JMIR Dermatology

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

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

Contents

Original Papers

XSL•FO
RenderX

Research Letters

Reviews

Case Report

XSL•FO
RenderX

Corrigenda and Addenda

Viewpoints

XSL•FO
RenderX

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

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

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

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

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

Abstract

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

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

XSL•FO
RenderX

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

Methods

Overview

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

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

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

Statistical Analysis

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

Ethical Considerations

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

Results

Cohort Description

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

XSL•FO
RenderX

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

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

[a]Topical corticosteroids and topical calcineurin inhibitors.

[b]Azathioprine, methotrexate, and mycophenolate mofetil.

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

[d]Dupilumab and tralokinumab.

Digital Readiness

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

XSL•FO
RenderX

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

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

Smartphone Photographs

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

XSL•FO
RenderX

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

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

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

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

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

XSL•FO
RenderX

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

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

Characteristics of Patients Who Took Smartphone Photographs of Skin Lesions

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

XSL•FO
RenderX

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

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

[a]OR: odds ratio.

XSL•FO
RenderX

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

[c]JAK: Janus kinase.

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

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

Discussion

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

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

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

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

Conflicts of Interest

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

References

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

XSL•FO
RenderX

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

Abbreviations

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

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

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

XSL•FO
RenderX

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

XSL•FO
RenderX

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

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

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

XSL•FO
RenderX

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

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

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

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

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

Methods

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

Device Fine-Tuning and Validation

Dataset

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

XSL•FO
RenderX

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

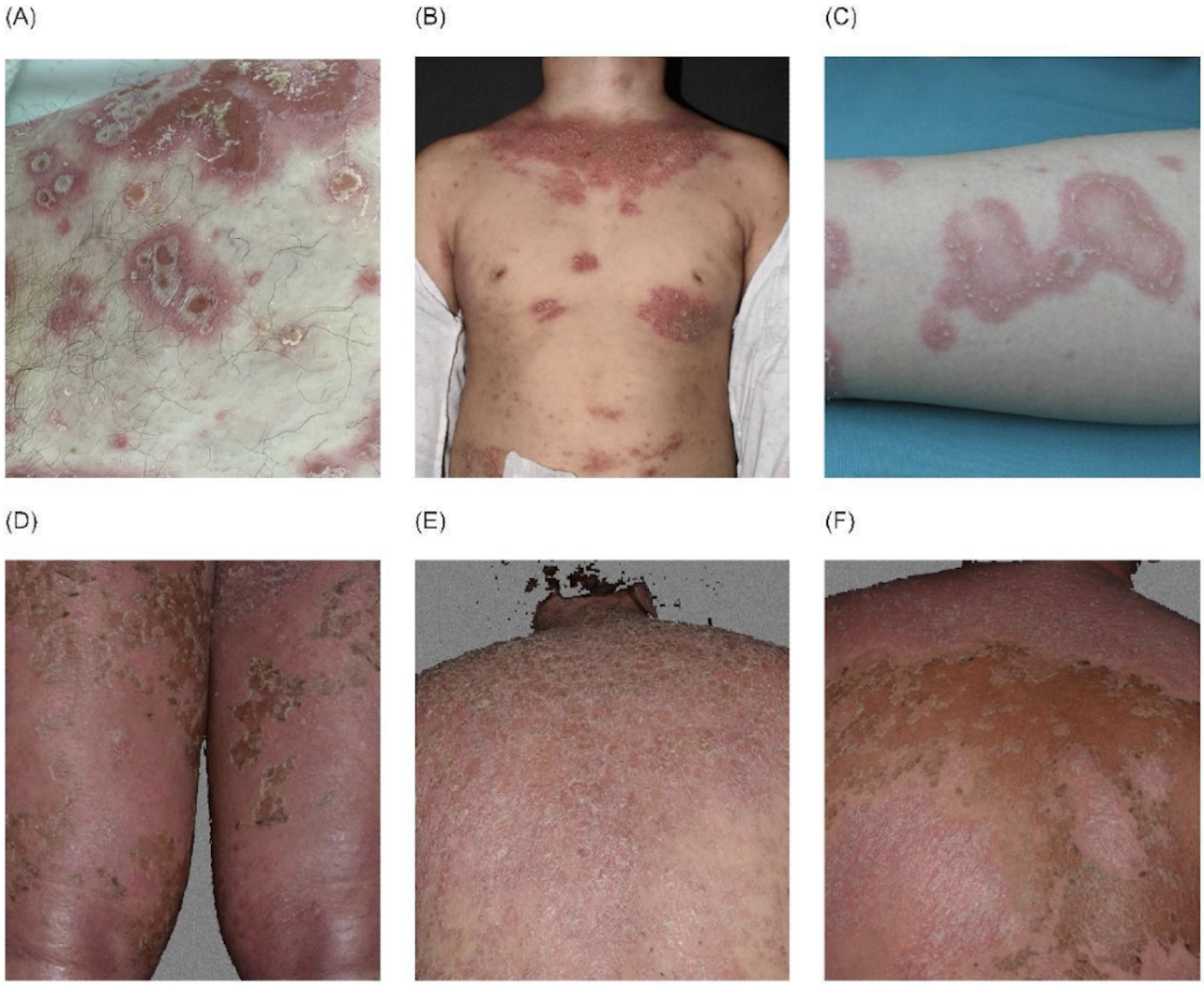


Table . Summary of patient demographics for the dataset of generalized pustular psoriasis images (N=4397).

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

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

Fine-Tuning

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

XSL•FO
RenderX

integrating the new GPP images with the existing validated dataset.

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

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

Validation

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

Clinical Study

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

Participant Selection and Training

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

Design

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

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

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

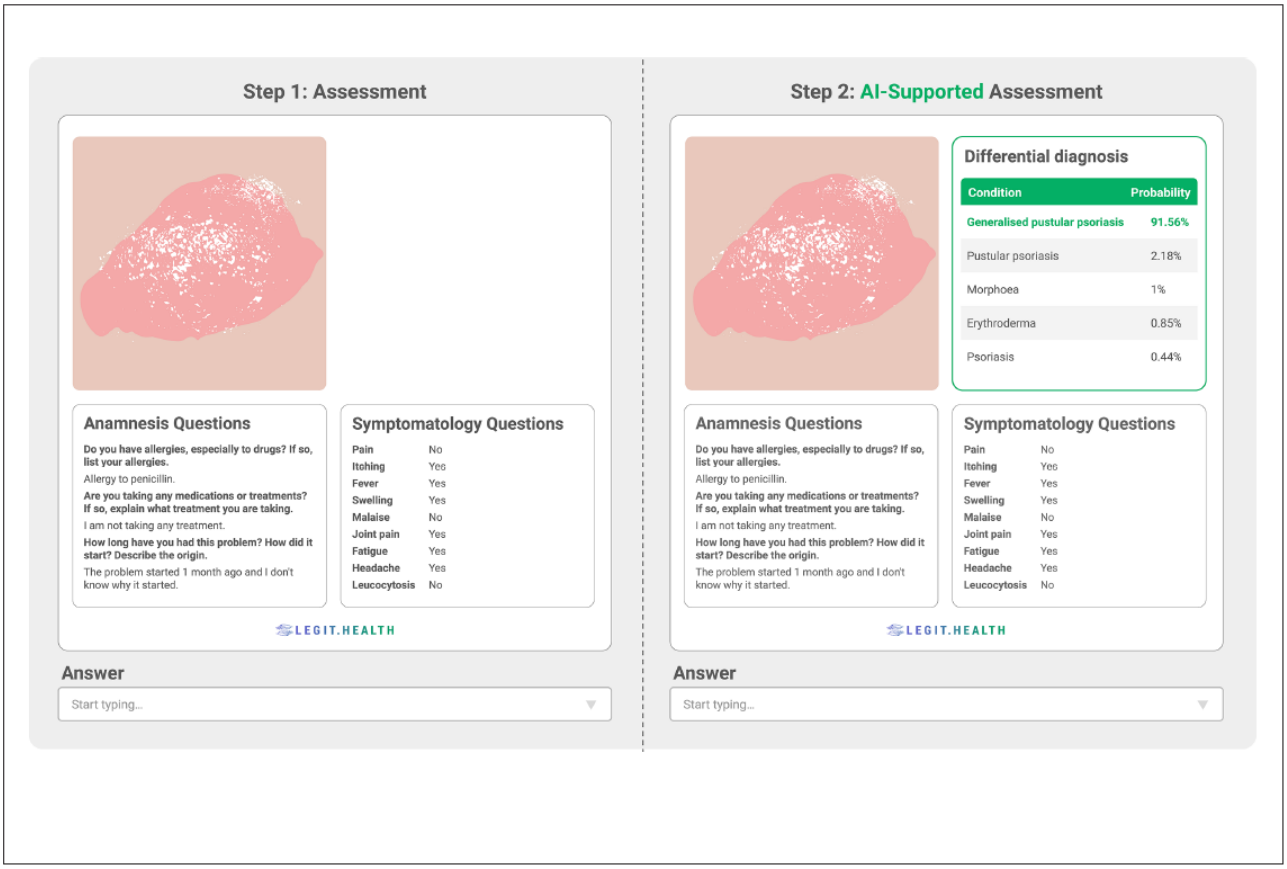
In both stages, participants had to make their choices from a list of 406 conditions, which could be easily explored by typing in either English or Spanish. This list, which was curated from Legit.Health's internal dataset, covered all pathologies encompassed by the device at different levels of detail,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

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

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

XSL•FO
RenderX

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



Ethical Considerations

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

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

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

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

Results

Device Performance

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

XSL•FO
RenderX

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

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

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

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

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

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

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

[c]Not applicable.

Clinical Evaluation

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

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

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

[a]HCP: health care practitioner.

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

XSL•FO
RenderX

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

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

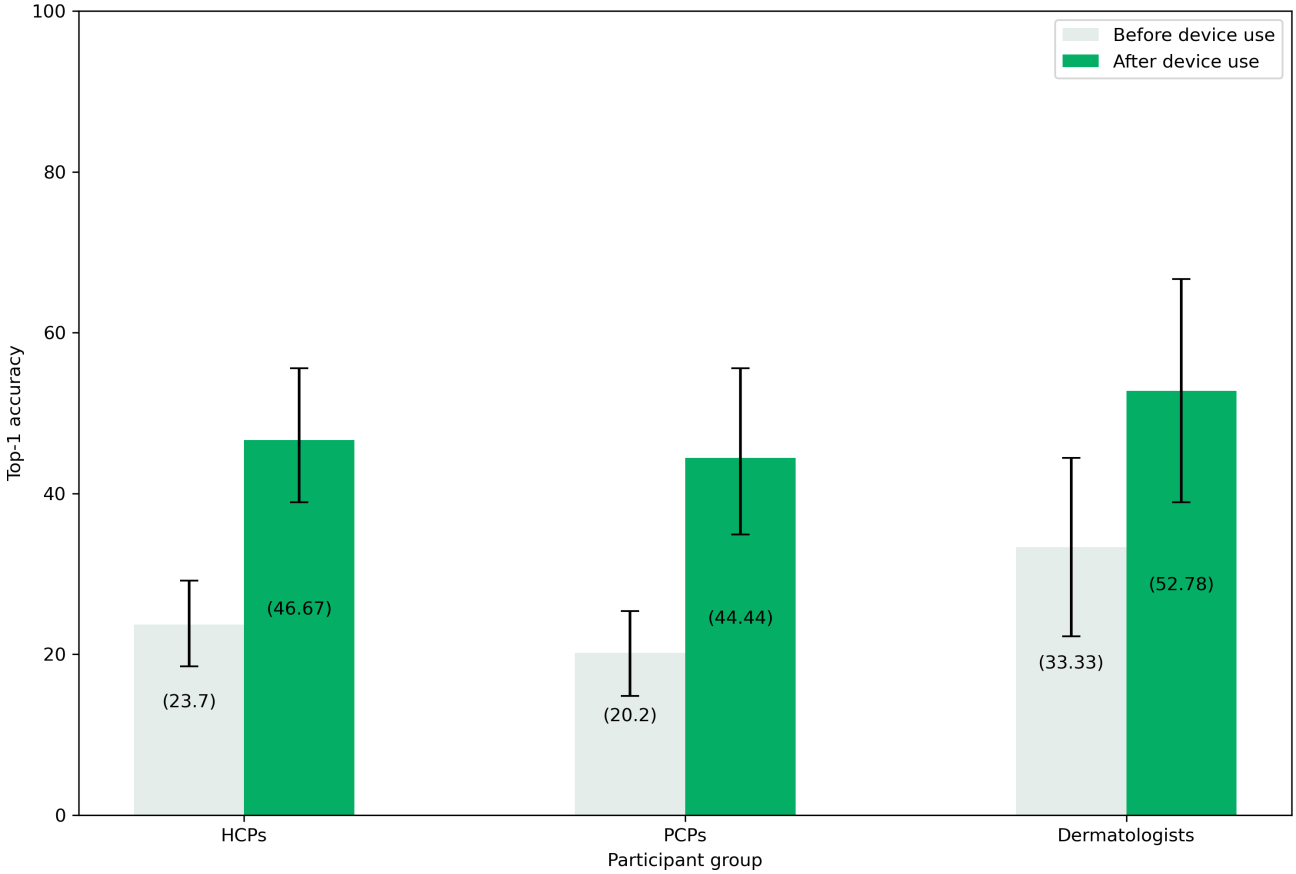
Table . Influence of the position of the correct condition on the ratio of changed diagnoses.

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

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

XSL•FO
RenderX

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



XSL•FO
RenderX

Table . Top-1 accuracy of all participants.

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

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

XSL•FO
RenderX

Table . Top-1 accuracy of primary care practitioners.

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

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

XSL•FO
RenderX

Table . Top-1 accuracy of dermatologists.

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

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

Discussion

Principal Findings

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

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

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

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

Limitations

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

XSL•FO
RenderX

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

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

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

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

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

Conclusions

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

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

Funding

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

Data Availability

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

Authors' Contributions

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

XSL•FO
RenderX

Conflicts of Interest

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

References

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

XSL•FO
RenderX

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

Abbreviations

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

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

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

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

XSL•FO
RenderX

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

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

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

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

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

Abstract

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

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

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

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

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

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

KEYWORDS

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

Introduction

Background

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

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

XSL•FO
RenderX

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

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

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

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

Literature Review

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

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

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

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

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

Methods

Participants

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

Exposure

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

Outcomes

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

XSL•FO
RenderX

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

Covariates

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

Statistical Analysis

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

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

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

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

Ethical Considerations

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

Results

Overview

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

XSL•FO
RenderX

Table . Characteristics of the study cohort [8].

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

XSL•FO
RenderX

Characteristics	Values
Yes	14,733 (4.72)

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

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

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

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

Depression

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

XSL•FO
RenderX

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

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

XSL•FO
RenderX

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

[a]aOR: adjusted odds ratio.

[b]Ref: reference.

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

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

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

Poor Mental Health Days

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

XSL•FO
RenderX

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

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

XSL•FO
RenderX

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

[a]aOR: adjusted odds ratio.

[b]Ref: reference.

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

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

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

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

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

Discussion

Principal Findings

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

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

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

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

XSL•FO
RenderX

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

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

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

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

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

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

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

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

Conclusions

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

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

XSL•FO
RenderX

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

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

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

Acknowledgments

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

Data Availability

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

Conflicts of Interest

None declared.

References

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

XSL•FO
RenderX

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

Abbreviations

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

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

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

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

XSL•FO
RenderX

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

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

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

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

Abstract

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

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

XSL•FO
RenderX

Methods

Study Design

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

Ethical Considerations

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

Data Collection

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

Independently Rated Measurements

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

Statistical Analysis

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

Results

Study Population

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

XSL•FO
RenderX

Table . Demographic characteristics.

