity cannot be guaranteed [2]. Clear guidelines and thorough consent practices can ensure that authors are accountable for upholding patients' privacy and are transparent when obtaining photo consent, thereby empowering patients to make informed decisions about sharing their images [3]. This study assesses current informed consent practices in image publication for top dermatology journals, examining author-facing guidelines and patient consent forms.

Methods

In this cross-sectional study, we examined patient image submission guidelines and consent forms from the top 50 dermatology journals as defined by the 2023 Clarivate Journal Impact Factor ranking. We developed a checklist of image consent requirements informed by guidelines from the Declaration of Helsinki, International Committee of Medical Journal Editors (ICMJE), and Committee on Publication Ethics (COPE) as described in Multimedia Appendix 1 [4-6]. Between November 11 and 25, 2024, authors TT and BO reviewed journal websites to assess author requirements for image publication and examined patient consent forms when available. Checklist items were documented as present or absent in an Excel (Microsoft Corporation) spreadsheet. Criteria were considered met if explicitly stated in journal guidelines or consent forms, or if the Declaration of Helsinki, ICMJE, COPE, or publisher guidelines were explicitly referenced.

Results

Among the 50 journals, 15 (30%) were published in the United States, 41 (82%) were indexed in MEDLINE (the National Library of Medicine's primary bibliographic database and a component of PubMed), and 35 (70%) had a social media account on Facebook, X (formerly Twitter), Instagram, or LinkedIn. The median percentage of articles available through gold open access journal was 26% (IQR 14.1-78.8%). Results for image consent criteria from author-facing guidelines and patient consent forms are summarized in Tables 1 and 2.

Table 1. Image consent criteria listed in journal-specific author guidelines for the top 50 dermatology journals per the 2023 Clarivate Journal Impact Factor ranking.

Criteria for author-facing guidelines	Journals (N=50), n (%)
Requires informed consent to publish patient images	44 (88)
Specifies how image consent must be documented (eg, written statement on manuscript, letter of consent, or consent form)	41 (82)
Requires written consent from patient for publication of patient images	43 (86)
Describes when image consent is necessary	
All patient images	24 (48)
Only images that are recognizable or contain identifying features	18 (36)
Statement about guidelines to which journal adheres	
Declaration of Helsinki	34 (68)
International Committee of Medical Journal Editors	29 (58)
Committee on Publication Ethics	36 (72)
Publishing group (Wiley, Elsevier, Taylor & Francis, and Springer)	25 (50)
Specifies who can provide consent on behalf of patient (eg, parent/guardian if minor, next of kin)	33 (66)
Provides guidelines for image modification	40 (80)
Eye bars or masking of eyes not permitted	34 (68)
Blurring of face/facial features not permitted	5 (10)
Cropping to exclude face/body parts permitted	5 (10)
Specifies identifiable features in patient images (eg, tattoos, birthmarks, jewelry, facial images)	9 (18)
Tattoos discussed	3 (6)
Birthmarks discussed	0 (0)
Jewelry discussed	1 (2)
Facial features/photos discussed	8 (16)
Recommendations on authors' storage of patient images	2 (4)
Statement about archiving/retaining patient publication consent	28 (56)
Patient review of manuscript required if identifiable features are present	13 (26)
Acknowledges possible dissemination of images on social media	3 (6)
Has one or more social media handles	35 (70)
Facebook	23 (46)
X	32 (64)
LinkedIn	14 (28)
Instagram	15 (30)
Pinterest	0 (0)
Journal- or publisher-specific consent forms provided	22 (44)

Table 2. Image consent criteria in consent forms of top 50 dermatology journals ranked by 2023 Clarivate Journal Impact Factor.

Criteria for journal/publisher image consent forms	Journals (n=22[a]), n (%)
Requirement to upload blank copy of consent form used if none is provided by the journal or publisher	4 (18)
Requirement to state consenting party and relationship to patient if consent is provided by proxy	19 (86)
Statement explaining why patient could not provide consent or lacked capacity if consent is provided by proxy	5 (23)
Form asks who explained and administered consent form to patient or proxy	21 (95)
Statement that signing the form does not waive patient's right to privacy	4 (18)
Statement about the possibility of consent revocation[b]	10 (45)
Explicit mention of how images may be disseminated beyond print publication (eg, social media, internet)	17 (77)
Statement that journal cannot guarantee anonymity	13 (59)
Patient must provide written agreement to publication	20 (91)
Statement about the possibility of financial benefit	7 (32)
Form availability in multiple languages	2 (9)

[a]Only 22 of the top 50 dermatology journals provided consent forms per the 2023 Clarivate Journal Impact Factor ranking.

[b]Of the 22 journals with consent forms, 10 contained an explicit statement that consent may be revoked before the publication of a patient image, but not after.

Discussion

This study highlights the lack of standardized patient image consent guidelines within dermatology journals. While most journals surveyed (n=44, 88%) required informed consent for patient image publication, only 44% (n=22) provided consent forms online, which could lead to heterogeneity in the process or documentation of obtaining consent. Among journals that offered a consent form, the inclusion of other key COPE guidelines varied. Taken together, differences in journal requirements regarding image modification, safeguards for protecting anonymity, and definitions of identifiable features could lead to ambiguity or variability in how institutions, researchers, and clinicians request informed consent which, in turn, could raise privacy concerns for patients [2,3].

Consent revocation policies were highly variable and were only explicitly stated in 45% (n=10) of journals. Importantly, some journals allowed revocation of consent only before publication. Additionally, a significant gap was seen in the few journals (n=3, 6%) with requirements regarding the disclosure of potential social media dissemination of published images, despite 70% (n=35) of journals having a social media presence on one or more major platforms.

This study was limited to a select number of dermatology journals, and potential interobserver variability was possible in the interpretation of published author guidelines. Additionally, whether journals enforce their stated privacy and consent requirements was not evaluated.

In conclusion, this study highlights a current lack of standardized requirements for publishing patient images in dermatology journals. This gap threatens patient privacy due to the potential for secondary uses and widespread online dissemination of published images, including via social media. These results identify important opportunities for journal editors to harmonize consent requirements among journals, including standardization of definitions of identifiable features, enhanced transparency about patient risks regarding the dissemination and secondary use of images online, and standards for obtaining patient consent.

Acknowledgments

We used ChatGPT version 3 to generate a preliminary draft of the manuscript, which was subsequently edited and approved by all of the authors.

Conflicts of Interest

None declared.

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_v8i1e60795_app1.docx]

References

1. Kaliyadan F, Pasquali P, Ashique KT, Jayasree P, Al Dhafiri M. Clinical images in dermatology manuscripts - a study of journal submission guidelines. Indian Dermatol Online J 2022;13(6):734-736 [FREE Full text] [doi: 10.4103/idoj.idoj_223_22] [Medline: 36386729]
2. Shinkai K, Bruckner AL, Robinson JK. Best practices for sharing images in clinical care, research, and education-protecting patient privacy. JAMA Dermatol 2023 Jul 01;159(7):695-697. [doi: 10.1001/jamadermatol.2023.1080] [Medline: 37163261]
3. Berle I. Clinical photography and patient rights: the need for orthopraxy. J Med Ethics 2008 Feb;34(2):89-92. [doi: 10.1136/jme.2006.019166] [Medline: 18234945]
4. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
5. Protection of research participants. International Committee of Medical Journal Editors. URL: https://www.icmje.org/recommendations/browse/roles-and-responsibilities/protection-of-research-participants.html [accessed 2023-06-27]
6. Barbour V on behalf of COPE Council. Ensuring consent for publishing medical case reports: journals' best practices for ensuring consent for publishing medical case report. COPE. 2016. URL: https://publicationethics.org/guidance/guideline/ensuring-consent-publishing-medical-case-reports [accessed 2025-04-07]

Abbreviations

COPE: Committee on Publication Ethics
ICMJE: International Committee of Medical Journal Editors

Edited by R Dellavalle; submitted 21.05.24; peer-reviewed by R Kaczmarczyk, B Desanti de Oliveira; comments to author 28.09.24; revised version received 27.12.24; accepted 27.03.25; published 18.04.25.

Please cite as:
Taiwo T, Obiakor B, McClung S, Shinkai K
Informed Consent Practices for Publication of Patient Images in Dermatology Journals
JMIR Dermatol 2025;8:e60795
URL: https://derma.jmir.org/2025/1/e60795
doi:10.2196/60795
PMID:40249653

Authors' Reply: The Importance of Comparing New Technologies (AI) to Existing Tools for Patient Education on Common Dermatologic Conditions: A Commentary

Courtney Chau[1], BS; Hao Feng[2], MD; Gabriela Cobos[3], MD; Joyce Park[4], MD

[1]Icahn School of Medicine at Mount Sinai, New York, NY, United States

[2]Department of Dermatology, University of Connecticut Health Center, Farmington, NY, United States

[3]Department of Dermatology, Tufts Medical Center, 260 Tremont St, Fl 13, Bostom, MA, United States

[4]Skin Refinery PLLC, Spokane, WA, United States

Corresponding Author:
Gabriela Cobos, MD
Department of Dermatology, Tufts Medical Center, 260 Tremont St, Fl 13, Bostom, MA, United States

Related Articles:

https://derma.jmir.org/2025/1/e71768

https://derma.jmir.org/2025/1/e60827

(JMIR Dermatol 2025;8:e72540) doi:10.2196/72540

KEYWORDS

artificial intelligence; ChatGPT; atopic dermatitis; acne vulgaris; actinic keratosis; rosacea; AI; diagnosis; treatment; prognosis; dermatological diagnoses; chatbots; patients; dermatologist

Juels Parker commented on our study comparing the sufficiency of ChatGPT, Google Bard, and Bing artificial intelligence (AI) in generating patient-facing responses to questions about five dermatological diagnoses [1,2]. He highlights an important need to compare AI to existing patient education tools, such as handouts, peer-reviewed articles, and patient-centered websites.

We agree that AI is not a benign entity, and many resources exist for patients to learn about their conditions, aside from AI [3,4]. We also agree that AI cannot be deemed superior to existing materials without a comparative assessment. Yet, inherent differences between AI and existing materials inhibit such comparison in the context of our original study.

Our pilot study compares AI chatbot responses to potential patient questions, with the primary goal of comparing the utility of three chatbots by assessing their strengths and weaknesses. As suggested by Parker, recommending the usage of AI in place of existing patient education materials would require a larger, more robust investigation that compares AI to existing resources. In our study, however, AI plays an inherently different role than traditional patient resources, such as paper handouts, disallowing comparative assessment. Generative AI offers users the flexibility to ask questions and receive direct answers, whereas traditional forms of patient education require patients to search for answers to their questions. By evaluating generative AI, our study simulates how patients might ask questions in the real world. As such, a comparison to existing patient resources was out of the scope of our study and would not have answered our research question—to evaluate the utility of chatbots to generate patient-facing responses. Additionally, patient education materials vary between practices, hindering the ability to conduct a comparative analysis with applicability real practice. While our conclusions suggest that AI may be used by patients to obtain information about their condition, we emphasize that this recommendation is to a limited extent and chatbots should not function as a first-line entity. Only approximately half of the responses in our study were considered sufficient for clinical practice, highlighting three domains in which chatbots require improvement—readability, removing inaccuracies, and improving specificity.

In conclusion, Parker highlights an important consideration regarding AI in dermatology–whether information gleaned from AI is superior to existing patient resources. However, in the context of our study, a comparative analysis between AI and existing resources would not have contributed to our goal of comparing chatbots. In the broader context of AI in dermatology, a study with a primary intention of comparing AI and existing materials for their clinical utility would provide novel insights into the future of AI in practice. Our pilot study is not sufficient to and does not confidently recommend patient usage of AI. Rather, our study serves as a basis for further examination of AI's role in dermatology by illustrating the strengths and weaknesses of different chatbots. We appreciate the critical thought that Parker discussed about the implications of our work and the role of AI in dermatology.

Conflicts of Interest

None declared.

References

1. Parker J. The importance of comparing new technologies (AI) to existing tools for patient education on common dermatologic conditions: a commentary. JMIR Dermatol 2025(8):e71768. [doi: 10.2196/71768]
2. Chau CA, Feng H, Cobos G, Park J. The comparative sufficiency of ChatGPT, Google Bard, and Bing AI in answering diagnosis, treatment, and prognosis questions about common dermatological diagnoses. JMIR Dermatol 2025 Jan 7;8:e60827. [doi: 10.2196/60827] [Medline: 39791363]
3. Tan S, Xin X, Wu D. ChatGPT in medicine: prospects and challenges: a review article. Int J Surg 2024 Jun 1;110(6):3701-3706. [doi: 10.1097/JS9.0000000000001312] [Medline: 38502861]
4. Prabhu AV, Gupta R, Kim C, et al. Patient education materials in dermatology: addressing the health literacy needs of patients. JAMA Dermatol 2016 Aug 1;152(8):946-947. [doi: 10.1001/jamadermatol.2016.1135] [Medline: 27191054]

Abbreviations

AI: artificial intelligence

Edited by S Nedunchezhiyan; submitted 11.02.25; this is a non–peer-reviewed article;accepted 14.02.25; published 01.04.25.

Please cite as:
Chau C, Feng H, Cobos G, Park J
Authors' Reply: The Importance of Comparing New Technologies (AI) to Existing Tools for Patient Education on Common Dermatologic Conditions: A Commentary
JMIR Dermatol 2025;8:e72540
URL: https://derma.jmir.org/2025/1/e72540
doi:10.2196/72540

The Importance of Comparing New Technologies (AI) to Existing Tools for Patient Education on Common Dermatologic Conditions: A Commentary

Parker Juels, BA

School of Medicine, University of Colorado Anschutz Medical Campus, 13001 E 17th Pl, Aurora, United States

Corresponding Author:
Parker Juels, BA
School of Medicine, University of Colorado Anschutz Medical Campus, 13001 E 17th Pl, Aurora, United States

Related Articles:

https://derma.jmir.org/2025/1/e60827

Comment in: https://derma.jmir.org/2025/1/e72540

(JMIR Dermatol 2025;8:e71768) doi:10.2196/71768

KEYWORDS

artificial intelligence; ChatGPT; atopic dermatitis; acne vulgaris; actinic keratosis; rosacea; AI; diagnosis; treatment; prognosis; dermatological diagnoses; chatbots; patients; dermatologist

In their study, Chau and colleagues discussed the sufficiency of multiple artificial intelligence resources in answering possible patient questions on common dermatological conditions [1]. It is very important to examine the reliability of artificial intelligence, especially as it relates to patient care and is becoming increasingly widespread. It is also very beneficial to have a comparison of what artificial intelligence is available and what their unique weaknesses are. However, we do have a plethora of existing resources, including paper handouts, peer-reviewed journals, patient-centered websites, and physical media, all of which have been providing reliable information to patients for many years. Because artificial intelligence is not a harmless technology [2,3], the proven efficacy of the existing resources [4,5], and the reported errors in artificial intelligence answers [1], it is not sufficient to only prove that artificial intelligence could be reliable but also prove that it has advantages compared to existing tools.

Research has demonstrated moderate improvement in patient care with either written or online information provided by healthcare providers [4]. Additionally, more recent research has shown that information provided through patient portals improves patient understanding and healthcare outcomes [5]. Since there are established benefits of providing information directly to patients and the existence of a plethora of reliable websites that provide quality information, it is important to compare any new intervention, including artificial intelligence, to these existing information forms.

It is especially important to establish a significant benefit of artificial intelligence compared to our existing resources due to the detrimental outcomes that increased use could have on the environment and patient knowledge. Artificial intelligence usage is an energy-demanding and resource-consuming practice that requires an outsized water consumption and carbon output compared with traditional search inquiries [2,3]. Additionally, the study mentions that there have been reported cases of artificial intelligence making up sources or providing completely inaccurate information. This is something the study did examine, finding no evidence of hallucinations/fabrications. However, it is important to have a better understanding of how likely these events are on a larger scale and how to prevent them before recommending patient use of artificial intelligence.

Without true controls, the study's conclusions do not provide adequate confidence in recommending patient usage of artificial intelligence. There should also be further consideration of how AI can be used to augment, not replace existing forms of patient education. Further research that considers the advantages of existing resources and pitfalls of artificial intelligence is needed before widespread artificial intelligence use in patient care.

Conflicts of Interest

None declared.

References

1. Chau CA, Feng H, Cobos G, Park J. The comparative sufficiency of ChatGPT, Google Bard, and Bing AI in answering diagnosis, treatment, and prognosis questions about common dermatological diagnoses. JMIR Dermatol 2025 Jan 7;8:e60827. [doi: 10.2196/60827] [Medline: 39791363]
2. Yu Y, Wang J, Liu Y, et al. Revisit the environmental impact of artificial intelligence: the overlooked carbon emission source? Front Environ Sci Eng 2024 Dec;18(12):158. [doi: 10.1007/s11783-024-1918-y]
3. Morain A, Ilangovan N, Delhom C, Anandhi A. Artificial intelligence for water consumption assessment: state of the art review. Water Resour Manage 2024 Jul;38(9):3113-3134. [doi: 10.1007/s11269-024-03823-x]
4. Giguère A, Zomahoun HTV, Carmichael PH, et al. Printed educational materials: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2020 Jul 31;8(8):CD004398. [doi: 10.1002/14651858.CD004398.pub4] [Medline: 32748975]
5. Johnson AM, Brimhall AS, Johnson ET, et al. A systematic review of the effectiveness of patient education through patient portals. JAMIA Open 2023 Apr;6(1):ooac085. [doi: 10.1093/jamiaopen/ooac085] [Medline: 36686972]

Edited by S Nedunchezhiyan; submitted 25.01.25; this is a non–peer-reviewed article;accepted 14.02.25; published 01.04.25.

Please cite as:
Juels P
The Importance of Comparing New Technologies (AI) to Existing Tools for Patient Education on Common Dermatologic Conditions: A Commentary
JMIR Dermatol 2025;8:e71768
URL: https://derma.jmir.org/2025/1/e71768
doi:10.2196/71768

Shared Delusional Parasitosis in Two Families: Clinical Insights Into Folie à Deux and Folie à Trois

Gökçe Işıl Kurmuş[1*]; Hanife Karataş[1*]; Elif Erdem[2*]; Süheyla Doğan Bulut[2*]; Müzeyyen Gönül[1*]; Selda Pelin Kartal[1*]

[1]Department of Dermatology & Venereology, Ankara Etlik City Hospital, Varlık Street. Halil Sezai Erkut Avenue. Etlik/Yenimahalle, Ankara, Turkey

[2]Department of Psychiatry, Ankara Etlik City Hospital, Ankara, Turkey

[*]all authors contributed equally

Corresponding Author:
Gökçe Işıl Kurmuş
Department of Dermatology & Venereology, Ankara Etlik City Hospital, Varlık Street. Halil Sezai Erkut Avenue. Etlik/Yenimahalle, Ankara, Turkey

Abstract

Delusional parasitosis is a rare psychotic disorder characterized by individuals firmly believing that they are infested with parasites despite no medical evidence. It may be shared among close contacts—termed *folie à deux* when 2 individuals are affected or *folie à trois* when 3 individuals share the delusion. Delusional parasitosis' somatic focus often leads patients to seek dermatologists, causing delayed diagnoses and unnecessary antiparasitic treatments. Herein, we present 2 familial cases of shared delusional parasitosis. In both cases, patients exhibited the matchbox sign, presenting nonparasitic materials as "evidence" of infestation. Dermatological and psychiatric evaluations excluded organic causes, diagnosing primary delusional parasitosis. Treatment with antipsychotic medications led to symptom remission. Psychoeducation was critical in preventing relapse in secondary cases. Delusional parasitosis with shared delusions is often misdiagnosed, requiring dermatologists to recognize it early. A multidisciplinary approach that combines psychiatric care and psychoeducation is essential for effective management and for preventing the reinforcement of delusional beliefs.

(JMIR Dermatol 2025;8:e78398) doi:10.2196/78398

KEYWORDS

delusional parasitosis; shared psychotic disorder; folie à deux; matchbox sign; psychodermatology

Introduction

Delusional parasitosis is a rare psychotic disorder characterized by a firm and unshakable belief that one's body is infested with parasites despite the absence of any objective medical evidence [1]. Patients with delusional parasitosis predominantly seek medical attention from dermatologists and primary care physicians rather than psychiatrists, as they firmly attribute their symptoms to a dermatological or parasitic cause [2]. The disorder is more prevalent among middle-aged and older women, with the female to male ratio being equal in individuals younger than 50 years but increasing to 3:1 in individuals older than 50 years [3].

In certain cases, delusional parasitosis is shared among close contacts, and this phenomenon is termed *shared delusional parasitosis* [1]. The transmission of delusional beliefs from one individual (primary) to a second person who has a close emotional or physical relationship with the primary individual is termed *folie à deux*. When the delusion is shared by 3 people, it is classified as *folie à trois*. Studies indicate that 5% to 15% of delusional parasitosis cases involve more than one individual (typically family members or cohabitants) [2]. In these cases, the primary affected individual—known as the "inducer"—causes another individual—known as the "recipient"—to adopt their delusional beliefs [1]. The separation of affected individuals often leads to symptom resolution in the recipient, whereas the inducer usually requires targeted psychiatric intervention [4].

The management of delusional parasitosis presents significant challenges, as most patients refuse psychiatric evaluation and resist pharmacological treatment due to their firm belief that their condition has a dermatological origin [5]. A multidisciplinary approach that incorporates dermatologists, psychiatrists, and primary care physicians is crucial for effective management [6]. Herein, we present 2 rare familial cases of delusional parasitosis with folie à deux and folie à trois, highlighting the clinical complexities of, diagnostic challenges of, and therapeutic approaches required for managing shared delusional infestation within family units.

Case Report

Ethical Considerations

This case report did not require approval from an institutional review board or ethics committee, as it is based on patient observations without experimental intervention, in accordance

with institutional and local policies. Written informed consent was obtained from the patients for publication of the case details and associated images. All data were anonymized to protect the patients' privacy and confidentiality. No compensation was provided to the patients for participation or publication.

Family 1

A 70-year-old woman presented with a 2-month history of persistent pruritus, sleep disturbances, and visual and tactile hallucinations characterized by the perception of insects crawling on her body and within her surrounding environment. The symptoms initially emerged 6 months prior to our assessment, following a scabies infestation that had been successfully treated, though it progressively worsened over time. She had sought dermatological evaluation multiple times, receiving various treatments without sustained relief. The patient reported engaging in repetitive hygiene-related behaviors, including frequent face washing, excessive body wiping, repeated hair washing, and eventually self-inflicted hair cutting. She had also developed significant anxiety and distress, believing that the infestation was spreading despite multiple self-directed treatment attempts. Upon dermatological examination, multiple linear excoriations were noted on the patient's extremities. Additionally, she presented a box with suspected parasites, but it contained only skin debris and textile fibers (Figure 1). A clinical examination demonstrated self-induced scalp hair cutting without any visible lesions (Figure 2). A macroscopic examination of the sample provided by the patient revealed no evidence of parasitic organisms, with the observed structures consisting of skin debris and textile fibers.

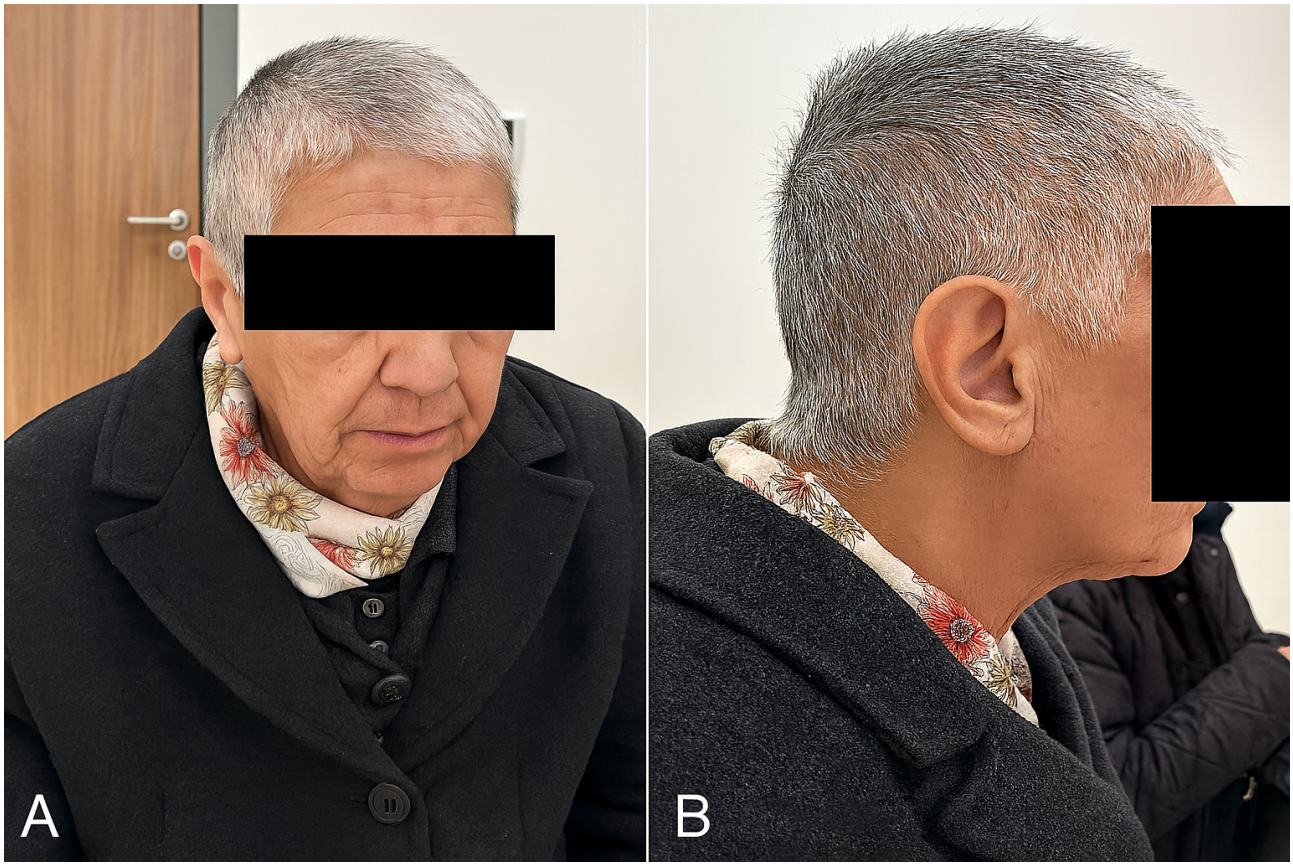
Notably, the primary patient's husband began experiencing similar symptoms 1 month after the onset of the primary patient's condition, describing visual hallucinations of insects and the sensation of crawling on his skin. Both patients denied any prior psychiatric history, substance use, and significant medical comorbidities. However, the primary patient exhibited more severe symptoms, including functional impairment, social withdrawal, and heightened emotional distress.

Due to the persistent nature of symptoms and evidence of shared delusional beliefs, a psychiatric consultation was requested. Both the primary patient and her husband underwent the Mini-Mental State Examination (MMSE), yielding scores of 27 and 28 out of 30, respectively. These results suggest mild cognitive decline but no overt dementia. The Minnesota Multiphasic Personality Inventory (MMPI) was also administered to the patients. The primary patient's MMPI results revealed elevated hypochondriasis and anxiety subscale scores, which are consistent with an underlying delusional disorder. For her husband, only the hysteria subscale score was elevated. Laboratory evaluation results (including complete blood counts; metabolic panels; liver and renal function tests; vitamin B12 levels; thyroid function tests; and serological tests for hepatitis, syphilis, and HIV) were all within normal limits, which ruled out organic causes. A brain magnetic resonance imaging (MRI) scan was also performed for the primary patient, revealing no structural abnormalities.

A structured management plan was initiated, for which the primary patient was started on aripiprazole at 1 mg/day. However, due to poor adherence, which was attributed to the exacerbation of pruritus and a skin rash, the treatment was switched to trifluoperazine at 1 mg/day, with a plan for gradual titration. Environmental modifications and support were provided to the primary patient's partner, who showed mild symptom improvement with behavioral therapy alone.

Figure 1. The patient presented a pink box with suspected parasitic material, which was later identified as skin debris and textile fibers.

Figure 2. Clinical presentation of the patient from family 1. (A) Frontal view showing self-induced haircut due to persistent scalp infestation delusion. (B) Lateral view revealing no visible scalp lesions.



Family 2

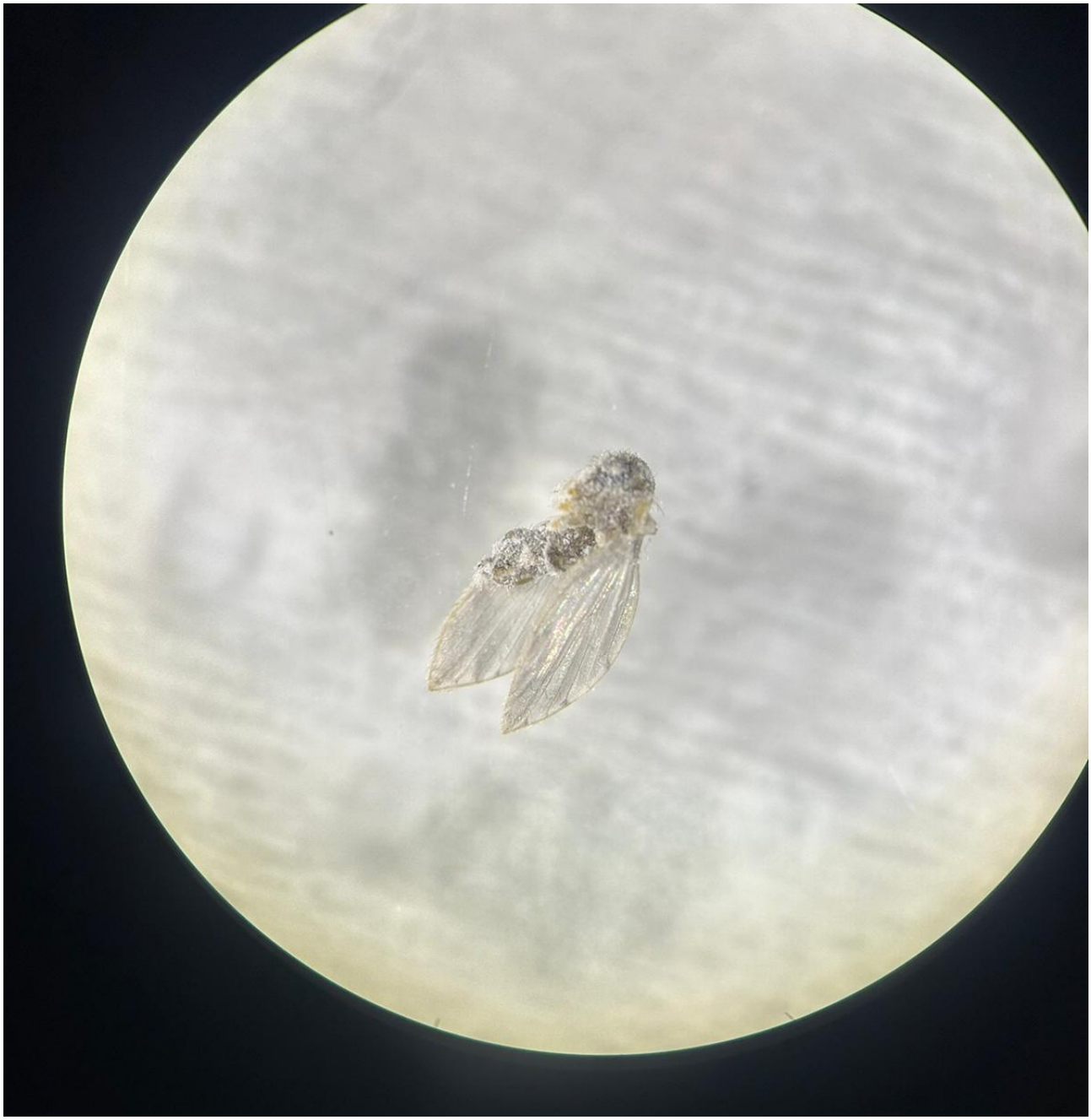
A 59-year-old married woman—a mother of two—presented with a 1-year history of persistent pruritus and the sensation of insects crawling on her body. She had initially sought dermatological evaluation multiple times, convinced that she had pediculosis or scabies; however, a clinical examination and laboratory investigations failed to confirm any parasitic infestation. The patient had undergone various empirical treatments, including antiparasitic shampoos and repeated courses of topical medications. Despite the lack of medical confirmation, she persistently self-administered these treatments. The onset of her symptoms was temporally associated with the concern that her son's friend, who had been recently released from prison, might have introduced parasites into their home. Despite reassuring explanations from multiple physicians, she continued to engage in compulsive hygiene behaviors, including daily house disinfection, meticulous ironing of clothes, and frequent bathing, leading to progressive social withdrawal.