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

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

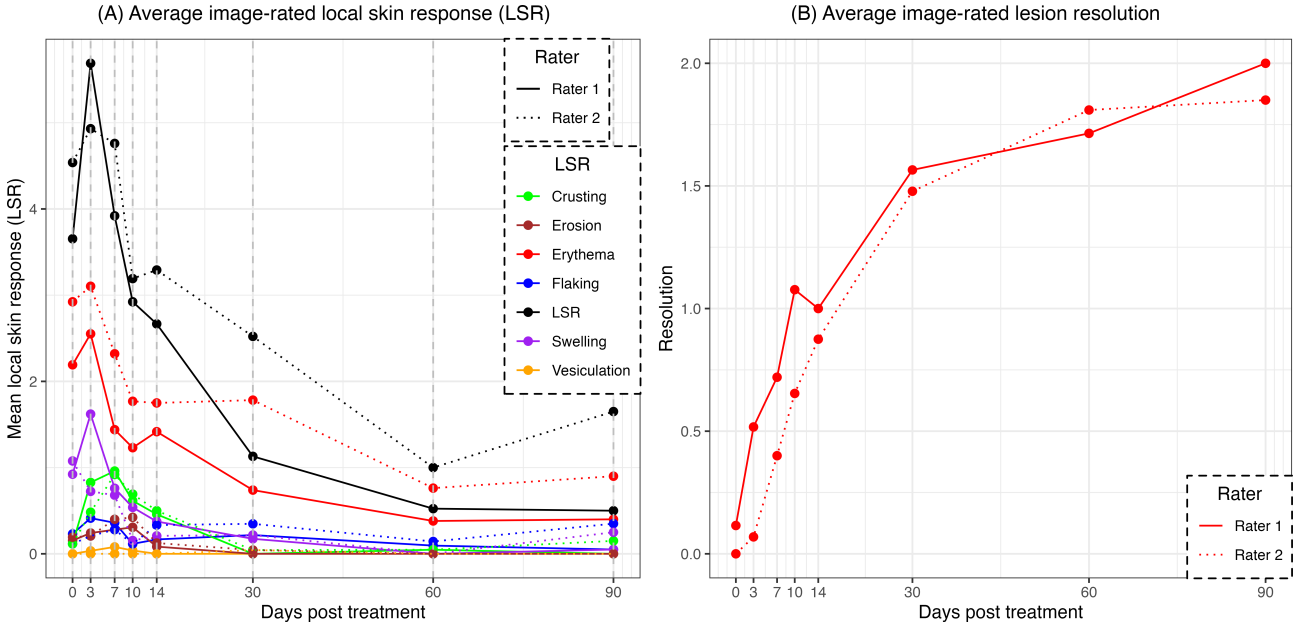
Cryosurgery Efficacy and Tolerability

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

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

XSL•FO
RenderX

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



Reliability of Image-Rated Metrics

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

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

XSL•FO
RenderX

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

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

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

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

[c]N/A: not applicable.

XSL•FO
RenderX

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

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

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

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

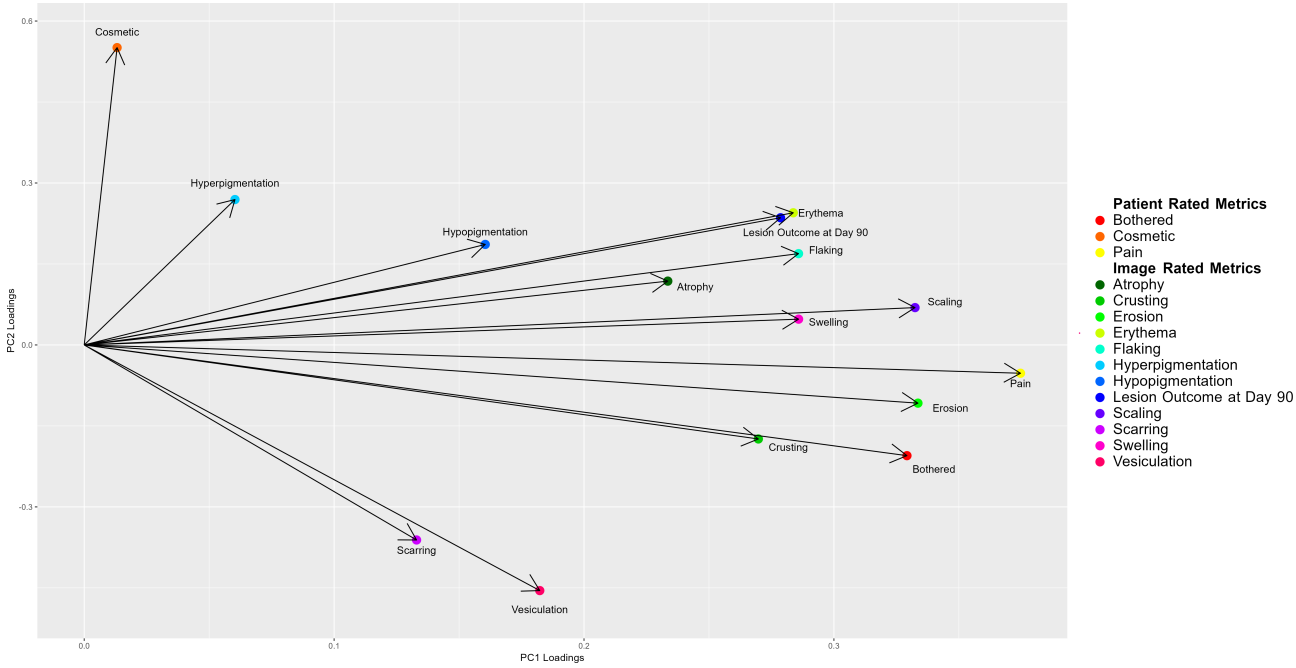
[c]N/A: not applicable.

Predictors of Lesion Response and Patient-Rated Side Effects

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

XSL•FO
RenderX

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



Discussion

Principal Findings and Comparison With Prior Work

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

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

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

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

XSL•FO
RenderX

Limitations

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

Conclusions

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

Acknowledgments

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

Funding

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

Data Availability

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

Authors' Contributions

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

Conflicts of Interest

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

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

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

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

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

References

XSL•FO
RenderX

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

Abbreviations

LSR: local skin response

XSL•FO
RenderX

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

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

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

XSL•FO
RenderX

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

XSL•FO
RenderX

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

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

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

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

Methods

Study Overview

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

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

To assess the frequency of side effects and other possible medication reactions associated with the use of compounded topical finasteride and minoxidil available via the Hims & Hers platform, we conducted a retrospective analysis of patient data collected during the course of routine clinical follow-up via the platform between April 1, 2021 and April 30, 2025. As this was an analysis of data gathered from individuals actively engaged in treatment, there was no control group.

XSL•FO
RenderX

Data Collection

The analysis included two sets of data. The first set of data consisted of responses to a follow-up check-in assessment sent to patients approximately 130 days following treatment initiation. The check-in queried patients about their treatment satisfaction and experience with side effects. To assess treatment satisfaction, patients were asked to indicate "yes" or "no" to the following prompt: "I'm happy with the way my treatment is working." To assess experience with side effects, patients were asked to respond "yes" or "no" to the following question: "Are you bothered by any side effects or other negative reactions from your treatment?" No other questions pertaining to side effects were included in the check-in.

The second set of data consisted of unprompted communications sent via in-app or web-based messaging from patients to their care team. Patients can send these unprompted messages at any time for review by the care team. These communications undergo continuous quality assurance by a clinical quality team that monitors patient messages in real-time for mention of side effects or other possible medication reactions and follows-up as appropriate. Their work includes validating the data to ensure that such events are appropriately recorded—for example, that the side effects and reactions reported are reported by patients in relation to one of the topical finasteride and minoxidil products highlighted in this analysis. Utilizing both sets of data ensured that all occurrences, both solicited and spontaneously reported by patients, were included in the analysis.

Statistical Analysis

Descriptive statistics using Google Colab (Mountain View, CA) were used to quantify the percentage of patients who reported satisfaction with treatment in their follow-up check-in, the percentage of patients who reported having been bothered by side effects or other negative reactions in their follow-up check-in, the percentage of patients who indicated experiencing a side effect or other possible medication reaction in messages to their care team, the percentage of patients who sought a higher level of care due to such a reaction, and the percentage of patients who discontinued treatment due to such a reaction. For results regarding the percentages of patients who reported treatment satisfaction and side effects in their follow-up check-in, the number of patients who completed a check-in is used as the sample size. For results regarding the percentage of patients who reported a side effect to their care team, the total number of patients prescribed a compounded topical finasteride product is used a the sample size. This is due to the fact that all patients had the ability to message their care team; thus, all patients can be included in the denominator.

Ethical Considerations

This study was approved by the WCG Institutional Review Board (Protocol 001, Review 20244102). All study procedures were conducted in accordance with the principles of the Declaration of Helsinki. The study protocol included a Waiver of Informed Consent, as all data analyzed were collected during the course of routine care and de-identified prior to analysis. Patients were not compensated for their participation in this study.

Results

Baseline Demographics

A total of 638,629 male patients with AGA received a prescription for a compounded topical finasteride and minoxidil product between April 1, 2021 and April 30, 2025. A total of 151,352 completed the follow-up check-in querying patients about their treatment satisfaction and experience with side effects.

The mean (SD) age of all patients who received a prescription for a compounded topical finasteride product (n=638,629) was 39.6 (11.9) years, while the mean (SD) age of those who completed the follow-up check-in (n=151,352) was 41.2 (11.8) years.

Treatment Satisfaction and Side Effects as Reported During Follow-Up Check-In

Overall, 121,615 (80.4%, n=151,352, 95% CI [80.2%, 80.6%]) patients who completed the follow-up check-in reported being satisfied with their treatment. A total of 4034 (2.7%, n=151,352, 95% CI [2.6%, 2.8%]) reported experiencing side effects.

Of the 151,352 patients who completed the follow-up check-in, 138,645 had been prescribed the 0.3% topical finasteride and 6% minoxidil spray; 10,774 had been prescribed the 0.3% topical finasteride, 7% minoxidil, 2.2% ketoconazole, and 0.2% biotin spray; and 1933 had been prescribed the 0.3% topical finasteride and 6% minoxidil serum. Table 1 outlines treatment satisfaction and the frequency of side effects reported by patients receiving each treatment.

Table . Treatment satisfaction and frequency of side effects reported by patients during follow-up check-ins.

	All topical finasteride treatments (n=151,352)	Topical finasteride (0.3%) and minoxidil (6%) spray (n=138,645)	Topical finasteride (0.3%), minoxidil (7%), ketoconazole (2.2%), and biotin (0.2%) spray (n=10,774)	Topical finasteride (0.3%) and minoxidil (6%) serum (n=1933)
Treatment satisfaction, n (%)	121,615 (80.4)	111,165 (80.2)	8900 (82.6)	1550 (80.2)
Experienced side effects, n (%)	4034 (2.7)	3716 (2.7)	251 (2.3)	67 (3.5)

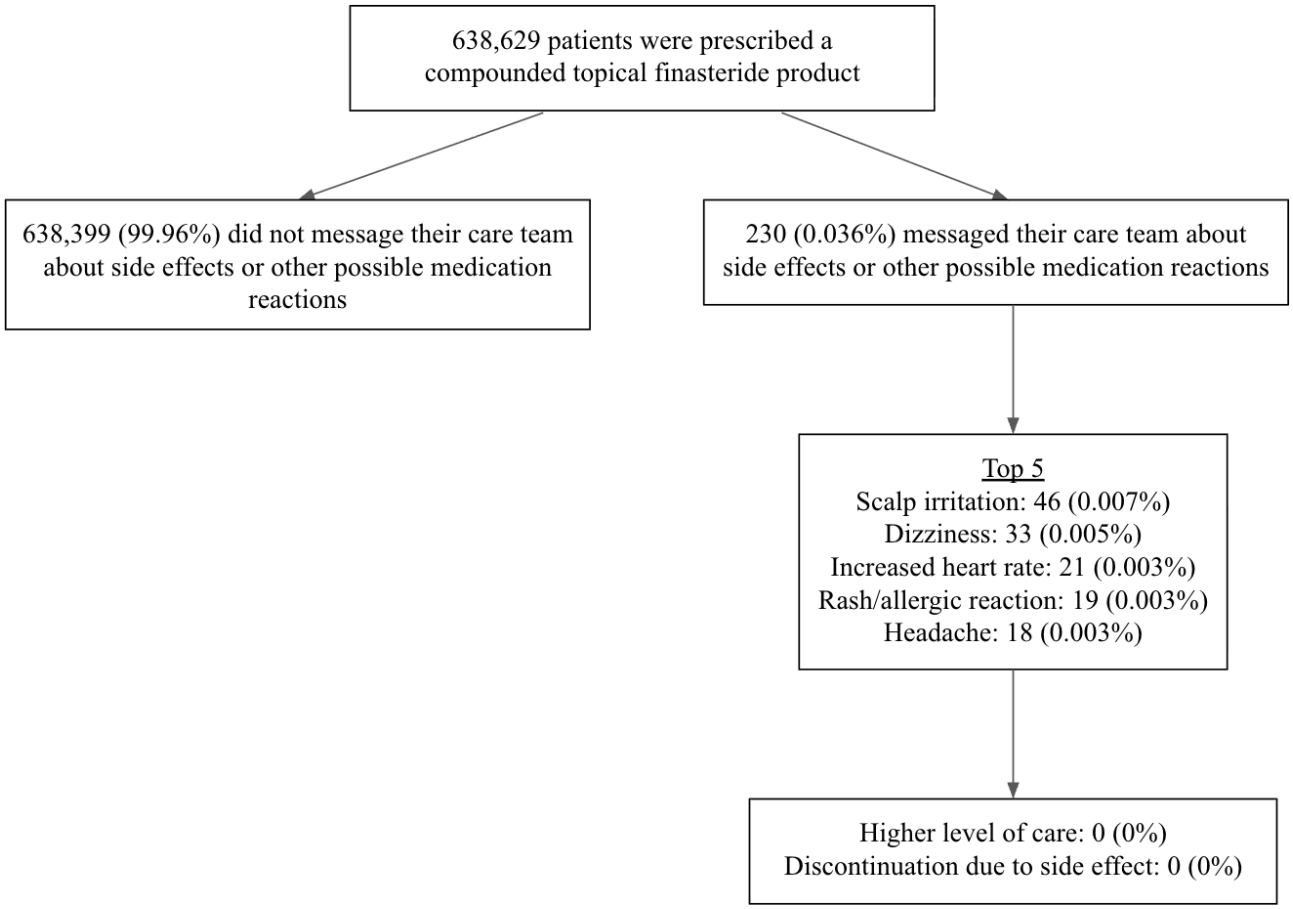
XSL•FO
RenderX

Side Effects and Other Possible Medication Reactions Reported in Patients' Communications to Their Care Team

Of the 638,629 patients prescribed a compounded topical finasteride and minoxidil product, 230 (0.04%, n=638,629, 95% CI [0.035%, 0.045%]) sent their care team messages concerning side effects or other possible medication reactions. The most commonly reported occurrences were scalp irritation (46/638,629, 0.007%, 95% CI [0.0064%, 0.0076%]), dizziness (33/638,629, 0.005%, 95% CI [0.0045%, 0.0055%]), increased heart rate (21/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]), rash or some allergic reaction (19/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]), and headache (18/638,629, 0.003%, 95% CI [0.0026%, 0.0035%]). Sexual side effects, specifically decreased libido and erectile dysfunction, were reported by 12/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Depression was reported by 13/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Anxiety was reported by 10/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]). Cognitive concerns were reported by 10/638,629 patients (0.002%, 95% CI [0.0017%, 0.0023%]).

No patients reported seeking a higher level of care (eg, emergency room or urgent care visit) related to a side effect or other possible medication reaction. No patients reported discontinuing treatment due to such an occurrence. During the study period, 1 spouse reported the death of a partner. Upon follow-up, no cause was identified and no causality was established. Figure 1 provides a summary of the side effects and other possible medication reactions reported by patients via messaging.

Figure 1. Schematic of side effects and other possible medication reactions reported by patients in messages to their care team.



Discussion

In this largest study of patient satisfaction and tolerability associated with the use of novel compounded formulations of topical finasteride and minoxidil, we found that 80% of those who completed a follow-up check-in reported satisfaction with treatment and less than 3% reported experiencing side effects. An additional 0.04% of patients sent their care team messages concerning side effects or other medication reactions. The most common reactions appeared to fall into one of two categories: (1) scalp irritation and rash, likely associated with the route of administration; (2) dizziness, increased heart rate, and headache, likely attributable to minoxidil acting as a vasodilator. Of note, sexual side effects, depression, anxiety, and cognitive concerns previously associated with oral finasteride were reported by just 0.002% of patients. There were no reports of "post-finasteride syndrome" [12].

Early clinical trials of 1 mg oral finasteride for the treatment of male AGA found that 3.8% of participants experienced adverse events possibly, probably, or definitely related to treatment, specifically decreased libido, erectile dysfunction, and ejaculation disorder, and 1.4% discontinued treatment due to such adverse events [13]. Trials of 2% topical minoxidil for the treatment of male AGA found that the most common adverse events were minor respiratory events such as colds and

XSL•FO
RenderX

respiratory infections (3.37% of participants), followed by dermatological reactions such as itching (1.94%) [14]. Trials of 5% topical minoxidil for the treatment of male AGA found that headache was the most frequently reported adverse drug reaction (1.7%), followed by dermatological reactions such as pruritus (1.1%) and rash (1.1%) [14].

A comparison of our findings to the findings of these historic studies reinforces the favorable tolerability profile of topical medications. Altogether, these results demonstrate that the novel compounded formulations of topical finasteride and minoxidil available to male patients with AGA via the Hims & Hers platform are associated with high satisfaction among patients and few reported side effects.

To date, few clinical trials have examined the use of topical finasteride in the treatment of male AGA [10,11,15]. A Phase III RCT by the Topical Finasteride Study Group in Europe found that 41.4% of participants reported treatment-emergent adverse events and 9.9% experienced treatment-related adverse events [10]. Another Phase III RCT in China found that 68.4% of participants reported treatment-emergent adverse events and 8.3% experienced treatment-related adverse events [11]. In both studies, the frequency of adverse events among participants using topical finasteride was similar to those using placebo. A retrospective study of 238 patients who received topical finasteride via a German direct-to-consumer teledermatology platform and completed a 6-week follow-up questionnaire found that 11.8% of patients reported adverse events after initiating the use of topical finasteride [15].

However, the aforementioned studies are methodologically limited by their relatively small sample sizes. This study, which included over 600,000 patients who were prescribed compounded topical finasteride in a real-world context, offers a much more robust and meaningful assessment of patient-reported satisfaction and tolerability associated with treatment.

There are limitations of this analysis. First, this was a retrospective analysis of data collected during the course of routine care and not an RCT, and therefore, we cannot confirm any causal relationships between patients' use of compounded topical finasteride and minoxidil and the reported outcomes. Second, we partly relied on data from an optional follow-up check-in questionnaire sent to patients approximately 130 days after treatment initiation. The rate of check-in completion was relatively low, with 23.7% of patients completing the check-in. This may indicate some selection bias, such that patients who were more engaged in or satisfied with their treatment may have been more likely to respond to the check-in and less likely to report side effects. Patients who reported side effects or other reactions to outside health care providers may not have been captured. Third, our reliance on retrospective data meant that we were unable to systematically examine other data of interest, such as the severity of and types of intervention sought for side effects and other medication reactions reported by patients.

However, our analysis also had several strengths. First, our sample size was impressive, with 638,629 patients prescribed a compounded topical finasteride product, all of whom had the ability to communicate with their care team at any time during the course of treatment, and 151,352 of whom completed the follow-up check-in that specifically queried patients about their experience with treatment and side effects. Second, our analysis utilized real-world data. The use of real-world data enables clinicians and researchers to better understand how patients experience treatment in their daily lives, thus increasing the generalizability of results. Third, in addition to relying on the optional follow-up check-in questionnaire to collect data on patient-reported side effects, we were also able to utilize unsolicited patient communications concerning side effects and other possible medication reactions. Having these additional data increased the likelihood that we were able to capture all occurrences reported by patients.

In conclusion, our analysis found that patients prescribed novel formulations of compounded topical finasteride and minoxidil for the treatment of AGA via a national telehealth platform tolerated the treatment well. The majority reported satisfaction with the treatment, and there were few reports of side effects. Future research should include RCTs to assess the efficacy, safety, and tolerability of topical finasteride. Together, this work may help provide more treatment options for those with AGA.

Acknowledgments

No generative AI or AI-assisted technologies were utilized in performing the study or preparing the manuscript.

Funding

This research was supported by Hims & Hers Health, Inc., which did not restrict data analysis, interpretation, or publication.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to commercial restrictions; however, they may be available from the corresponding author upon reasonable request.

Conflicts of Interest

JY, SM ME, and PC are full-time employees of Hims & Hers Health, Inc. JK and AM serve as advisors to Hims & Hers Health, Inc. Hims & Hers Health, Inc. had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

XSL•FO
RenderX

References

1. Lolli F, Pallotti F, Rossi A, et al. Androgenetic alopecia: a review. Endocrine 2017 Jul;57(1):9-17. [doi: 10.1007/s12020-017-1280-y] [Medline: 28349362]
2. Androgenetic alopecia. MedlinePlus. URL: https://medlineplus.gov/genetics/condition/androgenetic-alopecia/#resources [accessed 2026-01-13]
3. Aukerman EL, Jafferany M. The psychological consequences of androgenetic alopecia: a systematic review. J Cosmet Dermatol 2023 Jan;22(1):89-95. [doi: 10.1111/jocd.14983] [Medline: 35403805]
4. Olsen EA, Dunlap FE, Funicella T, et al. A randomized clinical trial of 5% topical minoxidil versus 2% topical minoxidil and placebo in the treatment of androgenetic alopecia in men. J Am Acad Dermatol 2002 Sep;47(3):377-385. [doi: 10.1067/mjd.2002.124088] [Medline: 12196747]
5. Nestor MS, Ablon G, Gade A, Han H, Fischer DL. Treatment options for androgenetic alopecia: efficacy, side effects, compliance, financial considerations, and ethics. J Cosmet Dermatol 2021 Dec;20(12):3759-3781. [doi: 10.1111/jocd.14537] [Medline: 34741573]
6. Andy G, John M, Mirna S, et al. Controversies in the treatment of androgenetic alopecia: the history of finasteride. Dermatol Ther 2019 Mar;32(2):e12647. [doi: 10.1111/dth.12647] [Medline: 30253001]
7. Gupta AK, Bamimore MA, Williams G, Talukder M. Finasteride use: evaluation of depression and suicide risk. J Cosmet Dermatol 2025 Mar;24(3):e70102. [doi: 10.1111/jocd.70102] [Medline: 40082195]
8. Neubauer Z, Ong MM, Lipner SR. No increased risk of mood disorders in male androgenetic alopecia patients treated with dutasteride and finasteride: a retrospective cohort database study. J Am Acad Dermatol 2025 Jul;93(1):197-199. [doi: 10.1016/j.jaad.2025.03.068] [Medline: 40158535]
9. Lee SW, Juhasz M, Mobasher P, Ekelem C, Mesinkovska NA. A systematic review of topical finasteride in the treatment of androgenetic alopecia in men and women. J Drugs Dermatol 2018 Apr 1;17(4):457-463. [Medline: 29601622]
10. Piraccini BM, Blume-Peytavi U, Scarci F, et al. Efficacy and safety of topical finasteride spray solution for male androgenetic alopecia: a phase III, randomized, controlled clinical trial. J Eur Acad Dermatol Venereol 2022 Feb;36(2):286-294. [doi: 10.1111/jdv.17738] [Medline: 34634163]
11. Zhou C, Yang B, Zeng H, et al. Efficacy and safety of topical finasteride spray solution in the treatment of Chinese men with androgenetic alopecia: a phase III, multicenter, randomized, double-blind, placebo-controlled study. Chin Med J (Engl) 2025 Mar 17. [doi: 10.1097/CM9.0000000000003495] [Medline: 40090937]
12. Diviccaro S, Melcangi RC, Giatti S. Post-finasteride syndrome: an emerging clinical problem. Neurobiol Stress 2020 May;12:100209. [doi: 10.1016/j.ynstr.2019.100209] [Medline: 32435662]
13. Propecia (finasteride): prescribing information. : Merck & Co, Inc; 2012 URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/020788s020s021s023lbl.pdf [accessed 2026-01-13]
14. Rogaine (minoxidil): product monograph. : Johnson & Johnson, Inc; 2017 URL: https://pdf.hres.ca/dpd_pm/00040937.PDF [accessed 2026-01-13]
15. Abeck F, Hansen I, Kött J, et al. Patient-reported outcomes of topical finasteride/minoxidil treatment for male androgenetic alopecia: a retrospective study using telemedical data. J Cosmet Dermatol 2024 Sep;23(9):2956-2963. [doi: 10.1111/jocd.16360] [Medline: 38713003]

ABBREVIATIONS

AGA: androgenetic alopecia
FDA: Food and Drug Administration
RCT: randomized controlled trial

Edited by R Dellavalle; submitted 23.Sep.2025; peer-reviewed by EM Hassan, J Jasso-Olivares; revised version received 08.Dec.2025; accepted 17.Dec.2025; published 22.Jan.2026.

Please cite as:
Yu J, Mochida S, Emery M, Carroll P, Ko J, Mostaghimi A
Patient Satisfaction, Side Effects, and Other Reactions Reported by Adult Men Prescribed Compounded Topical Finasteride via a National Telehealth Platform: Retrospective Analysis of Real-World Data
JMIR Dermatol 2026;9:e84676
URL: https://derma.jmir.org/2026/1/e84676
doi:10.2196/84676

© Jessica Yu, Sachie Mochida, Michele Emery, Patrick Carroll, Justin Ko, Arash Mostaghimi. Originally published in JMIR Dermatology (http://derma.jmir.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative

XSL•FO
RenderX

Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Community-Based Teledermatology for Urgent Suspected Skin Cancer: Health Economic Cost-Comparison and Discrete Event Simulation Study

Tim C H Hoogenboom[1], PhD; Pablo García Martínez[2], PhD; Piyush Mahapatra[1], MA, MBBS, FRCS; Nurul Ain Nizar[1], MPH

[1]Open Medical Ltd, 97-107 Uxbridge Rd, London, United Kingdom

[2]Department of Dermatology, Salford Care Organisation part of the Northern Care Alliance NHS Foundation Trust, Manchester, UK, Manchester, United Kingdom

Corresponding Author:
Nurul Ain Nizar, MPH
Open Medical Ltd, 97-107 Uxbridge Rd, London, United Kingdom

Abstract

Background: The increasing incidence and financial burden of skin cancer place immense pressure on the UK's National Health Service (NHS). Systemic challenges, including dermatologist shortages and long waiting lists, complicate timely assessment of skin lesions for patients under the urgent suspected cancer pathway. While teledermatology offers an innovative solution compared to traditional face-to-face appointments, standard teledermatology models still face limitations in addressing health care access barriers. Community-based decentralized models may reduce such barriers, but the cost and operational impact of such specific models remain largely underresearched.

Objective: This study evaluated the differences in financial cost to the NHS and patient waiting times at the Northern Care Alliance NHS Foundation Trust by comparing a community-based teledermatology model using Pathpoint eDerma against the Trust's standard-of-care for patients in the urgent suspected skin cancer pathway.

Methods: This study used an ambidirectional design involving 2 distinct analyses. The cost comparison analysis (CCA) compared costs incurred under the teledermatology model (intervention arm, n=563) against the Trust's standard care, represented by a synthetic comparator arm (n=4011). The discrete event simulation (DES) modeled the operational impact on patient waiting times over a 1-year period. Data for the intervention arm were collected prospectively from December 2022 to May 2023 for CCA and up to November 2023 for DES, while comparator data were collected retrospectively from September 2021 to December 2022. Publicly available resource costs were incorporated to ensure the robustness of the analyses.

Results: The community-based teledermatology model was associated with significant improvements in both cost to the NHS and patient waiting times. The CCA revealed a mean cost saving of £45 (£1=US $1.24) per referral (95% CI £22-£60; *P*<.001). This cost saving was associated with a 26% reduction in the proportion of patients requiring a full diagnostic biopsy, falling from 48% (1925/4011) in standard care to 22% (124/563) in the teledermatology model as well as time savings in face-to-face clinics and administration. Furthermore, the DES demonstrated that, on average, the teledermatology pathways decreased the time to reach a clinical diagnosis by 9.90 (95% CI 9.64-10.16) days; to communicate a diagnosis to patients by 54.18 (95% CI 50.76-57.61) days; and to reach a histopathological diagnosis by 62.8 (95% CI 59.76-65.83) days compared to standard care.

Conclusions: The implementation of the community-based teledermatology model appears to be a highly effective, cost-efficient strategy associated with shortened patient journeys. The intervention showed a faster initial triage phase, but the study identified the histopathology process as the next major systemic constraint that could deter further pathway efficiency. Achieving timely diagnosis for all patients, including those requiring diagnostic biopsies, will necessitate continued strategic investment in innovative technologies to accelerate this downstream process.

(JMIR Dermatol 2026;9:e86402) doi:10.2196/86402

KEYWORDS

teledermatology; urgent suspected cancer; cost-effectiveness; discrete event simulation; patient waiting times; community-based; NHS; National Health Service

XSL•FO
RenderX

Introduction

The incidence of skin cancer continues to rise steadily in the United Kingdom. The number of nonmelanoma skin cancer diagnoses surpasses the combined total of the 4 most prevalent nonskin cancers: breast, prostate, lung, and colorectal cancer [1], with malignant melanoma currently listed as the fifth most common cancer [2,3]. The increase in public awareness of skin cancers [4] and the lack of training for general practitioners (GPs) to confidently assess skin lesions [5] have also led to a surge in suspicious skin lesion referrals in the past decade [6]. Skin cancer places a significant financial strain on the National Health Service (NHS), with associated costs in England now projected to reach between £338 million (£1=US $1.24) and £465 million, a significant increase from the previous projection of over £180 million in 2020 [4]. Treatment costs for advanced skin cancer, particularly metastatic melanoma, can be substantial, exceeding £200,000 per case, while early-stage disease has a drastically lower health impact and treatment cost, highlighting the importance of early diagnosis both for the patient and the NHS [7,8].

The rising burden necessitates early detection and management, which, in turn, has placed immense pressure on dermatology services. The NHS is already grappling with significant resource constraints and a national shortage of specialist dermatologists [9]. With almost a quarter of consultant posts unfilled, over 380,000 people are waiting more than 18 weeks for a dermatology appointment [9]. The urgent suspected cancer (USC) pathway aims to provide timely access to care and minimize the risk of cancer progression following initial lesion identification [10]. The aim is achieved by prioritizing patients referred under this pathway to meet the 28-day “Faster Diagnosis Standard” (FDS) [10]; however, national statistics reveal a low conversion rate, indicating that many of these skin lesions are benign [11]. The prioritization of referrals under this pathway has strained resources and subsequently delayed care for patients with other serious, noncancerous conditions like eczema and psoriasis [12].

The systemic challenges have led to the growing adoption of teledermatology as an innovative solution to transform service delivery, although the traditional model of lesion assessment through face-to-face (F2F) clinic appointments remains common [10,13]. Teledermatology is most frequently delivered through a store-and-forward (SAF) model, which involves capturing and transmitting skin lesion images and clinical data on digital platforms for remote assessment by dermatologists [14]. A body of evidence has demonstrated the general effectiveness of this SAF model in addressing workforce constraints, reducing waiting times, and improving patient access [15,16].

However, there are operational limitations reported with the conventional SAF model, specifically within the image capture process [10,13,14,17]. Models relying on image capture performed at the secondary care provider level maintain high image quality but limit efficiency gains by requiring patients to travel to a centralized hospital facility [10,13,14,17]. Conversely, referrer (GP) image capture models often suffer from lower image quality, reduced standardization, and limited uptake due to existing constraints on GP capacity [10,13,14,17]. Consequently, the effectiveness of the traditional SAF model is still limited by existing challenges in health care access, including geographical distance and avoidant behaviors that discourage patients from visiting centralized secondary care facilities [10,13,14,17]. Single-site SAF models in particular are also limited in their ability to support regional strategic plans that seek to pool resources and support smaller sites.