Approximately 1 week after staying at the patient's home, her sister developed similar symptoms, reporting crawling sensations, frequent washing, and self-inflicted hair cutting. Shortly thereafter, the patient's niece, who had briefly visited the sister's home, also developed identical symptoms, suggesting a progressive shared delusional component within the family unit. Additionally, the patient provided a sample on a white napkin, which she insisted contained evidence of the parasites (Figure 3). A microscopic evaluation identified the presence of a fly within the sample (Figure 4).

None of the affected individuals had a history of psychiatric disorders, substance abuse, or alcohol consumption. Due to the worsening clinical picture and evidence of a transmitted delusional belief, a psychiatric evaluation was conducted. The Positive and Negative Syndrome Scale (PANSS) was administered, with the index patient scoring 69 (positive symptoms subscale score: 16; negative symptoms subscale score: 12; general psychopathology subscale score: 41), indicating significant psychotic symptoms. To exclude organic causes, a comprehensive laboratory workup (including complete blood counts; serum electrolytes; liver and renal function tests; fasting glucose; thyroid function tests; vitamin B12 levels; and serological tests for hepatitis, syphilis, and HIV) was performed, and all results were within normal limits. A brain MRI scan was also conducted for the primary patient, revealing no structural abnormalities.

Figure 3. The patient provided a white napkin containing alleged parasites, which, on examination, revealed a psychodid fly.

Figure 4. Microscopic view of the sample from Figure 3, showing a psychodid fly—a nonparasitic insect species.



She was referred to the psychiatry department for a consultation, and trifluoperazine was initiated at 5 mg/day, with a planned dose escalation. Over the course of hospitalization, her pruritic symptoms gradually diminished; however, she continued to express concerns about spreading the infestation to others, reflecting partial insight impairment. Consequently, her trifluoperazine dose was increased to 15 mg/day, and structured psychotherapy was introduced, focusing on cognitive restructuring and anxiety management techniques.

Psychoeducation sessions were also conducted for the patient's sister and niece to enhance their understanding of the disorder and minimize the reinforcement of shared delusional beliefs. They were advised on strategies for supporting the patient's recovery while avoiding behaviors that might reinforce the delusion.

By the fourth week of treatment, the patient exhibited a significant reduction in symptom severity, with her PANSS score improving to 50 (positive symptoms subscale score: 11; negative symptoms subscale score: 12; general psychopathology subscale score: 27). After marked improvements in functionality and symptom control, she was discharged with outpatient psychiatric follow-up appointments, continued pharmacotherapy, and ongoing family psychoeducation for preventing symptom recurrence in the shared delusional network.

Discussion

Delusional parasitosis is classified into primary, secondary, and organic forms. Primary delusional parasitosis occurs as an isolated delusional disorder without any underlying psychiatric or medical condition [7]. Secondary delusional parasitosis is

associated with psychiatric disorders, such as schizophrenia, major depressive disorder, dementia, anxiety disorders, and phobias [6,8]. The organic form of delusional parasitosis arises due to medical conditions, including hypothyroidism, anemia, diabetes mellitus, vitamin B12 deficiency, hepatitis, syphilis, and HIV infection [2]. Additionally, substance abuse, particularly cocaine use, has been implicated in triggering delusional parasitosis symptoms [5]. In our cases, neurocognitive evaluations, including the MMSE and MMPI, showed no significant cognitive deficits, suggesting a primary psychiatric etiology. Further, the laboratory outcomes were within normal limits and thus excluded organic reasons. As such, our cases were accepted as primary delusional parasitosis.

A key clinical hallmark of delusional parasitosis is the matchbox sign, that is, patients presenting dermatologists with small particles, such as dust, skin debris, or fibers, as "evidence" of their infestation [9]. In our cases, patients presented us with similar materials, including dust, fibers, and skin debris. In the second case, the patient presented a fly, which was identified as belonging to the family Psychodidae and subfamily Psychodinae (flies that do not harm humans), in addition to these materials. Patients with delusional parasitosis frequently experience tactile hallucinations, including sensations of crawling, stinging, or biting, which reinforce their conviction of infestation [10]. Consequently, they often engage in excessive hygiene practices, such as repeated washing, application of caustic substances, or compulsive skin scratching, which may lead to secondary skin damage, including excoriations, ulcerations, and irritant contact dermatitis [11]. The psychological burden associated with delusional parasitosis frequently results in social withdrawal; depression; and, in some cases, self-harm (as observed in our cases) [7].

Shared psychotic disorder (folie à deux) is a rare and complex psychiatric condition in which 2 or more individuals develop the same delusional beliefs [12]. Typically, a dominant individual (the primary case) has an underlying psychotic disorder, while the secondary individual adopts the same delusion through close emotional association and suggestibility. The primary case often presents with a chronic psychotic illness, such as a delusional disorder, whereas the secondary individual typically exhibits a more passive personality, lower self-esteem, and heightened susceptibility to suggestion [13].

Shared psychotic disorder is most commonly observed within nuclear families, with delusional beliefs typically being transmitted between spouses, between siblings, or between parents and children. This indicates that both genetic predisposition and psychosocial factors play a role in the disorder's etiology [13]. In our first case, the disorder manifested between a married couple; the wife, as the dominant individual, exhibited primary psychotic symptoms, and the husband later developed hallucinations and delusions that were similar to those of the wife. The second case involved a progressive transmission of delusional beliefs to the patient's sister and niece, which is consistent with folie à trois, wherein 3 individuals share the same fixed delusion.

Psychosocial stressors may act as triggers or maintaining factors of delusional parasitosis [6]. In our first case, the initial scabies infestation acted as a triggering factor for the development of delusional beliefs. In contrast, the delusion in the second case was precipitated by the patient's concern that her son's friend, who had been a guest in their home, might have introduced parasites into the household.

The management of shared delusional parasitosis is complex, requiring both pharmacological intervention and psychotherapeutic intervention [9]. Due to the lack of large-scale randomized controlled trials that focus on delusional parasitosis, treatment strategies rely primarily on case reports and small-scale studies [11]. Antipsychotic treatment remains the cornerstone; in particular, second-generation agents, such as risperidone and olanzapine, have demonstrated efficacy in reducing symptom severity [2]. However, patient compliance remains a significant obstacle [9].

In our first case, aripiprazole—an atypical antipsychotic drug—was initiated but was later switched to trifluoperazine—a first-generation antipsychotic—because of a skin rash. Adverse reactions, such as extrapyramidal symptoms, required adjunctive biperiden therapy, which led to symptom improvement. This finding is consistent with reports describing the necessity of extrapyramidal symptom monitoring in patients with delusional parasitosis who are on first-generation antipsychotics [1]. In the second case, the patient was treated with trifluoperazine and did not exhibit extrapyramidal symptoms.

An important consideration in shared delusional parasitosis cases is whether both the inducer and the recipient require treatment. Although the primary patient typically requires pharmacological intervention, the recipient may improve with separation and psychoeducation alone [12]. In our cases, behavioral therapy and psychoeducation were sufficient for secondary patients, who exhibited spontaneous symptom remission once the inducer underwent structured treatment. Psychoeducation also played a critical role in preventing the reinforcement of delusional beliefs among family members—a strategy emphasized in previous reports [2].

Conclusion

Shared delusional parasitosis (folie à deux and folie à trois) is a rare but clinically significant disorder that poses diagnostic challenges, particularly in cases involving family members or individuals in close relationships. Delusional parasitosis is often misdiagnosed as true parasitosis, leading to repeated antiparasitic treatments and delays in psychiatric intervention. Since dermatologists are often the first point of contact, it is crucial that they recognize delusional parasitosis early and facilitate psychiatric referrals. In our cases, patients underwent multiple dermatological consultations before receiving a psychiatric diagnosis, illustrating the need for greater awareness among dermatologists. Establishing a therapeutic alliance with patients is essential, as direct confrontation may lead to treatment refusal and the worsening of symptoms. As such, dermatologists should adopt a supportive approach and serve as a bridge between patients and psychiatric care to facilitate appropriate intervention. The management of delusional parasitosis requires close collaboration between dermatology and psychiatry professionals.

Conflicts of Interest

None declared.

References

1. Bewley AP, Lepping P, Freudenmann RW, Taylor R. Delusional parasitosis: time to call it delusional infestation. Br J Dermatol 2010 Jul;163(1):1-2. [doi: 10.1111/j.1365-2133.2010.09841.x] [Medline: 20645978]
2. Reich A, Kwiatkowska D, Pacan P. Delusions of parasitosis: an update. Dermatol Ther (Heidelb) 2019 Dec;9(4):631-638. [doi: 10.1007/s13555-019-00324-3] [Medline: 31520344]
3. Mumcuoglu KY, Leibovici V, Reuveni I, Bonne O. Delusional parasitosis: diagnosis and treatment. Isr Med Assoc J 2018 Jul;20(7):456-460. [Medline: 30109800]
4. Kim C, Kim J, Lee M, Kang M. Delusional parasitosis as 'folie a deux'. J Korean Med Sci 2003 Jun;18(3):462-465. [doi: 10.3346/jkms.2003.18.3.462] [Medline: 12808343]
5. Lepping P, Freudenmann RW. Delusional parasitosis: a new pathway for diagnosis and treatment. Clin Exp Dermatol 2008 Mar;33(2):113-117. [doi: 10.1111/j.1365-2230.2007.02635.x] [Medline: 18205853]
6. Szepietowski JC, Salomon J, Hrehorów E, Pacan P, Zalewska A, Sysa-Jedrzejowska A. Delusional parasitosis in dermatological practice. J Eur Acad Dermatol Venereol 2007 Apr;21(4):462-465. [doi: 10.1111/j.1468-3083.2006.01900.x] [Medline: 17373971]
7. Al-Imam AML. A systematic literature review on delusional parasitosis. Journal of Dermatology & Dermatologic Surgery 2016 Jan;20(1):5-14. [doi: 10.1016/j.jdds.2015.11.003]
8. Jerrom R, Mortimer H, Martin K, Siddiquee R, Bagchi D, Goulding JMR. A case series of shared delusional infestation: folie à deux revisited. Clin Exp Dermatol 2020 Jun;45(4):414-416. [doi: 10.1111/ced.14138] [Medline: 31729765]
9. Campbell EH, Elston DM, Hawthorne JD, Beckert DR. Diagnosis and management of delusional parasitosis. J Am Acad Dermatol 2019 May;80(5):1428-1434. [doi: 10.1016/j.jaad.2018.12.012] [Medline: 30543832]
10. Ahmad K, Ramsay B. Delusional parasitosis: lessons learnt. Acta Derm Venereol 2009;89(2):165-168. [doi: 10.2340/00015555-0587] [Medline: 19326002]
11. Gajbhiye A, Ali T, Aziz S, et al. Delusional parasitosis: a case series. Ind Psychiatry J 2023 Nov;32(Suppl 1):S258-S261. [doi: 10.4103/ipj.ipj_240_23] [Medline: 38370972]
12. Giam A, Tung YL, Tibrewal P, Dhillon R, Bastiampillai T. Folie à deux and delusional parasitosis. Asian J Psychiatr 2017 Aug;28:152-153. [doi: 10.1016/j.ajp.2017.04.027] [Medline: 28784373]
13. Galindo C, Galindo L. Folie à deux in the family environment: a case of shared delusion between a mother and her son. Cureus 2025 Feb 7;17(2):e78703. [doi: 10.7759/cureus.78703] [Medline: 39926628]

Abbreviations

MMPI: Minnesota Multiphasic Personality Inventory
MMSE: Mini-Mental State Examination
MRI: magnetic resonance imaging
PANSS: Positive and Negative Syndrome Scale

Edited by E Parker; submitted 02.06.25; peer-reviewed by CD Liberato, D Matmour; revised version received 14.07.25; accepted 15.07.25; published 07.08.25.

Please cite as:
Kurmuş GI, Karataş H, Erdem E, Doğan Bulut S, Gönül M, Kartal SP
Shared Delusional Parasitosis in Two Families: Clinical Insights Into Folie à Deux and Folie à Trois
JMIR Dermatol 2025;8:e78398
URL: https://derma.jmir.org/2025/1/e78398
doi:10.2196/78398

Cutaneous Atrophy Following Corticosteroid Injections for Tendonitis: Report of Two Cases

Rebecca Colwell[1], MD; Mitchell Gullickson[2], MD; Jonathan Cutlan[1], MD; Erik Stratman[1], MD

[1]Marshfield Clinic Health System, 1000 N Oak Ave, Marshfield, WI, United States

[2]School of Medicine and Health Sciences, University of North Dakota, Grand Forks, ND, United States

Corresponding Author:
Erik Stratman, MD
Marshfield Clinic Health System, 1000 N Oak Ave, Marshfield, WI, United States

Abstract

Cutaneous atrophy resulting from corticosteroid injections for musculoskeletal indications is an underrecognized adverse effect among orthopedists and dermatologists. We present two cases of cutaneous atrophy following corticosteroid injections for wrist tendonitis. Patients presenting with cutaneous atrophy following orthopedic corticosteroid injections may be misdiagnosed with linear morphea, atrophoderma, or vascular disorders and receive unnecessary workups and delays in appropriate management. Dermatologists play an essential role in the evaluation of these patients.

(JMIR Dermatol 2025;8:e67921) doi:10.2196/67921

KEYWORDS

lipoatrophy; cutaneous atrophy; corticosteroid; adverse effects; tendonitis; musculoskeletal

Introduction

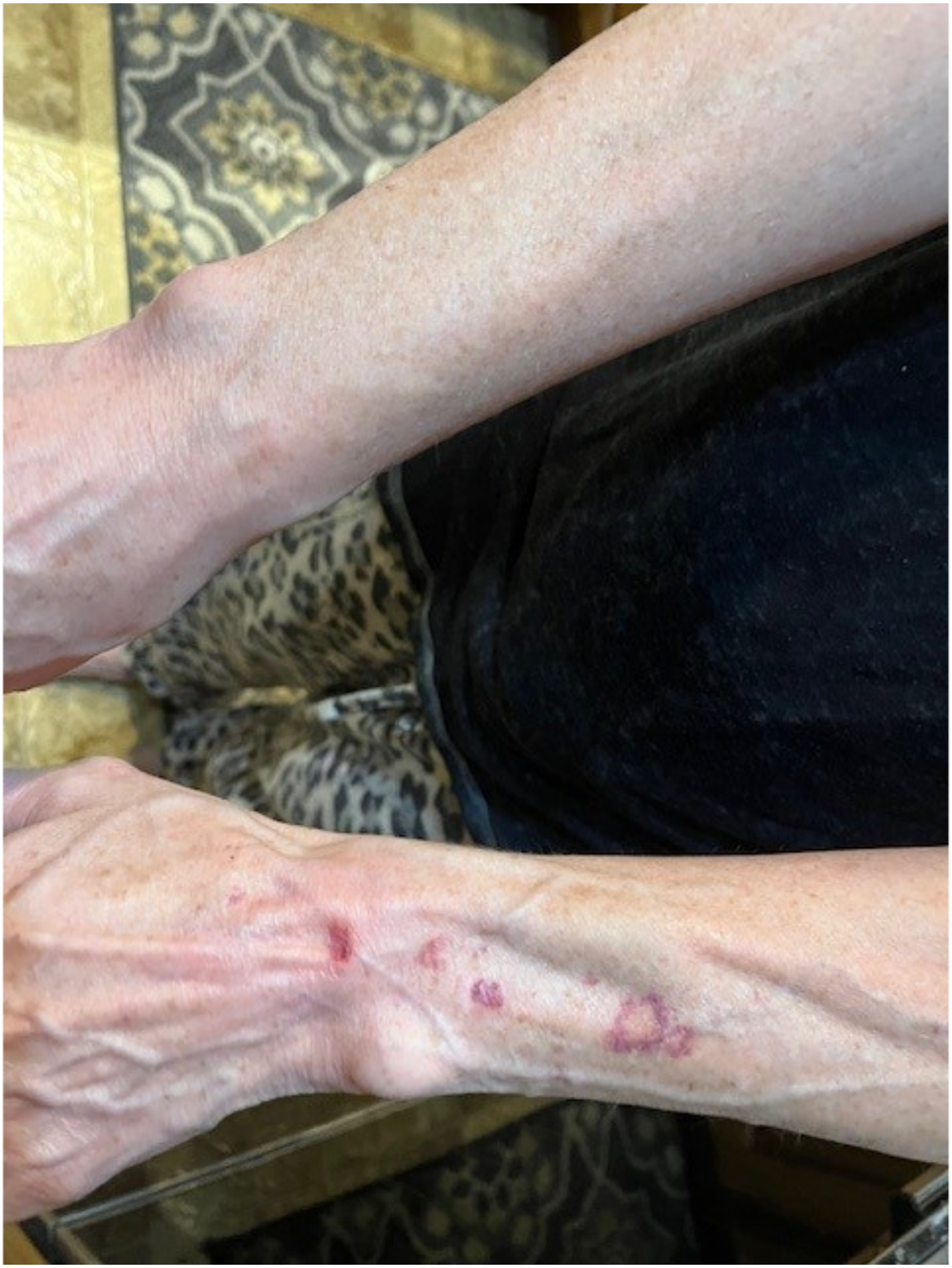
Injectable corticosteroids are commonly used to treat musculoskeletal conditions, including tendonitis [1]. Common adverse reactions to corticosteroid injections include atrophy, depigmentation, and cellulitis [1]. Skin depigmentation is a well-recognized adverse effect of corticosteroid injections, but atrophy is underrecognized. Atrophy typically manifests 2-4 months following the injection but may be delayed up to a year [2]. The pathophysiology of soft tissue atrophy and hypopigmentation is hypothesized to stem from macrophage-induced breakdown of adipose tissue, impaired function of melanocytes, and decreased synthesis of type I and type III collagen [3].

Dermatologists are familiar with the risks of cutaneous atrophy due to topical, intralesional, and intramuscular corticosteroid use from dermatologist-initiated treatments but may be less familiar with adverse effects associated with orthopedic uses. We present two patients with pronounced cutaneous atrophy of the injected wrists after corticosteroid injections for tendonitis.

Case 1

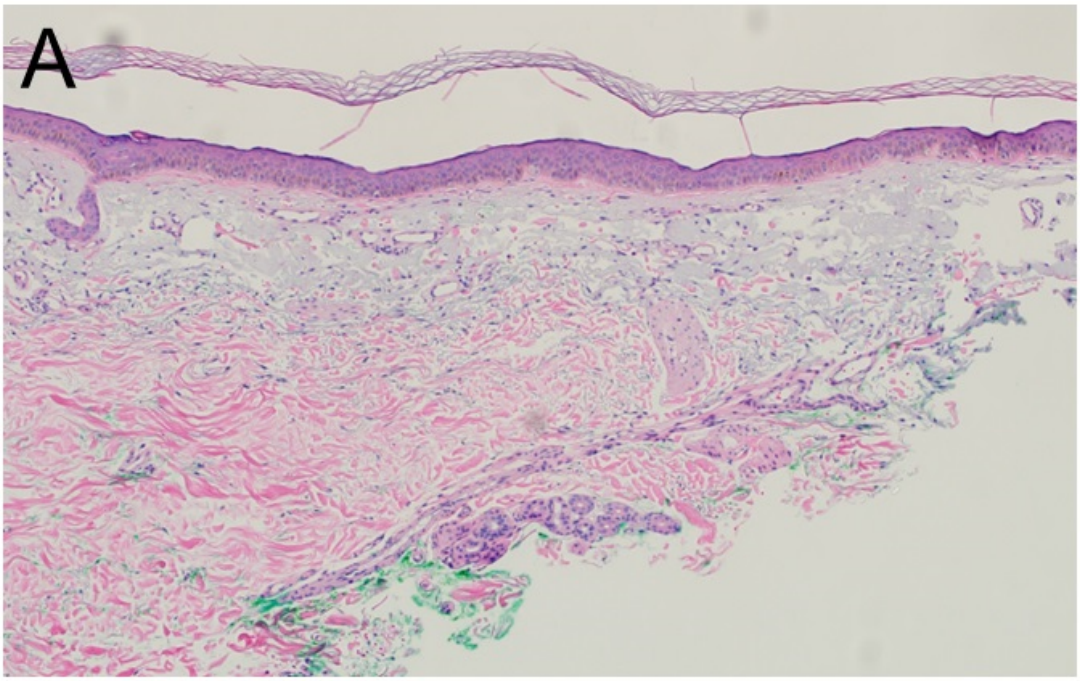
A 58-year-old woman presented to the Department of Dermatology for evaluation of skin fragility and discoloration over her left extensor forearm. Six months earlier, she received a 1-mL injection of a suspension of 0.5 mL of 40 mg/mL triamcinolone mixed with 0.5 mL of 1% lidocaine in the extensor carpi radialis brevis and extensor carpi radialis longus to treat wrist tendonitis. This initially relieved her pain, but 2 months following the injection, she noticed skin discoloration and soft tissue atrophy at her left distal forearm near the injection site. She noted proximal extension of the forearm atrophy. A physical examination revealed linear epidermal, dermal, and subcutaneous tissue atrophy; scattered ecchymoses; and cigarette-paper wrinkling of the skin on the left lateral wrist, forearm, and hand (Figure 1).

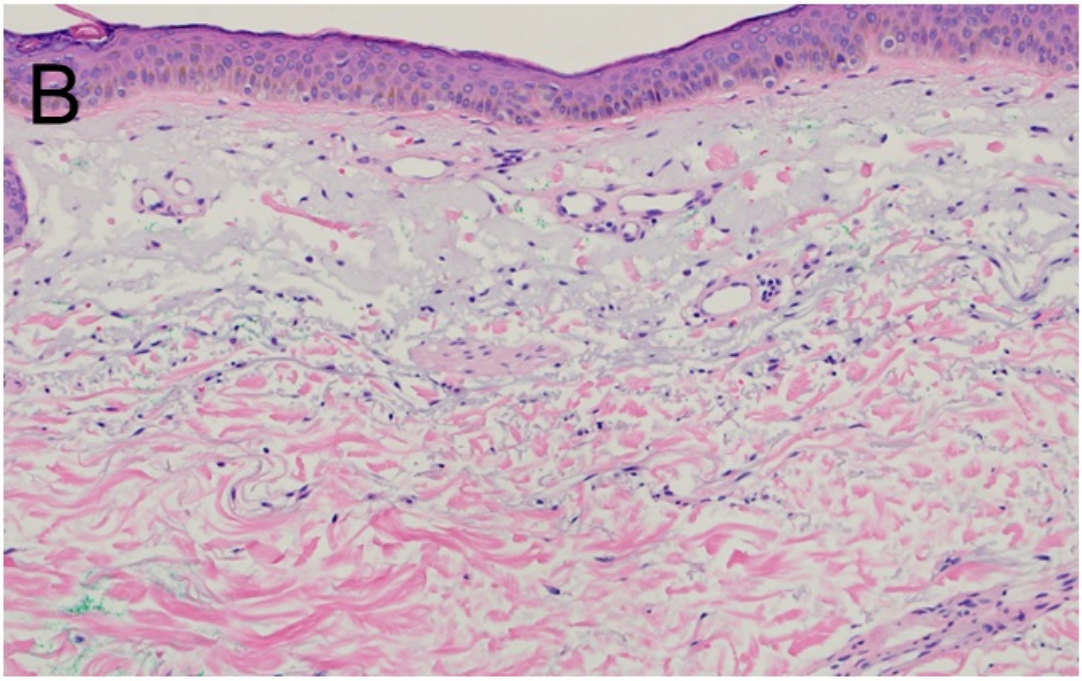
Figure 1. The left forearm shows prominent subcutaneous atrophy and purpura.



Neither the patient nor her care team connected the findings to previous wrist injections because the findings extended several centimeters proximal to the original injection sites. An electromyography test demonstrated no abnormalities. A 3-mm punch biopsy demonstrated mild epidermal atrophy, dermal elastosis, and slight vascular prominence (Figure 2).

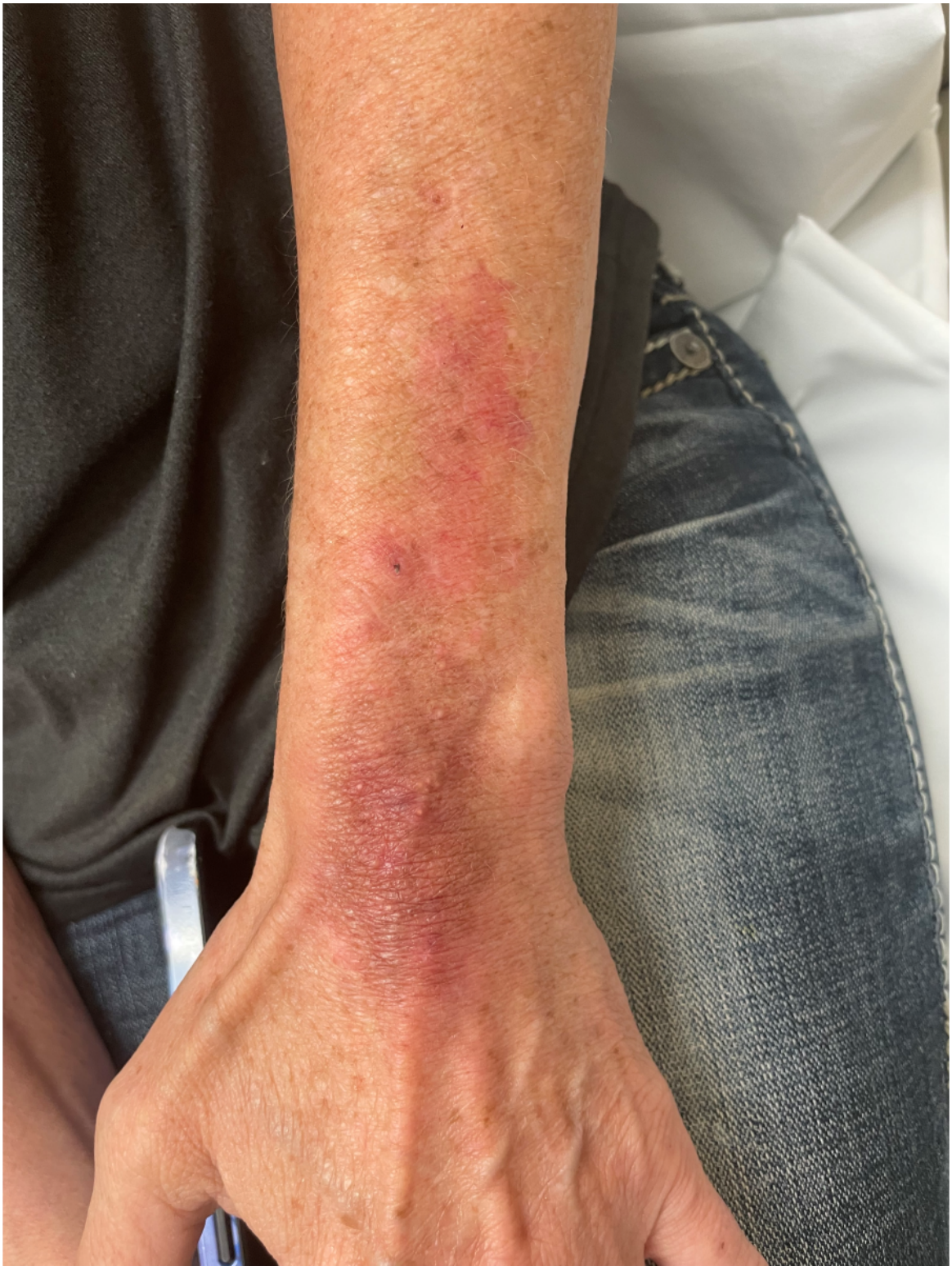
Figure 2. Epidermal atrophy, solar elastosis, and vascular prominence. A, Hematoxylin and eosin stain, original magnification x 100. B, Hematoxylin and eosin stain, original magnification x 200.





Active monitoring was chosen for management. One year later, the patient still experienced skin fragility, distal arm lipoatrophy, and wrist weakness. A physical examination showed persistent epidermal, dermal, and subcutaneous tissue atrophy including some telangiectasias, hemosiderin deposition, and cigarette-paper wrinkling of the skin over the left lateral wrist, proximal dorsal hand, and forearm (Figure 3).

Figure 3. Persistent subcutaneous atrophy, telangiectasias, and hemosiderin deposition over the left forearm.

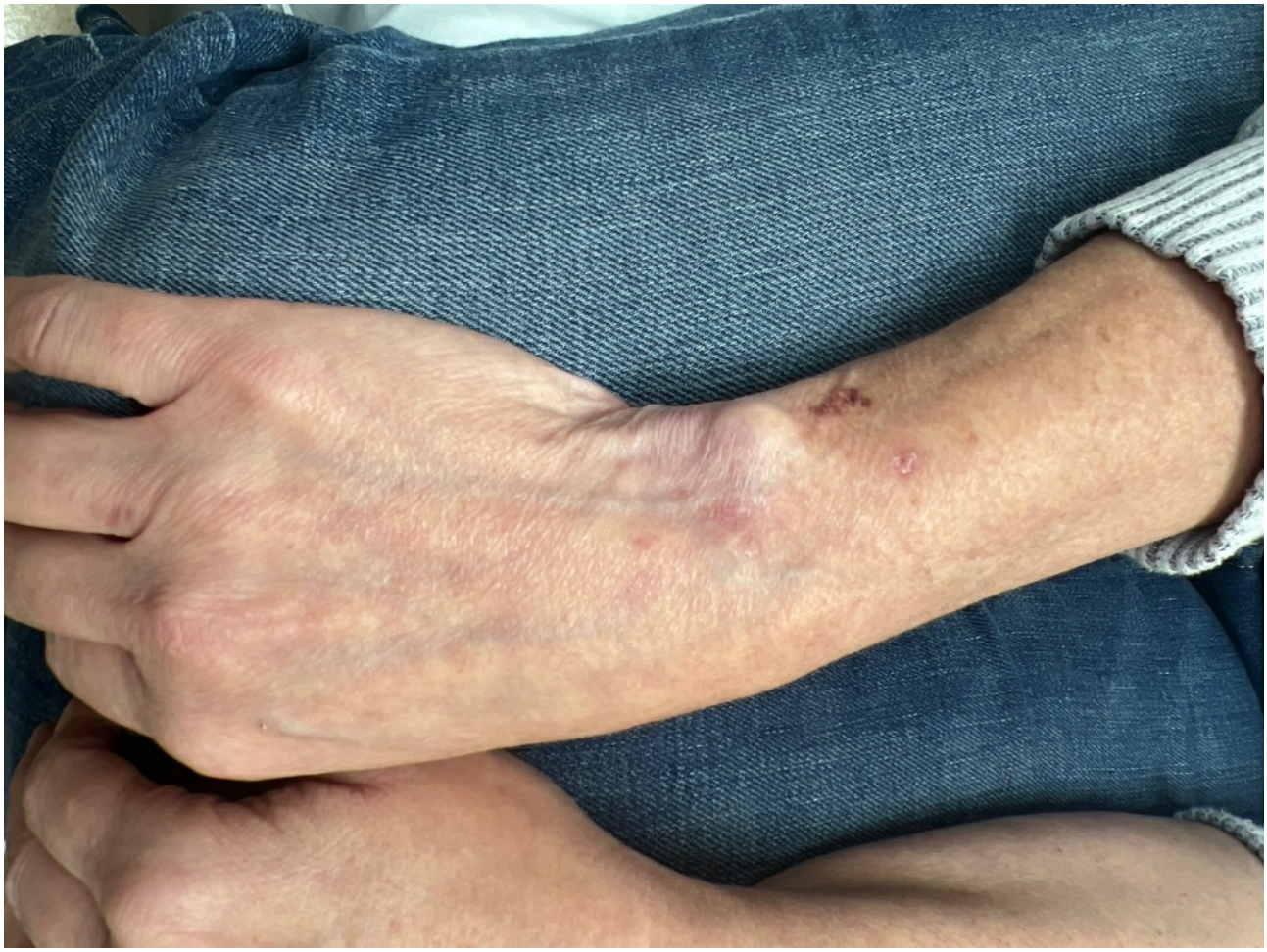


The patient consulted with the Department of Plastic Surgery for autologous fat grafting but declined further treatment.

Case 2

A 54-year-old woman presented with 3 months of painful, progressive purpura over the forearm skin associated with skin fragility. She received four separate 1-mL injections of a suspension of 0.5 mL of 40 mg/mL triamcinolone mixed with 0.5 mL of 1% lidocaine for extensor carpi ulnaris tendonitis and a partial triangular fibrocartilage complex tear. She noted her distal ulnar head was more prominent and the surrounding skin was hypopigmented. Pain, purpura, and skin fragility began shortly after the fourth injection, prompting a referral to the Department of Dermatology. A physical examination revealed epidermal, dermal, and subcutaneous tissue atrophy with overlying linear purpura (Figure 4).

Figure 4. The right wrist shows ulnar prominence secondary to subcutaneous atrophy with overlying purpura.



The patient was recommended to use over-the-counter topical vitamin C and E oils. The patient was offered serial saline injections but declined further treatment at the time of writing this report.

Ethical Considerations

Both patients provided written consent for their photographs and medical information to be published in print and online, with the understanding that this information may be publicly available.

Discussion

The timing and location of the patients' symptoms are most indicative of iatrogenic atrophy after corticosteroid injections. This is an uncommon but known adverse effect of these procedures. Clinicians injecting corticosteroids should advise patients of this risk in their informed consent, particularly when performing superficial injections. The unilateral proximal linear extension of the atrophy and dyspigmentation are often underrecognized as related to the therapeutic injection because the skin atrophy is so extensive and distant from the site of injection. This extension likely occurs secondary to venous or lymphatic diffusion of the insoluble microcrystalline steroid crystals [2]. A glossary of the dermatologic terms described in the report has been provided in Multimedia Appendix 1.

Clinicians can reduce the risks by choosing short-acting more soluble corticosteroids, avoiding injections with unnecessarily high concentrations or volumes of topical steroids, utilizing a 23- to 27-gauge needle to maximize delivery, and considering the use of point-of-care ultrasound—where available—to infiltrate anatomically discreet structures such as tendon sheaths [2].

There are few cases in the literature regarding the efficacy of the therapeutic options for steroid-induced lipoatrophy [2]. Current treatment options include autologous fat grafting, serial saline injections, autologous blood injections, and poly-l-lactic acid injections.

Autologous fat grafting is hypothesized to influence angiogenesis, improving soft tissue quality [4]. However, fat grafting does not address the dyspigmentation, telangiectasia, or epidermal fragility often observed with steroid-induced atrophy. Serial saline injections improve atrophy through the resuspension of steroid crystals, allowing for macrophage-mediated phagocytosis of these crystals [5-7]. Autologous blood injections stimulate cellular and humoral immune response factors such as vascular endothelial growth factor and hepatocyte growth factor [8,9]. Finally, for patients with more limited atrophy, poly-l-lactic acid injections can be administered, with the maximum improvement observed 6 months after injections [10].

Dermatologists managing patients with extensive iatrogenic atrophy cannot overlook the psychosocial impacts this may have. Given the high visibility of the body region affected, significant emotional toll can occur. Functional impairment and persistent weakness may also occur, occasionally with functional impairment and persistent weakness [1]. Medicolegal, workers' compensation, and risk management conversations are commonly needed, so dermatologists should be prepared for such inquiries and communications.

Extensive cutaneous wrist and forearm skin and soft tissue atrophy may occur in patients undergoing orthopedic wrist injections with corticosteroids. Dermatologists can play a pivotal role in identifying the cause, educating providers who perform these orthopedic procedures, and directing patients to the appropriate treatments.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Glossary of dermatologic terms.
[DOCX File, 16 KB - derma_v8i1e67921_app1.docx]

References

1. Brinks A, Koes BW, Volkers ACW, Verhaar JAN, Bierma-Zeinstra SMA. Adverse effects of extra-articular corticosteroid injections: a systematic review. BMC Musculoskelet Disord 2010 Sep 13;11:206. [doi: 10.1186/1471-2474-11-206] [Medline: 20836867]
2. Pace CS, Blanchet NP, Isaacs JE. Soft tissue atrophy related to corticosteroid injection: review of the literature and implications for hand surgeons. J Hand Surg Am 2018 Jun;43(6):558-563. [doi: 10.1016/j.jhsa.2018.03.004] [Medline: 29622410]
3. Oikarinen A, Haapasaari KM, Sutinen M, Tasanen K. The molecular basis of glucocorticoid-induced skin atrophy: topical glucocorticoid apparently decreases both collagen synthesis and the corresponding collagen mRNA level in human skin in vivo. Br J Dermatol 1998 Dec;139(6):1106-1110. [doi: 10.1046/j.1365-2133.1998.02646.x] [Medline: 9990383]
4. Coleman SR. Structural fat grafting: more than a permanent filler. Plast Reconstr Surg (1946) 2006 Sep;118(Suppl):108S-120S. [doi: 10.1097/01.prs.0000234610.81672.e7]
5. Daruwalla SB, Dhurat RS, Sharma A, Nitya MSN, Ghate S. Reversal of steroid-induced lipoatrophy with serial injections of isotonic saline in a child. Dermatol Ther 2020 May;33(3):e13354. [doi: 10.1111/dth.13354] [Medline: 32239676]
6. Shumaker PR, Rao J, Goldman MP. Treatment of local, persistent cutaneous atrophy following corticosteroid injection with normal saline infiltration. Dermatol Surg 2005 Oct;31(10):1340-1343. [doi: 10.1111/j.1524-4725.2005.31216] [Medline: 16188192]
7. Margulies SL, Morris A. Successful treatment of lipoatrophy with normal saline. JAAD Case Rep 2015 Nov;1(6):415-417. [doi: 10.1016/j.jdcr.2015.10.008] [Medline: 27051796]
8. Kutlu Ö, Ozdemir Cetinkaya P. Autologous blood injection for treatment of steroid atrophy: a case report. J Cosmet Dermatol 2021 Oct;20(10):3253-3256. [doi: 10.1111/jocd.14126] [Medline: 33811423]
9. Anitua E, Andía I, Sanchez M, et al. Autologous preparations rich in growth factors promote proliferation and induce VEGF and HGF production by human tendon cells in culture. J Orthop Res 2005 Mar;23(2):281-286. [doi: 10.1016/j.orthres.2004.08.015] [Medline: 15779147]
10. Brodell DW, Marchese Johnson S. Use of intralesional poly-L-lactic acid in the treatment of corticosteroid-induced lipoatrophy. Dermatol Surg 2014 May;40(5):597-599. [doi: 10.1111/dsu.12459] [Medline: 24533626]

Edited by R Dellavalle; submitted 24.10.24; peer-reviewed by D Zivanovic, G Whelan; revised version received 20.01.25; accepted 21.01.25; published 13.02.25.

Please cite as:
Colwell R, Gullickson M, Cutlan J, Stratman E
Cutaneous Atrophy Following Corticosteroid Injections for Tendonitis: Report of Two Cases
JMIR Dermatol 2025;8:e67921
URL: https://derma.jmir.org/2025/1/e67921
doi:10.2196/67921

Acral Persistent Papular Mucinosis in the United States: Case Series and Literature Review

Devin Miller[1,2], BS; Rachel Manci[1], MD; Jay Patel[1], MD; William Guo[1], MD; Daniel Lozeau[1], MD; James Briley[1], DO

[1]Department of Dermatology, Stony Brook University Hospital, 1320 Stony Brook Road, Building F Suite 200, Stony Brook, NY, United States
[2]Lewis Katz School of Medicine at Temple University, Philadelphia, PA, United States

Corresponding Author:
Devin Miller, BS
Department of Dermatology, Stony Brook University Hospital, 1320 Stony Brook Road, Building F Suite 200, Stony Brook, NY, United States

Abstract

Background: Acral persistent papular mucinosis (APPM) is a localized variant of lichen myxedematosus (LM) characterized by asymptomatic, flesh-colored papules primarily distributed on the hands and forearms. This chronic dermatosis, distinct from generalized mucinosis due to its lack of systemic involvement, remains underreported in medical literature.

Objective: In this study, we present two cases of APPM to the limited pool of documented cases in the United States, highlighting its emerging recognition.

Methods: This is a case series of two patients presenting with asymptomatic papular eruptions on the hands and wrists, consistent with the typical presentation of APPM. Diagnostic confirmation via biopsy revealed focal cutaneous mucinosis. Comprehensive laboratory evaluations, including serum and urine protein electrophoresis, showed no evidence of underlying gammopathy in either patient.

Results: Treatment modalities for APPM are limited and often ineffective. Unlike other forms of LM, APPM features are confined to skin lesions, posing primarily as a cosmetic concern with a favorable prognosis. Accurate diagnosis of this localized LM is crucial to differentiate it from the more severe, generalized form, scleromyxedema, which can have organ involvement and may become fatal. Notably, while spontaneous resolution is reported in LM, including discrete papular mucinosis, APPM typically persists without resolution even after extended follow-up.

Conclusions: These cases underscore the importance of recognizing APPM and advocating for broader awareness and exploration of its clinical variability, etiology, and management strategies. With increasing recognition, the understanding of APPM can be enhanced, paving the way for optimized management and improved outcomes for affected individuals.

(JMIR Dermatol 2025;8:e77714) doi:10.2196/77714

KEYWORDS

acral persistent papular mucinosis; dermoscopy; dermatology; lichen myxedematosus; mucin

Introduction

Acral persistent papular mucinosis (APPM) is a chronic, localized subtype of papular mucinosis, also known as lichen myxedematosus (LM). LM is characterized by lichenoid cutaneous manifestations, mucinous deposits, fibroblast proliferation, and dermal fibrosis. APPM is distinct in its localized nature, primarily affecting the extensor surfaces of the distal forearms and hands. Notably, it lacks systemic involvement and is not associated with the thyroid diseases seen in generalized forms [1]. To our knowledge, only six cases have been reported in the United States, with approximately 70 additional cases documented across Europe, North America, South America, and Asia, highlighting the limited available literature [1-5].

Classically, APPM presents as asymptomatic, solitary, white- or flesh-colored papules on the extensor surfaces of the hands, wrists, and dorsal forearms, ranging from 2 to 5 mm in size. These papules contain mucin deposits in the upper reticular dermis and often persist for years [1,6]. Nonetheless, uncommon outliers exist, with APPM-like mucinosis reported on the legs and chest [7-9]. Additionally, pruritic lesions have been reported in isolated cases [2]. These findings challenge the traditional assumption that APPM is an asymptomatic cutaneous condition limited to the forearms and hands. A potential genetic and environmental role has been suggested based on familial occurrences of APPM; however, the etiopathogenesis of the disease has yet to be explored extensively and remains uncertain [2,9].

Herein, we present two cases of APPM, helping to shed light on a condition currently underreported in the medical literature. Consent for the publication of all patient photographs and medical information is provided by the authors, stating that all patients gave consent for their photographs and medical information to be published in print and online versions and with the understanding that this information may be publicly available.

Case Descriptions

Case 1: A 64-year-old female patient with papular eruption on the hands (Figures 1 and 2).

Figure 1. Flesh-colored to slightly yellow firm flat-topped papules on the bilateral dorsal hands and wrists. A biopsy was obtained from the circled lesion on the left dorsal wrist.

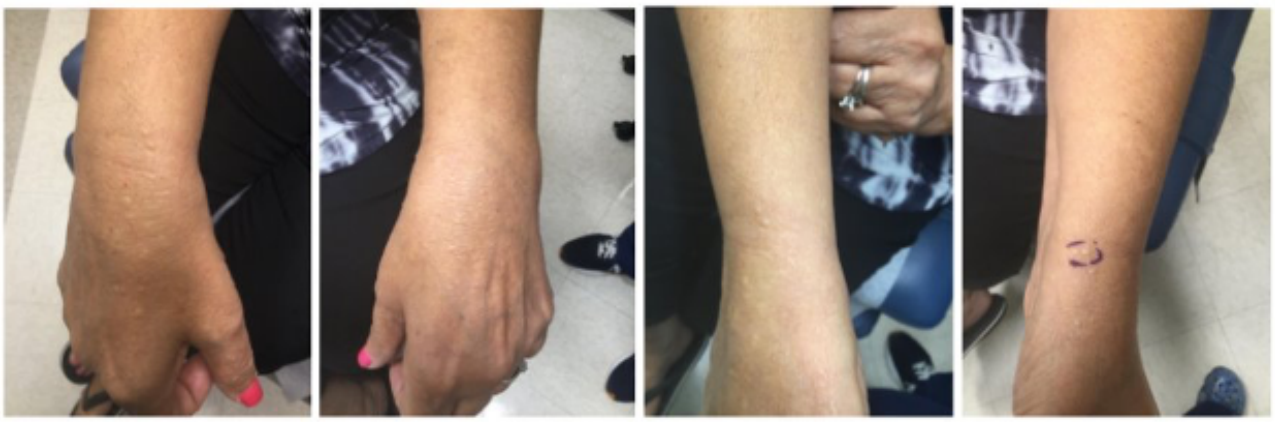
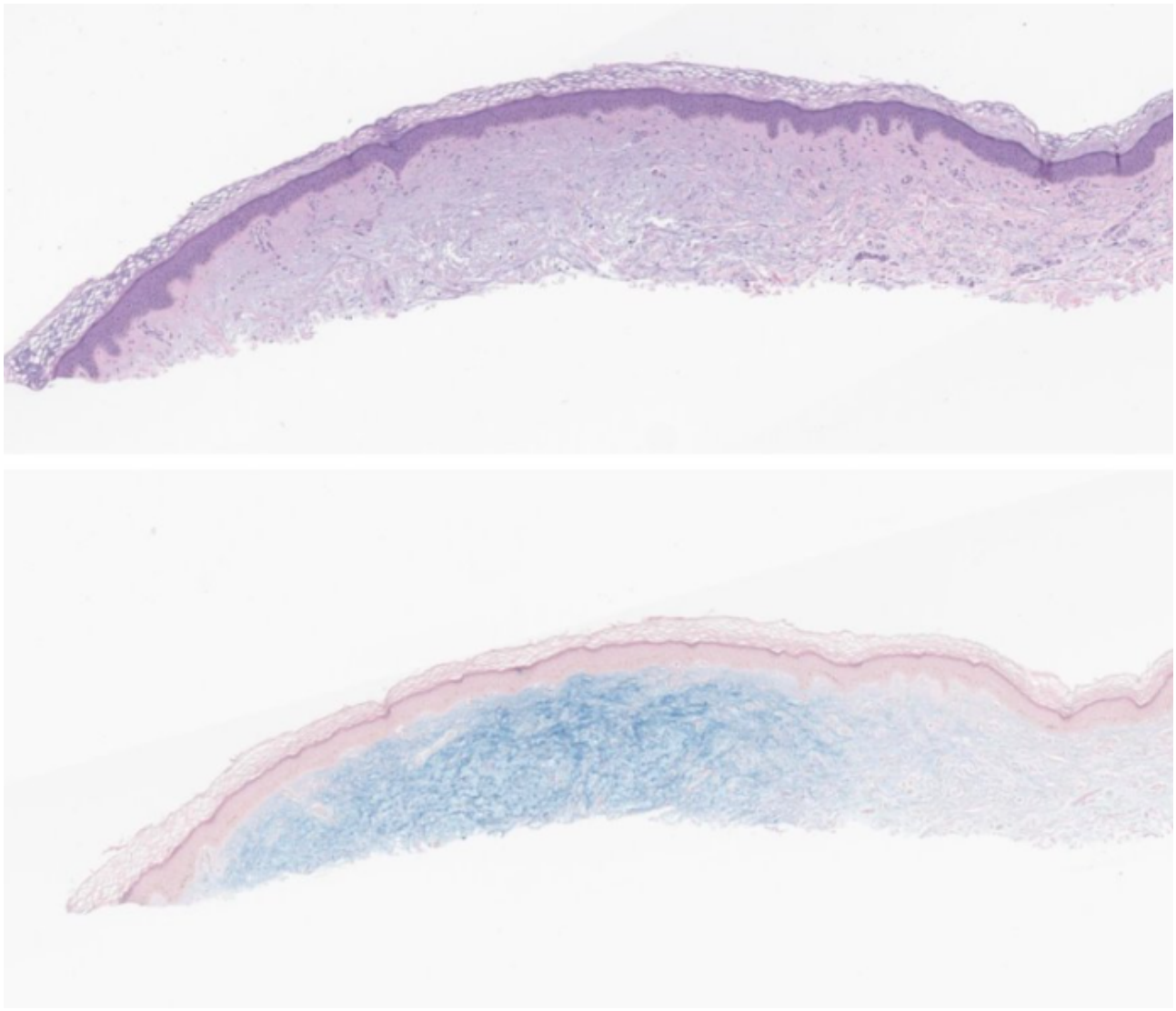


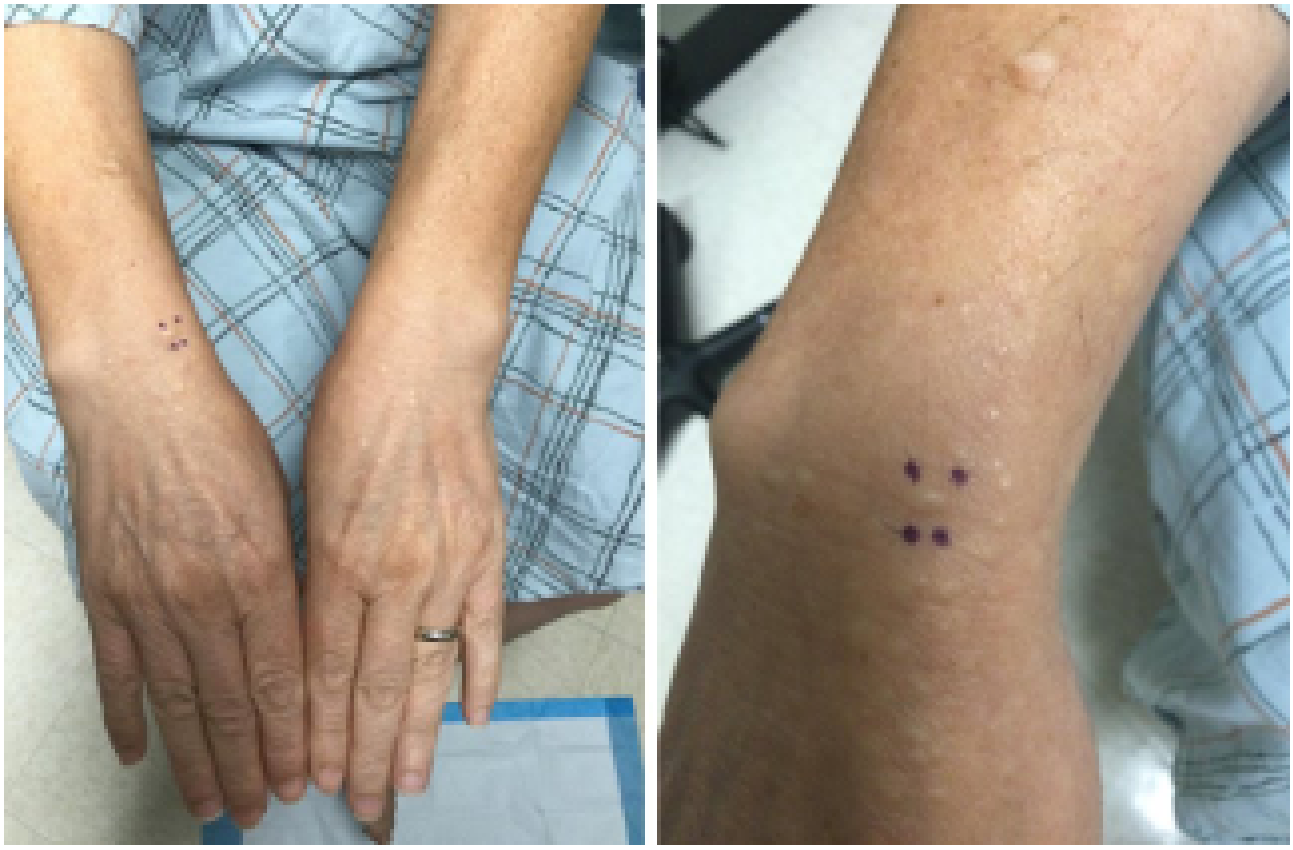
Figure 2. A shave biopsy with H&E (a) revealed deposition of blue-gray mucin within the superficial dermis, highlighted by colloidal iron stain, (b) Original magnification 40X.



A 64-year-old female patient with no significant past medical history presented with a several-year history of an intermittently pruritic papular eruption on both hands. Physical examination revealed multiple superficial flesh-colored papules ranging from 2 - 5mm in size on the bilateral dorsal hands, wrists, and distal forearms (Figure 1). Biopsy from the left wrist demonstrated focal cutaneous mucinosis, and a colloidal iron stain confirmed mucin deposition within the superficial dermis (Figure 2). Serum protein electrophoresis and urine protein electrophoresis were negative, aiding in ruling out an underlying monoclonal gammopathy. The patient was diagnosed with APPM and elected to defer any treatment. There has been no progression or spontaneous resolution of her condition to date.

Case 2: A 67-year-old male patient with papular eruption on the dorsal hands and wrists (Figure 3).

Figure 3. Flesh-colored to slightly yellow spongy papules on the bilateral dorsal hands and wrists. A biopsy was obtained from the circled lesion of the right dorsal wrist.



A 67-year-old male patient with no pertinent past medical history presented with a several-year history of stable papular eruptions on the dorsal hands, wrists, and distal forearms. The physical examination revealed focally scattered, flesh-colored, spongy papules measuring 2 - 5 mm in size. A biopsy from the right dorsal wrist revealed focal cutaneous mucinosis. Based on the clinicopathological correlation and lack of systemic involvement, a diagnosis of APPM was made. The patient deferred treatment and was subsequently lost to follow-up; to our knowledge, there was no progression or spontaneous resolution of his condition during the observed period.

Discussion

We present two additional cases of APPM, adding to the six previously reported cases in the United States literature [2]. This may reflect either a rising recognition of the condition or significant underreporting to date.

Notably, both of our patients were predominantly asymptomatic and had no underlying medical conditions, consistent with prior reports suggesting that APPM is not associated with systemic disease or an underlying gammopathy [10]. Unlike other forms of LM, APPM is a skin-limited condition with a favorable prognosis.

Accurate diagnosis of APPM is crucial to differentiate it from generalized LM, scleromyxedema, which can have organ involvement and may be fatal without proper diagnosis and treatment. Diagnostic features of APPM include the presence of ivory to flesh colored papules ranging 2 - 5mm in size, female predominance, persisting without spontaneous resolution, and the absence of systemic disease overlap or associated gammopathy [10]. Histologically, APPM is characterized by focal, well-circumscribed mucin in the papillary and mid dermis, sparing the Grenz zone, with the absence or variations of fibroblast proliferation [10]. Unlike other forms of LM, including discrete papular mucinosis, which may resolve spontaneously, APPM generally persists over time, as observed in our cases [2].

A variety of treatment strategies for APPM have been described in the literature. Topical and intralesional corticosteroids have shown minimal to no clinical improvement [2]. Tacrolimus 0.1% ointment has been postulated as a potential treatment option for LM by inhibiting tumor necrosis factor (TNF)-α secretion and transforming growth factor (TGF)-β-induced collagen synthesis, although only a partial response has been reported in the literature [2,9]. Destructive modalities, such as electrofulguration, have demonstrated efficacy in lesion resolution, albeit with mild scarring [2].

In conclusion, our case series highlights the importance of recognizing APPM and adds to the 70 documented cases worldwide, including now eight from the United States (Table 1). As APPM remains an underreported entity in the medical literature, these cases serve to enhance awareness and encourage further exploration into its clinical variability, etiology, genetic predispositions, and optimal management strategies. Importantly, our cases provide additional evidence to support accurate diagnostic approaches that help distinguish APPM from more severe forms of LM, such as scleromyxedema. Proper

diagnosis can help prevent unnecessary treatment and testing. Increased recognition of APPM will ultimately enhance understanding of the condition and guide better management, leading to improved outcomes for affected patients.

Table . Summary of the APPM cases reported from clinics in the US [2,10].

Study name	Author and year	Patient demographics	Clinical features	Histological findings
Acral persistent papular mucinosis: a distinctive dermal mucinosis. Case reported at the meeting of the American Academy of Dermatology, San Antonio, Texas	Berbaum 1987[a] [11]	N/A[b]	N/A	N/A
Acral persistent papular mucinosis	Fosko 1992 [12]	40-year-old-female	Back of hands, extensor aspect of wrists Developingx1 yr	N/A
Flesh-colored papules on the wrists of a 61-year-old man	Kineston 2004 [13]	61-year-old-male	Back of wrists and hands gradual increase in #x5 yrs	N/A
Acral persistent papular mucinosis	Harris 2004 [10]	55-year-old-female	Back of hands, extensor aspect of wrists and forearms; Increasing in #x5 yrs	Mild epidermal thinning with intact structure; widened dermal collagen spacing. Alcian blue staining revealed defined mucin deposits in upper/mid-reticular dermis, sparing the grenz zone. Hyaluronidase digestion confirmed hyaluronic acid. Scattered fibrocytes and mast cells present [10]
'Spreading bumps' on hands of a Native American	Sebastian 2008 [14]	62-year-old-male	Dorsa of hands, wrists and extensor forearms slowly spreading	N/A
Treatment of acral persistent papular mucinosis using an Erbium-YAG[c] laser	Graves 2015 [15]	60-year-old-female	Dorsal hands	A tissue sample taken from a lesion on the right dorsal hand showed localized mucin accumulation when examined with colloidal iron stain, aligning with features of acral persistent papular mucinosis.
Our manuscript (Case 1)	__[e]	64-year-old-female	Dorsal hands and wrists x several years	A shave biopsy from the left wrist with H&E[d] (Figure 2) revealed deposition of blue-gray mucin within the superficial dermis, highlighted by colloidal iron stain (Figure 2).
Our manuscript (Case 2)	__[e]	67-year-old-male	Dorsal hands and wrists	A biopsy from the right dorsal wrist revealed focal cutaneous mucinosis.

[a]No full text was available.

[b]N/A: not available.

[c]YAG: Yttrium Aluminum Garnet.

[d]H&E: Hematoxylin and Eosin.

[e]Not applicable.

Acknowledgments

These cases have not been published previously.

Conflicts of Interest

None declared.

References

1. Rongioletti F, Ferreli C, Atzori L. Acral persistent papular mucinosis. Clin Dermatol 2021;39(2):211-214. [doi: 10.1016/j.clindermatol.2020.10.001] [Medline: 34272012]
2. Luo DQ, Wu LC, Liu JH, Zhang HY. Acral persistent papular mucinosis: a case report and literature review. J Dtsch Dermatol Ges 2011 May;9(5):354-359. [doi: 10.1111/j.1610-0387.2010.07594.x] [Medline: 21122068]
3. Mori A, Shinkuma S, Mitsui Y, Ogawa K, Miyagawa F, Asada H. Acral persistent papular mucinosis: a case report and summary of 24 Japanese cases. J Dermatol 2021 Oct;48(10):1574-1578. [doi: 10.1111/1346-8138.16043] [Medline: 34169556]
4. Toh JJH, Goh NSG, Wang DY. A rare case of acral persistent papular mucinosis. Clin Case Rep 2020 Feb;8(2):344-346. [doi: 10.1002/ccr3.2639] [Medline: 32128185]
5. Park YJ, Shin HY, Choi WK, Lee AY, Lee SH, Hong JS. Solitary acral persistent papular mucinosis nodule: a case report and summary of eight Korean cases. World J Clin Cases 2023 May 6;11(13):3086-3091. [doi: 10.12998/wjcc.v11.i13.3086] [Medline: 37215408]
6. Rongioletti F, Rebora A, Crovato F. Acral persistent papular mucinosis: a new entity. Arch Dermatol 1986 Nov;122(11):1237-1239. [Medline: 3777968]
7. Navarrete-Dechent C, Bajaj S, Marghoob A, González S, Jaque A. Acral persistent papular mucinosis (APPM): dermoscopy of an uncommon disease. J Am Acad Dermatol 2017 Feb;76(2S1):S10-S11. [doi: 10.1016/j.jaad.2016.03.046] [Medline: 28087015]
8. Madan V, Gangopadhyay M, Dawn G. Multiple asymptomatic papules on the legs. Clin Exp Dermatol 2007 Nov;32(6):773-774. [doi: 10.1111/j.1365-2230.2007.02490.x] [Medline: 17953645]
9. Jun JY, Oh SH, Shim JH, Kim JH, Park JH, Lee DY. Acral persistent papular mucinosis with partial response to tacrolimus ointment. Ann Dermatol 2016 Aug;28(4):517-519. [doi: 10.5021/ad.2016.28.4.517] [Medline: 27489445]
10. Harris JE, Purcell SM, Griffin TD. Acral persistent papular mucinosis. J Am Acad Dermatol 2004 Dec;51(6):982-988. [doi: 10.1016/j.jaad.2004.07.002] [Medline: 15583597]
11. Berbaum M, Litt JZ, Bass J. Acral persistent papular mucinosis: a distinct dermal mucinosis. 1987 Presented at: Meeting of the American Academy of Dermatology; Dec 8, 2025; San Antonio, Texas.
12. Fosko SW, Perez MI, Longley BJ. Acral persistent papular mucinosis. J Am Acad Dermatol 1992 Dec;27(6 Pt 1):1026-1029. [doi: 10.1016/s0190-9622(08)80277-0] [Medline: 1479089]
13. Kineston DP, Willard RJ, Krivda SJ. Flesh-colored papules on the wrists of a 61-year-old man. Arch Dermatol 2004 Jan;140(1):121-126. [doi: 10.1001/archderm.140.1.121-a]
14. Sebastian S, Zlotoff B, DeSantis E. Spreading bumps' on hands of a Native American. Clin Exp Dermatol 2008 Mar;33(2):225-227. [doi: 10.1111/j.1365-2230.2007.02621.x]
15. Graves MS, Lloyd AA, Ross EV. Treatment of acral persistent papular mucinosis using an Erbium-YAG laser. Lasers Surg Med 2015 Aug;47(6):467-468. [doi: 10.1002/lsm.22368]

Abbreviations

APPM: acral persistent papular mucinosis
LM: lichen myxedematosus

Edited by R Dellavalle; submitted 27.05.25; peer-reviewed by E Akasaka, F Rongioletti; revised version received 07.08.25; accepted 08.08.25; published 16.09.25.