To mitigate the challenges inherent in conventional teledermatology, we implemented a community-based teledermatology model, using the Pathpoint eDerma digital platform (Open Medical Ltd). This model was specifically designed to ensure high-quality image capture delivered by a dedicated health care assistant while providing timely and accessible care closer to patients’ homes. This approach aligns directly with the NHS’s strategic vision to shift “from hospital to community,” thereby improving access, quality, standardization, and uptake of the teledermatology service [18]. The platform’s data architecture and integration capabilities are also future-proofed to support delivery of more comprehensive regional care models [19].

Currently, the literature lacks a robust, real-world evaluation of a community-based model within the urgent suspected skin cancer pathway in the NHS. This study sought to address this gap by providing a comprehensive analysis of the cost implications and operational efficiency of the community-based teledermatology model from a health care system standpoint. In this study, we compare a community-based teledermatology model using Pathpoint eDerma against the trust’s standard of care for patients in the urgent suspected skin cancer pathway to evaluate the associated cost implications for the Northern Care Alliance NHS Foundation Trust (NCAFT) and patients’ waiting times.

Methods

Study Setting

The study was conducted within the dermatology department at NCAFT. This trust is one of the largest NHS providers in the country, which delivers health care to over 1 million people across Salford, Oldham, Rochdale, and Bury [20]. As part of a wider teledermatology project launched in December 2022, NCAFT established its first community-based teledermatology service. This model used a community diagnostic center (CDC) in Bury, allowing patients to have their skin lesions photographed closer to home rather than at the traditional location, Salford Royal Hospital. The study population included patients referred by their GPs to the urgent suspected skin cancer pathway (formerly known as the 2-week wait [2WW]), who presented with 2 or fewer suspicious lesions.

Study Design

This study used a comparative, ambidirectional design to evaluate a teledermatology model against local standard care. It consisted of 2 distinct analyses: a cost comparison analysis (CCA) to evaluate the budgetary implications for the NHS trust and a discrete event simulation (DES) to model the differences in patient waiting times. The study is considered ambidirectional

XSL•FO
RenderX

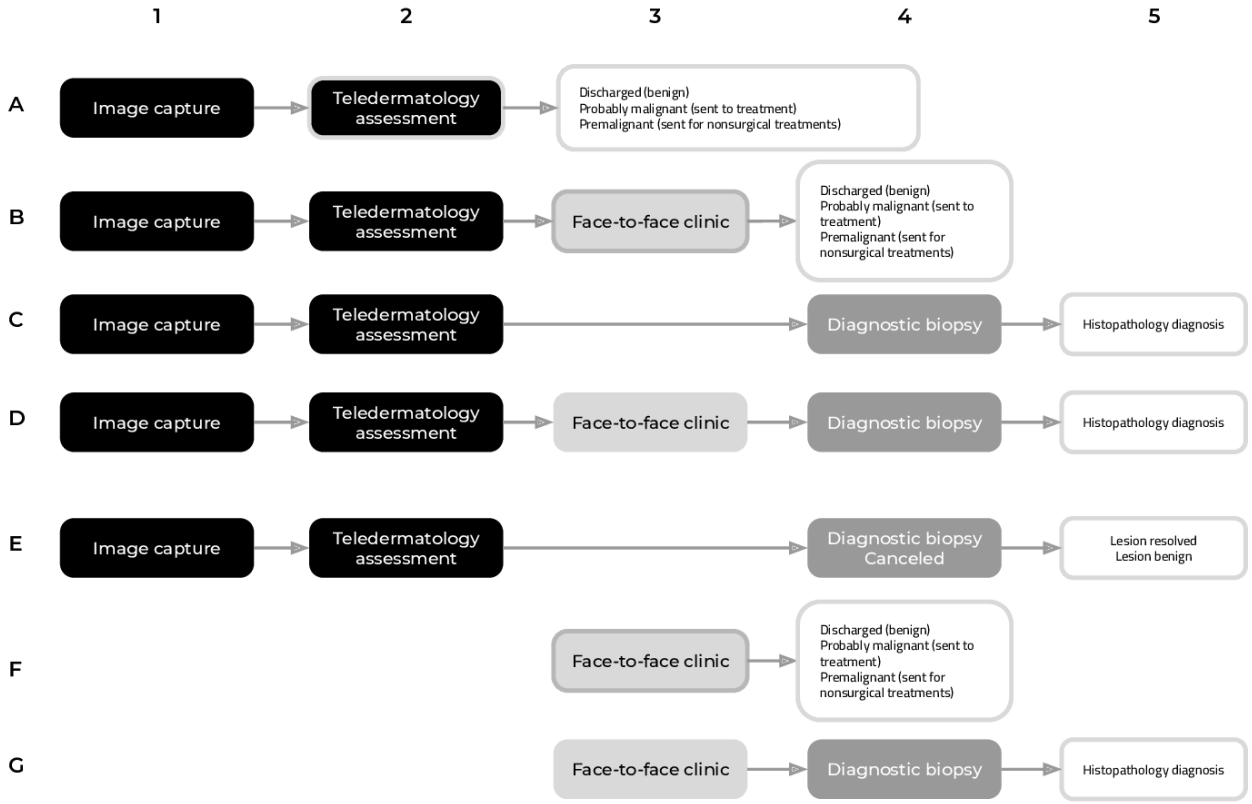
as data for the intervention arm were collected prospectively, while data for the comparator arm were collected retrospectively from historical records. Different time horizons were used in the analyses and are detailed in the data collection section.

Intervention and Comparator Arms

Each arm of the study comprised several distinct pathways representing the patient journey from referral to diagnosis. Under the intervention arm, a patient's journey began with an image capture appointment at the CDC. There, a trained health care assistant captured high-quality macroscopic and dermatoscopic images of the lesion(s) and uploaded them, along with referral notes and a digitally completed patient questionnaire, to the eDerma platform. A consultant dermatologist then remotely reviewed this information to determine the clinical outcome. This process defined 5 distinct intervention pathways (A-E).

In contrast, the comparator arm reflected the local standard-of-care pathways (F and G), where patients attended a F2F clinical appointment with a consultant dermatologist for an in-person dermoscopic examination. Following this consultation, the patient was either discharged (pathway F) or scheduled for a diagnostic biopsy (pathway G). The outcome of a canceled diagnostic biopsy (pathway E) is also a potential event for patients following the standard-of-care F2F pathway, though less frequent and was not observed in the dataset used for this study (Figure 1).

Figure 1. Potential pathways under the intervention and comparator arms.



Data Collection

For the CCA, prospective data for the teledermatology model's intervention arm were sourced from the Pathpoint eDerma platform, spanning from December 2022 to May 2023. Conversely, retrospective data for the standard care comparator arm were collected from historical NCAFT records from September 2021 to December 2022 and supplemented with public NHS protocols. This process created a synthetic control group for the comparator arm, enabling a direct cost comparison between the 2 pathways. Due to the historical nature of the records, this cohort was reconstructed using a triangulated audit approach. A manual review of 70 F2F clinical cases was conducted to determine granular pathway distributions (F and G), which were then calibrated against a larger trust-level dataset of over 5000 urgent referrals to establish representative waiting times. The comparator arm's period was also chosen to ensure it reflected a post-lockdown health care environment. National NHS data indicated that by late 2021, USC referrals for skin lesions had returned to, or exceeded, prepandemic volumes [21]. Consequently, the analysis assumes that patient reluctance to seek treatment was not a significant factor in the diagnostic timelines observed during the comparator period. Additionally, we included resource utilization data, such as staff labor and procedure costs (eg, diagnostic biopsies), which were adjusted to 2023 values for inflation using the Personal Social Services Research Unit manual [22] and the NHS National Cost Collection schedule [23].

The same retrospective data were then repurposed for the DES to inform the operational parameters of the "as-is" model, effectively defining the simulation's structure and serving as the baseline for comparison. To increase the real-world accuracy of the simulation, we integrated additional data to define specific

XSL•FO
RenderX

parameters. These included patient pathway proportions and communication methods from the eDerma platform, patient inflow rates from NHS England's Cancer Waiting Times Statistics [21], and staff numbers estimated from NHS Hospital & Community Health Service monthly workforce statistics [24]. The simulation of the intervention arm was further strengthened by incorporating an additional 6 months of prospective data, extending its data collection period up to November 2023.

Statistical Analyses

To ensure the study's independence and meet the requirements of the teledermatology project's funders, the CCA and DES were conducted by an external health technology consultancy, Health Tech Enterprise. Although the analyses were performed independently, the authors of this study provided the necessary datasets and contextual information. They also reviewed the methodology to ensure it accurately reflected the specific teledermatology project under evaluation.

The health economic CCA compared and quantified the budgetary impact of the teledermatology model against standard care from the perspective of the NHS. The analysis focused on direct medical costs incurred from the point of referral to the point of diagnosis, including health care staff labor (Bands 3, 5, 9, and Consultant levels), diagnostic biopsy procedures, and technology platform licensing. A health economic model was constructed in Microsoft Excel (Microsoft 365, v2310; Microsoft Corp) to determine the average cost per referral for both the intervention and comparator arms. This was followed by an incremental analysis to account for the SEs associated with health care staff involvement.

To test whether the difference in mean costs was statistically significant, an unpaired *t* test was performed. Further, we conducted a probability sensitivity analysis with 1000 model iterations and a deterministic sensitivity analysis to account for uncertainty in input parameters and evaluate how changes in key parameters, such as the number of biopsies, affected the results.

Moreover, we performed a DES to evaluate the impact on patient waiting times. The skin cancer screening pathways were modeled as an agent-based DES using SIMUL8 software (Professional Edition) over a 1-year time horizon. The model's capacity was based on the availability of health care staff and accounted for the seasonality of patient inflow. The simulation was run for five 1-year instances, with each patient "agent" entering the model upon referral and exiting once a diagnosis was communicated.

The simulation measured 3 key end points, representing different waiting periods: the average time from referral to clinical diagnosis, the average time from referral to histopathological diagnosis (for patients requiring a diagnostic biopsy), and the average time to final diagnosis communication.

Model Assumptions

The CCA was built upon several assumptions regarding patient characteristics and resource utilization to enable a direct cost comparison between the intervention and comparator pathways. It was assumed that patient characteristics were the same in both the intervention and comparator arms. For the standard-of-care (comparator) arm, all consultations were modeled as F2F. Specific durations were assigned to clinical activities; the initial F2F consultation with a consultant dermatologist was set at 20 minutes, while a follow-up clinic appointment after the eDerma assessment was set at 15 minutes. The resource cost model assumed that a photography appointment, conducted by an NHS Band 3 health care assistant at a local GP practice, took 20 minutes, and the remote eDerma assessment by a consultant dermatologist took approximately 4.97 minutes. Furthermore, it was assumed that any dermoscopy examination performed in the comparator arm was included within the F2F consultation time, eliminating the need for an additional photography assessment.

Similarly, the DES relied on several assumptions to model patient flow and system capacity under controlled conditions. The simulation modeled high-risk patients referred via the 2WW pathway with 2 or fewer lesions, reflecting current clinical practice and ensuring a realistic assessment of the current adoption of eDerma technology in the health care setting. This approach provides a baseline understanding of the technology's impact under current conditions and practices.

The simulation also assumed 100% patient retention (no dropouts) and no patient "no shows" for F2F appointments. The model also excluded patient mortality. Operational capacity was modeled based solely on health care staff availability, assuming fixed staffing levels throughout the simulation and excluding equipment or facility constraints. Furthermore, labor was scheduled with staff available on weekdays (8:30 AM to 6:30 PM) and 70% of staff available for reduced weekend shifts, with the exception of GP. Procedures were allocated based on a simplified first-come, first-served basis and patient risk level, disregarding individual patient availability.

Ethical Consideration

The study used anonymized standard care data. The project was reviewed by the NCAFT Research and Innovation Department (22HIP55) and was determined not to require further ethical review. Specifically, the project was registered as a Health Improvement Project (22HIP55) and was determined by the Research and Innovation office not to require NHS Health Research Authority review. This determination was made in accordance with national regulations, including the Data Protection Act [25], the General Data Protection Regulation [26], and applicable Health Research Authority [27] guidance. As this was a retrospective study using anonymized routinely collected data, the requirement for formal informed consent was not required, as confirmed by the NCA Research and Innovation department. For the Patient-Reported Experience Measure surveys, patients were informed of the evaluation's purpose, and participation was entirely voluntary. No compensation or financial incentives were provided to the patients participating in the surveys. To ensure patient privacy, all data were deidentified before being used for analysis under the formal Data Sharing and Processing Agreements. Furthermore, all researchers were issued formal Letters of Access mandating strict adherence to the trust's information governance and

XSL•FO
RenderX

confidentiality policies, and all methods in this study adhered to the ethical principles outlined in the Declaration of Helsinki.

Results

Description of Dataset

This study's dataset included 563 urgent skin lesion referrals (so-called "2WW" referrals) managed via the intervention arm. A synthetic comparator arm of 4011 referrals was generated based on referral-to-diagnosis times from 2 years of urgent skin cancer referrals (>5000 cases), supplemented with more granular information regarding the methods of diagnosis (F and G in Figure 1) that were available from 70 cases. The same data distributions were subsequently used to conduct the DES.

Within the intervention arm cohort, 324 out of 563 (57.5%) were female participants, and this cohort had a mean age of 61.5 (SD 17.6) years at the time of referral (Table 1), consistent with the higher risk of skin cancer in older populations resulting from accumulated, lifelong UV exposure [1,4,10]. While granular demographic data for the synthetic comparator arm were unavailable, clinical comparability between the 2 arms was maintained by ensuring that both cohorts consisted of adult patients referred via the USC pathway, presenting with 2 or fewer suspicious lesions at the trust. Furthermore, a parallel health inequality assessment, conducted independently of the CCA and DES but within the same trust and clinical pathway, confirmed that the local patient population remains sociodemographically consistent. The NCA cohort typically reflects a more socioeconomically advantaged profile, with a median Index of Multiple Deprivation decile of 6 (IQR 3-8) and a mode of 8.

Table . Sociodemographic and clinical characteristics of the intervention arm (N=563).

Characteristic	Value
Sex, n (%)	
Female	324 (57.5)
Male	239 (42.5)
Age (y), mean (SD)	61.5 (17.6)
Referral pathway	Urgent suspected cancer with 2 or fewer suspicious lesions

Cost Comparison Analysis

The analysis demonstrates substantial cost savings for the NCAFT through the implementation of the community-based teledermatology model compared to standard care (Table 2). The probability sensitivity analysis showed a mean savings of £45 (95% CI £22-£60) per referral with the eDerma model. This translates to a potential overall savings of £25,251 (95% CI £12,462-£34,002) for the NCAFT from December 2022 to May 2023. An unpaired *t* test confirmed that the difference in mean costs between the 2 arms was statistically significant (P<.001). Based on 1000 iterations, a cost reduction per referral was observed in 94% of cases.

A subanalysis further revealed the mechanism behind these savings by assessing the distribution of referrals and their associated unit costs (Table 3). The highest costs were incurred in pathways involving a full diagnostic biopsy, specifically pathways C, D, and G. The community-based eDerma model led to an 18% reduction in referrals needing a biopsy, as the percentage of patients in biopsy-reliant pathways fell from 49% (1965/4011) in the comparator arm (pathway G) to 31% (175/563) in the intervention arm (pathways C and D combined). A 1-way deterministic sensitivity analysis supported this finding by showing that if the eDerma system could offset the need for full biopsies, it could lead to a mean cost savings of up to £135 per referral.

Table . Potential cost savings via the eDerma intervention arm.

Economic parameter	Cost savings[a] (comparator arm–eDerma)	95% CI
Potential cost savings per referral	£45	£22-£60
Potential overall cost savings	£25,251	£12,462-£34,002

[a]£1=US $1.24.