Please cite as:
Miller D, Manci R, Patel J, Guo W, Lozeau D, Briley J
Acral Persistent Papular Mucinosis in the United States: Case Series and Literature Review
JMIR Dermatol 2025;8:e77714
URL: https://derma.jmir.org/2025/1/e77714
doi:10.2196/77714

Evaluating Artificial Intelligence Models in Dermatology: Comparative Analysis

Aneri Bhargav Patel[1,2], BS; William Driscoll[3], BS; Conan H Lee[2], BS; Cameron Zachary[4], MD; Nicole M Golbari[5], MD, MPH; Janellen Smith[4], MD

[1]Department of Dermatology, University of California, Irvine, 850 Health Sciences Road, 2nd floor, Irvine, CA, United States

[2]School of Medicine, University of California, Davis, Sacramento, CA, United States

[3]New York Genomics, New York, NY, United States

[4]Department of Dermatology, University of California, Irvine, 850 Health Sciences Road, 2nd floor, Irvine, CA, United States

[5]Department of Dermatology, NYU Langone Health, New York, NY, United States

Corresponding Author:
Aneri Bhargav Patel, BS
Department of Dermatology, University of California, Irvine, 850 Health Sciences Road, 2nd floor, Irvine, CA, United States

Abstract

DermGPT demonstrated strong potential for improving answer clarity and conciseness in dermatology-related queries, while ChatGPT provided more robust source citations, enhancing trust in evidence-based responses.

(JMIR Dermatol 2025;8:e74040) doi:10.2196/74040

KEYWORDS

ChatGPT; DermGPT; artificial intelligence; dermatology; education; LLM; large language model

Introduction

Large language models (LLMs) like OpenAI's GPT-4o use transformer architectures with self-attention to process and generate human-like responses. ChatGPT, developed by OpenAI [1], enhances a GPT-4 model with reinforcement learning from human feedback, filtering inappropriate content [2]. These models predict the next word based on prior context. Trained on vast internet data, they can address diverse topics, including dermatology. However, LLMs may "hallucinate," producing plausible but incorrect information [3,4], limiting clinical utility.

DermGPT [5], developed under the Palo Alto Medical Foundation, is tailored for dermatology. Beyond drafting notes and authorizations, it answers dermatology questions using a GPT base model enhanced by a research database. By sourcing answers from this material and showing citations, DermGPT aims to reduce hallucinations and better support dermatologists [6]. We compared its responses to those of ChatGPT.

Methods

Overview

ChatGPT was selected for its popularity and prior evidence of superiority in dermatology-related tasks. A double-blind study found dermatologists preferred ChatGPT over Google's Bard for patient handouts [7]. ChatGPT 4o was used. DermGPT's only available model was used.

Two dermatology residents, CZ and NMG, authored a list of questions posed to each LLM (Multimedia Appendix 1). Three questions to which DermGPT did not provide a response were excluded as nonevaluable item pairs. The two models' answers for a given question were paired and assigned as A or B using a computer-generated randomization list. Any identifiable metadata such as formatting was cleared. The survey was distributed to dermatologists at the University of California, Irvine, and the University of California, Davis, via email and QR codes. Survey takers were informed that both responses were produced by LLMs, but they were blinded to which model produced which response. They were asked to choose their preferred answers based on quality—specifically, which answer they thought would be best suited for patient care or was most accurate.

The rating options were as follows:

- Model A better
- Model B better
- Equal quality
- Both inadequate

Statistical analysis was conducted using SAS OnDemand for Academics (version 9.4). χ^2 tests ($P<.05$) assessed significance. Interrater reliability was not prespecified and not assessed; ratings were aggregated at the item level.

Ethical Considerations

This study used a voluntary, anonymous survey of physicians and residents. According to institutional and national guidelines,

the project did not require institutional review board review because no identifiable information was collected and the study posed minimal risk.

Participants provided implied consent by completing the survey after being informed of its purpose and their ability to withdraw at any time. No compensation was provided. The survey responses were analyzed in aggregate to ensure anonymity and privacy in accordance with institutional standards. The study followed the ethical principles of the Declaration of Helsinki, adhered to Committee on Publication Ethics guidelines**,** and met all institutional requirements for minimal-risk survey research.

Results

Overview

Of 64 dermatology faculty and 30 residents across the University of California, Irvine, and the University of California, Davis, we received a total of 19 responses, comprising 13 attending physicians and 6 residents or fellows. This corresponds to an overall response rate of approximately 20%.

Which LLM's Answer Was Better: ChatGPT or DermGPT?

Overall, DermGPT's answers (48.1%) were preferred over ChatGPT's (28.4%); the χ^2 test was significant with P=.04 (P<.05). In the attending group, DermGPT's answers were preferred (93/195, 47.7%) over ChatGPT's (56/195, 28.7%). Likewise, in the resident group, DermGPT's answers were preferred (44/90, 48.9%) versus ChatGPT (25/90, 27.8%) (Table 1).

Table . User-preferred artificial intelligence answer.

Group/responses	ChatGPT		DermGPT		Other		Total answers	
	Values, n (%)	Percentage of total responses	Values, n (%)	Percentage of total responses	Values, n (%)	Percentage of total responses	Responses	Percentage of total responses
Attending	56 (28.7)	19.6	93 (47.7)	32	46 (23.6)	16.1	195	68.4
Resident	25 (27.8)	8.8	44 (48.9)	15.4	21 (23.3)	7.4	90	31.6
Total	81	28.4	137	48.1	67	23.5	285	100

[a]χ^2 test: P=.04.

Which LLM's References Were Better: ChatGPT or DermGPT?

Overall, ChatGPT references (46%) were preferred over DermGPT (23.5%; χ^2_2=1.385; P=.50). In the attending group, ChatGPT references were also preferred (94/195, 48.2%) over DermGPT (45/195, 23.1%). Likewise, in the resident group, ChatGPT references were preferred (37/90, 41.1%) versus DermGPT (22/90, 24.4%) (Table 2).

Table . Overall preference for references.[a]

Group	ChatGPT		DermGPT		Other		Total answers	
	Values, n (%)	Percent of total responses	Values, n (%)	Percent of total responses	Values, n (%)	Percent of total responses	Responses	Percent of total responses
Attending	94 (48.2)	33	45 (23.1)	15.8	56 (28.7)	19.6	195	68.4
Resident	37 (41.1)	13	22 (24.4)	7.7	31 (34.4)	10.9	90	31.6
Total	131	46	67	23.5	87	30.5	285	100

[a]χ^2_2=1.385; P=.50.

Discussion

Principal Results

Out of 195 responses, users generally preferred DermGPT's answers, while ChatGPT was favored for its reference citations (Table 2). DermGPT's concise and well-phrased responses made it accessible for quick clinical reference. However, 3 questions were excluded because DermGPT issued disclaimers instead of direct answers, recommending consultation with a dermatologist or guidelines. The multimedia appendices show the results tabulated from SAS as well as the questionnaire and responses (Multimedia Appendices 1-6).

ChatGPT consistently cited reputable references such as the *Journal of the American Academy of Dermatology* and the *Journal of the American Medical Association*, contributing to user trust and perceived academic rigor. Although DermGPT offers clarity, ChatGPT's strong sourcing enhances credibility. These results suggest the potential for a hybrid model that combines both strengths.

Limitations

Our study was constrained by a small rater sample (n=19) and multiple ratings per rater and per question. As a result, *P* values should be interpreted as exploratory rather than confirmatory. The sample may not represent all dermatology clinicians, limiting generalizability. Subgroup patterns were underpowered.

Comparison With Prior Work

Several studies have compared LLMs to each other and to humans. He et al [8] found GPT-4 sometimes produced inaccurate, nonindividualized responses to laboratory-related queries. Iannantuono et al [9] compared ChatGPT-4, ChatGPT-3.5, and Google Bard in immunooncology, stressing the need for expert verification. Fernández-Pichel et al [10] found LLMs answered 80% of health questions accurately, though results were sensitive to prompt phrasing. This is the first study comparing ChatGPT and DermGPT for dermatologic responses.

Conclusions and Future Directions

Future research should include models like Claude and Gemini, expand sample size, and explore combining DermGPT's brevity with ChatGPT's sourcing. These results highlight the importance of balancing clarity and citation in artificial intelligence–assisted medical tools.

Acknowledgments

We are grateful to the dermatology residents, attendings, and fellows at the University of California, Irvine, and University of California, Davis, who took the time to take our survey and make this study possible. We used the generative AI tool ChatGPT by OpenAI and DermGPT to generate answers and references for our survey, which we analyzed. The original answers and questions posed have been made available in Multimedia Appendix files.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study process.
[PNG File, 48 KB - derma_v8i1e74040_app1.png]

Multimedia Appendix 2
Survey questionnaire.
[DOCX File, 4134 KB - derma_v8i1e74040_app2.docx]

Multimedia Appendix 3
SAS results, part 1.
[DOCX File, 79 KB - derma_v8i1e74040_app3.docx]

Multimedia Appendix 4
SAS results, part 2.
[DOCX File, 15 KB - derma_v8i1e74040_app4.docx]

Multimedia Appendix 5
SAS results, part 3.
[DOCX File, 15 KB - derma_v8i1e74040_app5.docx]

Multimedia Appendix 6
Survey questions and answers.
[DOCX File, 32 KB - derma_v8i1e74040_app6.docx]

Multimedia Appendix 7
Comments from survey takers.
[DOCX File, 7 KB - derma_v8i1e74040_app7.docx]

References

1. ChatGPT. OpenAI. 2024. URL: https://openai.com/chatgpt [accessed 2025-11-17]
2. Chaudhari S, Aggarwal P, Murahari V, et al. RLHF deciphered: a critical analysis of reinforcement learning from human feedback for LLMs. ArXiv. Preprint posted online on Apr 12, 2024. [doi: 10.48550/arXiv.2404.08555]

3. Blank IA. What are large language models supposed to model? Trends Cogn Sci (Regul Ed) 2023 Nov;27(11):987-989. [doi: 10.1016/j.tics.2023.08.006]
4. Berberette E, Hutchins J, Sadovnik A. Redefining “hallucination” in LLMs: towards a psychology-informed framework for mitigating misinformation. ArXiv. Preprint posted online on Feb 1, 2024. [doi: 10.48550/arXiv.2402.01769]
5. DermGPT. 2024. URL: https://dermgpt.com [accessed 2025-11-17]
6. Kamnagar R. AI-based DermGPT focuses on improved clinic productivity. Pract Dermatol. 2024. URL: https://practicaldermatology.com/topics/practice-management/ai-based-dermgpt-focuses-on-improved-clinic-productivity/23990/ [accessed 2025-11-17]
7. 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]
8. He Z, Bhasuran B, Jin Q, et al. Quality of answers of generative large language models versus peer users for interpreting laboratory test results for lay patients: evaluation study. Journal of Medical Internet Research. Preprint posted online on Apr 17, 2024. [doi: 10.2196/preprints.56655]
9. Iannantuono GM, Bracken-Clarke D, Karzai F, Choo-Wosoba H, Gulley JL, Floudas CS. Comparison of large language models in answering immuno-oncology questions: a cross-sectional study. Oncologist 2024 May 3;29(5):407-414. [doi: 10.1093/oncolo/oyae009] [Medline: 38309720]
10. Fernández-Pichel M, Pichel JC, Losada DE. Evaluating search engines and large language models for answering health questions. NPJ Digit Med 2025 Mar 10;8(1):153. [doi: 10.1038/s41746-025-01546-w] [Medline: 40065094]

Abbreviations

LLM: large language model

Edited by R Dellavalle; submitted 16.03.25; peer-reviewed by ARD Luzuriaga, B Tasci; revised version received 15.10.25; accepted 02.11.25; published 04.12.25.

Please cite as:
Patel AB, Driscoll W, Lee CH, Zachary C, Golbari NM, Smith J
Evaluating Artificial Intelligence Models in Dermatology: Comparative Analysis
JMIR Dermatol 2025;8:e74040
URL: https://derma.jmir.org/2025/1/e74040
doi:10.2196/74040

Evaluating the Readability of Pediatric Neurocutaneous Syndromes–Related Patient Education Material Created by a Custom GPT With Retrieval Augmentation

Nneka Ede[1,2], MSE, MD; Robyn Okereke[3], MD

[1]Department of Biomedical Engineering, Cockrell School of Engineering, The University of Texas at Austin, 1500 Red River Street, Austin, TX, United States

[2]Dell Medical School, The University of Texas at Austin, Austin, TX, United States

[3]Department of Dermatology, Oregon Health and Sciences University, Portland, OR, United States

Corresponding Author:
Nneka Ede, MSE, MD
Department of Biomedical Engineering, Cockrell School of Engineering, The University of Texas at Austin, 1500 Red River Street, Austin, TX, United States

Abstract

In our study, we developed a GPT assistant with a custom knowledge base for neurocutaneous diseases, tested its ability to answer common patient questions, and showed that a GPT using retrieval augmentation generation can improve the readability of patient educational material without being prompted for a specific reading level.

(JMIR Dermatol 2025;8:e59054) doi:10.2196/59054

KEYWORDS

ChatGPT; large language model; LLMs; natural language processing; NLP; machine learning; artificial intelligence; generative AI; application programming interface; API; OpenAI; neurocutaneous syndromes; cutaneous; skin; dermatology; patient education; educational; GPT assistant; custom GPT; readability; answer; response; health education

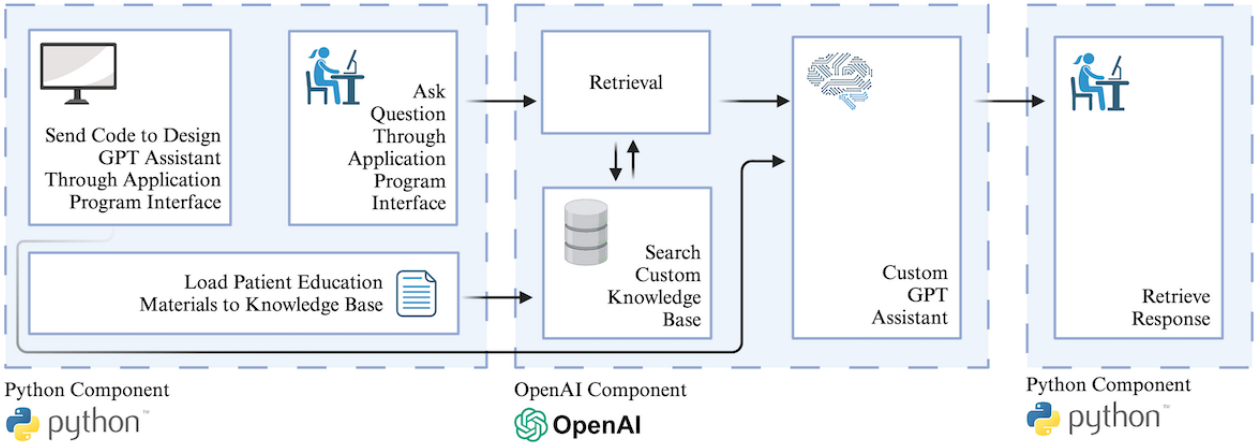
Introduction

Children with rare diseases and their families often face the challenge of understanding information regarding such diseases, and educational material is often written above the American Medical Association's recommended sixth-grade level [1,2]. GPTs can create patient education materials, but their readability often exceeds readers' comprehension levels [3-5]. GPT assistants are custom GPTs that can use retrieval augmentation generation (RAG) to access specific knowledge [6]. This study aims to evaluate a GPT assistant's ability to provide readable patient information on pediatric neurocutaneous syndromes in comparison to ChatGPT-4.

Methods

A GPT assistant was developed by using Python and OpenAI's application program interface (API; Figure 1). It was not programmed to answer questions at a specific reading level. Clinician and patient educational materials on four neurocutaneous diseases—tuberous sclerosis complex, neurofibromatosis type 1, neurofibromatosis type 2, and Sturge-Weber syndrome—were integrated into the configuration, with readability ranging from the eighth-grade level to the collegiate level, including sources like UpToDate and Johns Hopkins Medicine.

Figure 1. Flow diagram of the creation of the GPT assistant and how it functions. This figure was created in BioRender [7].



Five frequently asked patient and caregiver questions surrounding etiology, diagnosis, and management for each of the four diseases were asked to the GPT assistant, with and without a prompt for a response at a sixth-grade reading level (Multimedia Appendix 1). This process was repeated in ChatGPT-4. To minimize overoptimization of the models as questions were asked, no data were cached between API requests, and chat history and training were disabled. Readability was assessed by averaging the following eight readability formulas: Automated Readability Index, Flesch Reading Ease Formula, Gunning Fog Index, Flesch-Kincaid Grade Level Formula, Coleman-Liau Index, SMOG (Simple Measure of Gobbledygook) Index, Linsear Write Formula, and FORCAST Readability Formula (Multimedia Appendix 2) [8]. Two-tailed *t* tests and an ANOVA were used for comparison. Response accuracy was assessed via the OpenFactCheck Python package [9] and then confirmed by the authors (Multimedia Appendix 3).

Results

The overall average reading level of information generated without any specific prompting for a reading level was 11.4 (SD 2.04) for the custom GPT assistant and 15.41 (SD 2.0) for ChatGPT-4 (Table 1), revealing that the use of a GPT assistant with a knowledge base of patient educational material improved readability by approximately 4 reading levels (t_{35}=−6.02; P<.001). When prompted to answer questions at a sixth-grade reading level, the custom GPT assistant and ChatGPT-4 had average reading levels of 8.8 (SD 0.83) and 9.5 (SD 1.28), respectively, revealing a 0.7 difference in reading level (t_{38}=−2.05; P=.047). The combined use of a GPT assistant and reading level prompt resulted in the best performance ($F_{3,73}$=61.74; P<.001; Multimedia Appendix 4).

Table . Average of readability scores for responses generated by the custom GPT assistant without a prompt for reading level, by ChatGPT-4 without a prompt for reading level, by the custom GPT assistant with a prompt for a sixth-grade reading level, and by ChatGPT-4 with a prompt for a sixth-grade reading level. The average reading grade level is an average of 8 common readability formulas.

Metrics	Custom GPT assistant	ChatGPT-4	Custom GPT assistant + prompted reading level	ChatGPT-4 + prompted reading level
Average reading grade level, mean (SD)	11.40 (2.04)	15.41 (2.0)	8.80 (0.83)	9.50 (1.28)
Automated Readability Index, mean (SD)	11.68 (2.54)	16.60 (2.45)	9.30 (1.00)	10.04 (1.62)
Flesch Reading Ease, mean (SD)	49.95 (14.84)	23.41 (12.47)	74.65 (5.39)	69.70 (7.34)
Gunning Fog Index, mean (SD)	13.93 (2.51)	18.41 (2.57)	10.23 (1.05)	10.84 (1.72)
Flesch-Kincaid Grade Level, mean (SD)	10.79 (2.32)	15.32 (2.20)	7.56 (0.98)	8.24 (1.45)
Coleman-Liau Index, mean (SD)	11.70 (2.63)	16.07 (2.14)	8.21 (0.94)	9.21 (1.29)
SMOG[a] Index, mean (SD)	10.09 (1.84)	13.37 (1.77)	6.69 (0.92)	7.38 (1.37)
Linsear Write score, mean (SD)	11.88 (2.68)	16.09 (2.73)	10.35 (1.49)	10.83 (2.07)
FORCAST readability	10.85 (1.18)	12.10 (0.74)	8.99 (0.46)	9.28 (0.76)

[a]SMOG: Simple Measure of Gobbledygook.

Discussion

The GPT assistant provided more readable responses about pediatric neurocutaneous diseases than ChatGPT-4 when no reading level was specified and when a reading level was prompted. Using the GPT assistant with a reading level prompt achieved the best results, suggesting that when a GPT assistant accesses educational materials with a variety of reading levels, readability improves. However, specifying a reading level in ChatGPT-4 resulted in better performance than the GPT assistant without a reading level prompt. Furthermore, there is only a small difference in reading level between models when a comprehension level is prompted, indicating that this action enhances readability, though this is not always intuitive for users. GPT assistants provide another avenue for improving readability, with or without a reading level prompt.

This study also indicated that the caliber of data used when designing a GPT directly influences model results. Poor data quality affects machine learning models' performance. In the context of readability, poor quality equates to resources with high reading levels. RAG in a GPT assistant allows access to materials with lower reading levels, thereby improving responses without the need for specific prompts. Recent research has determined that RAG improves patient information accuracy and reduces GPT hallucinations; our results show that it can also improve readability [10,11]. If all documents were at a sixth-grade level, readability may improve further; however, more research is needed to determine this.

GPT assistants have the potential to give pediatric dermatology patients and their families another modality for learning and asking questions about the conditions they face—one that is more understandable than ChatGPT alone. Furthermore, GPT assistants may enable clinicians to fine-tune information produced by a GPT specifically for their patient population. GPT assistants with a knowledge base incorporating easy-to-read material can better aid physicians in providing patient- and caregiver-level information, with or without a specific reading level prompt, when compared to ChatGPT-4 alone. A limitation of this study is the limited number of questions assessed. However, this study provides a foundation for larger-scale future research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prompts input into the GPT assistant and ChatGPT.
[DOCX File, 7 KB - derma_v8i1e59054_app1.docx]

Multimedia Appendix 2
Readability formula definitions.

[DOCX File, 13 KB - derma_v8i1e59054_app2.docx]

Multimedia Appendix 3
Supplemental methods and results for response accuracy.
[DOCX File, 14 KB - derma_v8i1e59054_app3.docx]

Multimedia Appendix 4
ANOVA results.
[DOCX File, 158 KB - derma_v8i1e59054_app4.docx]

References

1. D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med 2001 Jul;155(7):807-812. [doi: 10.1001/archpedi.155.7.807] [Medline: 11434848]
2. COST Action BM1105, Badiu C, Bonomi M, et al. Developing and evaluating rare disease educational materials co-created by expert clinicians and patients: the paradigm of congenital hypogonadotropic hypogonadism. Orphanet J Rare Dis 2017 Mar 20;12(1):57. [doi: 10.1186/s13023-017-0608-2] [Medline: 28320476]
3. Ayre J, Mac O, McCaffery K, et al. New frontiers in health literacy: using ChatGPT to simplify health information for people in the community. J Gen Intern Med 2024 Mar;39(4):573-577. [doi: 10.1007/s11606-023-08469-w] [Medline: 37940756]
4. Jin JQ, Dobry AS. ChatGPT for healthcare providers and patients: practical implications within dermatology. J Am Acad Dermatol 2023 Oct;89(4):870-871. [doi: 10.1016/j.jaad.2023.05.081] [Medline: 37315798]
5. Golan R, Ripps SJ, Reddy R, et al. ChatGPT's ability to assess quality and readability of online medical information: evidence from a cross-sectional study. Cureus 2023 Jul 20;15(7):e42214. [doi: 10.7759/cureus.42214] [Medline: 37484787]
6. Assistants API overview. OpenAI. URL: https://platform.openai.com/docs/assistants/overview?context=with-streaming [accessed 2025-05-06]
7. Ede N. Figure 1. Flow diagram of the creation of the GPT assistant and how it functions. BioRender. 2025. URL: https://app.biorender.com/citation/685d9abbd45f2d774558aed4 [accessed 2025-06-26]
8. Readability formulas. URL: https://readabilityformulas.com/ [accessed 2025-05-05]
9. Iqbal H. Hasaniqbal777/openfactcheck: v0.3.0. Zenodo. 2024 Aug 22. URL: https://zenodo.org/records/13358665 [accessed 2025-06-17]
10. Miao J, Thongprayoon C, Suppadungsuk S, Garcia Valencia OA, Cheungpasitporn W. Integrating retrieval-augmented generation with large language models in nephrology: advancing practical applications. Medicina (Kaunas) 2024 Mar 8;60(3):445. [doi: 10.3390/medicina60030445] [Medline: 38541171]
11. Yang R, Ning Y, Keppo E, et al. Retrieval-augmented generation for generative artificial intelligence in health care. Npj Health Syst 2025 Jan 25;2:2. [doi: 10.1038/s44401-024-00004-1]

Abbreviations

API: application program interface
RAG: retrieval augmentation generation
SMOG: Simple Measure of Gobbledygook

Edited by I Brooks; submitted 31.03.24; peer-reviewed by D Chrimes, H Khan, H Mondal, TA El-Hafeez; revised version received 07.05.25; accepted 28.05.25; published 16.07.25.

Please cite as:
Ede N, Okereke R
Evaluating the Readability of Pediatric Neurocutaneous Syndromes–Related Patient Education Material Created by a Custom GPT With Retrieval Augmentation
JMIR Dermatol 2025;8:e59054
URL: https://derma.jmir.org/2025/1/e59054
doi:10.2196/59054

Patterns of Public Interest in Lipomas and Lipoma-Removal Procedures: Google Trends Analysis

Keenan Duggal, BA

NYU Langone Health, 550 1st Ave, New York, NY, United States

Corresponding Author:
Keenan Duggal, BA
NYU Langone Health, 550 1st Ave, New York, NY, United States

Abstract

Background: Lipomas are benign tumors composed of encapsulated adipocytes. Although relatively common, uncertainty remains about the population-level prevalence, the etiology, and the degree of public interest in lipomas and associated removal procedures.

Objective: The spatiotemporal patterns of public interest in lipomas and lipoma removal procedures were characterized.

Methods: Google Trends data that report the relative search volume (RSV) of Google queries pertaining to lipomas and their removal procedures at national and international levels were analyzed. To contextualize these trends, the RSV for lipomas was compared to that of several other common dermatological conditions in the United States.

Results: In the United States, lipomas have consistently generated lower levels of public interest than other common dermatological conditions, but interest in the condition has been rising since the mid-2010s. Across the world, public interest in lipomas appears to be the highest in pockets of Eastern Europe, whereas in the United States, relative interest has been higher in Midwestern and Southern states. In addition, the interest in lipoma removal procedures has risen steadily from 2004 to the present, with particularly high RSVs coming from Southwestern states

Conclusions: Dermatologists and plastic surgeons should be aware of the increasing public interest in lipomas and lipoma-removal procedures. Clinical awareness is especially important in states with an elevated interest in lipomas and their associated removal procedures.

(JMIR Dermatol 2025;8:e62993) doi:10.2196/62993

KEYWORDS

lipoma; fatty tumor; adipocyte; public interest; Google Trends

Introduction

Lipomas are benign adipocytic tumors [1,2]. While lipomas most commonly present on patients as solitary entities, it has been estimated that multiple lipomas develop in approximately 5% of affected patients [3]. In such cases, multiple fatty tumors can result from several rare medical disorders including familial multiple lipomatosis, Dercum disease, Madelung disease, and Gardner syndrome [1,4]. In general, the etiology of lipomas is unknown, but it has been postulated that soft-tissue injuries and genetic mutations might both play a role in their formation [5-8].

Lipomas can be unsightly to patients, can cause pain depending on their location [9,10], and can occasionally grow to sufficiently large sizes to disrupt quality of life [11,12]. Additionally, lipomas can have clinical presentations similar to more serious malignant liposarcoma tumors [13,14]. For these reasons, patients sometimes elect to have their lipoma(s) removed and biopsied. Generally, removal proceeds through intralesional injections of lipolytic agents, localized liposuction, or surgical excision [1,4].

While lipomas are estimated to affect around 1% of the population [1], the precise prevalence is difficult to estimate due to the elective nature of treatment and the nonexistence of universal screening modalities [15]. Compounding this uncertainty, very little is known about the degree of public awareness of lipomas and interest in associated treatments. These knowledge deficits restrict healthcare providers' holistic understanding of the condition [16], which in turn might impede their ability to communicate with affected patients. To mitigate these knowledge gaps, Google Trends data were analyzed for the topic "Lipoma" and the query "Lipoma Removal" across time and several geographic scales. Because Google captures the vast majority of search engine traffic, the data provide a reasonable proxy for the totality of online public interest [17].

Methods

To characterize public interest in lipomas, Google Trends data describing the national (United States) and international relative search volume (RSV) for the medical condition "Lipoma" between January 1, 2004 and May 21, 2024 were downloaded.

Google Trends RSV data range from 0 to 100 and describe the relative frequency of a Google search, normalized to account for underlying spatiotemporal variation in internet usage. By using the topic "Lipoma," the reported RSV reflects an aggregate value for all of the searched keywords/terms that "share the same concept in any language" [18]. In addition to temporal patterns, geographic patterns in lipoma-related searches were also assessed.

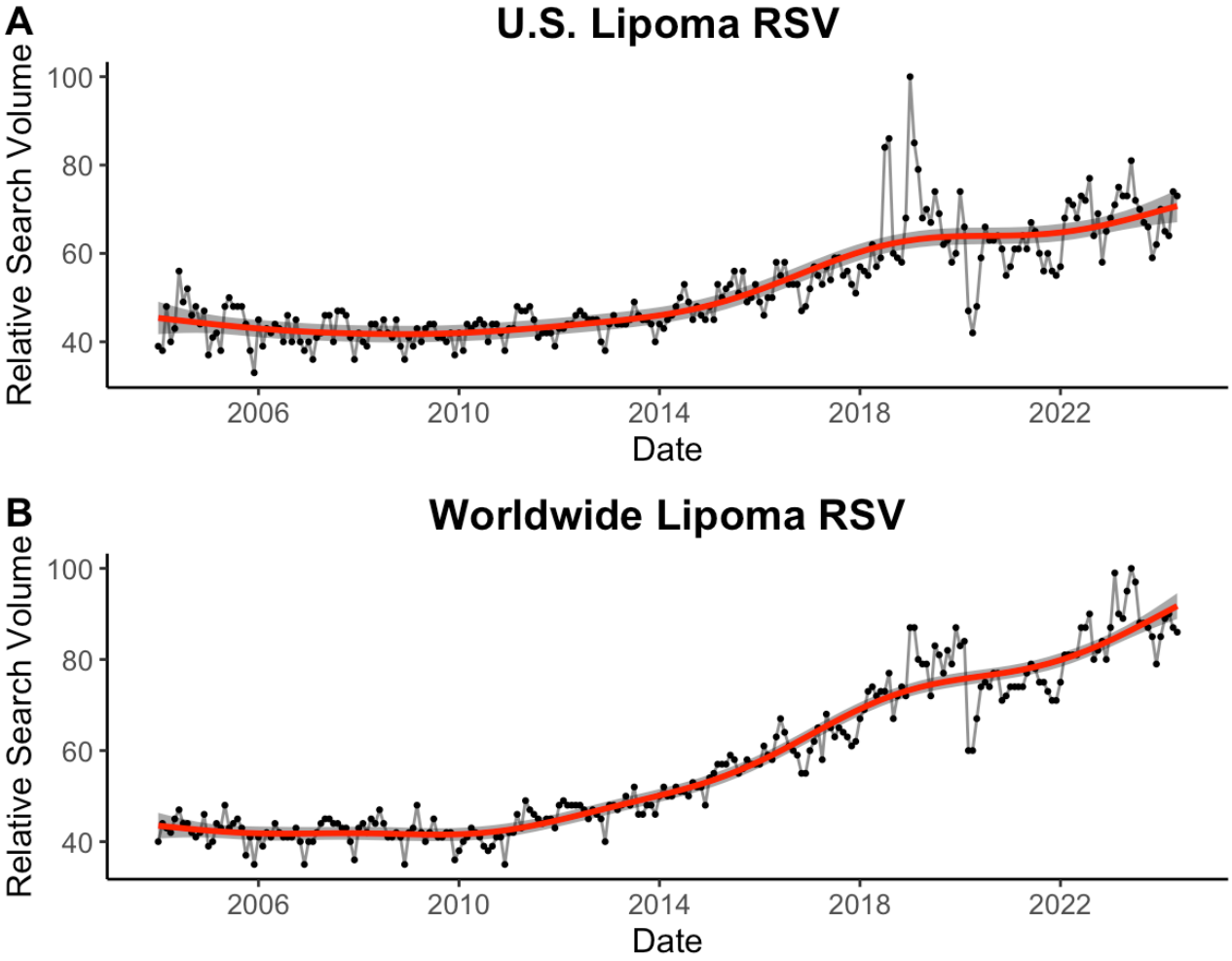
To situate public interest in "lipomas" within its broader dermatological context, the Google Trends "compare" feature was used to juxtapose the RSV of the topic "lipoma" against that of 4 other common dermatological conditions (queried equivalently as topics), which were selected to capture a range of prevalent skin conditions that can prompt dermatological consultation: atopic dermatitis, psoriasis, skin cancer, and rosacea [19,20].

To track public interest pertaining specifically to the "treatment" of lipomas, temporal and geographic trends were assessed using the search query "Lipoma removal." I first verified that this query captures relevant public interest by analyzing the lists of "Related Topics" and "Related Queries." For this search, two temporal windows were used for geographic comparisons: a long-term summary (January 1, 2004 to May 21, 2024) and a short-term summary (May 21, 2019 to May 21, 2024).

Results

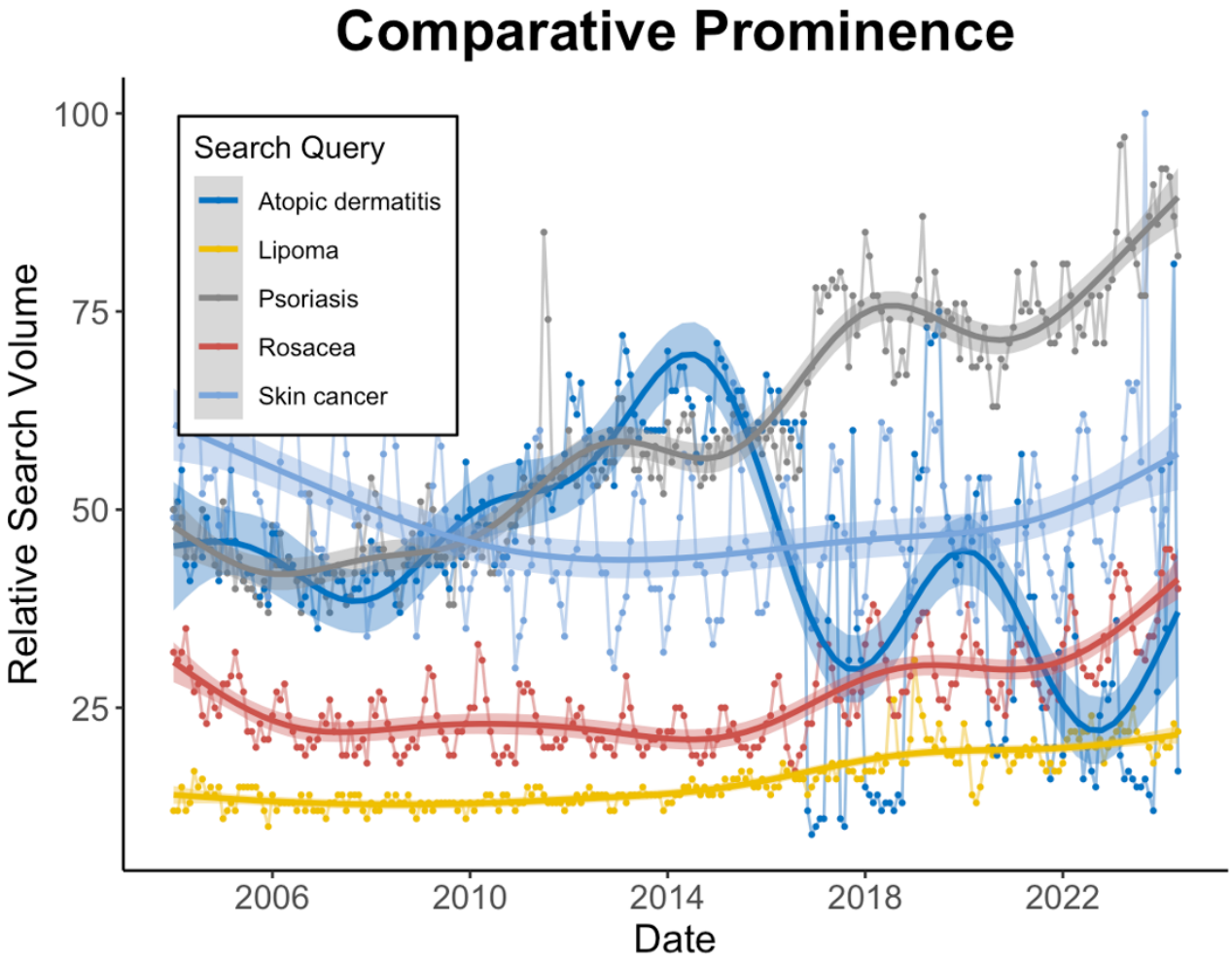
Over the past 20 years, there has been a substantial increase in public interest in lipomas in the United States and across the globe (Figure 1). Evanescent spikes in these time series are likely the product of transient popular cultural coverage related to the condition and/or significantly disruptive societal events (eg, the COVID-19 pandemic).

Figure 1. Temporal trends in public interest in lipomas. The monthly relative search volume (RSV) of lipoma-related search queries made on Google (A) in the United States and (B) around the world between January 1, 2004, and May 21, 2024, are plotted with generalized additive models and associated confidence intervals overlaid.



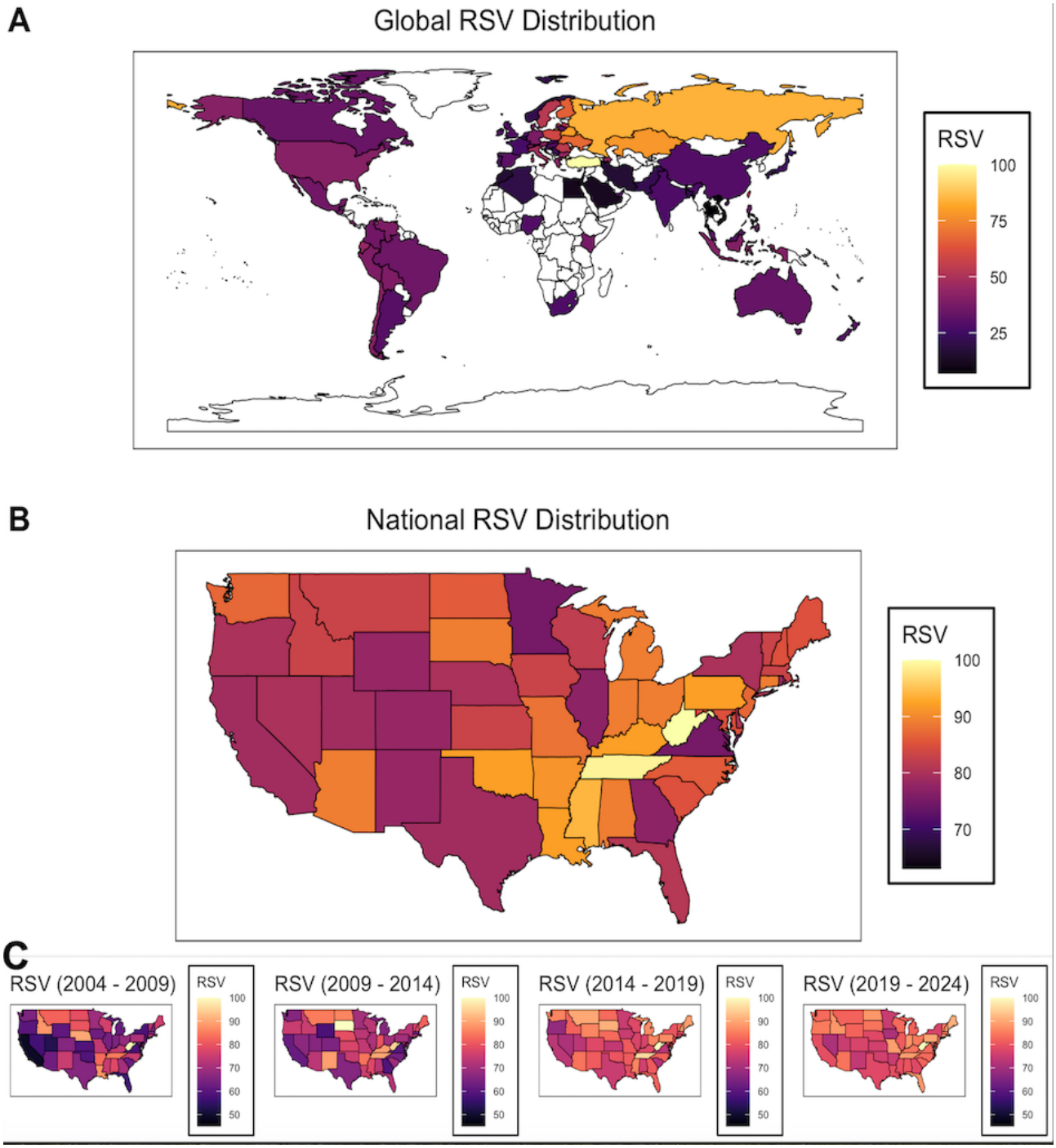
Over the past 20 years, public interest in lipomas was consistently lower than that in 4 other common dermatological conditions: atopic dermatitis, psoriasis, skin cancer, and rosacea (Figure 2).

Figure 2. Comparative public interest in dermatological conditions. The Google Trends relative search volumes (RSVs) between January 1, 2004 and May 21, 2024 for 5 dermatological conditions are plotted on the same scale. Generalized additive models and associated confidence intervals describe smoothed trends over time.



To characterize geographic regions with relatively elevated long-term interest in lipomas, the 20-year average country-level RSVs for the lipoma topic were examined. Lipoma-associated interest was the highest in pockets of Eastern Europe and Asia (Figure 3A). Specifically, the countries with the highest RSV for the lipoma condition were Turkey, Russia, Belarus, Kazakhstan, and Ukraine, with RSVs of 100, 84, 80, 78, and 68, respectively. Comparatively, the United States had an RSV of 41 for the same query.

Figure 3. Geographic patterns of public interest in lipomas. (A-B) The relative search volume (RSV) of the Google search topic "lipoma" was compared across (A) countries and (B) contiguous states in the United States and averaged over the temporal period of January 1, 2004 to May 21, 2024. (C-F) The geographic distribution of "lipoma" RSV was compared in 5-year intervals in the contiguous states of the United States.

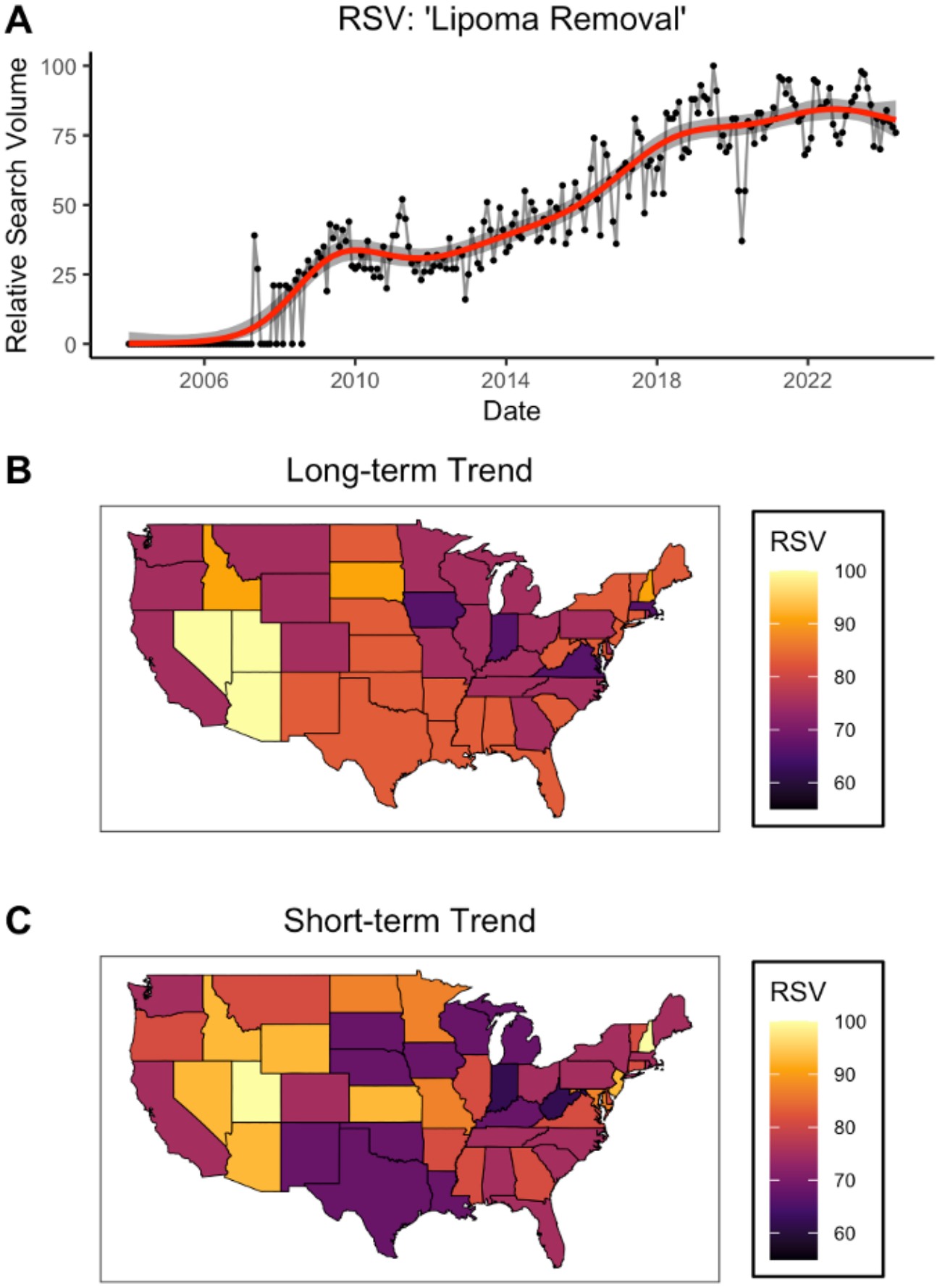


Next, average state-level RSVs over the same period were assessed in the United States. Here, public interest in lipomas was the highest in Southern and Midwestern states (Figure 3B). Specifically, West Virginia, Tennessee, Mississippi, Kentucky, Pennsylvania, Louisiana, and Oklahoma each had RSVs greater than 92.

Finally, to evaluate if and how the geographic patterns of interest have changed over time in the United States, the average RSV of Google's lipoma topic was assessed over 5-year increments from January 1, 2004, to January 1, 2024. While the spatial distribution of RSV remained relatively stable, the regional homogeneity generally increased over time (Figure 3C-F).

There was a clear and consistent increase in Google searches containing both of the terms "lipoma" and "removal" from the mid-2000s onwards, interrupted only by a transient decline during the COVID-19 pandemic (Figure 4A). On average, interest in lipoma removal has been relatively higher in non-coastal Western states (Figure 4B). This long-term spatial pattern is largely consistent with that of the most recent 5-year period (Figure 4C).

Figure 4. Public interest in lipoma removal (A) 20-year trend of monthly elative search volume (RSV) values for the Google search query “lipoma removal”. (B) Geographic distribution over the period of January 01, 2004 to May 21, 2024. (C) Geographic distribution over the period of May 21, 2019 to May 21, 2024.



Discussion

This study leveraged Google Trends data to show that public interest in lipomas and lipoma-associated topics, while comparatively lower than that for other common dermatological conditions, has increased both in the United States and internationally over the past 20 years. Additionally, the interest in lipoma removal procedures has increased substantially over the past 20 years in the United States. While the descriptive nature of the analyses and claims presented here makes Google Trends a reasonable and sufficient source of data, the limitations inherent to Google Trends data must still be recognized (for instance that keyword selection can introduce bias, and that the RSV does not reflect the total search volume) [16,18,21].

While increasing interest in lipomas and lipoma-removal procedures could be explained by an increasing prevalence of the condition, consistent with increasing population-level adiposity [22], it is more likely that it is the result of increasing online health information-seeking behavior and increasing public health literacy [23,24]. Irrespective of the cause, increasing public interest in lipomas indicates a need for clinicians to be vigilant, prepared to encounter the condition, and comfortable with educating their patients about it.

The decrease in public interest in lipomas and associated "treatments" observed during the COVID-19 pandemic is reflective of a larger phenomenon wherein elective/aesthetic surgeries declined in popularity [25]. While it is logical that online interest in lipomas stalled during a period when other health concerns became more salient, it is surprising that there was no dramatic surge in interest in lipomas and lipoma-removal procedures, which was observed for similar cosmetic conditions and procedures, following the relaxation of public health guidelines [26]. Moving forward, it will be valuable to continue to monitor public interest relating to lipomas and other dermatological conditions to assess future trends and ultimately to inform healthcare practitioners of relevant patient interest [16,21].

Data Availability

All data associated with this manuscript are publicly available on https://trends.google.com/trends/.

Conflicts of Interest

None declared.

References

1. Charifa A, Azmat CE, Badri T. Lipoma Pathology: StatPearls Publishing; 2024. URL: http://www.ncbi.nlm.nih.gov/books/NBK482343/ [accessed 2024-05-21]
2. Nakao K, Onishi F, Kiyama M, Minabe T. The growth factors of subcutaneous benign lipoma: consideration from anatomical position of occurrence. Plast Reconstr Surg Glob Open 2022 Sep;10(9):e4524. [doi: 10.1097/GOX.0000000000004524] [Medline: 36168603]
3. Mejía Granados DM, de Baptista MB, Bonadia LC, Bertuzzo CS, Steiner CE. Clinical and molecular investigation of familial multiple lipomatosis: variants in the HMGA2 gene. Clin Cosmet Investig Dermatol 2020;13:1-10. [doi: 10.2147/CCID.S213139] [Medline: 32021365]
4. Kolb L, Yarrarapu SNS, Ameer MA, Rosario-Collazo JA. Lipoma: StatPearls Publishing; 2024. URL: http://www.ncbi.nlm.nih.gov/books/NBK507906/ [accessed 2024-05-21]
5. Aust MC, Spies M, Kall S, Jokuszies A, Gohritz A, Vogt P. Posttraumatic lipoma: fact or fiction? Skinmed 2007;6(6):266-270. [doi: 10.1111/j.1540-9740.2007.06361.x] [Medline: 17975353]
6. Panagopoulos I, Gorunova L, Agostini A, Lobmaier I, Bjerkehagen B, Heim S. Fusion of the HMGA2 and C9orf92 genes in myolipoma with t(9;12)(p22;q14). Diagn Pathol 2016 Feb 9;11:22. [doi: 10.1186/s13000-016-0472-8] [Medline: 26857357]
7. Genuardi M, Klutz M, Devriendt K, Caruso D, Stirpe M, Lohmann DR. Multiple lipomas linked to an RB1 gene mutation in a large pedigree with low penetrance retinoblastoma. Eur J Hum Genet 2001 Sep;9(9):690-694. [doi: 10.1038/sj.ejhg.5200694] [Medline: 11571558]
8. Kanojia D, Dakle P, Mayakonda A, et al. Identification of somatic alterations in lipoma using whole exome sequencing. Sci Rep 2019 Oct 7;9(1):14370. [doi: 10.1038/s41598-019-50805-w] [Medline: 31591430]
9. Al-Mnayyis A, Al Sharie S, Araydah M, Talafha M, Haddad F. Parosteal lipoma of the forearm. Medicine (Balt) 2021;100(46):e27876. [doi: 10.1097/MD.0000000000027876]
10. Pirri C, Stecco C, De Caro R, Foti C, Özçakar L. Radiating upper limb pain due to a large subcutaneous lipoma: fascial sono-palpation. Pain Med 2020 Dec 25;21(12):3721-3723. [doi: 10.1093/pm/pnaa081] [Medline: 32299098]
11. Allen B, Rader C, Babigian A. Giant lipomas of the upper extremity. Can J Plast Surg 2007;15(3):141-144. [doi: 10.1177/229255030701500308] [Medline: 19554145]
12. Silistreli OK, Durmuş EU, Ulusal BG, Oztan Y, Görgü M. What should be the treatment modality in giant cutaneous lipomas? Review of the literature and report of 4 cases. Br J Plast Surg 2005 Apr;58(3):394-398. [doi: 10.1016/j.bjps.2004.09.005] [Medline: 15780237]

13. Natella R, Varriano G, Brunese MC, et al. Increasing differential diagnosis between lipoma and liposarcoma through radiomics: a narrative review. Explor Target Antitumor Ther 2023;4(3):498-510. [doi: 10.37349/etat.2023.00147] [Medline: 37455823]
14. Enzinger FM, Winslow DJ. Liposarcoma. Virchows Arch Path Anat 1962;335(4):367-388. [doi: 10.1007/BF00957030]
15. Campanacci M. Lipomas. In: Campanacci M, editor. Bone and Soft Tissue Tumors: Clinical Features, Imaging, Pathology and Treatment: Springer; 1999:983-999. [doi: 10.1007/978-3-7091-3846-5_68]
16. Tijerina JD, Morrison SD, Nolan IT, Vail DG, Nazerali R, Lee GK. Google Trends as a tool for evaluating public interest in facial cosmetic procedures. Aesthet Surg J 2019 Jul 12;39(8):908-918. [doi: 10.1093/asj/sjy267] [Medline: 30304356]
17. Global search engine desktop market share 2024. Statista. URL: https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ [accessed 2024-05-21]
18. 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]
19. Barankin B, Rosacea GL. Two common oculocutaneous disorders. Can Fam Physician 2002;48(4):721-724. [Medline: 12046367]
20. Landis ET, Davis SA, Taheri A, Feldman SR. Top dermatologic diagnoses by age. Dermatol Online J 2014 Apr 16;20(4):22368. [Medline: 24746305]
21. Tijerina JD, Morrison SD, Vail DG, Lee GK, Nazerali R. The utility of Google Trends data for analyzing public interest in breast procedures. Ann Plast Surg 2019 May;82(5S Suppl 4):S325-S331. [doi: 10.1097/SAP.0000000000001806] [Medline: 30870175]
22. Wang Y, Beydoun MA, Min J, Xue H, Kaminsky LA, Cheskin LJ. Has the prevalence of overweight, obesity and central obesity levelled off in the United States? Trends, patterns, disparities, and future projections for the obesity epidemic. Int J Epidemiol 2020 Jun 1;49(3):810-823. [doi: 10.1093/ije/dyz273] [Medline: 32016289]
23. Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. Cog Soc Sci 2017 Jan 1;3(1):1302785. [doi: 10.1080/23311886.2017.1302785]
24. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. Health Promot Int 2018 Oct 1;33(5):901-911. [doi: 10.1093/heapro/dax015] [Medline: 28369557]
25. COVIDSurg Collaborative. Elective surgery cancellations due to the COVID-19 pandemic: global predictive modelling to inform surgical recovery plans. Br J Surg 2020 Oct;107(11):1440-1449. [doi: 10.1002/bjs.11746] [Medline: 32395848]
26. Aktas EH, Balci UD, Karacaoglu E. COVID pandemic aftermath: changing dynamics on cosmetic and aesthetic surgery demands. Aesthet Plast Surg 2023 Aug;47(4):1658-1665. [doi: 10.1007/s00266-022-03231-9] [Medline: 36715726]

Abbreviations

RSV: relative search volume

Edited by R Dellavalle; submitted 06.06.24; peer-reviewed by A Edalatpour, Q Tan; revised version received 09.12.24; accepted 10.12.24; published 17.01.25.

Please cite as:
Duggal K
Patterns of Public Interest in Lipomas and Lipoma-Removal Procedures: Google Trends Analysis
JMIR Dermatol 2025;8:e62993
URL: https://derma.jmir.org/2025/1/e62993
doi:10.2196/62993

Dermal Filler Use in Patients Undergoing Chemotherapy and Radiation Therapy

Mia Panlilio[1], BA; Elizabeth Tchernogorova[2], BA; Binh Minh Nguyen[1], BA; Sana Khan[1], BS; Leslie Torgerson[1], MD

[1]College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Road, Parker, CO, United States

[2]College of Osteopathic Medicine, Rocky Vista University, Ivins, UT, United States

Corresponding Author:
Mia Panlilio, BA
College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Road, Parker, CO, United States

Abstract

Dermal fillers have gained increasing popularity for their ability to enhance facial symmetry, restore volume, and improve skin texture. However, their use in patients with cancer undergoing active chemotherapy and radiation therapy poses unique challenges, as these treatments can alter both the safety profile and efficacy of filler procedures. Chemotherapy can interfere with normal wound healing and immune responses, warranting a more cautious and individualized approach when considering dermal fillers in this population. Although rare, dermal fillers have been associated with adverse outcomes in a limited number of diseases, including cellulitis, autoimmune/inflammatory syndrome induced by adjuvants, and a possible predisposition to malignancy. Other effects include localized inflammatory, systemic hypersensitivity, and delayed granulomatous formation, and these could be more severe in patients undergoing antineoplastic therapy. Furthermore, chemotherapy is often paired with adjuvant radiation therapy in cancer treatment, making it important to note the potential changes radiation can have on the skin. More research is needed to examine the direct interactions of chemotherapy and radiation on various filler materials injected within the skin, and how these can alter one's risk of adverse effects. The lack of research on this topic further emphasizes that clinicians should thoroughly educate patients receiving chemotherapy and adjuvant radiation treatment about the heightened potential risks associated with dermal filler injections and treatment regimens should be planned accordingly to minimize any adverse events.

(JMIR Dermatol 2025;8:e76898) doi:10.2196/76898

KEYWORDS

dermal; filler; chemotherapy; cosmetic; aesthetic; radiation

Introduction

Dermal fillers, which are minimally invasive cosmetic treatments, have gained widespread popularity for their ability to restore volume, soften wrinkles, and enhance facial contours. Various filler types are available, including hyaluronic acid-based products, calcium hydroxylapatite, poly-L-lactic acid, polymethyl methacrylate, and silicone, each offering unique properties and applications [1,2].

Hyaluronic acid-based fillers are one of the most popular types that are used for fine lines and wrinkles, lasting about 6 to 12 months before naturally getting absorbed by the body [1]. Calcium hydroxylapatite fillers are a thicker consistency product used for volume restoration in deeper lines and wrinkles, and they typically last 12 months or longer [1]. Poly-L-lactic acid fillers are also popular, and are used to treat deeper facial wrinkles, with results lasting more than 2 years [1]. Polymethyl methacrylate is a physical filler that comes in the form of a microsphere, or tiny ball, that remains under the skin indefinitely for long term support, and it also contains collagen which provides structure and firmness [1]. Silicone fillers have been used in some parts of the world for volume restoration; however, silicone filler products currently are not approved by the Food and Drug Administration (FDA) in the United States and have fallen out of favor due to the high occurrence of filler migration, swelling and tissue damage, and severe inflammatory reactions to the material [1].

The rising use of fillers can be attributed to their effectiveness, convenience, and the growing demand for non-surgical aesthetic enhancements. However, these treatments are not without risks. Contraindications include allergies, infections, and certain medical conditions that could heighten the risk of infection or foreign body reactions, such as immunosuppressive or autoimmune conditions. Of particular concern is the use of dermal fillers in patients actively undergoing chemotherapy and radiation therapy. Chemotherapy and radiation can compromise immune function and tissue repair mechanisms, potentially altering both the safety and effectiveness of filler procedures. These therapies may disrupt normal wound healing and immune surveillance, necessitating a more cautious and individualized approach when considering dermal filler use in oncology patients. This viewpoint aims to evaluate the current literature

on the safety and potential adverse effects of dermal filler use in patients undergoing chemotherapy and radiation therapy.

Adverse Effects of Fillers

With the increase in popularity and accessibility of aesthetic treatments, there has been significant growth in the research sector. In 2024, the United States ranked #1 for the most published articles on soft tissue filler injections globally [3]. However, there remains a significant lack of research for the effects of dermal filler injections on patients undergoing chemotherapy and radiation therapy. Dermal fillers can cause adverse effects such as localized inflammatory, systemic hypersensitivity, delayed granulomatous formation, cellulitis, and autoimmune/inflammatory syndrome induced by adjuvants (ASIA) [4-9]. Additionally, while generally considered safe, dermal fillers, particularly those composed of hyaluronic acid (HA), introduce exogenous compounds into the tissue environment [10]. Endogenous HA is highly concentrated in some malignant neoplasms and its interaction with the CD44 receptor is implicated in tumor growth and poor prognosis [10]. While there is theoretical concern that exogenous HA fragments may contribute to a pro-tumor microenvironment, direct evidence linking fillers to new malignancy is lacking.

Case Reports in Oncology Patients

While there has been minimal research conducted on the topic of dermal filler reactions related to chemotherapy and radiation therapy, there are a few case studies describing reactions of dermal fillers in relation to these treatments, as summarized in Table 1.

One case report was about a 53-year-old woman with advanced non-small cell lung cancer who received carboxymethyl cellulose and polycaprolactone microsphere injection in her anterior neck 15 months previously [11]. She developed nodules and hardened folds along her sternocleidomastoid muscle bilaterally while receiving a course of 5 nivolumab infusions [11]. These nodules and hardened folds were later determined to be foreign body reactions to the filler. The patient also developed autoimmune colitis during the nivolumab infusions and was administered systemic glucocorticoids, which resolved both the autoimmune colitis and foreign body reaction to the filler. This case exemplifies how patients with preexisting fillers have an increased risk of developing a foreign body reaction when initiating chemotherapy. Additionally, the administration of glucocorticoids and subsequent resolution of the foreign body reaction provides some support towards using glucocorticoids to treat these reactions in this patient population.