XSL•FO
RenderX

Table . Referral percentage and unit costs in each pathway.

Arms and associated pathways	Referral percentage (%)	Unit cost (£)[a]
Intervention arm (eDerma)		
A	44	55
B	21	219
C	17	371
D	14	542
E	3	137
Comparator arm (standard-of-care)		
F	51	163
G	49	364

[a]£1=US $1.24.

Discrete Event Simulation

The DES used a slightly different pathway distribution due to the additional 6 months data for the intervention arm. The simulation demonstrated a similar and even more pronounced effect, with the biopsy rate falling from 48% (1925/4011) (pathway G) in standard care to 22% (pathway C and D) in the teledermatology model (Table 4).

Table . Pathway distribution used in the discrete event simulation.

Arms and associated pathways	Referral percentage (%)
Intervention arm	
A	53
B	24
C	5
D	17
E	1
Comparator arm	
F	52
G	48

Referral to Diagnosis Communication

To determine the overall average waiting times for the eDerma and standard-of-care arms, a weighted average was calculated by combining the average time for each pathway (A-E in the intervention arm vs F and G in the comparator arm) with the proportion of patients using each communication method. Finally, an incremental analysis was performed to compare the average waiting times of the 2 arms.

The methods of diagnosis communication included letters, emails via the eDerma platform, telephone calls, and F2F appointments. Complete results on the average time for each communication method and the weighted average of each pathway are detailed in Multimedia Appendix 1.

The analysis of weighted average times per pathway reveals that pathway A in the intervention arm has the shortest average time from referral to diagnosis communication, at 8 days. In contrast, the pathways involving diagnostic biopsies, pathways C and D from the intervention arm and pathway G from the comparator arm, show substantially longer weighted average times of 52.4 and 131.1 days, respectively.

The overall weighted average time for the intervention arm was 18.97 (SE 0.92) days. Conversely, the standard care arm had a weighted average time of 73.16 (SE 1.48) days, illustrating a more extended duration (Table 5).

Table . Comparison of the overall weighted average time from referral to diagnosis communication.

Referral to diagnosis (d)	Intervention arm	Comparator arm
Weighted average time	18.97	73.16
Weighted standard error	0.92	1.48
Maximum time	109.5	252.3

XSL•FO
RenderX

Referral to Clinical Diagnosis

The DES analysis demonstrated that the eDerma arm reached clinical diagnosis faster, with a mean of 7.38 (95% CI 7.24-7.52) days from the initial referral. In the comparator arm, the mean time to clinical diagnosis was longer, at 17.29 (95% CI 17.07-17.50) days.

Referral to Histopathological Diagnosis

For patients undergoing a biopsy, the histopathological diagnosis occurred, on average, within 66.42 (95% CI 65.33-67.50) days in the eDerma arm (Table 6). This was significantly shorter than the comparator arm, where the average waiting period for a histopathological diagnosis was 129.21 (95% CI 126.38-132.05) days (Table 6).

A subsequent incremental analysis highlighted the impact of the community-based teledermatology model in reducing patients' waiting times. On average, the teledermatology pathways were associated with a decrease in the time required to establish and communicate a skin cancer diagnosis to a patient by 54.18 (95% CI 50.76 to 57.61) days. Furthermore, the average waiting time for a clinical diagnosis and histopathological diagnosis was reduced by 9.90 (95% CI 9.64-10.16) days and 62.80 (95% CI 59.76-65.83) days, respectively. These findings are summarized in Table 7.

Table . Comparison of referral to clinical diagnosis and referral to histopathological diagnosis results.

Parameter	Intervention arm (teledermatology model)	Comparator arm (standard care)
Referral to clinical diagnosis (d)		
Mean (SE)	7.38 (0.07)	17.29 (0.11)
95% CI	7.24-7.52	17.07-17.50
Max	31.67	50.28
Referral to histopathological diagnosis (d)		
Mean (SE)	66.42 (0.55)	129.21 (1.45)
95% CI	65.33-67.50	126.38-132.05
Max	109.41	248.37

Table . Summary of discrete event simulation analysis.

Incremental analysis Δ (waiting time in current care–waiting time in eDerma care)	Mean (95% CI)
Average time to diagnosis communication (d)	54.18 (50.76-57.61)
Average time to clinical diagnosis (d)	9.90 (9.64-10.16)
Average time to histopathological diagnosis (d)	62.80 (59.76-65.83)

Discussion

Principal Findings

The study demonstrated that the community-based teledermatology model, facilitated by Pathpoint eDerma, was associated with significant cost savings of £45 per referral and a 54.18-day reduction in diagnosis communication. This efficiency is driven by the operational shift from a centralized secondary care workflow to a decentralized, community-level approach. By moving imaging to local CDCs, the model effectively diverted referrals away from hospital-based bottlenecks and addressed the "upstream" consultant-capacity shortages, allowing dermatologists to triage cases remotely.

Another driver of these efficiencies was the 26% reduction in the proportion of patients requiring biopsy-reliant pathways, falling from 48% (1925/4011) in standard care to 22% (124/563) in the teledermatology model. The analysis demonstrates that diagnostic biopsies are the most expensive and resource-intensive stage of the urgent skin cancer journey. By facilitating definitive clinical triage through high-quality imaging, the teledermatology model minimizes unnecessary surgical interventions, potentially leading to cost savings of up to £135 per referral when biopsies are offset.

The model successfully mitigates the initial bottleneck associated with consultant capacity, effectively managing high referral volumes. Our DES highlighted that this upstream optimization has, in turn, revealed a downstream constraint within the histopathology service. The massive disparity in waiting times between pathways requiring a biopsy and those that do not highlights a critical resource limitation at the histopathology level. This constraint now represents the primary rate-limiting step impacting referral-to-diagnosis time. Therefore, to realize further gains in operational efficiency and to ensure that the 28-day FDS is consistently achieved, future service improvement initiatives should be directed toward optimizing the capacity and workflow of the histopathology reporting pathway.

Operational solutions, such as the outsourcing of histopathology services, are currently performed in NCAFT following the completion of this study to increase reporting capacity. Looking forward, this constraint could also potentially be resolved through the adoption of technologies like digital and computational pathology [28,29]. Digital pathology (eg, whole

XSL•FO
RenderX

slide imaging) can enable remote review by pathologists [28]. Alternatively, computational pathology, including artificial intelligence, could accelerate diagnostic throughput by providing intelligent triage to prioritize high-risk cases and automated quantification to speed up time-intensive tasks like measuring tumor margins [29].

Policy Implications

The success of this model has significant implications for NHS health system planning. The findings support the NHS England Teledermatology Roadmap [10] by demonstrating that decentralized models can successfully achieve national targets that remain elusive under traditional F2F care. For these models to be scalable, future policies must prioritize high standards of digital interoperability where platforms can securely span the entire referral journey, enabling swift information sharing between GP, patients, and secondary care providers. Furthermore, the transition "from hospital to community" aligns with the NHS 10-year vision [18], suggesting that future capital investment should be directed toward community imaging infrastructure to sustain these efficiency gains.

Limitations and Recommendations

As an ambidirectional study, the use of nonoverlapping time horizons between the study arms introduces a potential risk of temporal confounding. Observed improvements in wait times and biopsy rates may have been partially influenced by systemic shifts in NHS dermatology protocols following the October 2022 guidelines. Nonetheless, it is essential to distinguish between the national mandate provided by these guidelines and the operational execution enabled by the intervention. Although the guidelines formalize the 28-day FDS, national dermatology waiting lists have remained at historic highs due to chronic consultant shortages [9]. Furthermore, national data indicate that by late 2021, USC referrals had already reached record-high volumes [21], suggesting that the comparator period (September 2021 to December 2022) represents a stable baseline of high systemic operational pressure. The 54.18-day reduction in diagnosis communication suggests an operational impact that far exceeds typical year-on-year service fluctuations, indicating that the digital model addressed these persistent pressures more effectively than the traditional F2F standard of care.

Another methodological limitation was the exclusion of false-negative rates and their associated long-term treatment costs from the economic model. While a longitudinal follow-up for this specific cohort was not feasible, SAF teledermatology possessed an established safety profile in clinical literature, demonstrating diagnostic accuracy and sensitivity comparable to traditional F2F triage [10,14,15,30]. To ensure a transparent and unbiased economic evaluation, these potential downstream costs were excluded consistently across both study arms. Consequently, the reported £45 mean cost saving represents a conservative estimate of the direct, front-end budgetary impact on the trust.

The DES relied on approximations for parameters such as histopathological processing times and professional review durations, which may not perfectly mirror real-world variability. Additionally, the simulation did not account for missed appointments or fluctuating staff availability, and the single-site nature of the study may limit the immediate generalizability of these findings to other regions. To address these constraints, future research could use multisite longitudinal designs, such as randomized stepped-wedge trials, to assess the long-term sustainability of these cost savings. Given that various teledermatology models are already deployed across the United Kingdom, a collaborative effort to prospectively collect standardized data for comparative health economic evaluations may represent the most pragmatic next step.

Furthermore, the use of a F2F clinic as the comparator, as opposed to a direct comparison with a traditional single-site SAF teledermatology service, was a deliberate choice to evaluate the model against current local standards. It is crucial to note that the community-based model offers inherent strategic advantages not captured by wait-time and financial metrics alone, such as improved patient access [12], reduced patient costs [31], and its unique suitability for deploying region-wide collaborative care networks that single-site models may struggle to support.

From a clinical perspective, this study focused on patients in the USC pathway with 2 or fewer suspicious lesions, a cohort primarily consisting of benign cases or early-stage malignancies. While this pathway is key for early detection, individuals with late-stage skin cancer are less likely to access care through this specific route. Therefore, further research is required to understand the journeys of patients with advanced disease to identify unique barriers and develop targeted interventions for timely access.

Ultimately, the community-based teledermatology model was developed to resolve NHS dermatology capacity constraints and to serve as a robust mitigation strategy for traditional access barriers. While the primary focus of this study was on the budgetary and operational implications, findings from a parallel Health Inequality Assessment at the same trust confirmed that diagnostic timelines remained equitable across all age and deprivation groups within this model. This suggests that decentralizing care to local CDCs mitigates the travel and mobility burdens that traditionally hinder older populations at higher risk of skin cancer [1,4,10] and those from more deprived socioeconomic backgrounds [31].

Conclusion

Our results suggest that a community-based teledermatology model, using the Pathpoint eDerma platform, can provide an efficient and cost-effective pathway for urgent suspected skin cancer referrals, benefiting both the health care system and patients. By being associated with a 26% reduction in biopsy-reliant pathways and an observed 54.18-day reduction in communication times, the model offers a viable strategy to address critical resource constraints within the NHS. Future practical interventions should focus on expanding community-based imaging infrastructure to reduce travel burdens for high-risk older populations, while exploring the potential of digital pathology and other operational solutions to resolve the remaining downstream bottlenecks identified in this analysis.

XSL•FO
RenderX

Acknowledgments

The authors would like to extend their sincere gratitude to the Open Medical Patient and Public Involvement group for their invaluable insights and feedback. Their shared experiences, particularly regarding patient waiting times, underscored the importance of this study from a patient's perspective. They also thank Atlas Naqvi and Vasilii Tereshchenko, who played an instrumental role in setting up the teledermatology service at the Northern Care Alliance National Health Service (NHS) Foundation Trust.

Funding

This work was commissioned and funded by the NHS Cancer Programme, with the support of Small Business Research Initiative for Healthcare and the NHS Accelerated Access Collaborative. The views expressed in the publication are those of the authors and not necessarily those of the NHS Cancer Programme or its stakeholders (grant SBRIC01P3030).

Data Availability

The datasets used and analyzed for the current study, along with the full discrete event simulation report that includes additional subanalyses, are available from the corresponding author upon reasonable request.

Authors' Contributions

TCHH conceptualized the study in partnership with Health Tech Enterprise and the Patient and Public Involvement and Engagement committee, collected and prepared datasets for analysis, reviewed the methodology and results, and contributed to the manuscript's write-up. NAN performed a comprehensive literature review, prepared the original manuscript, and was responsible for continuous refinement of the manuscript up to the point of publication. PGM reviewed the manuscript draft, offered practical insights into the functioning of the dermatology department at the trust, and potential external influences. PM obtained study funding. All authors have critically reviewed and accepted the final format of the manuscript.

Conflicts of Interest

Authors NAN, TCHH, and PM declare a competing interest in their employment by Open Medical Ltd.

Multimedia Appendix 1
Average time from referral to diagnosis per communication methods and weighted average time per pathway.
[PDF File, 85 KB - derma_v9i1e86402_app1.pdf]

References

1. Non-melanoma skin cancer statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-melanoma-skin-cancer [accessed 2025-10-01]
2. Melanoma skin cancer statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer#melanoma_stats0 [accessed 2025-10-01]
3. Chen K, Liu X. Global burden of skin cancer and its subtypes: a comprehensive analysis from 1990 to 2021 with projections to 2040. Front Public Health 2025;13:1610661. [doi: 10.3389/fpubh.2025.1610661] [Medline: 40977764]
4. Skin cancer campaign. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/awareness-and-prevention/be-clear-on-cancer/skin-cancer-campaign [accessed 2025-10-23]
5. Shivakumar H, Chanda UL, Onwuchekwa O. Evaluating the diagnostic accuracy and challenges of the two-week wait referral pathway for skin cancers in primary care. Cureus 2025 Jan;17(1):e77410. [doi: 10.7759/cureus.77410] [Medline: 39811725]
6. Early cancer diagnosis data hub. Early Diagnosis Hub. URL: https://crukcancerintelligence.shinyapps.io/EarlyDiagnosis/ [accessed 2025-10-23]
7. Goon PKC, Greenberg DC, Igali L, Levell NJ. Predicted cases of U.K. skin squamous cell carcinoma and basal cell carcinoma in 2020 and 2025: horizon planning for National Health Service dermatology and dermatopathology. Br J Dermatol 2017 May;176(5):1351-1353. [doi: 10.1111/bjd.15110] [Medline: 27714753]
8. Cutting VAT from sunscreen could save the NHS over £128 million per year in cancer treatment costs. Melanoma Focus. URL: https://melanomafocus.org/news-blog/cutting-vat-from-sunscreen-could-save-the-nhs-over-128-million-per-year-in-cancer-treatment-costs/ [accessed 2025-10-01]
9. Dermatology overview. : Getting It Right First Time (GIRFT) and NHS England & NHS Improvement; 2021 URL: https://www.gettingitrightfirsttime.co.uk/wp-content/uploads/2021/11/Dermatology-overview.pdf [accessed 2025-10-01]
10. A teledermatology roadmap: implementing safe and effective teledermatology triage pathways and processes. National Health Service England. URL: https://www.england.nhs.uk/long-read/