Another case report was about a woman in her 60s who received dermal filler 25 years prior to initiating chemotherapy; she developed painless facial nodules after 2 courses of ipilimumab treatment [12]. The nodules were excised and later confirmed to be a foreign body reaction to filler via histology. Like the previous case, this again highlights the increased risk of foreign body reactions in patients with preexisting fillers who are undergoing chemotherapy and shows the risk is still heightened years after filler injection.

In a case report published in 2019, a 52-year-old man with preexisting HA fillers in his cheeks, undergoing cetuximab and radiation for glossotonsillary malignancy, developed significant inflammation and edema at the filler sites within hours of his first cetuximab dose [5]. This acute reaction was distinct from typical cetuximab-induced rashes and subsided rapidly after the fillers were dissolved. This exemplifies a more atypical and rapid presentation of a chemotherapy-induced foreign body reaction to fillers and highlights how filler dissolution could be used to resolve the reaction.

Table . Summary of the presented case reports.

Patient characteristics	Filler type	Cancer therapy	Time between filler injection and reaction	Adverse events	Outcomes
53-year-old woman with NSCLC[a]	Carboxymethyl cellulose and polycaprolactone microsphere	Nivolumab	15 months	Bilateral nodules and indurated folds along the sternocleidomastoid muscles, later found to be a foreign body reaction to the filler	Patient also developed autoimmune colitis from nivolumab treatment and was treated with systemic glucocorticoids, which also resolved the filler foreign body reaction
61-year-old woman with superficial spreading melanoma in the left scapular region	Unspecified synthetic filler material	Ipilimumab	25 years	Painless granulomatous facial nodules found to be a foreign body reaction to the filler	All granulomatous nodules were excised
52-year-old man with glossotonsillary malignancy	Hyaluronic acid	Cetuximab and radiation therapy	2 hours	Inflammation and edema around filler sites	Acute reaction rapidly resolved following filler dissolution
43-year-old woman with chronic myeloid leukemia	Hyaluronic acid	Imatinib mesylate	Filler injected during chemotherapy treatment	None mentioned	Successful treatment

[a]NSCLC: non-small cell lung cancer.

One case report in Korea described "successful" HA filler injection in a patient actively undergoing imatinib mesylate treatment for chronic myeloid leukemia and was presented in 2019 [13]. However, researchers only followed the patient for 10 weeks after filler treatment, so possible reactions occurring after 10 weeks were not described. The lack of longer-term follow-up therefore makes this case a questionable example of a truly "successful" filler treatment.

Other case studies available mostly describe foreign body reactions occurring after longer time periods, such as the case report describing a reaction 25 years after dermal filler injection [12]. Additionally, there are no studies specifying filler reactions with certain types of chemotherapy agents. For example, imatinib mesylate is a tyrosine kinase inhibitor, whereas ipilimumab is a CTLA-4 inhibitor. The differences in each drug's mechanism of action could impact the type of a reaction a patient experiences based on the immunological pathways affected. This makes it difficult for physicians and scientists to determine the true long-term impacts and efficacy of dermal filler on patients undergoing chemotherapy treatment without further research on reactions between specific filler materials and chemotherapies of different drug classes.

Effects of Radiation on the Skin

In a prospective cohort study done in 2014, the effects of radiotherapy were studied to see what changes occur within the skin when radiation is used for the treatment of breast cancer [14]. After radiotherapy, the irradiated breast showed a notable decrease in skin hydration, an increase in skin pH, increase in pigmentation, and increase in cutaneous blood flow. Radiotherapy is also known to damage skin barrier function because it induces apoptosis and necrosis of epidermal cells, thus decreasing the production of natural moisturizing factors and intercellular lipids [6]. It also causes an alkalinization of the stratum corneum, which is a layer of the skin that favors bacterial and fungal proliferation, thus increasing the risk of infection [14]. When it comes to possible interactions radiation therapy can have in patients with existing fillers, there are no current clinical studies available.

Specific Patient Risks

Patients undergoing chemotherapy or those who are immunosuppressed face unique challenges when considering dermal fillers. Chemotherapy can significantly impact the immune system, increasing the risk of infections and other adverse effects. The American Society of Clinical Oncology (ASCO) and the Infectious Diseases Society of America (IDSA) highlight that patients with cancer-related immunosuppression are particularly vulnerable to infections due to neutropenia and other factors [15]. Neutropenia, a common side effect of chemotherapy, reduces the body's ability to fight infections, making patients more susceptible to bacterial and fungal infections. This increased susceptibility can lead to adverse effects such as cellulitis, delayed wound healing, and atypical inflammatory reactions following dermal filler injections [15].

A study presented at the 2024 ASCO Annual Meeting assessed the safety and quality of life in oncology patients receiving dermal fillers during active cancer therapy. The study found that fillers can improve quality of life and aesthetic outcomes, such as improving measured quality of life scores, making patients appear less ill, sad, or distracted, and about 81% of patients in the study reported feeling satisfied following injections and would plan future aesthetic treatments [4]. Despite these reported improvements, there were minor dermal side effects in 13 out of the 127 patients included in the study, and 1 patient who developed a delayed inflammatory reaction [4]. Additionally, a systematic review noted that immunosuppressed patients, including those on chemotherapy, may have an increased risk of adverse effects such as filler granulomas and infectious complications [1,9].

Recommendations for Clinicians

Chemotherapy and radiation can cause tissue volume loss, skin dryness, sensitivity, and changes in pigmentation [2,6]. Cosmetic treatments such as dermal fillers can help patients with cancer feel more comfortable with their appearance, boosting self-esteem, contributing to a sense of increased psychological well-being and improving their quality of life [2]. Clinicians should therefore acknowledge and proactively address these crucial psychosocial benefits. However, this presents a dual consideration: while offering aesthetic and psychological advantages, these interventions simultaneously carry an elevated risk of adverse effects within patient populations undergoing chemotherapy and radiation therapies. Therefore, clinicians must thoroughly educate patients about these potential interactions and complications.

Consultations regarding dermal filler treatment in this specific patient population calls for a discussion involving the patient's oncologist, dermatologist, and injector. This multi-disciplinary discussion is crucial to obtain a comprehensive analysis of the patient to determine the safest and most effective cosmetic plan. To guide this process, a pre-treatment checklist should include (1) a review of the patient's cancer treatment timeline, (2) analysis of the patient's immune status or susceptibility to infections, (3) assessment of skin integrity (especially in irradiated areas), (4) documentation of any prior filler use, and (5) a clear discussion of risks, benefits, and procedure timing relative to immunosuppressive therapy.

The treating provider should also emphasize the importance of using aseptic techniques during the procedure and closely monitor patients post-injection to promptly manage any adverse reactions.

Conclusion

When considering dermal filler use in patients undergoing chemotherapy and radiation therapy, a tailored, risk-informed approach is essential. An in-depth consultation between the patient, the patient's oncologist, and their injector should involve a thorough analysis of the patient's skin, as well as answering questions regarding what kind of cancer treatment they are receiving, and how recent was their last treatment with

chemotherapy or radiation. Given the altered immune function and changes to the skin's structural integrity caused by chemotherapy and radiation, this population faces an elevated risk of infection, delayed wound healing, and immune-mediated filler complications. While the therapeutic benefits of cosmetic restoration, such as improved self-image and quality of life, are meaningful, they must be balanced against these unique risks.

Currently, there is minimal research that has been done to show exactly what adverse effects can occur from dermal filler treatments in patients who are actively undergoing chemotherapy or radiation therapy. Furthermore, the included case reports vary in their level of clinical detail, histological confirmation, follow-up duration, use of fillers, and their combinations with cancer therapies. This heterogeneity limits the ability to compare cases systematically or draw consistent conclusions. Some papers fail to specify the exact type, brand, or batch of filler used, which hinders the ability to assess biocompatibility or degradation in immunocompromised hosts. The proposed recommendations in this viewpoint are extrapolated from what is currently known about how chemotherapy and radiation therapy affect the skin, and the adverse effects of dermal filler treatment. The first filler agent approved by the FDA for cosmetics was Zyderm in 1981 [16]. There have been hundreds of new agents developed between 1981 and 2025, and research on long-term effects of varying fillers are developing with time [16].

More research is needed to examine the direct interactions of chemotherapy and radiation on various filler materials injected within the skin, and how these interactions can increase or decrease the risk of adverse effects. Future investigation should prioritize comparative studies that analyze risks among patients receiving different classes of chemotherapies (eg, immune checkpoint inhibitors vs tyrosine kinase inhibitors) and explore whether treatment timing (active therapy or remission) affects the likelihood of adverse filler-related events. Additional multi-center, prospective studies should also be done to address the gaps identified.

Authors' Contributions

MP, ET, and BMN wrote the original manuscript and provided data for Table 1. SK and LT contributed to manuscript reviewing and editing. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

References

1. Guo J, Fang W, Wang F. Injectable fillers: current status, physicochemical properties, function mechanism, and perspectives. RSC Adv 2023 Aug 4;13(34):23841-23858. [doi: 10.1039/d3ra04321e] [Medline: 37577103]
2. Tejero P, Pinto H, editors. Aesthetic Treatments for the Oncology Patient, 1st edition: CRC Press; 2020. URL: https://doi.org/10.1201/9780203728888 [accessed 2025-08-25] [doi: 10.1201/9780203728888]
3. Liao ZF, Cong LY, Li FW, et al. The research trend of soft tissue filler injection from 2000 to 2022: a bibliometric and visualized analysis. Plast Reconstr Surg Glob Open 2024 Feb;12(2):e5579. [doi: 10.1097/GOX.0000000000005579] [Medline: 38313594]
4. Shachar E, Hasson E, Arzi O, et al. Aesthetics treatment with injectable fillers in patients with cancer: safety and quality of life during active antineoplastic therapy. JCO 2024 Jun 1;42(16_suppl):e23245-e23245. [doi: 10.1200/JCO.2024.42.16_suppl.e23245]
5. Pathmanathan S, Dzienis M. Cetuximab associated dermal filler reaction. BMJ Case Rep 2019 Aug 26;12(8):e228882. [doi: 10.1136/bcr-2018-228882] [Medline: 31451454]
6. Proietti I, Skroza N, Mambrin A, et al. Aesthetic treatments in cancer patients. Clin Cosmet Investig Dermatol 2021;14:1831-1837. [doi: 10.2147/CCID.S342734] [Medline: 34898993]
7. Owczarczyk-Saczonek A, De Boulle K. Hyaluronic acid fillers and ASIA syndrome: case studies. Clin Cosmet Investig Dermatol 2023;16:2763-2771. [doi: 10.2147/CCID.S419716] [Medline: 37818199]
8. Heydenrych I, De Boulle K, Kapoor KM, Bertossi D. The 10-point plan 2021: updated concepts for improved procedural safety during facial filler treatments. Clin Cosmet Investig Dermatol 2021;14(779-814):779-814. [doi: 10.2147/CCID.S315711] [Medline: 34276222]
9. Edwards PC, Fantasia JE. Review of long-term adverse effects associated with the use of chemically-modified animal and nonanimal source hyaluronic acid dermal fillers. Clin Interv Aging 2007;2(4):509-519. [doi: 10.2147/cia.s382] [Medline: 18225451]
10. Hill A, McFarlane S, Johnston PG, Waugh DJJ. The emerging role of CD44 in regulating skeletal micrometastasis. Cancer Lett 2006 Jun 8;237(1):1-9. [doi: 10.1016/j.canlet.2005.05.006] [Medline: 15979783]
11. Syunyaeva Z, Kahnert K, Kauffmann-Guerrero D, Huber RM, Tufman A. Dermal filler injections mimic tumor activity during immune checkpoint inhibition. Respiration 2018;95(5):362-363. [doi: 10.1159/000487601] [Medline: 29621778]
12. Bisschop C, Bruijn MS, Stenekes MW, Diercks GFH, Hospers GAP. Foreign body reaction triggered by cytotoxic T lymphocyte-associated protein 4 blockade 25 years after dermal filler injection. Br J Dermatol 2016 Dec;175(6):1351-1353. [doi: 10.1111/bjd.14674] [Medline: 27086747]

13. Sung Y, Kim MH, Cho E. Successful hyaluronic acid filler injection in a chronic myeloid leukemia patient taking imatinib mesylate. J Cosmet Laser Ther 2019;21(4):243-244. [doi: 10.1080/14764172.2018.1525745] [Medline: 30285513]
14. Hu SCS, Hou MF, Luo KH, et al. Changes in biophysical properties of the skin following radiotherapy for breast cancer. J Dermatol 2014 Dec;41(12):1087-1094. [doi: 10.1111/1346-8138.12669] [Medline: 25354814]
15. Taplitz RA, Kennedy EB, Bow EJ, et al. Antimicrobial prophylaxis for adult patients with cancer-related immunosuppression: ASCO and IDSA clinical practice guideline update. J Clin Oncol 2018 Oct 20;36(30):3043-3054. [doi: 10.1200/JCO.18.00374] [Medline: 30179565]
16. Matton G, Anseeuw A, De Keyser F. The history of injectable biomaterials and the biology of collagen. Aesthetic Plast Surg 1985;9(2):133-140. [doi: 10.1007/BF01570345] [Medline: 3895848]

Abbreviations

ASCO: American Society of Clinical Oncology
ASIA: autoimmune/inflammatory syndrome induced by adjuvants
FDA: Food and Drug Administration
HA: hyaluronic acid
IDSA: Infectious Diseases Society of America

Edited by R Dellavalle; submitted 03.05.25; peer-reviewed by A Ayatollahi, G Kroumpouzos; revised version received 06.07.25; accepted 07.07.25; published 09.09.25.

Please cite as:
Panlilio M, Tchernogorova E, Nguyen BM, Khan S, Torgerson L
Dermal Filler Use in Patients Undergoing Chemotherapy and Radiation Therapy
JMIR Dermatol 2025;8:e76898
URL: https://derma.jmir.org/2025/1/e76898
doi:10.2196/76898

Teledermatology to Support Self-Care in Chronic Spontaneous Urticaria

Laura Schuehlein[1], PhD; Martin Peters[1], MBA; Graham Jones[2], PhD

[1]Novartis AG, Basel, Switzerland

[2]Tufts University Medical Center, 800 Washington St, Boston, MA, United States

Corresponding Author:
Graham Jones, PhD
Tufts University Medical Center, 800 Washington St, Boston, MA, United States

Abstract

Chronic spontaneous urticaria (CSU) is an autoimmune prompted skin disorder, whose hallmarks include the unpredictable onset of hives and itch. Symptom duration typically exceed 6 weeks, and flares can occur for up to 5 years or longer if untreated, impacting potentially any area of the body. The absence of obvious triggers and the variation in onset frequency often delays formal diagnosis which on average is approximately 2 years from first presentation. Initial standard of care is the use of low through to higher strength antihistamines in the first instance, with eventual escalation to prescription anti-inflammatory agents and potentially biologics once patients are under managed care. The societal impacts of delays in diagnosis are marked, with data suggesting CSU impacts up to 1% of the population, primarily of working age and with twice the prevalence in women. Herein, we advocate for the deployment of smartphone imaging and generative artificial intelligence technology to improve detection and early management of CSU through integrated self-care approaches. Such approaches embodying the tenets of P4 personalized medicine could have sustained impact on the disease through awareness campaigns, reducing the burden on the dermatology community and facilitating earlier access to curative therapeutic interventions.

(JMIR Dermatol 2025;8:e81830) doi:10.2196/81830

KEYWORDS

dermal imaging; chronic spontaneous urticaria; diagnosis; teledermatology; self care

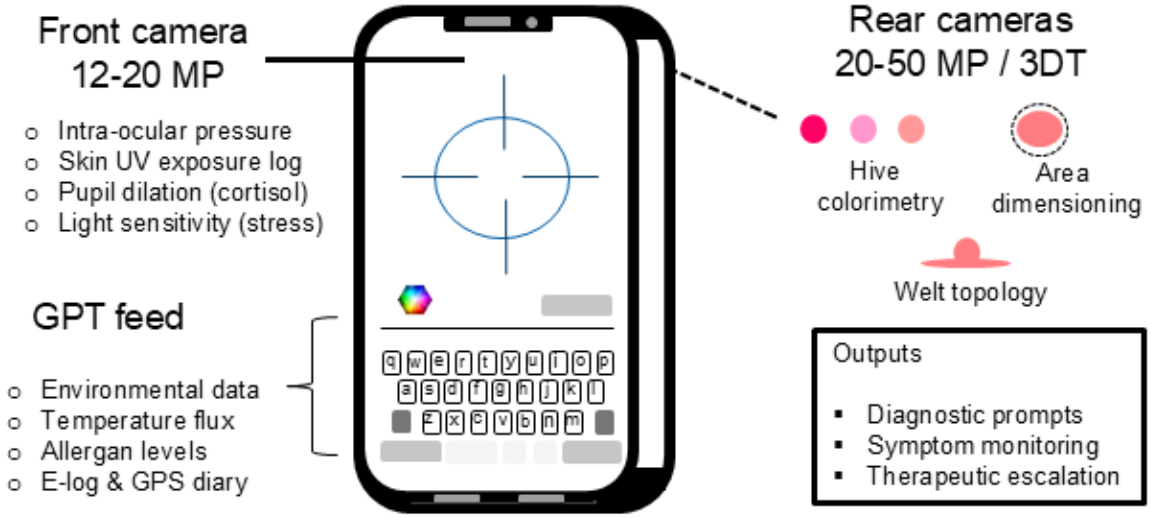
Introduction

Dermatological conditions are estimated to afflict nearly 2 billion people globally, yet because of historical shortages of expert dermatologists, the majority of cases are addressed by general practitioners, resulting in lowered diagnostic accuracy and delays in patients receiving optimized care. Chronic spontaneous urticaria (CSU) is representative of this conundrum, with patients typically addressing symptoms periodically and episodically with antihistamine agents, which provide short-term relief from symptoms but do not address the root cause. With powerful disease modifying prescription agents now available, it becomes imperative that patients' symptomatic presentations are documented, then adequately assessed to allow escalation as soon as practicable [1-6]. Given the rich structural and topological elements associated with symptoms (eg, hives, itch), there is natural interest in the use of imaging technologies coupled with computational tools to improve diagnosis [7]. In one recent teledermatological study involving over 16,000 cases, a deep learning system was deployed on photographic images and demonstrated high diagnostic accuracy [8]. The rapid evolution of mobile health technologies has now resulted in standardized methods for sharing of personal health images using smartphone technologies and has resulted in dedicated support apps for CSU such as CRUSE [9]. A key to the long-term deployment of and adherence to these solutions lies at the intersect of behavioral and social sciences [10], and we explore tactics herein which may lead to more widespread adoption.

Discussion

The impact of digital health on modern medicine is being witnessed in multiple areas, most dramatically, where patients obtain medical-related information on symptoms and potential remedies [11]. Though quality and validity of this information varies greatly, there is increased awareness surrounding skin-related diseases together with methods, typically using smartphone cameras, to assess visual characteristics to aid diagnosis [12]. In the case of CSU, some of the visible hallmarks are hives and welts whose appearance, locations, dimensions, and topology are potentially useful indicators in diagnosis, self-care, and outcome assessments (Figure 1) [4].

Figure 1. Specimen CSU relevant passive and active monitoring. CSU: chronic spontaneous urticaria.



Although some way-off from providing clinical grade characterization, these features could be useful in (1) raising awareness and prompting a patient to seek medical care, (2) serving as time-stamped reference points to track disease progression, and (3) confirming any observed treatment effects, for example, from topical adjuvants, over the counter symptomatic agents or prescribed disease modifying therapies. If used with precision and guidance, such information could become a key component in the patient's electronic health record, similar to how patient-recorded outcomes or self-assessment questionnaires have utility as part of a patient's composite case history. Achieving desired medical grade standard will require active dialog with professional dermatological societies, including standardized formats, rules and metrics for image capture and storage, procedures to validate images against controls, verification and authentication that it is associated with the patient, and informed consenting for upload and sharing of data [13]. Nonetheless, rapid progress is being observed and there is every reason to believe that such approaches will be embraced by the medical community and become integral components of telehealth based care of dermatological conditions in the future.

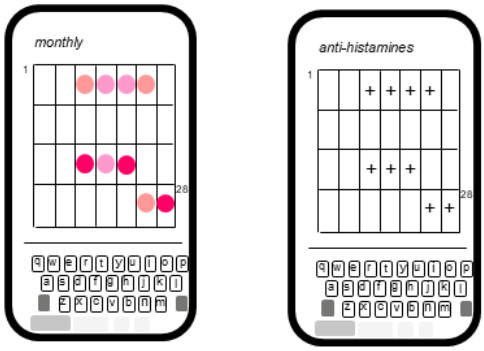
While these efforts progress, there is need to consider additional aspects of self-care which can help improve adherence to best practices and promote lifelong engagement with patients beyond addressing visual symptoms. A key to this approach is to consider behavioral and environmental factors which impact the patient on a perpetual basis, to drive engagement at multiple touch points. Ideally these can be achieved using a single device, and it turns out that the modern smartphone offers multiple options in this regard. For example, staple environmental factors including allergen levels in the atmosphere, ambient temperature, barometric pressure and relative humidity are all potential contributing factors in triggering immune responses [14]. Precision data, linked to GPS coordinates, can be provided on a real-time basis to smartphones, and artificial intelligence (AI or GPT based alerts could be directed to the patient—eg, on entering an area with high levels of specific pollens. This could be of additional relevance if recent dermatological images of CSU-related features (Figure 1) are stored on the smartphone and used to prompt alerts, for example, need for symptomatic therapy dose escalation. There may also be potential for pre-emptive action to avoid disease triggering. The environmental-based data could alert a patient to reconsider an action (eg, walking in a particular area) but the front camera of the smartphone could also be of utility. The sensors of this camera could be engaged automatically (eg, unlocking via retina scan) and have the ability to measure numerous features including intraocular pressure [15] and pupil dilation [16]. Both metrics correlate with elevated levels of plasma cortisol, which itself is a biomarker for stress and a potential trigger for CSU [17]. Likewise, light sensitivity and heightened UV exposure to facial skin are also potential immune-mediated triggers and could be detected and measured through the front camera sensor (Figure 1) [18]. Taken in sum, there are multiple opportunities to frame CSU patient self-management through the medium of the smartphone and in doing so, provide an effective mechanism for risk avoidance, real-time education, and longitudinal disease monitoring.

Integrating Into Managed Care

Successfully implementation of a smartphone-based tracking tool into managed care for CSU patients will require careful consideration of the patient-provider interface. Whether managed by a specialist dermatologist or primary care physician, there are a number of common scoring tools available including the Urticaria Activity Score (UAS), Urticaria Control Test (UCT), and Chronic Urticaria Quality of Life Questionnaire (CU-Q2oL) [19]. These scales can provide key insights to disease history and impact and inform treatment plans. This said, their deployment is not widespread or uniform, as there is a natural tendency for providers to treat presented visual symptoms, initially with antihistamines. Since a high percentage of patients are refractory to antihistamine treatment administered at the standard dose, there is potential for patients to disengage from care, resulting in impact to quality of life. The envisioned app could provide, on demand, a digital log of any visual

symptoms, a patient provided log of impact on quality of life, and potentially a log of medication history from which to assess impact of both therapeutics themselves and dosage (Figure 2). Such digital records could also serve to inform insurer and payer claims by offering accurate insight to impact on quality of life, potentially supporting decisions on dose escalations or the need to consider switching to alternate or higher efficacy medications to control symptoms and co-morbidities [20]. Although patient reported outcomes constitute components of existing diagnostic tools [19], there may be potential for higher levels of engagement if provisioned on the patient's personal device, in an app which also imports personal images and medication logs. There may also be the potential to keep patients engaged in their care in the absence of flares or symptoms, for example, by careful design of features in the app which celebrate 'days clear of symptoms' through prompts. Clearly, if approached appropriately, there is potential for such an app to strengthen the interactions between patient and provider, based on longitudinally captured data and where the patient becomes more proactive in care of their condition. More specifically, given the limited time available for typical patient-physician consults, the availability of this digital data could help maximize productivity and the drive to improved outcomes.

Figure 2. Combining images and medication log with activities of daily living data to assess control.



Symptom	Daily	Weekly	Occasional	Impacting work?
Hives			X	
Itch		X		
Sleep loss	X			X
Fatigue	X			
Anxiety	X			X
Depression		X		
Angioedema			X	

The Road to Implementation

The evolution of digital tools for patient self-management continues apace and in the case of CSU, the smartphone offers high potential to become a core component and contributor to patients electronic health records. In order to realize the myriad benefits and opportunities outlined herein, the following are anticipated next steps:

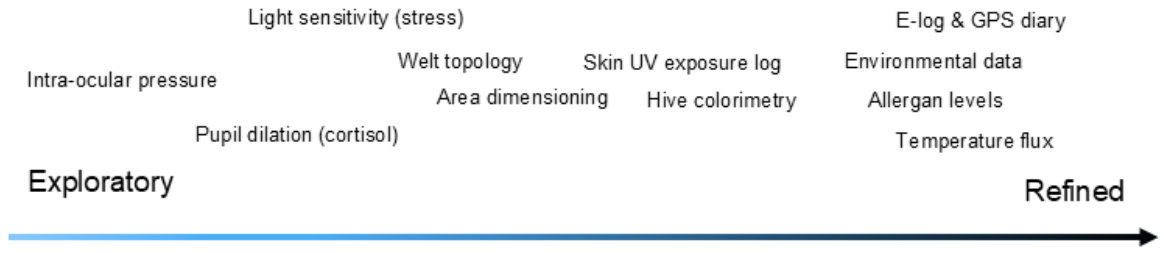
- Alignment between CSU patient advocacy groups and technology developers to optimize existing tools and design next generation apps and software development kits which can be customized around the individual patient
- Discussion with medical professional societies on required features, sensors, and software as medical device (SaMD) grade algorithms needed in devices and patients on which features would be of highest relevance to them
- Developing a route for integration with existing physician tools such as validated urticaria scores from which to calibrate device outputs [19]
- Technological refinement of some of the sensors and algorithms to realize medical grade precision
- Development of standardized frameworks for data capture, storage, and sharing respectful of General Data Protection Regulations
- The development of additional teledermatology media for patients (blogs, resources, apps) and the medical community (eg, scholarly journals, scientific conferences)

It is anticipated that many of the features and concepts outlined could be offered through patient support networks provisioned by developers of symptomatic and disease modifying therapies, typically through native apps or web portals. However, a key to the approach advocated herein is to more fully engage the patient on their terms, through their lived experiences in a highly personalized manner. There are two potential pathways for introduction of the features described – either as a consumer or patient support tool or through a more refined, clinical grade instrument which falls under the aegis of a regulated SaMD provisions. In either case the relative maturity of the sensors, algorithms, and their clinical validation are key considerations and there is optimism in this regard as the many features proposed are constantly evolving and refining. Although feeds on environmental data are now commonplace globally [21], and colorimetric analysis of dermatologic features is becoming more precise [22,23], many of the desired features will benefit from concerted effort for analytical refinement and clinical validation (Figure 3). Considerable progress has been made in feature dimensioning [24], and topology [8,25,26], and there is every reason to be optimistic that with collaborative efforts these can be refined for CSU applications. Likewise, metrics on ocular light sensitivity continue to evolve through measures of luminance [27,28], as do approaches to assessing pupil dilation using smartphones [29]. Even the more ambitious potential applications such as monitoring intro-ocular pressure can be accomplished with the aid of external tonometers, suggesting that future generations of devices might incorporate such functionality when miniaturized [30]. Regardless of whether such features are available in consumer products of SaMD variants, the route to implementation must also be mindful of the extant need for continual calibration of device sensors. A number of approaches to device calibration have been suggested [22,31], and systematic procedures on design

of appropriate clinical trials [32]. Finally, and importantly, learnings from the ground breaking work which went in to development of existing support apps such as CRUSE can help guide adoption of appropriate controls in terms of data privacy, and provide culturally relevant learnings from the different regions where it is deployed globally [9]. In terms of clinical development, integrating learnings from the device through standardized reporting frameworks such as COSORT AI and TRIPOD AI [33,34] will allow predictive models to evolve and may help benefit related dermatological research.

The combination of smartphone sensors coupled with powerful generative AI features, both of which are trained and evolved around the individual patient themselves has a high probability of achieving the goals articulated herein and we encourage the dermatology community to embrace this vision with enthusiasm. Accordingly, we encourage active dialog in the CSU community and reader base of this journal to help realize this opportunity.

Figure 3. Path to maturity and validation of digital technology measures.



Acknowledgments

The authors are employees of Novartis Pharmaceuticals. Views expressed are those of the authors and Novartis Pharmaceuticals had no influence on the content of the submitted manuscript. References cited were selected using search engines (SciFinder, Google Scholar) on the basis of scientific relevance, with priority afforded to most recent exemplars. Generative AI was not used in any portion of the manuscript.

Conflicts of Interest

None declared.

References

1. Zuberbier T, Abdul Latiff AH, Abuzakouk M, et al. The international EAACI/GA²LEN/EuroGuiDerm/APAAACI guideline for the definition, classification, diagnosis, and management of urticaria. Allergy 2022 Mar;77(3):734-766. [doi: 10.1111/all.15090] [Medline: 34536239]
2. Saini SS. Chronic spontaneous urticaria: etiology and pathogenesis. Immunol Allergy Clin North Am 2014 Feb;34(1):33-52. [doi: 10.1016/j.iac.2013.09.012] [Medline: 24262688]
3. Maurer M, Weller K, Bindslev-Jensen C, et al. Unmet clinical needs in chronic spontaneous urticaria. A GA. Allergy 2011 Mar;66(3):317-330. [doi: 10.1111/j.1398-9995.2010.02496.x] [Medline: 21083565]
4. Friedman A, Kwatra SG, Yosipovitch G. A practical approach to diagnosing and managing chronic spontaneous urticaria. Dermatol Ther (Heidelb) 2024 Jun;14(6):1371-1387. [doi: 10.1007/s13555-024-01173-5] [Medline: 38758422]
5. Fricke J, Ávila G, Keller T, et al. Prevalence of chronic urticaria in children and adults across the globe: Systematic review with meta-analysis. Allergy 2020 Feb;75(2):423-432. [doi: 10.1111/all.14037] [Medline: 31494963]
6. Yosipovitch G, Biazus Soares G, Mahmoud O. Current and emerging therapies for chronic spontaneous urticaria: a narrative review. Dermatol Ther (Heidelb) 2023 Aug;13(8):1647-1660. [doi: 10.1007/s13555-023-00972-6] [Medline: 37386330]
7. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. Per Med 2013;10(6):565-576. [doi: 10.2217/pme.13.57] [Medline: 25342952]
8. 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]
9. Neisinger S, Sousa Pinto B, Ramanauskaite A, et al. CRUSE® -An innovative mobile application for patient monitoring and management in chronic spontaneous urticaria. Clin Transl Allergy 2024 Jan;14(1):e12328. [doi: 10.1002/clt2.12328] [Medline: 38282190]
10. Boucher A, Peters M, Jones GB. How digital solutions might provide a world of new opportunities for holistic and empathic support of patients with hidradenitis suppurativa. Dermatol Ther (Heidelb) 2024 Aug;14(8):1975-1981. [doi: 10.1007/s13555-024-01234-9] [Medline: 39042318]

11. Lu Q, Schulz PJ. Physician perspectives on internet-informed patients: systematic review. J Med Internet Res 2024 Jun 6;26:e47620. [doi: 10.2196/47620] [Medline: 38842920]
12. Ouellette S, Rao BK. Usefulness of smartphones in dermatology: a US-based review. Int J Environ Res Public Health 2022 Mar 17;19(6):3553. [doi: 10.3390/ijerph19063553] [Medline: 35329240]
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. Saini SS, Kaplan AP. Chronic spontaneous urticaria: the devil's itch. J Allergy Clin Immunol Pract 2018;6(4):1097-1106. [doi: 10.1016/j.jaip.2018.04.013] [Medline: 30033911]
15. Schwartz B, Seddon JM. Increased plasma cortisol levels in ocular hypertension. Arch Ophthalmol 1981 Oct;99(10):1791-1794. [doi: 10.1001/archopht.1981.03930020665008] [Medline: 7295128]
16. Langer K, Jentsch VL, Wolf OT. Cortisol promotes the cognitive regulation of high intensive emotions independent of timing. Eur J Neurosci 2022 May;55(9-10):2684-2698. [doi: 10.1111/ejn.15182] [Medline: 33709613]
17. Tomaszewska K, Słodka A, Tarkowski B, Zalewska-Janowska A. Neuro-immuno-psychological aspects of chronic urticaria. J Clin Med 2023 Apr 26;12(9):3134. [doi: 10.3390/jcm12093134] [Medline: 37176575]
18. Salminen A, Kaarniranta K, Kauppinen A. Photoaging: UV radiation-induced inflammation and immunosuppression accelerate the aging process in the skin. Inflamm Res 2022 Aug;71(7-8):817-831. [doi: 10.1007/s00011-022-01598-8] [Medline: 35748903]
19. Weller K, Groffik A, Church MK, et al. Development and validation of the Urticaria Control Test: a patient-reported outcome instrument for assessing urticaria control. J Allergy Clin Immunol 2014 May;133(5):1365-1372. [doi: 10.1016/j.jaci.2013.12.1076] [Medline: 24522090]
20. Konstantinou GN, Podder I, Konstantinou G. Mental health interventions in refractory chronic spontaneous urticaria: a call to expand treatment guidelines. Cureus 2025 Mar;17(3):e81443. [doi: 10.7759/cureus.81443] [Medline: 40303526]
21. Johnston FH, Wheeler AJ, Williamson GJ, et al. Using smartphone technology to reduce health impacts from atmospheric environmental hazards. Environ Res Lett 2018 Apr 1;13(4):044019. [doi: 10.1088/1748-9326/aab1e6]
22. Xuan Y, Barry C, Antipa N, Wang EJ. A calibration method for smartphone camera photophlethysmography. Front Digit Health 2023;5:1301019. [doi: 10.3389/fdgth.2023.1301019] [Medline: 38075521]
23. Dugonik B, Golob M, Marhl M, Dugonik A. Optimizing digital image quality for improved skin cancer detection. J Imaging 2025 Mar 31;11(4):107. [doi: 10.3390/jimaging11040107] [Medline: 40278023]
24. Liu C, Fan X, Guo Z, Mo Z, Chang EIC, Xu Y. Wound area measurement with 3D transformation and smartphone images. BMC Bioinformatics 2019 Dec 18;20(1):724. [doi: 10.1186/s12859-019-3308-1] [Medline: 31852433]
25. Nightingale RC, Ross MT, Cruz RLJ, Allenby MC, Powell SK, Woodruff MA. Frugal 3D scanning using smartphones provides an accessible framework for capturing the external ear. J Plast Reconstr Aesthet Surg 2021 Nov;74(11):3066-3072. [doi: 10.1016/j.bjps.2021.03.131] [Medline: 34088646]
26. Rudari M, Breuer J, Lauer H, et al. Accuracy of three-dimensional scan technology and its possible function in the field of hand surgery. Plast Reconstr Surg Glob Open 2024 Apr;12(4):e5745. [doi: 10.1097/GOX.0000000000005745] [Medline: 38655101]
27. Zalewski S, Skarżyński K. The photometric testing of high-resolution digital cameras from smartphones-a pilot study. Sensors (Basel) 2024 Oct 29;24(21):6936. [doi: 10.3390/s24216936] [Medline: 39517833]
28. Chapter 3 Correlation of luminance with human light sensitivity. In: Information Visualization, 3rd edition: Elsevier; 2013. [doi: 10.1016/B978-0-12-381464-7.00003-X]
29. Barry C, de Souza J, Xuan Y, Holden J, Granholm E, Wang EJ. At-home pupillometry using smartphone facial identification cameras. Presented at: Proc SIGCHI Conf Hum Factor Comput Syst; Apr 29, 2022. [doi: 10.1145/3491102.3502493]
30. Rahimi-Nasrabadi H, Moore-Stoll V, Tan J, et al. Luminance contrast shifts dominance balance between ON and OFF pathways in human vision. J Neurosci 2023 Feb 8;43(6):993-1007. [doi: 10.1523/JNEUROSCI.1672-22.2022] [Medline: 36535768]
31. Patonis P. A comparative study on the use of smartphone cameras in photogrammetry applications. Sensors (Basel) 2024 Nov 15;24(22):7311. [doi: 10.3390/s24227311] [Medline: 39599088]
32. Choi E, Long V, Phan P, et al. Protocol for a double-blinded randomised controlled trial and process evaluation of a digital psychotherapeutic app in Singapore to improve symptom burden in patients with dermatological problems. BMJ Open 2025 Jul 25;15(7):e098266. [doi: 10.1136/bmjopen-2024-098266] [Medline: 40713034]
33. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020 Sep;26(9):1364-1374. [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
34. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 18;385:q902. [doi: 10.1136/bmj.q902] [Medline: 38636956]

Abbreviations

AI: artificial intelligence

CSU: Chronic spontaneous urticaria
SaMD: software as medical device

Edited by O Yilmaz; submitted 04.08.25; peer-reviewed by MS Janney, V Long; revised version received 10.09.25; accepted 11.10.25; published 13.11.25.

Please cite as:
Schuehlein L, Peters M, Jones G
Teledermatology to Support Self-Care in Chronic Spontaneous Urticaria
JMIR Dermatol 2025;8:e81830
URL: https://derma.jmir.org/2025/1/e81830
doi:10.2196/81830

Dermatologic Research in Displaced Populations: Importance, Challenges, and Proposed Solutions

Derek Maas[1], MBA, BS; Jackleen S Marji[2], MD, PhD

[1]New York University Grossman School of Medicine, New York, NY, United States

[2]The Ronald O. Perelman Department of Dermatology, New York University Grossman School of Medicine, 335 West 52nd St, 6th Floor, New York, NY, United States

Corresponding Author:
Jackleen S Marji, MD, PhD
The Ronald O. Perelman Department of Dermatology, New York University Grossman School of Medicine, 335 West 52nd St, 6th Floor, New York, NY, United States

Abstract

Displaced populations face complex dermatologic challenges. Contributing factors include low immunization rates, poor sanitation, crowded living conditions, and physical abuse. Chronic inflammatory conditions and infectious diseases, including fungal infections and scabies, are prevalent. Research is crucial to reduce the spread of disease, improve care in these populations, and develop sustainable frameworks for long-term dermatologic health care in crisis settings. The paucity of dermatologist support in this field exacerbates the issue. Ethical considerations include nonmaleficence and culturally sensitive practices, and proposed solutions include trauma-informed care training, advocacy for equitable research funding, teledermatology, and the development of shared international screening guidelines. Further research is essential to enhance dermatologic care for displaced populations.

(JMIR Dermatol 2025;8:e64828) doi:10.2196/64828

KEYWORDS

displaced persons; dermatology; dermatologic research; refugees; internally displaced persons; asylum seekers; skin diseases; epidemiology; health services accessibility; trauma-informed care; communicable diseases; vaccination; telemedicine; mental health services; posttraumatic stress disorder; ethical research; health care disparities; global health; infectious; scabies; fungal infections; bacterial infections; war-related injuries; health policy; health care delivery; scars; genital diseases; mental health; research design; cultural competency; informed consent

Background

For the purpose of this article, the term “displaced person” refers specifically to individuals who have been forcibly displaced due to conflict, persecution, violence, or disasters, including refugees, internally displaced persons, and asylum seekers. The global displacement crisis has led to the forced migration of 122.6 million individuals as of 2024, which increased from 59.2 million in 2014 [1]. Of these individuals, 71% are hosted in low- to middle-income countries, and 40% are children, many of whom encounter significant dermatologic health issues [1,2]. Displaced populations have complex health care needs. Dermatologic conditions, not frequently prioritized in acute care settings, represent significant disease burden and often serve as visible markers of hygiene-related issues, systemic illness, or infectious outbreaks [3]. Infectious diseases such as malaria, measles, acute respiratory infections, and diarrheal illnesses are among the major causes of morbidity and mortality [4]. Along with malnutrition (particularly in children), these problems account for the majority of deaths among displaced persons [4]. Mental health disorders are also prevalent due to the severe psychological stress associated with displacement [5]. This complex interplay of organ systems and the transient nature of the communities pose challenges to conducting dermatologic research, but such targeted research is crucial for understanding and addressing the skin care needs of these individuals [6]. Without it, health care providers lack the necessary epidemiologic data to design interventions and allocate resources effectively. Dermatologists are needed to assist in developing tailored management strategies [2]. We propose practical solutions to improve the mutual benefit of this research.

Dermatologic Conditions in Displaced Populations

A review of skin diseases in displaced populations [3] notes an increased frequency of cutaneous conditions in the scarce literature available, reporting a prevalence between 18.7% and 96.2% [7-10]. High rates of cutaneous conditions are due to several factors. War and conflict severely damage health care infrastructure [2]. Of surveyed respondents in the Syrian conflict zone, an endemic area of leishmaniasis, 12% knew that they could access treatment at hospitals or health centers, and less than a quarter had heard of the disease’s vector, the tsetse fly [11]. These findings helped drive educational initiatives in the

community [11]. This is one example of dermatologists and infectious disease specialists collaborating to guide targeted education to at-risk populations and front-line providers.

Chronic inflammatory conditions are prevalent and often overlooked in displaced populations. Four studies [7-9,12] aligned with this paper's definition of displaced persons. Migrants in Maltese reception centers (n=2216) had rates of "contact dermatitis and other eczema" of 4.8% [7]. A Jordanian refugee camp study (n=288) reported "dermatitis/eczema" at rates of 33.8%, while a study of migrants on the Mediterranean coast in Italy (n=6188) found rates of dermatitis at 7.5% [8,9]. Further, a diagnosis in the category "dermatitis and eczema" encompassed 21.3% of 380 Rohingya refugees living in the Kutupalong camp in Bangladesh [12]. Other forms of chronic dermatologic conditions have rarely been differentiated in the literature, but the Jordanian and Rohingya population studies specifically identified "papulosquamous disorders," occurring at rates of 6.9% and 2.9% [8,12]. Research on the management of these conditions in the setting of displacement is a potential area for growth. For example, comparing the effectiveness of various forms of barrier repair could guide the improved preparation of front-line health care providers.

Furthermore, informal settlements of displaced persons include diverse subpopulations with varying immunization levels; the seroprevalence often does not reach levels that confer herd immunity [2]. This, combined with poor sanitation and crowded housing conditions, leads to the rapid spread of contagious and vaccine-preventable diseases like measles and varicella [2,3]. Prior studies have found infectious diseases to represent 20.8% to 72% of skin conditions in displaced persons [7-9,12]. However, rates vary depending on classification criteria and potential diagnostic overlap. Viral infection rates fall between 0.7% and 8.5%, while fungal and bacterial infections have been reported from 7.9% to 49% and 3.2% to 11.2%, respectively [7-9,12]. Understanding local rates of communicable conditions is crucial to developing targeted vaccination efforts.

The process of displacement itself often forces migrants into extreme conditions, with many forced to travel in small boats [2,3]. A common location of arrival for these vessels is south of the Mediterranean Sea, where skin diseases commonly seen include secondary bacterial infections, scabies, deep abscesses, and tissue necrosis [3]. It is well established that scabies is particularly pervasive, with rates ranging from 3.5% to 58% [7-9,12]. A 2007 review studied the efficacy of various scabies treatments in refugee camps, highlighting the success of mass ivermectin administration [13]. The study demonstrated the importance of research in developing effective community interventions.

Current findings show wide variation in the rates of skin manifestations, and more research is needed for effective treatment and prevention. Future studies should further delineate rates of communicable infections by region while spreading the focus to chronic inflammatory conditions.

Challenges and Ethical Considerations

The backbone of research is ethical practice, and important aspects include nonmaleficence, beneficence, justice, and respect for persons. However, these mainstays are often not adequate for the complexities of vulnerable populations [6]. Access is fraught with difficulties due to safety concerns; many displaced persons lack the legal right to work or reside in their host country and consequently are transiently located with increased risk of arrest and detention [1,6]. For these reasons, individuals may show reluctance to engage in research and be concerned with confidentiality. Furthermore, participants often lean on researchers as a form of support, leading to conflicts of interest and trouble with the informed consent process, which may already be difficult to understand [6]. These considerations underscore the integration of culturally sensitive practices that foster an environment of open communication with balanced power dynamics [2,6]. Clinicians might be hesitant to study these populations in the first place due to anticipated difficulties in securing research approval and funding, given the inequitable distribution of academic funding and resources toward high-income countries [14]. Finally, geopolitical instability further complicates research efforts, as ongoing conflict and the displacement of health care workers hinder the implementation of structured studies; the politicization of global health and power imbalances in research partnerships only serves to exacerbate this challenge [15].

Additional barriers to research specifically apply to dermatology. Notably, screening guidelines for skin diseases vary internationally; the lack of shared guidelines poses a challenge to the design of systematic research on cutaneous conditions and the consistent provision of care [2,3]. It is also important to consider that the process of forced displacement often involves physical abuse and torture [2]. Scars, ecchymoses, and genital lesions are associated with trauma and are seen at high rates in displaced populations [2]. While investigators have infrequently distinguished conditions associated with trauma, scarring was found in migrants living in Maltese reception centers at a rate of 9.5% [3]. The spotlight that dermatologic research can place on cutaneous conditions has the potential to be a trigger that could retraumatize study participants, leading to posttraumatic stress disorder and other adverse mental health conditions [2].

Proposed Solutions

The risk of retraumatization in displaced populations makes nonmaleficence an ethical consideration of utmost importance [6]. To minimize the potential for psychological harm, clinicians working with these groups should be trained in trauma-informed care. Trauma-informed care training teaches the recognition of actions that could trigger memories of past traumatic events or add new traumatic experiences and requires that clinicians overcome the time constraint barrier of working in humanitarian settings (Table 1) [16]. A facet of this training involves understanding how and where to access mental health resources, which may be difficult during displacement [16]. The use of trauma-informed practices is of particular importance when it

comes to dermatologic conditions because of their visibility and frequent direct association with physical trauma.

Table . Proposed solutions and their potential impacts on improving dermatologic research with displaced populations.

Proposed solution	Potential impact	Explanation	Major obstacles
Teledermatology	↑ Continuity of research ↓ Spread of disease	Telemedicine platforms reduce the transmission of infections, provide consistent access to dermatologic expertise, and enhance data collection.	• Limited internet access • Lack of digital literacy
Shared international screening guidelines	↑ Standardization of research	Tailored protocols ensure consistency in research methodologies, improve the comparability of data, and aid in the development of targeted interventions.	• Geopolitical instability • Variability in health care infrastructure and regulation across countries
Advocacy for equitable research funding	↑ Continuity of research ↑ Availability of data	Advocacy efforts would help secure equitable global funding for research with vulnerable populations, strengthening the research process.	• Entrenched funding inequities favoring institutions in high-income countries
Trauma-informed care training	↓ Psychological harm	Clinician education on trauma-informed care reduces emotional stress, enhances trust, and improves patient cooperation in research studies.	• High clinician workload in humanitarian settings

Because of a frequent lack of access to primary care, displaced individuals often present urgently with dermatologic conditions, which can make management difficult and worsen the prognosis [2]. Delivering care close to the patient's community through community-based models is one way to combat this deficiency, and specialist training of front-line health care providers (including the World Health Organization, Red Cross, the United Nations High Commissioner for Refugees, and Doctors Without Borders) may allow for earlier diagnosis and treatment [2].

For complex cases, teledermatology has emerged as a potential solution for the shortage of trained dermatologists working in this field [2,3]. In a population with high rates of communicable skin disease, teledermatology also limits the spread of infection [2,3]. Integration of systems to conduct medical care virtually would also address the lack of consistent access to hidden populations, enabling continuity of care regardless of the patient's location [2,3]. Virtual platforms can facilitate improved understanding, confidentiality, and engagement with patient-centered, multimedia, interactive informed consent [17]. However, these platforms require stable internet access, compatible devices, and digital literacy for maximum effectiveness [17].

More research is needed to test the efficacy of standardized care models on dermatologic outcomes. Expanding the scope of these investigations requires the development of national screening guidelines for skin diseases in migrants and displaced persons, a task complicated by nation-specific differences in health care infrastructure and regulation (Table 1). To address the lack of resources, researchers should advocate for equitable global funding by raising awareness about the importance of research in vulnerable populations (Table 1). Clinicians looking to secure support can also form collaborative partnerships with agencies like the United Nations High Commissioner for Refugees and Doctors Without Borders.

Advancing dermatologic care for displaced populations requires an approach informed by ethical practices and cultural sensitivity. By addressing the unique challenges faced by displaced individuals, such as their legal uncertainties, high rates of infectious disease, and elevated potential for retraumatization, clinicians can work to develop more effective research strategies. Proposed solutions include advocacy for equitable research funding, development of uniform international screening protocols, use of teledermatology, and the integration of trauma-informed care into dermatologic services (Table 1). Further research is essential, and dermatologists must work with community health systems to craft and optimize care models.

Conflicts of Interest

None declared.

References

1. Figures at a glance. United Nations High Commissioner for Refugees. 2022. URL: https://www.unhcr.org/us/about-unhcr/who-we-are/figures-glance [accessed 2025-04-23]
2. Knapp AP, Rehmus W, Chang AY. Skin diseases in displaced populations: a review of contributing factors, challenges, and approaches to care. Int J Dermatol 2020 Nov;59(11):1299-1311. [doi: 10.1111/ijd.15063] [Medline: 32686140]

3. Padovese V, Knapp A. Challenges of managing skin diseases in refugees and migrants. Dermatol Clin 2021 Jan;39(1):101-115. [doi: 10.1016/j.det.2020.08.010] [Medline: 33228854]
4. Famine-affected, refugee, and displaced populations: recommendations for public health issues. MMWR Recomm Rep 1992 Jul 24;41(RR-13):1-76. [Medline: 1326713]
5. Porter M, Haslam N. Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. JAMA 2005 Aug 3;294(5):602-612. [doi: 10.1001/jama.294.5.602] [Medline: 16077055]
6. Pillai V, Mosier-Mills A, Khoshnood K. Methodological and ethical challenges in research with forcibly displaced populations. In: Allotey P, Reidpath D, editors. The Health of Refugees: Public Health Perspectives from Crisis to Settlement: Oxford Academic; 2019. [doi: 10.1093/oso/9780198814733.003.0011]
7. Padovese V, Egidi AM, Melillo Fenech T, et al. Migration and determinants of health: clinical epidemiological characteristics of migrants in Malta (2010-11). J Public Health (Oxf) 2014 Sep;36(3):368-374. [doi: 10.1093/pubmed/fdt111] [Medline: 24277779]
8. Saikal SL, Ge L, Mir A, et al. Skin disease profile of Syrian refugees in Jordan: a field-mission assessment. J Eur Acad Dermatol Venereol 2020 Feb;34(2):419-425. [doi: 10.1111/jdv.15909] [Medline: 31498503]
9. Di Meco E, Di Napoli A, Amato LM, et al. Infectious and dermatological diseases among arriving migrants on the Italian coasts. Eur J Public Health 2018 Oct 1;28(5):910-916. [doi: 10.1093/eurpub/cky126] [Medline: 30010744]
10. Kibar Öztürk M. Skin diseases in rural Nyala, Sudan (in a rural hospital, in 12 orphanages, and in two refugee camps). Int J Dermatol 2019 Nov;58(11):1341-1349. [doi: 10.1111/ijd.14619] [Medline: 31498882]
11. Tamiru HF, Mashalla YJ, Mohammed R, Tshweneagae GT. Cutaneous leishmaniasis a neglected tropical disease: community knowledge, attitude and practices in an endemic area, Northwest Ethiopia. BMC Infect Dis 2019 Oct 16;19(1):855. [doi: 10.1186/s12879-019-4506-1] [Medline: 31619180]
12. Khan SS, Padovese V, Maurer TA, et al. A skin disease and needs assessment analysis of the displaced Rohingya population in the Kutupalong refugee camp, Bangladesh. Clin Exp Dermatol 2020 Dec;45(8):1051-1054. [doi: 10.1111/ced.14310] [Medline: 32460353]
13. Strong M, Johnstone P. Interventions for treating scabies. Cochrane Database Syst Rev 2007 Jul 18(3):CD000320. [doi: 10.1002/14651858.CD000320.pub2] [Medline: 17636630]
14. Binagwaho A, Allotey P, Sangano E, Ekström AM, Martin K. A call to action to reform academic global health partnerships. BMJ 2021 Nov 1;375:n2658. [doi: 10.1136/bmj.n2658] [Medline: 34725093]
15. Persaud A, Bhat PS, Ventriglio A, Bhugra D. Geopolitical determinants of health. Ind Psychiatry J 2018;27(2):308-310. [doi: 10.4103/ipj.ipj_71_18] [Medline: 31359990]
16. Wilson C, Pence DM, Conradi L. Trauma-informed care. In: Bailey D, Mizrahi T, editors. Encyclopedia of Social Work: Oxford University Press; 2013. [doi: 10.1093/acrefore/9780199975839.013.1063]
17. Abujarad F, Alfano S, Bright TJ, et al. Building an informed consent tool starting with the patient: the patient-centered virtual multimedia interactive informed consent (VIC). AMIA Annu Symp Proc 2017:374-383. [Medline: 29854101]

Edited by R Alhusayen; submitted 07.08.24; peer-reviewed by A Ching, MA El-Hamd, Y Elsanousi; revised version received 17.03.25; accepted 30.03.25; published 05.05.25.

Please cite as:
Maas D, Marji JS
Dermatologic Research in Displaced Populations: Importance, Challenges, and Proposed Solutions
JMIR Dermatol 2025;8:e64828
URL: https://derma.jmir.org/2025/1/e64828
doi:10.2196/64828

Investigating Experiences With Scarring Among Transgender and Gender Diverse People: Mixed Methods Study

Nora Yanyi Sun[1]; Kanika Kamal[2,3], MD; Alex Sogomon Keuroghlian[4,5], MD, MPH

[1]Harvard College, Harvard University, 373 Quincy Mail Center 58 Plympton St, Cambridge, MA, United States

[2]Harvard Medical School, Cambridge, MA, United States

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

[4]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

[5]Department of Behavioral Health, Fenway Health, Boston, MA, United States

Corresponding Author:
Nora Yanyi Sun
Harvard College, Harvard University, 373 Quincy Mail Center 58 Plympton St, Cambridge, MA, United States

Abstract

Background: Scarring has been shown to have adverse health effects on marginalized patient groups. However, experiences of scarring among transgender and gender diverse (TGD) people have not yet been thoroughly characterized.

Objective: This study aimed to investigate the impacts of scarring related to gender-affirming care and other causes among TGD people.

Methods: Anonymous data were extracted from Reddit, a popular online platform organized into "subreddit" groups based on identities and interests. A combined total of 604 posts and comments that explicitly reference physical scarring were extracted from r/FtM, a subreddit for transmasculine people (449 posts and comments) and r/MtF, a subreddit for transfeminine people (155 posts and comments). Applying inductive thematic analysis, all posts and comments were coded and codes were sorted into overarching themes.

Results: Among the 604 posts and comments, the scars most discussed were secondary to gender-affirming care procedures, including mastectomy (n=338 posts and comments), hormone administration (n=102 posts and comments), and hair removal (n=38 posts and comments). Nongender-affirming care-related scars, such as those due to self-harm (n=43 posts and comments), were discussed less often. A total of five overarching themes emerged through thematic analysis: (1) concerns about physical outcomes related to scarring; (2) psychological distress related to scarring; (3) societal perceptions of scarring; (4) strategies to prevent, conceal, and minimize scarring; and (5) positive experiences with scarring.

Conclusions: For TGD people, scar complications, visibility, and permanence represent major concerns. While many TGD people ultimately accept scarring as an unavoidable consequence, scarring both related and unrelated to gender-affirming care can present a significant psychosocial stressor for TGD people. Scarring can result in physical health complications, gender dysphoria, and negative body image; visible scarring is also a barrier for TGD people who wish to blend into society. Clinicians should improve communication regarding scarring outcomes and scar-care procedures. Future research should focus on the development of scar prevention, care, and reduction techniques for TGD people.

(JMIR Dermatol 2025;8:e62714) doi:10.2196/62714

KEYWORDS

transgender; gender-affirming care; scarring; body image; gender; qualitative study; psychosocial; content; online forum; online; physical scarring; thematic analysis; scar; visibility; stress; barrier; marginalized patients; scar visbility; reddit; mastectomy; gender dysphoria; emotional trauma; psychosocial experience; trauma

Introduction

"Transgender and gender diverse" (TGD) is an umbrella term for people whose gender identities differ from societal expectations related to the sex they were assigned at birth. Some TGD people may choose to undergo gender-affirming surgery or pursue other forms of gender-affirming care at any point in their lives to alleviate gender dysphoria and better align their physical characteristics with their gender identity. Gender-affirming care is considered medically necessary and associated with numerous mental health benefits, including significantly decreased psychological distress and suicidality [1].

Over the past 2 decades, the number of gender-affirming surgeries among TGD people has increased. Approximately 9000 gender-affirming surgeries now occur annually across the

United States [1,2], thus likely resulting in an increased prevalence of surgical scarring. Although appearance and permanence vary broadly by surgical techniques [3], scarring remains a concern for some TGD people.

In addition to gender-affirming surgery, gender-affirming care more broadly, including nonsurgical interventions such as estrogen and testosterone injections and patches, may also result in significant scarring when used for long-term therapy [4]. In addition, gender-affirming hormones may cause increased acne scarring [5,6]. Furthermore, gender-affirming hormones may lead to greater visibility of pre-existing scars and predisposition to scarring due to changes in skin composition [7]. Similarly, hair removal procedures may result in scarring if not performed properly [8]. Finally, due to the unique combination of stressors experienced by TGD populations, higher rates of self-injury, a coping mechanism among some TGD people, may also be a cause of scarring [9].

Scarring may reveal TGD identities to other people and thereby interfere with goals of blending into society, thus potentially resulting in significant stress for TGD people [3]. Furthermore, studies have shown that scarring may cause significant mental health burden, especially in marginalized and minoritized populations; for example, scarring is associated with adverse psychological outcomes among breast cancer survivors and people of color [3]. Within minority stress theory, minoritized populations face a variety of external and internal stressors. As scarring is often perceived as a negative physical trait by society, bearing visible scars is an additional form of marginalization, and multiply marginalized people may experience superadditive stressors, exclusion, and discrimination, explaining the increased psychological distress observed among marginalized people with scars [10].

Furthermore, TGD people are uniquely vulnerable to body dissatisfaction because they experience gender dysphoria and increased body scrutiny from others. Per body objectification theory, increased awareness of one's body, especially physical characteristics that do not align with personal or societal goals, may lead to body shame, which is linked to increased psychological distress and disordered eating [11]. Thus, compared with cisgender people, TGD people may be especially vulnerable to body dissatisfaction due to their scars.

The impacts of scars on TGD people have not been well-characterized in the literature. Currently, most studies investigate scarring among TGD people through a surgical and dermatologic rather than psychological and societal lens [12]. Also, most studies have limited their scope to scarring directly caused by traditionally defined gender-affirming surgery [13]. Consistent with objectification theory, these limited studies with TGD populations have suggested that scarring may lead to gender dysphoria, body image dissatisfaction, and symptoms of both depression and anxiety (eg, breast augmentation surgery scars may reduce self-esteem or lead to negative self-perception) [14].

Although some studies mention scarring due to nonsurgical gender-affirming practices, such as chest binding, these studies focus on the overall impacts of these gender-affirming practices rather than specifically centering the impacts of scars on TGD people [15].

As physical appearance plays an important role in body image and self-perception [16], scarring that is not explicitly connected with gender-affirming care may represent a substantial portion of TGD people's scars and has the potential to affect them in unique ways that warrant study.

Understanding TGD people's experiences with scarring is necessary to motivate and inform innovation in scar treatments and supportive psychological care. Furthermore, treatment of scarring related or unrelated to gender-affirming care is currently deemed cosmetic and not medically necessary by most insurance companies; it is therefore not covered under most insurance policies [3]. An improved understanding of the impact of scarring for TGD people could support the medical and psychological necessity of scar treatment, and potentially more inclusive revisions to these policies.

Reddit is a popular online platform where users can enter long-form posts, comment on posts, and respond to comments. Due to the anonymous nature of the platform, it attracts a large number of TGD people to discuss personal health matters. Reddit forums have been studied to glean a better understanding of mental health [17,18] and sexual health [19], as well as LGBT (lesbian, gay, bisexual, and transgender) health [20]. The Reddit website is organized into subreddits centered around a specific subject, connecting people with similar identities or interests in dialogue.

"r/MtF" ("male-to-female") and "r/FtM" ("female-to-male") communities on Reddit are a pair of online subreddits for TGD people that host approximately 470,000 total users. In this study, we investigated the experiences of physical scarring among TGD people by analyzing data extracted from these 2 online forums and characterized the common experiences of scarring in these communities by identifying 6 common themes of the threads.

We note that the subreddits names "MtF" and "FtM" are based on an outdated, binary conception of gender [21] and inappropriately emphasize the sex assigned at birth. Therefore, this study may fail to adequately represent the experiences of TGD people with scarring, including nonbinary people, as well as experiences of TGD people who do not feel comfortable using these subreddits. From here onwards, this paper will refer to r/MtF as the transfeminine subreddit and r/FtM as the transmasculine subreddit.

Methods

Data Collection

Threads–defined as posts plus any comments on those posts–in the transfeminine and transmasculine community subreddits were screened by entering the keyword "scar" in the subreddit search function to extract threads that contain this term, including alternative forms of the word, such as "scarring." This method has been used in previous studies analyzing Reddit data [17].

As Reddit's search feature caps the number of threads that can be returned at a given time, we initially manually downloaded the 70 most recent qualifying threads from each subreddit on December 21, 2023. Threads that did not mention scarring or only mentioned scars in a metaphorical context were marked for removal from the dataset. Following removal of posts and comments that did not meet the study's search criteria, a total of 449 posts and comments from the transmasculine subreddit and 155 posts and comments from the transfeminine subreddit were coded.

Based on the concept of information power, an alternative to data saturation, no further threads were downloaded after the initially extracted posts because analyzed threads were information-dense, sometimes with hundreds of relevant comments; Furthermore, the study had a focused goal to examine effects of scarring among TGD people that was fulfilled by the initial analysis [22,23].