XSL•FO
RenderX

a-teledermatology-roadmap-implementing-safe-and-effective-teledermatology-triage-pathways-and-processes/ [accessed 2025-10-01]
11. Diagnostic waiting times and activity. National Health Service England. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/ [accessed 2025-09-09]
12. Referral optimisation for people with skin conditions. : National Health Service England; 2022 URL: https://www.england.nhs.uk/wp-content/uploads/2022/09/B1149-referral-optimisation-for-people-with-skin-conditions.pdf [accessed 2025-10-01]
13. Quality standards for teledermatology. : British Association of Dermatologists; 2022 URL: https://cdn.bad.org.uk/uploads/2022/02/29200021/Teledermatology-Quality-Standards.pdf [accessed 2025-10-01]
14. Jiang SW, Flynn MS, Kwock JT, Nicholas MW. Store-and-forward images in teledermatology: narrative literature review. JMIR Dermatol 2022;5(3):e37517. [doi: 10.2196/37517] [Medline: 35891983]
15. López-Liria R, Valverde-Martínez M, López-Villegas A, et al. Teledermatology versus face-to-face dermatology: an analysis of cost-effectiveness from eight studies from Europe and the United States. Int J Environ Res Public Health 2022 Feb 22;19(5):2534. [doi: 10.3390/ijerph19052534] [Medline: 35270227]
16. McKoy K, Halpern S, Mutyambizi K. International teledermatology review. Curr Dermatol Rep 2021;10(3):55-66. [doi: 10.1007/s13671-021-00333-6] [Medline: 34341713]
17. Service guidance and standards for the use of teledermatology. : British Association of Dermatologists; 2024 URL: https://cdn.bad.org.uk/uploads/2024/12/16150109/Service-Guidance-and-Standards-for-Teledermatology-2024.pdf [accessed 2025-10-01]
18. The NHS long term plan. : National Health Service England; 2019 URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20230418155402/https:/www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ [accessed 2025-10-01]
19. New digital platform set to transform urgent suspected skin cancer referrals in Lancashire and South Cumbria. Lancashire and South Cumbria Cancer Alliance. URL: https://www.lancashireandsouthcumbria.icb.nhs.uk/our-work/canceralliance/latest-news/new-digital-platform-set-transform-urgent-suspected-skin-cancer-referrals-lancashire-and-south-cumbria [accessed 2025-10-17]
20. Northern Care Alliance NHS Foundation Trust. URL: https://www.northerncarealliance.nhs.uk/about-us [accessed 2025-03-27]
21. 2022-23 monthly provider cancer waiting times statistics. National Health Service England. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/historic-cwt-data-before-september-2023/monthly-prov-cwt/2022-23-monthly-provider-cancer-waiting-times-statistics/ [accessed 2024-01-16]
22. Jones KC, Weatherly H, Birch S, et al. Unit costs of health and social care 2022 manual. : Personal Social Services Research Unit, University of Kent at Canterbury; 2022 URL: https://kar.kent.ac.uk/100519/1/Unit%20Costs%20of%20health%20and%20Social%20Care%202022%20%28amended%2013%20July%202023%29.pdf [accessed 2026-03-17] [doi: 10.22024/UNIKENT/01.02.100519]
23. National cost collection for the NHS. National Health Service England. URL: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ [accessed 2023-03-15]
24. NHS workforce statistics - July 2023. National Health Service England. 2023. URL: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/july-2023 [accessed 2023-12-15]
25. Data Protection Act 2018. legislation.gov.uk. URL: https://www.legislation.gov.uk/ukpga/2018/12/contents [accessed 2026-03-10]
26. European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). : Publications Office of the European Union URL: http://data.europa.eu/eli/reg/2016/679/oj [accessed 2026-03-19]
27. Is my study research? Health Research Authority. URL: http://www.hra-decisiontools.org.uk/research/ [accessed 2026-03-10]
28. Zarella MD, McClintock DS, Batra H, et al. Artificial intelligence and digital pathology: clinical promise and deployment considerations. J Med Imaging (Bellingham) 2023 Sep;10(5):051802. [doi: 10.1117/1.JMI.10.5.051802] [Medline: 37528811]
29. Kim I, Kang K, Song Y, Kim TJ. Application of artificial intelligence in pathology: trends and challenges. Diagnostics (Basel) 2022 Nov 15;12(11):2794. [doi: 10.3390/diagnostics12112794] [Medline: 36428854]
30. Chuchu N, Dinnes J, Takwoingi Y, et al. Teledermatology for diagnosing skin cancer in adults. Cochrane Database Syst Rev 2018 Dec 4;12(12):CD013193. [doi: 10.1002/14651858.CD013193] [Medline: 30521686]
31. Nizar NA, Farooki R, Mahapatra P, Halpern S, Hoogenboom TCH. Patient cost analysis of a community-based teledermatology service versus conventional outpatient appointments in East Kent: a retrospective study through a societal lens to reduce health inequalities. BMC Health Serv Res 2024 Dec 21;24(1):1632. [doi: 10.1186/s12913-024-12112-7] [Medline: 39709482]

Abbreviations

CCA: cost comparison analysis

XSL•FO
RenderX

CDC: community diagnostic center
DES: discrete event simulation
F2F: face-to-face
FDS: Faster Diagnosis Standard
GP: general practitioner
NCAFT: Northern Care Alliance NHS Foundation Trust
NHS: National Health Service
SAF: store-and-forward
USC: urgent suspected cancer

Edited by O Yilmaz; submitted 23.Oct.2025; peer-reviewed by H Mahler, S Heydari; revised version received 15.Feb.2026; accepted 25.Feb.2026; published 06.Apr.2026.

Please cite as:
Hoogenboom TCH, Martínez PG, Mahapatra P, Nizar NA
Community-Based Teledermatology for Urgent Suspected Skin Cancer: Health Economic Cost-Comparison and Discrete Event Simulation Study
JMIR Dermatol 2026;9:e86402
URL: https://derma.jmir.org/2026/1/e86402
doi:10.2196/86402

© Tim C H Hoogenboom, Pablo García Martínez, Piyush Mahapatra, Nurul Ain Nizar. Originally published in JMIR Dermatology (http://derma.jmir.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication at http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Over- and Underfunded Diseases by Comparing National Institutes of Health Funding for Skin Disease Research With US Skin Disease Burden According to 2021 Global Burden of Disease Data: Cross-Sectional Analysis

Aileen Park[1], BS; Emily Woolhiser[2,3], DO; Hannah Riva[4], MPAS; Leo Wan[5], BA; Haaris Kadri[1], BA; Elizabeth Lamberty[1*], BA; Parker Juels[1*], BA; Sandra Jaroonwanichkul[6], BA; Madison Reed[7], BS; Catherine Hegedus[8], BS; Dana Chen[9], BS; Danielle Duffle[8], BS, MS; Jessica Kirk[8], BS; Sydney Christensen[2], BA; Emma Shelby[1], BA; Robert Dellavalle[10], MD, PhD, MSPH

[1]School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[10]Department of Dermatology, School of Medicine, University of Minnesota, Phillips-Wangensteen Building, 516 Delaware St SE Suite 1-400, Minneapolis, MN, United States

[2]Kansas City University College of Osteopathic Medicine, Kansas City, MO, United States

[3]Memorial Hospital System, Pembroke Pines, FL, United States

[4]Paul L. Foster School of Medicine, Texas Tech University, El Paso, TX, United States

[5]West Virginia School of Osteopathic Medicine, Lewisburg, WV, United States

[6]School of Medicine, University of Missouri–Kansas City, Kansas City, MO, United States

[7]Kirksville College of Osteopathic Medicine, A.T. Still University, Kirksville, MO, United States

[8]Rocky Vista University College of Osteopathic Medicine, Parker, CO, United States

[9]A.T. Still University School of Osteopathic Medicine, Mesa, AZ, United States

[*]these authors contributed equally

Corresponding Author:
Robert Dellavalle, MD, PhD, MSPH
Department of Dermatology, School of Medicine, University of Minnesota, Phillips-Wangensteen Building, 516 Delaware St SE Suite 1-400, Minneapolis, MN, United States

Abstract

Background: Understanding the burden of various skin diseases can help guide funding allocation for skin disease research. A 2015 cross-sectional study found a partial correlation between US skin disease burden according to the 2010 Global Burden of Disease (GBD) study and National Institutes of Health (NIH) funding in 2012-2013.

Objective: This study aims to identify trends, correlations, and disparities in US skin disease burden and NIH research funding allocation using the latest data from the GBD 2021 and NIH funding data from the fiscal years 2021-2022.

Methods: A cross-sectional analysis was conducted to compare the disability-adjusted life years for 15 skin conditions from the GBD 2021 with NIH funding for these conditions in 2021-2022. Data were sourced from the GBD Results tool and the NIH RePORTER database.

Results: NIH funding for skin disease research and US skin disease burden according to the GBD 2021 were partially correlated, with several outliers. Malignant skin melanoma and pruritus were relatively overfunded, while psoriasis and urticaria were relatively underfunded.

Conclusions: Disease burden is just one of the many important factors that must be considered when allocating resources, including funding to encourage research efforts to improve patient outcomes and positively impact public health.

(JMIR Dermatol 2026;9:e71468) doi:10.2196/71468

KEYWORDS

epidemiology; burden of disease; disability-adjusted life years; research funding; melanoma; psoriasis; dermatitis

XSL•FO
RenderX

Introduction

The Global Burden of Disease (GBD) study aims to quantify worldwide health losses due to a wide variety of illnesses and injuries [1]. Disease burden is one of many important factors guiding decisions on policy development, disease prevention initiatives, and research funding allocation [1,2]. The GBD study quantifies disease burden using disability-adjusted life years (DALYs), a measure that accounts for both mortality due to disease (years of life lost) and years lived with decreased health and quality of life (years lived with disability; YLDs) [1]. GBD also accounts for the severity of disability (defined by any short-term or long-term loss of health) attributed to the variety of illnesses and injuries included in the study by factoring disability weights into the calculation of YLDs [1,3].

Skin conditions are ubiquitous worldwide and affect millions each year. As a result, dermatology continues to be a consistently innovative field that makes large strides in patient care thanks to a heavy research focus. Public funding is a major contributor to research and innovation in this field. In 2015, Hagstrom and colleagues [4] conducted a cross-sectional study that found a partial correlation between US skin disease burden according to the GBD 2010 and National Institutes of Health (NIH) funding in the fiscal years 2012 - 2013, identifying over- and underfunded diseases. Following this study, there have been major changes to the funding of dermatology research, with a 14.7% inflation-adjusted increase in research funding from 2015 to 2019 and fluctuations in funding after the COVID-19 pandemic [5,6]. This study reinvestigates the relationship between US skin disease burden using the latest GBD 2021 data and NIH funding data for 2021 - 2022.

Methods

Overview

A cross-sectional analysis was conducted to compare DALYs for the 15 skin conditions included in the GBD 2021 with NIH funding for these conditions in 2021 - 2022. Data were sourced from the GBD Results tool [1] and the NIH RePORTER database [7]. The search parameters used in GBD Results to obtain DALY metrics for all 15 aforementioned skin disease categories in the US were as follows: measure="DALYs," metric="number," location="United States of America," age="all ages," sex="both," and year="2021." DALY metrics were specifically gathered for the United States to facilitate a direct comparison between the US-specific burden of skin diseases measured by DALYs and funding allocated by the NIH in the United States for skin disease research.

To compile a comprehensive list of NIH-funded grants awarded for skin disease research during fiscal years 2021-2022, a total of 15 queries were entered into the NIH RePORTER database, with each query corresponding to one of the GBD skin disease categories. The following parameters were used to conduct all 15 of these search queries: fiscal year="2021 and 2022," text search logic="advanced," and limit project search="project title, project terms, and project abstracts." In the Text Search box, all *International Classification of Diseases, 10th Revision* codes categorized by the GBD 2021 under one specific skin disease category were strung with "AND," "OR," or "NOT" as determined necessary to capture all relevant NIH-funded grants.

All titles and abstracts of the grants obtained from NIH RePORTER were manually screened by two independent reviewers to determine inclusion versus exclusion (they were included if the grant studied any 1 of the 15 skin disease categories described by the GBD 2021). Following independent review, inclusion and exclusion decisions were cross-examined to identify conflicting decisions. A third reviewer served as a tie-breaker to resolve any discrepancies as needed.

Statistical analysis was performed assuming that the proportion of DALYs attributed to a disease should be the same as the proportion of NIH skin disease funding it receives (ie, if a specific disease is responsible for 25% of all US skin disease DALYs, that disease should receive 25% of all NIH skin disease funding). A one-to-one trendline was used to visualize this relationship and identify outliers representing relatively over- and underfunded skin diseases. An "observed-to-expected" ratio was calculated by dividing the true amount of funding a disease received by the amount of funding a disease could be expected to receive assuming a one-to-one relationship between DALYs and funding.

Ethical Considerations

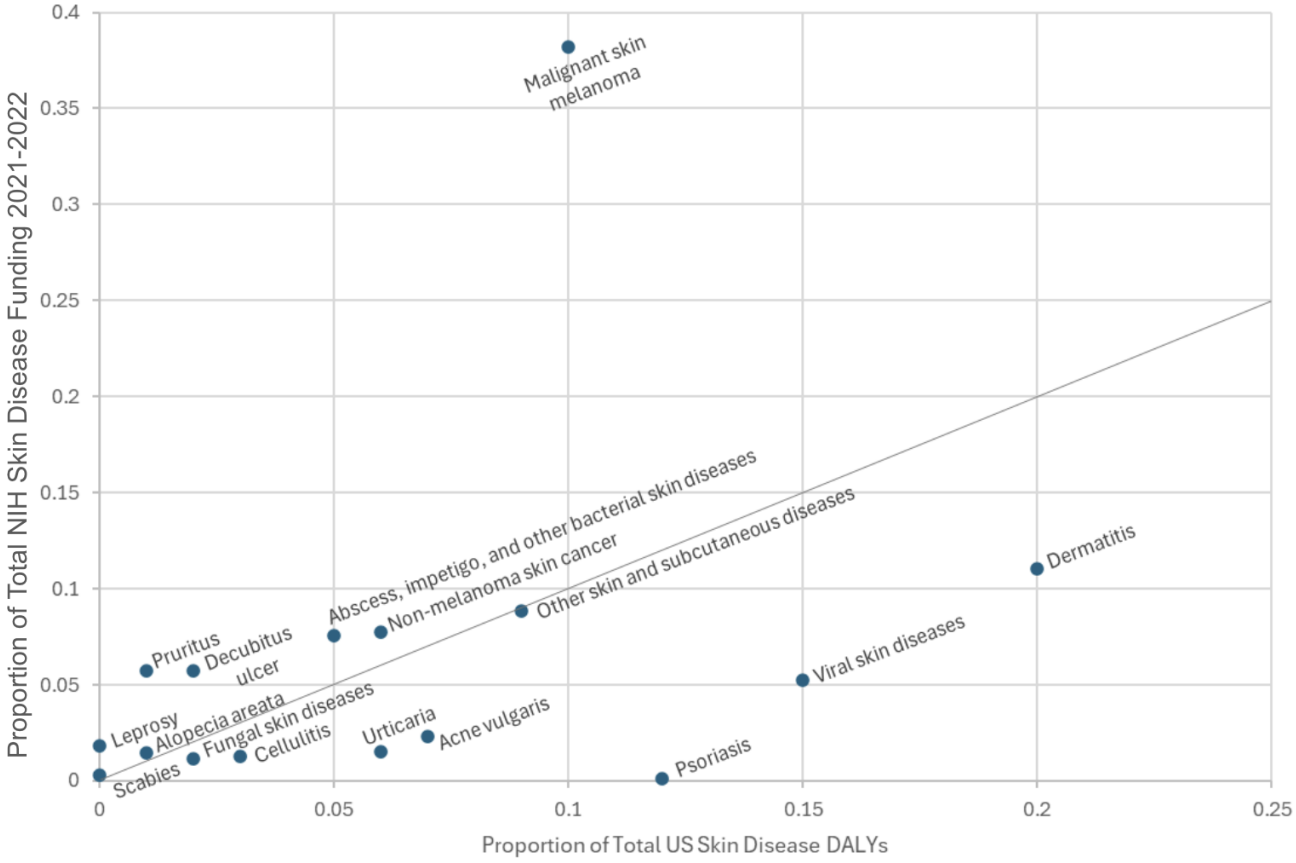
This study was exempt from review by the institutional review board, and no patient or participant consent was required or obtained, as this study did not constitute human subjects research and used publicly available data.

Results

Our analysis revealed a positive correlation between the percentage of total US skin disease DALYs in 2021 and the percentage of total NIH skin disease funding in 2021 - 2022. The correlation coefficient between these two data points was 0.3167 (95% CI 0.053626-0.579774). There were several key outliers when comparing DALYs to funding, indicating that certain skin diseases were relatively over- or underfunded in comparison to their proportion of total disease burden. Pruritus and malignant melanoma received 445% and 392% of the proportion of funding expected by their proportion of DALYs (Table 1). Other relatively overfunded diseases include leprosy, decubitus ulcers, bacterial skin diseases, and nonmelanoma skin cancer (Figure 1, Table 1).