Data Analysis

We performed inductive thematic analysis, as described by Braun and Clarke [24], by applying a nonpositivist "big Q" approach that acknowledged reflexivity and aimed at interpreting the qualitative data. The study team consisted of 3 authors, including 2 cisgender women and a nonbinary person; all members of the study team also identified as queer, a person of color, or both.

Before coding, each post and comment included in the analysis was assigned a row in a spreadsheet. Codes were then assigned to each row in this spreadsheet. Two authors then independently coded the first 14 posts and related comments in each dataset on their own copy of the spreadsheet, at which point both authors believed that they understood the data sufficiently to develop an initial codebook. The codes were then reviewed jointly by both authors to remove duplicated codes and combine the rest of the codes until consensus was reached on a working codebook. During the discussion process, the coders used reflexivity practices, such as reflecting on potential personal biases and verbalizing any assumptions made during coding.

The remainder of the threads was then coded by 1 author using the working codebook, amending the codebook as necessary. Coding also involved noting the type of scar in each post and comments to contextualize the results. The senior author provided clinical perspectives and critical suggestions throughout coding. Through discussion with the senior author, the coders organized the codes into a set of 6 overarching themes.

Ethical Considerations

The transfeminine and transmasculine subreddits, including all threads analyzed, are located within the public domain. On the Reddit platform, users are identified only by usernames, which are generally not connected to any identifying information. During the coding process as well as in this article, all user data were completely deidentified, with no usernames recorded. This study received a "Not Human Subjects Research Determination" from the Harvard Longwood Area Institutional Review Board because the authors are not able to ascertain the identities of the users whose posts are included in the analysis. To further preserve user anonymity, quantitative data are only presented in aggregated form, and we have conducted slight and judicious rephrasing of quotes while preserving meaning to prevent the original posts from being retrieved [25].

Results

Types of Scarring

Among posts and comments in the transmasculine subreddit, mastectomy scars were highly represented, comprising 75.3% (338/449) of posts and comments. Second to mastectomy scars were scars due to testosterone injection, comprising 17.4% (78/449) of posts and comments. These 2 types of scars represented a combined total of 92.6% (416/449) of all posts and comments in the transmasculine subreddit (Table 1).

Table . Types of scars in the transmasculine subreddit.

Type of scars	Count (n=449), n (%)
Mastectomy scars	338 (75.3)
Scarring due to testosterone administration	78 (17.4)
Self-harm scars	5 (1.1)
Other scars unrelated to gender-affirming care	15 (3.3)
Other scars related to gender-affirming care	8 (1.8)

In comparison with the transmasculine subreddit, sources of scars discussed on the transfeminine subreddit were more divergent. Self-harm scars (38 posts and comments, 24.5%), scarring due to hair removal procedures (29 posts and comments, 18.7%), and scarring due to estrogen administration (24 posts and comments, 15.5%) were the most common types of scars discussed. Scarring that is directly attributed to gender-affirming procedures comprised a majority of the threads (86/155, 55.5%), with a broad variety of both surgical and nonsurgical procedures mentioned, including hair removal, vaginoplasty, breast augmentation, facial feminization surgery, orchiectomy, and genital tucking (Table 2).

Table . Types of scars in the transfeminine subreddit.

Scar type	Count (n=155), n (%)
Self-harm scars	38 (24.5)
Scarring due to hair removal procedures	29 (18.7)
Scarring due to estrogen administration (eg, patches and injection)	24 (15.5)
Orchiectomy scars	11 (7.1)
Vaginoplasty scars	9 (5.8)
Acne scars	8 (5.2)
Stretch marks	6 (3.9)
Other scars related to gender-affirming care	13 (8.4)
Other scars unrelated to gender-affirming care	17 (10.9)

Qualitative Themes

Five overarching themes describing various experiences with scarring emerged through inductive thematic analysis: (1) concerns about physical outcomes related to scarring; (2) psychological distress related to scarring; (3) societal perceptions of scarring; (4) strategies to prevent, conceal, and minimize scarring; and (5) positive experiences with scarring. All 5 themes were identified in both subreddits and both in and beyond the context of gender-affirming care procedures (Table 3).

Table . Frequencies of codes and themes identified.

Themes and subthemes	r/FtM[a] (n=551), n (%)	r/MtF[b] (n=165), n (%)
Theme 1: concerns about physical outcomes related to scarring	83 (14.4)	62 (35.8)
Interference of pre-existing scars with gender-affirming care	6 (1)	19 (11)
Desire for specific scar appearance	20 (3.5)	7 (4.1)
Period of decreased mobility due to scarring	15 (2.6)	0 (0)
Scar sensitivity	7 (1.2)	2 (1.2)
Scar tissue build-up due to nonsurgical gender-affirming care	7 (1.2)	9 (5.2)
Hormone absorption on scars	13 (2.2)	0 (0)
Heterogeneity in scar healing	9 (1.5)	1 (0.6)
Scar evolution	2 (0.3)	12 (6.9)
Encouraging communication with physician regarding physical health concerns related to scarring	4 (0.7)	12 (6.9)
Theme 2: psychological distress related to scarring	19 (3.3)	28 (16.2)
Anxiety about increased predisposition to scarring	1 (0.2)	5 (2.9)
Anxiety about possible scar appearance	6 (1.0)	9 (5.2)
Hesitation to pursue gender-affirming care due to worries regarding scar permanence	5 (0.8)	3 (1.7)
Negative body image due to scarring	7 (1.2)	6 (3.4)
Scars as source of gender dysphoria	0 (0)	5 (2.8)
Theme 3: societal perceptions of scarring	131 (22.8)	2 (1.2)
Fear of identity revelation due to gender-affirming care scars	20 (3.5)	2 (1.2)
Transphobic views of scarring as mutilation	14 (2.4)	0 (0)
Exaggerated representations of scarring	97 (16.8)	0 (0)
Theme 4: strategies to prevent, conceal, and minimize scarring	260 (45.2)	62 (35.8)
Inquiry for scar prevention advice	6 (1.0)	0 (0)
Sharing scar prevention experience	9 (1.5)	11 (6.3)
Inquiry for scar care advice	6 (1.0)	3 (1.7)
Uncertainty about scar care due to poor communication from gender-affirming care clinician	27 (4.7)	3 (1.7)
Sharing scar care experiences	24 (4.2)	0 (0)
Scar concealment with clothing	31 (5.4)	3 (1.7)
Tattooing scars	5 (0.8)	1 (0.6)
Lasering scars	19 (3.3)	2 (1.1)
Concealing gender-affirming surgery scar origin	72 (12.5)	33 (19.1)
Scar revision surgery	55 (9.6)	5 (2.9)
Encouraging communication with physician regarding preventing, concealing, and minimizing scarring	6 (1.0)	1 (0.5)
Theme 5: positive experiences with scarring	82 (14.3)	19 (10.9)

Themes and subthemes	r/FtM[a] (n=551), n (%)	r/MtF[b] (n=165), n (%)
Preference for gender-affirming care despite scarring	57 (9.9)	1 (0.5)
Low visibility of scars	1 (0.2)	10 (5.7)
Scar improvement over time	9 (1.57)	0 (0)
Acceptance of and pride in scarring	15 (2.61)	8 (4.6)

[a]r/FtM: transmasculine subreddit.

[b]r/MtF: transfeminine subreddit.

Theme 1: Concerns About Physical Outcomes Related to Scarring

Posts discussed physical health concerns related to existing or future scars that were both related and unrelated to gender-affirming care (Table 4). Posts discussed the physical health implications of scarring, such as how to manage scarring-related complications, what the expected course of recovery and outcomes of scarring are, and whether scarring would make them ineligible for certain gender-affirming procedures. Many asked for advice on how to treat scars following gender-affirming surgery due to lack of effective communication from their care teams, which sometimes resulted in divergent responses. Posts shared concerns regarding the impact of pre-existing scars on eligibility, feasibility, and effectiveness of gender-affirming care. Posts also expressed concern that scarring near treatment sites for gender-affirming surgery or hormone therapies may render the treatment not physically feasible or effective for them. Posts also described users' experiences of and asked for advice regarding severe scarring due to improper technique administering gender-affirming hormones with injections or patches.

Table . Concerns about physical outcomes related to scarring.

Sample codes	Selected quotes
Desire for specific scar appearance	• *The shape I prefer for my own operation is scars straight across.* [Transmasculine subreddit, mastectomy scars]
Period of decreased mobility due to scarring	• *I couldn't stay after returning to my old job because even after 8 weeks I couldn't handle the heavy lifting. I still can't do it. I am unemployed right now and the scars are making it harder to find a job.* [Transmasculine subreddit, mastectomy scars] • *Don't move your arms or carry heavy objects for several weeks or months because this can stretch out the scars.* [Transmasculine subreddit, mastectomy scars]
Scar tissue build-up due to nonsurgical gender-affirming care (eg, gender-affirming hormones patches and gender-affirming hormones injections)	• *I've been using hormone patches for about 5 - 6 months now. I always shave and clean the area before application, but no matter how I take off the patches, it causes bleeding and scarring. I don't know what to do. I'm almost out of safe areas to apply the patches because of the scar tissue.* [Transfeminine subreddit, scarring due to estrogen administration]
Interference of pre-existing scars with gender-affirming care	• *Will having scars on my chest make it hard for me to get top surgery in the future? I have some scars on my chest below my collarbones. These scars aren't severe, but I'm worried they could cause problems.* [Transmasculine subreddit, mastectomy scars] • *I have many scars on the front of my thighs, which are not wounds. Would this affect testosterone absorption at all?* [Transmasculine subreddit, unspecified scars unrelated to gender-affirming care]

Theme 2: Psychological Distress Related to Scarring

Before receiving gender-affirming care, posts expressed anxiety regarding the appearance, visibility, and permanence of scars that may result from these procedures (Table 5). Specifically, users expressed anxiety about uncertainty and heterogeneity in scarring outcomes and changes in scar appearance over time, and some posts expressed that users had increased desire for visible scarring. Posts sometimes specified that concern and confusion about scarring outcomes had resulted from poor communication from a gender-affirming surgeon or other clinician.

Table . Psychological distress related to scarring.

Sample codes	Selected quotes
Anxiety about possible scar appearance	• *But I started thinking, what if the surgery pulls at my skin or something else goes wrong and I get some outer scar?* [Transfeminine subreddit, vaginoplasty scars] • *I am currently waiting for my top surgery to be scheduled. My physician said my scar will be a little rounded with the cuts. I am worried that the scar might be too rounded, and I won't like the appearance.* [Transmasculine subreddit, mastectomy scars] • *I wanted to know about scar placement and nipple placement, if he's worried about dog ears or if I might need revision later, but he talks so fast I could barely get a word in or process everything he was telling me.* [Transmasculine subreddit, mastectomy scars]
Negative body image due to scarring	• *I hate almost everything about my body: My face is too masculine and I hate the scars I've given myself.* [Transfeminine subreddit, self-harm scars] • *I think I have the unluckiest batch of genes yet… 7'8," looks like Kaiser Wilhelm II, and a boxy face with scars.* [Transfeminine subreddit, unspecified scars unrelated to gender-affirming care] • *I love showers, so dreading them because I have to see my scars only made me more upset.* [Transmasculine subreddit, mastectomy scars]

Theme 3: Societal Perceptions of Scarring

When exchanging advice regarding scar care and concealment, many posts mentioned users' motivations for altering scar appearance to improve social acceptance, representing the third theme of the threads (Table 6). While many motivations were internal, like those mentioned in the previous theme, some posts described motivations related to societal perceptions of scarring. Specifically, posts shared concerns about being "clocked"—recognized as TGD—due to prominent surgery scars. Some posts also expressed concerns about receiving transphobic comments that characterize their gender-affirming surgery scars as evidence of self-mutilation, as well as general aversion among family, friends, and community members toward scars of any origin.

Table . Societal perceptions of scarring.

Sample codes	Exemplary quotes
Fear of identity revelation to other people due to gender-affirming care scars	• *Does anyone have any tips about what I could say if someone clocked the scars on my vagina or clocked me in general? Even though I believe I'm stealth, I still worry a lot that I'll get clocked wherever I go.* [Transfeminine subreddit, vaginoplasty scars] • *I am having top surgery soon, and I'm super excited for it, but I'm really not looking forward to having visible scars. I don't want to get clocked by people, especially with trans scars becoming more recognized.* [Transmasculine subreddit, mastectomy scars]
Transphobic views of scarring as mutilation	• *My mother says I'm paying thousands of dollars to mutilate myself.* [Transmasculine subreddit, mastectomy scars] • *My entire family is vehemently against "mutilating" a healthy body part.* [Transmasculine subreddit, mastectomy scars]

Theme 4: Strategies to Prevent, Conceal, and Minimize Scars

Users made posts and comments exchanging advice on topics surrounding changing the appearance of scars—both related and unrelated to gender-affirming care, which represents the fourth theme of the threads (Table 7). Oftentimes, posts shared and asked for advice on scar care techniques that reduced and concealed scarring. Specifically, many users on the transmasculine subreddit exchanged advice for mastectomy scar aftercare so that the healed scar lines would be aesthetically pleasing or minimally visible. Various posts also discussed ways to conceal scarring. Popular methods for concealing a variety of scars included tattoos (especially over mastectomy scars),

make-up, body hair, clothing, and laser treatments for scar revision.

Table . Strategies to prevent, conceal, and minimize scars.

Sample codes	Exemplary quotes
Sharing scar prevention experience and advice	• *I was advised to wear a binder or ace bandages to help keep the chest shape in line.* [Transmasculine subreddit, mastectomy scars]
	• *After surgery, do not carry heavy objects for several weeks or months as that can stretch out the scars.* [Transmasculine subreddit, mastectomy scars]
Scar concealment with clothing	• *I have a scar on my chest that I hate showing off. Are there any bras out there that cover up the sternum?* [Transfeminine subreddit, other scars related to gender-affirming care]
Tattooing scars	• *Does anyone know how long you should wait until you can tattoo your top surgery scars? If anyone has tattoos on their scars, how was it? Any advice would be great, thank you.* [Transmasculine subreddit, mastectomy scars]
	• *I lost a lot of definition and pigment during healing, and I think I would be happier with my nipples if they were enhanced by a tattoo artist.* [Transmasculine subreddit, mastectomy scars]
Concealing gender-affirming care scar origin	• *If they ask what your scars are, you could say you had gynecomastia, which is where a biological male grows breast tissue due to a hormone imbalance. You could say you had weight loss surgery.* [Transmasculine subreddit, mastectomy scars]

Theme 5: Positive Experiences Related to Scarring

Despite expressing negative views about scarring, posts communicated a strong preference for scarring over the alternative of not receiving gender-affirming care (Table 8). Although many posts expressed frustration at the failure of modern surgical and dermatologic techniques to eliminate scarring, no posts in the sample describe regretting receiving gender-affirming care. Some posts reported satisfaction with scar appearance and even expressed acceptance of and pride in scars. A few posts in the transmasculine subreddit described associating scarring with masculinity and finding scars to be a source of gender euphoria. Some users shared that they viewed scarring as a valuable part of the TGD experience or described scarring as a source of bonding over shared experience within TGD communities.

Table . Positive experiences related to scarring.

Sample codes	Selected quotes
Preference for gender-affirming care despite scar visibility	• *Obviously, the presence of scars is preferable to having dysphoria.* [Transmasculine subreddit, mastectomy scars]
	• *I know I am very privileged to be able to get top surgery, and I would definitely rather have scars than not have surgery.* [Transmasculine subreddit, mastectomy scars]
Scar improvement over time	• *Just know that most scars fade really well, especially after 3+ years.* [Transmasculine subreddit, mastectomy scars]
	• *My skin has come back from some pretty bad times with scarring before and even though I still relapse a bit it does give me hope that my skin is capable of clearing up again eventually.* [Transmasculine subreddit, other scars unrelated to gender-affirming care]
Acceptance of and pride in scarring	• *I refuse to be ashamed of my suicide scars, and I believe you'll have the power to dismiss the bad feelings about your scars one day, too.* [Transfeminine subreddit, scarring due to suicide attempt]
	• *I love the connection I feel with other trans people, especially other transmascs, when we compare surgical scars, share resources, and share old garments or pillows that we don't need anymore.* [Transmasculine subreddit, other scars related to gender-affirming care]

Discussion

Principal Findings

TGD people are at increased risk of having scars from a variety of causes, both related and unrelated to gender-affirming care. Some TGD people see their scars as a physical representation of their resilience and authentic journey, while for other TGD people, scarring presents a psychosocial burden.

This study shows that scars may generate stress and anxiety for TGD people in various ways. Most directly, many TGD people experience stressful physical health concerns related to scarring. These include but are not limited to (1) how to properly care for scars; (2) how to administer gender-affirming interventions, such as gender-affirming hormones or hair removal, in a manner that minimizes injury and scarring; and (3) whether scarring and associated injuries may render certain gender-affirming procedures infeasible for them to pursue.

The prevalence of these concerns highlights a need for safer, standardized scar care procedures. A variety of scar care strategies, many of which have not been clinically validated and some of which may not be safe or effective, were exchanged on the subreddit. Furthermore, posts expressed frustration and stress regarding the divergent nature of recommendations that users obtained on the internet, from peers, and from their care teams.

In addition, clinicians should fully and clearly address potential physical health concerns related to scarring, such as how to administer gender-affirming interventions in a safe manner that minimizes scarring, and how pre-existing scars may impact eligibility for subsequent gender-affirming treatments. Beyond creating psychological distress for TGD people, suboptimal communication with care teams may contribute to adverse physical consequences, as some posts reported experiences of severe scar tissue build-up and scar infection.

Aside from physical health concerns, posts also shared concerns about scar shape and visibility. These concerns stemmed from both internalized sources, such as negative body image and gender dysphoria due to scarring, and external sources, such as transphobic reactions from other people against physical scars [26].

In a number of instances, posts attributed these concerns to suboptimal communication with the TGD person's care team. In particular, posts shared experiences of clinicians providing inadequate information on scarring outcomes and scar care procedures. Users also reported not having the opportunity or the comfort to ask questions. Specific obstacles to effective communication with care teams included clinicians who spoke too quickly or unclearly, who were not accessible outside of appointments, and who were dismissive of concerns. These findings are consistent with published literature that found that the providers who were most positively reviewed by TGD patients had in-depth clinical knowledge of gender-affirming care [27], which is demonstrated at least in part by adequately explaining interventions and addressing any patient concerns.

Several posts described anxiety due to uncertainty about how to deal with scarring and interest in hearing peer advice about gender-affirming procedure selection and scar minimization strategies. In addition to providing accounts of personal experiences, commenters on these posts frequently offered words of encouragement and support. Interactions on the subreddits, despite the large number of posts and anonymity, were almost always positive. These findings are consistent with previous literature characterizing how the lived experience of peers is key in facilitating decision-making and providing social support for TGD people undergoing gender-affirming care [28].

While the majority of posts mentioned the impact of scarring on self-perception, some posts also cited external stressors, such as stigma against scars and gender-affirming care scars as a potential obstacle to identity concealment. Several posts expressed that the permanence and visibility of scarring was a deterrent to pursuing gender-affirming care. Other posts stated that family members or friends refused to support pursuing gender-affirming care due to the scarring that the treatments might cause, specifically mentioning transphobic comments that associated scarring from gender-affirming surgery with "mutilation."

Importantly, potential negative psychosocial consequences of scarring from gender-affirming care are not a valid argument against providing gender-affirming interventions. Many posts explicitly prefaced their negative sentiments toward scarring with a strong desire for receiving gender-affirming interventions over not having the scars. Posts primarily described frustration at the limitations of current treatment techniques in minimizing scarring, and sometimes also at the lack or the cost of scar-minimizing treatments. The negative psychosocial impacts of scarring documented in this study highlight the necessity for continued improvement in scar-minimizing procedures.

Finally, several posts did not share negative views of scarring, described satisfaction with scar outcomes, or expressed finding their scars to be either a source of gender euphoria or a way to bond with other TGD people. These findings suggest that preferences for scar visibility and appearance are heterogeneous and influenced by a complex interplay of psychosocial factors, similar to TGD people's diverse desires and priorities with regard to different types of gender-affirming interventions [29].

Limitations

This study is limited by a Reddit-based sample of experiences that may not be representative of the experiences of all TGD people. Reddit users are predominantly white young adults based in the United States [30]. Furthermore, certain themes may be overrepresented or underrepresented as a consequence of the Reddit algorithm, which may boost certain types of threads to the forefront of users' pages and thus result in these topics receiving disproportionate attention within the sample of threads analyzed in this study.

The majority of posts analyzed in this study expressed negative or neutral experiences with scarring. A growing body of literature suggests that a significant portion of TGD people may take pride in their scars and even derive gender euphoria from them [31]. Although it has been well-established that TGD

people have diverse body goals, many of which may not involve passing as cisgender, these sentiments were underexpressed in the data analyzed in this study, potentially because people who have negative experiences related to scarring are more likely to seek out the subreddit to vent or ask for advice. Thus, the prevalence results should be considered in the context of the limitations of analyzing organically occurring data.

Overall, this exploratory study suggests that there can be significant negative psychological impacts associated with scarring from various sources among TGD people. Future research ought to focus on further analyzing psychosocial experiences associated with specific types of scarring mentioned in this study in order to gain greater depth of understanding and develop effective scar treatment options.

Conclusions

TGD people on Reddit describe a wide variety of experiences with scarring. TGD people discussed scars secondary to gender-affirming care interventions, such as mastectomy and hair removal, and scarring not explicitly related to gender-affirming care, as in instances of self-harm and acne. Some TGD people expressed acceptance of and appreciation for their scars; for other TGD people, scarring represented a significant source of psychosocial stress. Internally, scarring may trigger complex feelings of gender dysphoria and trauma; externally, scarring may present an obstacle to identity concealment and render TGD people vulnerable to stigmatization and transphobic hostility.

These concerns underscore the need for improvement of scar minimization procedures and standardization of safer and culturally responsive scar care. Reddit threads also highlight a need for clearer and more thorough communication from gender-affirming care teams, as well as the potential positive impact of TGD peers in facilitating decision-making and providing social support. Future studies should analyze the psychosocial impacts of specific types of scarring among TGD people with greater depth to better inform medical innovation and health care policy.

Authors' Contributions

NYS conceived the idea for this project, collected and coded the threads, and drafted the manuscript. KK acted as the second independent coder and helped establish the initial codebook. ASK supervised the study and provided intellectual contributions and edits to the manuscript.

Conflicts of Interest

ASK declares royalties as editor of a McGraw Hill textbook on transgender and gender diverse health care and of an American Psychiatric Association textbook on gender-affirming psychiatric care. The authors declare no competing financial interests.

References

1. Lane M, Ives GC, Sluiter EC, et al. Trends in gender-affirming surgery in insured patients in the United States. Plast Reconstr Surg Glob Open 2018 Apr;6(4):e1738. [doi: 10.1097/GOX.0000000000001738] [Medline: 29876180]
2. Mani VR, Valdivieso SC, Hanandeh A, Kalabin A, Ramcharan A, Donaldson B. Transgender surgery - knowledge gap among physicians impacting patient care. Curr Urol 2021 Mar;15(1):68-70. [doi: 10.1097/CU9.0000000000000002] [Medline: 34084125]
3. Kamal K, Li JJ, Keuroghlian AS. Addressing the physical and mental impacts of postsurgical scarring among transgender and gender diverse people. LGBT Health 2023 May;10(4):259-262. [doi: 10.1089/lgbt.2022.0308] [Medline: 36796004]
4. Yu T, Song J, Yang S, Li J, Chen X, Yang J. Local steroid hormone injections into hypertrophic scars resulted in depression of the lesion site and radiated linear depigmentation and atrophy surrounding the lesion: a case report. J Cosmet Dermatol 2022 Oct;21(10):4703-4706. [doi: 10.1111/jocd.15019] [Medline: 35460306]
5. Moustakli E, Tsonis O. Exploring hormone therapy effects on reproduction and health in transgender individuals. Medicina (Kaunas) 2023 Nov 29;59(12):2094. [doi: 10.3390/medicina59122094] [Medline: 38138197]
6. Yeung H, Kahn B, Ly BC, Tangpricha V. Dermatologic conditions in transgender populations. Endocrinol Metab Clin North Am 2019 Jun;48(2):429-440. [doi: 10.1016/j.ecl.2019.01.005] [Medline: 31027550]
7. Rosenberg S, Tilley PJM, Morgan J. “I Couldn’t Imagine My Life Without It”: Australian trans women’s experiences of sexuality, intimacy, and gender-affirming hormone therapy. Sex Cult 2019 Sep;23(3):962-977. [doi: 10.1007/s12119-019-09601-x]
8. Lim SPR, Lanigan SW. A review of the adverse effects of laser hair removal. Lasers Med Sci 2006 Sep;21(3):121-125. [doi: 10.1007/s10103-006-0377-y] [Medline: 16816888]
9. Butler C, Joiner R, Bradley R, et al. Self-harm prevalence and ideation in a community sample of cis, trans and other youth. Int J Transgend 2019;20(4):447-458. [doi: 10.1080/15532739.2019.1614130] [Medline: 32999629]
10. Balsam KF, Molina Y, Beadnell B, Simoni J, Walters K. Measuring multiple minority stress: the LGBT people of color microaggressions scale. Cultur Divers Ethnic Minor Psychol 2011 Apr;17(2):163-174. [doi: 10.1037/a0023244] [Medline: 21604840]
11. Meneguzzo P, Zuccaretti D, Tenconi E, Favaro A. Transgender body image: weight dissatisfaction, objectification & identity - complex interplay explored via matched group. Int J Clin Health Psychol 2024;24(1):100441. [doi: 10.1016/j.ijchp.2024.100441] [Medline: 38292830]

12. Marks DH, Awosika O, Rengifo-Pardo M, Ehrlich A. Dermatologic surgical care for transgender individuals. Dermatol Surg 2019 Mar;45(3):446-457. [doi: 10.1097/DSS.0000000000001718] [Medline: 30789503]
13. Mañero I, Arno AI, Herrero R, Labanca T. Cosmetic revision surgeries after transfeminine vaginoplasty. Aesthetic Plast Surg 2023 Feb;47(1):430-441. [doi: 10.1007/s00266-022-03029-9] [Medline: 36002774]
14. Garg SP, Weissman JP, Reddy NK, et al. Patient-reported outcomes of scar impact: comparing of abdominoplasty, breast surgery, and facial surgery patients. Plast Reconstr Surg Glob Open 2022 Oct;10(10):e4574. [doi: 10.1097/GOX.0000000000004574] [Medline: 36246077]
15. Pehlivanidis S, Anderson JR. A scoping review of the literature exploring experiences in the trans and gender diverse community with chest binding practices. International Journal of Transgender Health 2024:1-27. [doi: 10.1080/26895269.2024.2316691]
16. McGuire JK, Doty JL, Catalpa JM, Ola C. Body image in transgender young people: findings from a qualitative, community based study. Body Image 2016 Sep;18:96-107. [doi: 10.1016/j.bodyim.2016.06.004] [Medline: 27352103]
17. Park JYE, Howren AM, Davidson E, De Vera MA. Insights on mental health when living with rheumatoid arthritis: a descriptive qualitative study of threads on the Reddit website. BMC Rheumatol 2020 Nov 26;4(1):62. [doi: 10.1186/s41927-020-00163-2] [Medline: 33292866]
18. Zhu J, Yalamanchi N, Jin R, Kenne DR, Phan N. Investigating COVID-19's impact on mental health: trend and thematic analysis of Reddit users' discourse. J Med Internet Res 2023 Jul 12;25:e46867. [doi: 10.2196/46867] [Medline: 37436793]
19. Uberoi M, Abdulcadir J, Ohl DA, Santiago JE, Rana GK, Anderson FWJ. Potentially under-recognized late-stage physical and psychosexual complications of non-therapeutic neonatal penile circumcision: a qualitative and quantitative analysis of self-reports from an online community forum. Int J Impot Res 2023 May;35(3):234-241. [doi: 10.1038/s41443-022-00619-8] [Medline: 36274189]
20. Saha K, Kim SC, Reddy MD, et al. The language of LGBTQ+ minority stress experiences on social media. Proc ACM Hum Comput Interact 2019 Nov;3(CSCW):32935081. [doi: 10.1145/3361108] [Medline: 32935081]
21. Hyde JS, Bigler RS, Joel D, Tate CC, van Anders SM. The future of sex and gender in psychology: five challenges to the gender binary. Am Psychol 2019;74(2):171-193. [doi: 10.1037/amp0000307] [Medline: 30024214]
22. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by Information Power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
23. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
24. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
25. Ford E, Shepherd S, Jones K, Hassan L. Toward an ethical framework for the text mining of social media for health research: a systematic review. Front Digit Health 2020;2:592237. [doi: 10.3389/fdgth.2020.592237] [Medline: 34713062]
26. Mathews A, Costa B, Mikkola A, Harcourt D. Scars: How Our Wounds Make Us Who We Are': Improving appearance-based stigma, conceptualisation of beauty and body esteem through a documentary. Scars Burn Heal 2023;9:20595131231205398. [doi: 10.1177/20595131231205398] [Medline: 38022894]
27. Valente PK, Paine EA, Mellman W, Rael CT, MacCrate C, Bockting WO. Positive patient-provider relationships among transgender and nonbinary individuals in New York City. Int J Transgend Health 2023;24(2):247-262. [doi: 10.1080/26895269.2022.2136814] [Medline: 37114109]
28. Kia H, Kenney KA, Abramovich A, Ferlatte O, MacKinnon KR, Knight R. "Nowhere else to be found": Drawing on peer support experiences among transgender and gender-diverse people to substantiate community-driven gender-affirming care. Soc Sci Med 2023 Dec;339:116406. [doi: 10.1016/j.socscimed.2023.116406]
29. Tristani-Firouzi B, Veith J, Simpson A, Hoerger K, Rivera A, Agarwal CA. Preferences for and barriers to gender affirming surgeries in transgender and non-binary individuals. Int J Transgend Health 2022;23(4):458-471. [doi: 10.1080/26895269.2021.1926391] [Medline: 36324882]
30. Proferes N, Jones N, Gilbert S, Fiesler C, Zimmer M. Studying Reddit: a systematic overview of disciplines, approaches, methods, and ethics. Social Media + Society 2021 Apr;7(2):20563051211019004. [doi: 10.1177/20563051211019004]
31. Roh JH, Jung I, Suh Y, Kim MH. A potential association between COVID-19 vaccination and development of Alzheimer's disease. QJM 2024 Oct 1;117(10):709-716. [doi: 10.1093/qjmed/hcae103] [Medline: 38806183]

Abbreviations

LGBT: lesbian, gay, bisexual, and transgender
r/FtM: transmasculine subreddit
r/MtF: transfeminine subreddit
TGD: transgender and gender diverse

Edited by R Rodriguez, R Dellavalle; submitted 29.05.24; peer-reviewed by JR Anderson, N Döring; revised version received 25.02.25; accepted 26.02.25; published 06.06.25.

Please cite as:
Sun NY, Kamal K, Keuroghlian AS
Investigating Experiences With Scarring Among Transgender and Gender Diverse People: Mixed Methods Study
JMIR Dermatol 2025;8:e62714
URL: https://derma.jmir.org/2025/1/e62714
doi:10.2196/62714

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org