XSL•FO
RenderX

Figure 1. Scatterplot comparing proportion of National Institutes of Health (NIH) skin disease funding received in 2021 - 2022 with the proportion of total US skin disease disability-adjusted life years (DALYs) according to the 2021 Global Burden of Disease study.



XSL•FO
RenderX

Table . Comparison of disability-adjusted life year (DALY) rank from Global Burden of Disease GBD 2010 and 2021 study data, comparison of National Institutes of Health (NIH) funding in fiscal years 2012 - 2013 (data from Hagstrom et al [4]) and 2021-2022 (data from the current analysis), and the percentage of total US skin DALYs (in 2021) and NIH skin disease funding (in 2021 - 22).

Category	US DALY rank in 2021	Proportion of total US skin disease DALYs in 2021, %	US DALY rank in 2010[a]	NIH funding rank in 2021 - 2022	Proportion of total NIH skin disease funding in 2021 - 2022, %	NIH funding rank in 2012 - 2013[a]	Observed-to-expected ratio for funding[b]
Pruritus	13	1.29	5	6	5.74	6	4.45
Malignant skin melanoma	4	9.75	3	1	38.19	1	3.92
Decubitus ulcer	11	1.87	8	7	5.74	11	3.07
Abscess, impetigo, and other bacterial skin diseases	8	4.90	13	5	7.56	9	1.54
Nonmelanoma skin cancer	6	6.10	2	4	7.73	2	1.27
Alopecia areata	12	1.36	11	12	1.48	13	1.09
Other skin and subcutaneous diseases	—[c]	9.30	—	3	8.87	3	0.95
Scabies	14	0.38	14	15	0.30	16	0.8
Dermatitis	1	19.98	1	2	11.03	5	0.55
Fungal skin diseases	10	2.17	9	14	1.14	10	0.53
Cellulitis	9	3.42	12	13	1.26	12	0.37
Viral skin diseases	2	14.61	6	8	5.22	4	0.36
Acne vulgaris	5	6.99	4	9	2.29	14	0.33
Urticaria	7	5.80	7	11	1.51	15	0.26
Psoriasis	3	12.10	10	16	0.09	7	0.0082
Leprosy[d]	15	0	15	10	1.84	8	—[d]

[a]Data obtained from Hagstrom et al [4].

[b]Percentage of funding vs percentage of DALYs.

[c]Not applicable.

[d]Ratio of funding proportion to DALY proportion could not be calculated for leprosy, as the proportion of DALYs for leprosy was 0.

Conversely, psoriasis, fungal skin diseases, cellulitis, urticaria, acne vulgaris, viral skin diseases, and dermatitis were underfunded. Notably, psoriasis received only 0.82% of the funding expected by its disease burden (Table 1). Funding for scabies, alopecia areata, and the "other skin/subcutaneous diseases" category appeared well matched to their disease burden, receiving between 80% to 110% of the funding predicted by their respective DALYs (Figure 1, Table 1).

Discussion

Principal Findings

This study reinvestigated the relationship between US skin disease burden and NIH skin disease research funding using the latest GBD 2021 data and NIH funding data from fiscal years 2021 - 2022. Compared to Hagstrom et al's [4] 2015 study, many of the same trends in relative over- and underfunding of skin diseases were observed. For example, malignant melanoma remains the most significantly overfunded skin disease relative to its disease burden (Table 1) [4]. Nonmelanoma skin cancer and leprosy also remain overfunded, while dermatitis, acne vulgaris, urticaria, fungal skin diseases, and cellulitis remain underfunded (Table 1) [4]. Interestingly, pruritus and decubitus ulcers, previously underfunded in 2015, now appear to be relatively overfunded (Table 1) [4]. Funding for psoriasis was well matched to its disease burden in 2015, but in our updated analysis, psoriasis is the most underfunded skin disease category. Similarly, viral skin diseases were well funded in 2015 and now appear underfunded (Table 1) [4].

It is important to consider disease burden when allocating research funding to ensure adequate resources are being directed toward diseases with the most significant impact. Dedicating more resources toward high-burden diseases can improve individual health and quality of life by driving the development

XSL•FO
RenderX

of innovative treatments and can also provide long-term economic benefits by reducing health care costs and increasing overall workforce productivity.

In addition to disease burden, many other factors also significantly impact resource prioritization and funding allocation. For example, more research funding is likely to be allocated to diseases with strong public awareness and advocacy campaigns, such as malignant skin melanoma. Funding is also likely influenced by disease curability and the potential for therapeutic innovation. The NIH may also prioritize funding for diseases with lower incidence or prevalence but higher mortality (ie, metastatic melanoma, metastatic nonmelanoma skin cancer) rather than diseases with lower mortality but higher incidence or prevalence (ie, dermatitis and acne vulgaris).

Limitations

It is important to keep in mind that using data strictly from the GBD study and the NIH does not fully capture all of the nuances of US skin disease burden and research funding. An important limitation of this analysis, similar to Hagstrom et al's [4] prior study, is the exclusion of industry research funding by pharmaceutical companies and other nongovernmental entities from NIH funding data [4,7]. The NIH is the largest source of public funding for biomedical research; however, a significant portion of research funding also comes from nonprofits, philanthropic organizations, and private industry [8]. Therefore, while a disease may appear underfunded relative to its disease burden using GBD and NIH data alone, additional research funding from nongovernmental agencies may be filling this perceived gap in resource allocation. For instance, although our analysis showed that psoriasis received significantly less funding from the NIH relative to its disease burden, substantial funding from pharmaceutical companies has driven the development of innovative new drugs (ie, IL-23 and IL-17 inhibitors) that have transformed the treatment of psoriasis in recent years [9]. Similarly, previous reviews have cited US $22,291,506 in nonprofit funding for dermatology research in 2019 alone and US $9.3 billion dollars of private equity investment in dermatology health care and research between 2011 and 2021 [10,11].

Conclusions

Given the wide variety of factors that must be considered in order to optimally allocate research funding, several guidelines may help ensure that funding is prioritized for research efforts that will guide clinical practice, improve patient outcomes, and positively impact public health. In addition to prioritizing high-burden diseases, prioritizing funding for translational research can help expedite the incorporation of knowledge gained from basic science research into clinical practice and patient care. Periodically evaluating the real-world impact of funded research using metrics including patient outcomes and cost-efficacy can also help ensure that funding is being distributed to research that is meaningfully impacting clinical practice. Increased funding for conditions that are impacting our patients will allow innovative solutions that improve patient quality of life. With these guidelines in mind, disease burden can easily be incorporated as one of the many important factors that should be used to inform research funding allocation, clinical practice guidelines, and health policy.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

RD is the Editor-in-Chief of *JMIR Dermatology* but was not involved in the selection of this manuscript for publication.

References

1. Institute for Health Metrics and Evaluation. Global Burden of Disease interactive data visuals. 2024. URL: https://vizhub.healthdata.org/gbd-results [accessed 2025-12-16]
2. Yakupu A, Aimaier R, Yuan B, et al. The burden of skin and subcutaneous diseases: findings from the Global Burden Of Disease study 2019. Front Public Health 2023;11:1145513. [doi: 10.3389/fpubh.2023.1145513] [Medline: 37139398]
3. Salomon JA, Vos T, Hogan DR, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 2012 Dec;380(9859):2129-2143. [doi: 10.1016/S0140-6736(12)61680-8]
4. Hagstrom EL, Patel S, Karimkhani C, et al. Comparing cutaneous research funded by the US National Institutes of Health (NIH) with the US skin disease burden. J Am Acad Dermatol 2015 Sep;73(3):383-391. [doi: 10.1016/j.jaad.2015.04.039] [Medline: 26051697]
5. Price KN, Collier EK, Atluri S, Hsiao JL, Shi VY. National Institutes of Health dermatology funding trends 2015-2019. J Invest Dermatol 2021 Jan;141(1):232-235. [doi: 10.1016/j.jid.2020.05.089] [Medline: 32464149]
6. Joshi A, Gawey L, Harrison M, et al. National Institutes of Health dermatology funding trends: 2020-2024. J Invest Dermatol 2025 Oct;145(10):2629-2632. [doi: 10.1016/j.jid.2025.03.004] [Medline: 40118168]
7. RePORTER. National Institutes of Health. URL: https://reporter.nih.gov/advanced-search [accessed 2025-12-10]
8. Grants & funding. National Institutes of Health. URL: https://www.nih.gov/grants-funding [accessed 2024-11-03]
9. Subramonian A, Walter M. Newer Biologics for the Treatment Of Plaque Psoriasis: Canadian Agency for Drugs and Technologies in Health; 2021. URL: http://www.ncbi.nlm.nih.gov/books/NBK585965/ [accessed 2024-11-03]

XSL•FO
RenderX

10. Agarwal A, Orlow SJ. Nonprofit funding in dermatology: a cross-sectional analysis in the United States, 2015 to 2019. JAAD Int 2023 Sep;12:162-164. [doi: 10.1016/j.jdin.2023.06.001] [Medline: 37404247]
11. Venkatesh K, Nambudiri VE. Characterization of venture capital investments in dermatology: a cross-sectional analysis, 2011 to 2021. J Am Acad Dermatol 2023 Mar;88(3):665-667. [doi: 10.1016/j.jaad.2022.06.1177] [Medline: 35779641]

Abbreviations

DALY: disability-adjusted life year
GBD: Global Burden of Disease
NIH: National Institutes of Health
YLD: year lived with disability

Edited by W Guo; submitted 25.Jan.2025; peer-reviewed by D Hitchcock, J Huang, M Gasmi; revised version received 17.Oct.2025; accepted 04.Dec.2025; published 06.Jan.2026.

Please cite as:
Park A, Woolhiser E, Riva H, Wan L, Kadri H, Lamberty E, Juels P, Jaroonwanichkul S, Reed M, Hegedus C, Chen D, Duffle D, Kirk J, Christensen S, Shelby E, Dellavalle R
Identifying Over- and Underfunded Diseases by Comparing National Institutes of Health Funding for Skin Disease Research With US Skin Disease Burden According to 2021 Global Burden of Disease Data: Cross-Sectional Analysis
JMIR Dermatol 2026;9:e71468
URL: https://derma.jmir.org/2026/1/e71468
doi:10.2196/71468

© Aileen Park, Emily Woolhiser, Hannah Riva, Leo Wan, Haaris Kadri, Elizabeth Lamberty, Parker Juels, Sandra Jaroonwanichkul, Madison Reed, Catherine Hegedus, Dana Chen, Danielle Duffle, Jessica Kirk, Sydney Christensen, Emma Shelby, Robert Dellavalle. Originally published in JMIR Dermatology (http://derma.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dermatologic Conditions and Incident Anxiety in Young Adults: Propensity Score–Matched Retrospective Cohort Study

Isabella Zai[1]; Adrian Zai[2], MPH, MD, PhD

[1]Westwood High School, Westwood, MA, United States

[2]Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, 368 Plantation Street, Worcester, MA, United States

Corresponding Author:
Adrian Zai, MPH, MD, PhD
Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, 368 Plantation Street, Worcester, MA, United States

Abstract

Background: Dermatologic conditions such as acne, psoriasis, and dermatitis commonly affect young adults and may contribute to psychological distress. While prior studies have suggested an association between skin disease and anxiety, longitudinal population-level evidence in young adults remains limited.

Objective: This study aimed to examine the association between common dermatologic conditions and the incidence of anxiety among young adults using a large electronic health record–based cohort.

Methods: We conducted a retrospective cohort study using the TriNetX Research Network, including young adults aged 18 to 22 years with and without a qualifying dermatologic diagnosis between 2019 and 2020. The index date was defined as the first dermatologic diagnosis for exposed individuals and a qualifying ambulatory visit for controls. Individuals with a prior diagnosis of anxiety were excluded. Propensity score matching was used to balance demographic characteristics and ambulatory visit history between cohorts. Incident anxiety diagnoses were assessed at 1, 3, and 5 years following the index date. Cumulative incidence, absolute risk differences, risk ratios, and time-to-event analyses were evaluated.

Results: After propensity score matching, 169,720 individuals were included in each cohort. Young adults with dermatologic conditions consistently exhibited a higher incidence of anxiety across all time points. At 1 year, anxiety occurred in 3.9% (5838/149,464) of individuals in the dermatologic condition group compared with 3.4% (3237/95,020) of controls. By 3 years, incidence increased to 11.8% (17,559/149,464) of individuals vs 10.4% (9906/95,020) of controls, and by 5 years to 16.8% (25,184/149,464) of individuals vs 15.5% (14,712/95,020) of controls. Absolute risk differences widened over time, from 0.50 to 1.40 percentage points. Time-to-event analyses demonstrated a modest but consistent increase in hazard, with hazard ratios ranging from 1.12 to 1.14 (all $P<.001$).

Conclusions: In this large, propensity score–matched cohort of young adults, common dermatologic conditions were associated with a small but persistent increase in the incidence of anxiety over time. Although absolute differences were modest, the consistency of findings across multiple analytic approaches highlights the importance of considering psychological well-being as part of comprehensive care for young adults with dermatologic conditions.

(JMIR Dermatol 2026;9:e90820) doi:10.2196/90820

KEYWORDS

dermatologic diseases; anxiety disorders; young adult; electronic health records; propensity score matching

Introduction

Dermatologic conditions are among the most common reasons for ambulatory health care visits in adolescents and young adults [1,2]. Disorders such as acne, psoriasis, and dermatitis frequently emerge or worsen during late adolescence and early adulthood, a developmental period characterized by significant psychological, social, and identity-related transitions [3-6]. Although the physical manifestations of skin disease are readily apparent, their potential mental health sequelae are less consistently recognized in routine clinical care [7-13].

Prior studies have suggested an association between dermatologic conditions and adverse mental health outcomes, including anxiety [7,14,15]. However, much of this literature relies on cross-sectional surveys, clinic-based samples, or patient-reported measures of psychological distress [15-19]. While these approaches provide important insights into symptom burden and quality of life, they are limited in their ability to establish temporal relationships or estimate population-level risk [20,21]. In addition, many studies focus on adults or pediatric populations, with comparatively less attention to young adults navigating the transition between adolescent and adult care [5,22-25].

XSL•FO
RenderX

Electronic health record (EHR) data offer an opportunity to examine longitudinal associations between dermatologic diagnoses and subsequent mental health outcomes at scale [26-28]. Large, multi-institutional EHR networks allow for the identification of incident diagnoses, the alignment of outcomes to clinically meaningful index events, and the evaluation of risk over extended follow-up periods [29-31]. Such data can complement survey-based and qualitative work by providing estimates of real-world diagnostic patterns and temporal trends across diverse care settings.

Despite increasing interest in psychodermatology, few studies have leveraged longitudinal EHR data to quantify the risk of incident anxiety following dermatologic diagnoses in young adults. In particular, there is limited evidence characterizing how anxiety diagnoses accumulate over time after an initial dermatologic diagnosis, or how this risk compares with peers without dermatologic conditions when baseline differences in demographics and health care use are accounted for.

The objective of this study was to evaluate the association between common dermatologic diagnoses and subsequent incident anxiety among young adults aged 18 to 22 years using a large federated EHR network. We conducted a propensity score–matched cohort study comparing cumulative incidence and time to anxiety diagnosis at 1, 3, and 5 years of follow-up. By focusing on incident anxiety and aligning outcomes to clearly defined index events, this study aimed to provide population-level estimates of anxiety risk following dermatologic diagnoses during a critical developmental period.

Methods

Study Design and Data Source

We conducted a retrospective matched cohort study using deidentified EHR data accessed through the TriNetX Research Network, a federated platform comprising multiple health care organizations across the United States [32]. The network contains longitudinal inpatient and outpatient data, including demographics, diagnoses, and encounter information. All analyses were performed within the secure analytic environment provided by TriNetX using deidentified data in accordance with institutional policies.

The study was reviewed by the institutional review board and determined to be exempt.

Study Population

The study population included young adults aged 18 to 22 years with clinical encounters recorded between January 1, 2019, and December 31, 2020. The 2019 to 2020 index period was selected to allow sufficient longitudinal follow-up of up to 5 years within the available TriNetX data while capturing a contemporary cohort of young adults receiving routine care. This period was not selected to coincide with the COVID-19 pandemic but rather reflects the most recent interval permitting extended follow-up. To ensure engagement with the health care system, patients were required to have at least one qualifying ambulatory visit during the study period.

Patients were followed longitudinally for the development of incident anxiety diagnoses through December 31, 2025, allowing for up to 5 years of potential follow-up.

Exposure Definition and Index Date

The exposed cohort consisted of patients with a diagnosis of a dermatologic condition, defined by the presence of *International Classification of Diseases, Tenth Revision* (*ICD-10*) diagnosis codes corresponding to common dermatologic disorders. The index date for exposed patients was defined as the date of the first qualifying dermatologic diagnosis during the study period.

The comparison cohort consisted of patients without any recorded dermatologic diagnosis at any time prior to or during the study period. For these patients, the index date was defined as the date of the first qualifying ambulatory visit occurring within the same calendar period.

Outcome Definition

The primary outcome was incident anxiety, defined as a new diagnosis of an anxiety disorder recorded after the index date, identified using *ICD-10* diagnosis codes for anxiety disorders.

To focus on incident anxiety, patients with any recorded anxiety diagnosis during the 1 year before the index date were excluded at the cohort definition stage. In addition, outcome analyses excluded patients with anxiety diagnoses occurring prior to the start of each analytic time window to ensure that all patients included in the risk set were at risk for incident anxiety during the specified follow-up period.

Outcome time windows began 1 day after the index date to reduce misclassification of anxiety diagnoses recorded during the same clinical encounter as the dermatologic diagnosis or index visit.

Covariates

Baseline covariates included age, sex, race, and ethnicity, as well as measures of health care use prior to the index date, including the presence of ambulatory visits. Covariates were selected a priori based on clinical relevance and their potential role as confounders of the association between dermatologic diagnoses and subsequent anxiety.

Anxiety diagnoses were not included as matching covariates because patients with prior anxiety were excluded from the study population, and postindex variables were not incorporated into propensity score estimation.

Propensity Score Matching

To reduce confounding, propensity score matching was performed to create balanced cohorts of patients with and without dermatologic diagnoses. Propensity scores were estimated using logistic regression based on baseline demographic characteristics and health care use measures.

Patients were matched one-to-one using nearest neighbor matching without replacement. Balance between cohorts was assessed using standardized mean differences, with values less than 0.1 indicating acceptable balance.

All outcome analyses were conducted using the matched cohorts.

XSL•FO
RenderX

Follow-Up and Time Horizons

Patients were followed from the index date until the earliest of the following events: first recorded anxiety diagnosis, end of the specified follow-up window, or last available encounter in the dataset.

Separate analyses were conducted using follow-up windows of 1, 3, and 5 years following the index date. As follow-up duration increased, fewer patients contributed data to later time horizons due to censoring at the end of available follow-up.

Statistical Analysis

Cumulative incidence of anxiety was calculated for each cohort at 1, 3, and 5 years following the index date. Absolute risk differences and relative risks with 95% CIs were reported.

Time to incident anxiety was additionally evaluated using Kaplan-Meier methods, with differences between cohorts assessed using the log-rank test. Hazard ratios with 95% CIs were estimated to quantify relative differences in the rate of anxiety diagnosis over time. Kaplan-Meier analyses were performed as supportive time-to-event analyses to assess consistency with cumulative incidence findings.

All statistical tests were 2-sided, and statistical significance was defined as $P<.05$.

Sensitivity Analyses

Sensitivity analyses were conducted by comparing effect estimates across multiple follow-up horizons to assess the robustness and temporal consistency of findings.

Ethical Considerations

This study was reviewed by the Institutional Review Board of UMass Chan Medical School and determined to be exempt from human subjects research (STUDY00002783). The research used deidentified EHR data accessed through the TriNetX Research Network. Because the data were deidentified prior to analysis and posed minimal risk to participants, no direct identifiers were available to the investigators, participants were not contacted, and no attempts were made to reidentify individuals, informed consent was not required in accordance with the Common Rule provision for exempt secondary research [33,34] and the HIPAA (Health Insurance Portability and Accountability Act) Privacy Rule standard for deidentification.

Data were analyzed within the secure TriNetX analytics environment, and results were reported only in aggregate form. No participant compensation was provided, as this study did not involve direct participant contact.

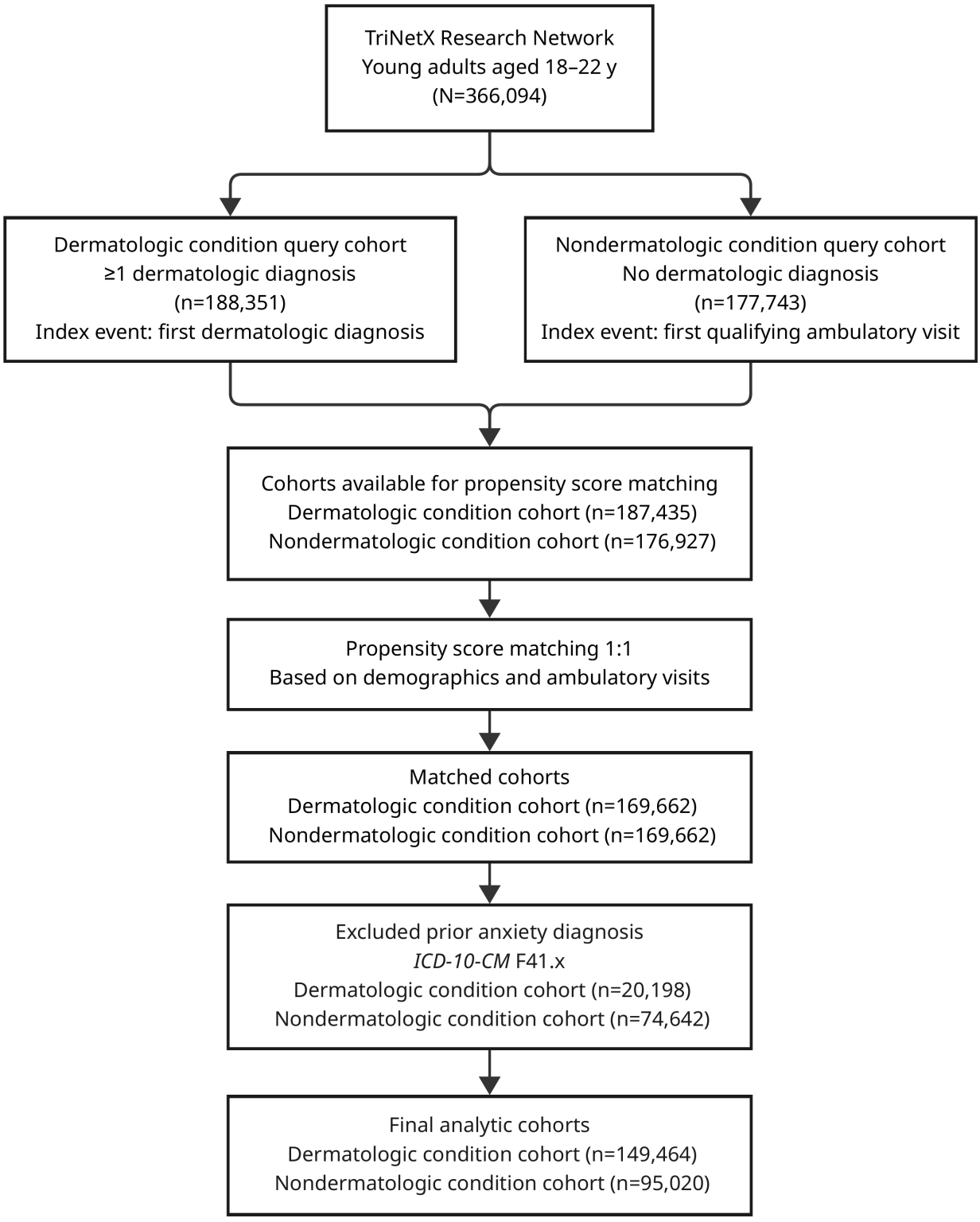
Results

Cohort Assembly and Baseline Characteristics

Figure 1 summarizes cohort assembly. Among young adults aged 18 to 22 years with a qualifying dermatologic diagnosis between January 1, 2019, and December 31, 2020, eligible individuals were matched 1:1 to controls without dermatologic diagnoses using propensity scores. After matching, 169,720 individuals were retained in each cohort.

XSL•FO
RenderX

Figure 1. Cohort selection flow diagram. *ICD-10-CM*: *International Classification of Diseases, Tenth Revision, Clinical Modification*.



Baseline demographic characteristics before and after matching are shown in Table 1. Prior to matching, modest differences were observed across race, ethnicity, sex, and ambulatory visit history. After propensity score matching, covariates were well balanced between cohorts, with standardized mean differences below 0.01 for all matched variables.

XSL•FO
RenderX

Table . Baseline characteristics of propensity score–matched young adults with and without dermatologic diagnoses[a].

Characteristic	Dermatology cohort (n=169,662), n (%)	Nondermatology cohort (n=169,662), n (%)	Standardized mean difference
Sex			
Female	96,440 (56.8)	96,257 (56.7)	0.0018
Male	73,139 (43.1)	73,325 (43.2)	0.0019
Race and ethnicity			
American Indian or Alaska Native	1018 (0.6)	10,542 (0.6)	0.0024
Asian	9289 (5.5)	9356 (5.5)	0.0017
Black or African American	27,461 (16.2)	27,451 (16.2)	0.0001
Hispanic or Latino	23,930 (14.1)	23,610 (13.9)	0.0052
Native Hawaiian or Other Pacific Islander	886 (0.5)	9198 (0.5)	0.0026
Not Hispanic or Latino	112,308 (66.2)	112,696 (66.4)	0.0047
White	107,599 (63.4)	107,415 (63.3)	0.0023
Prior ambulatory visit before index	98,656 (58.1)	98,495 (58.1)	0.0019

[a]Cohorts were matched 1:1 using propensity score matching based on sex, race, ethnicity, and prior ambulatory health care use. All standardized mean differences were less than 0.01, indicating excellent postmatching balance.

Young adults aged 18 to 22 years were identified from the TriNetX Research Network between 2019 and 2020. Patients were categorized based on the presence or absence of a qualifying dermatologic diagnosis (including acne, psoriasis, or dermatitis or eczema), and those with a prior diagnosis of anxiety were excluded. Propensity score matching was applied to create balanced cohorts with and without dermatologic conditions. The final matched cohorts were followed for incident anxiety diagnoses at 1-, 3-, and 5-year intervals after the index date.

Incident Anxiety at 1, 3, and 5 Years

Table 2 presents the cumulative incidence of newly diagnosed anxiety disorders at 1, 3, and 5 years following the index date. Across all time horizons, young adults with dermatologic conditions experienced a consistently higher incidence of anxiety compared with matched controls. The absolute difference between cohorts increased over time, while the relative increase in risk remained modest across follow-up periods.

Figure 2 illustrates these cumulative incidence estimates over time, demonstrating a parallel increase in anxiety diagnoses in both cohorts with a persistently higher absolute risk among individuals with dermatologic conditions.

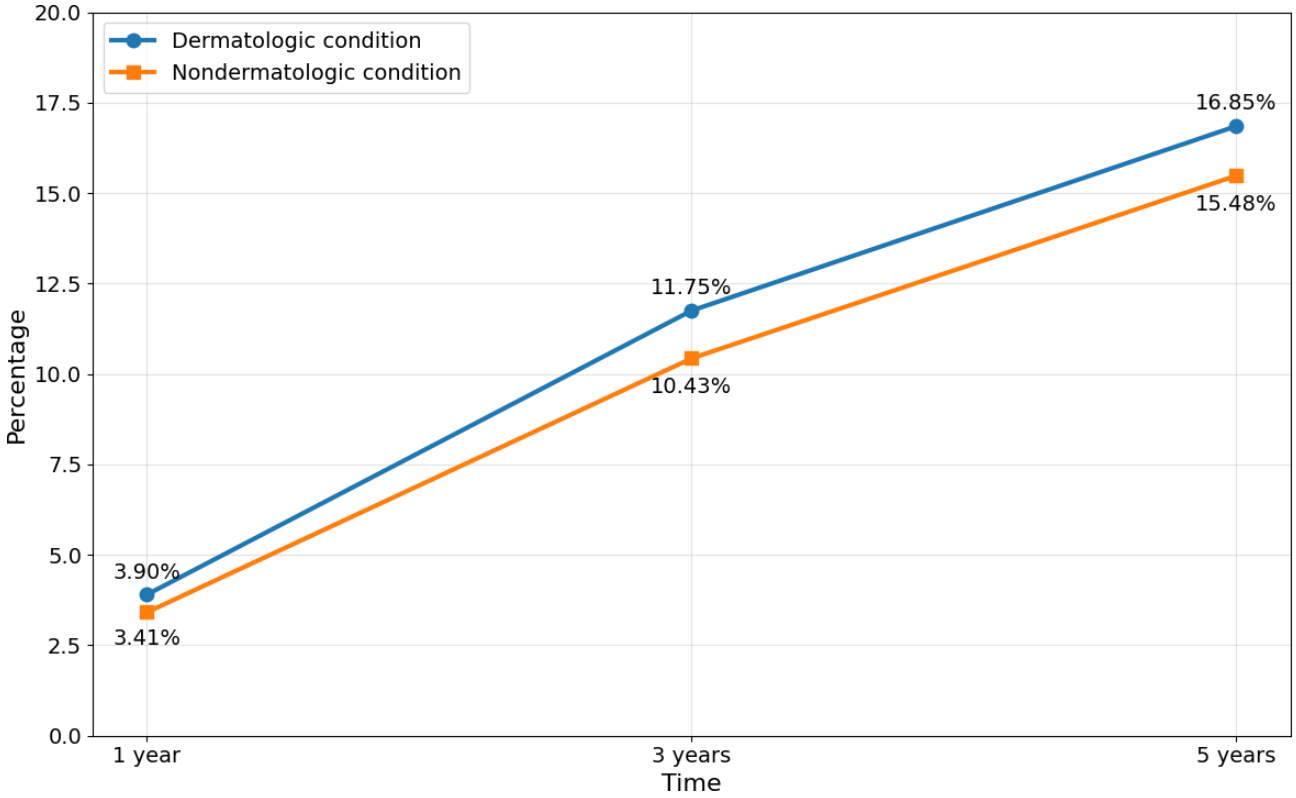
Table . Risk of incident anxiety among young adults with vs without dermatologic conditions[a].

Follow-up period and cohort	Incident anxiety, n (%)	Risk difference (%; 95% CI)	Risk ratio (95% CI)	Odds ratio (95% CI)
1 year				
With a dermatologic condition (n=149,464)	5838 (3.9)	0.50 (0.30 - 0.70)	1.147 (1.099 - 1.196)	1.153 (1.103 - 1.204)
Without a dermatologic condition (n=95,020)	3237 (3.4)	Reference	Reference	Reference
3 years				
With a dermatologic condition (n=149,464)	17,559 (11.8)	1.32 (1.07 - 1.58)	1.127 (1.101 - 1.153)	1.144 (1.114 - 1.174)
Without a dermatologic condition (n=95,020)	9906 (10.4)	Reference	Reference	Reference
5 years				
With a dermatologic condition (n=149,464)	25,184 (16.8)	1.40 (1.10 - 1.70)	1.088 (1.068 - 1.109)	1.106 (1.082 - 1.131)
Without a dermatologic condition (n=95,020)	14,712 (15.5)	Reference	Reference	Reference

[a]Patients with a prior diagnosis of anxiety before the index date were excluded. Analyses were conducted after propensity score matching.

XSL•FO
RenderX

Figure 2. Cumulative incidence of anxiety among young adults with and without dermatologic conditions. Cumulative incidence of new-onset anxiety at 1, 3, and 5 years after the index date among propensity score–matched young adults aged 18 to 22 years with dermatologic conditions diagnosed between 2019 and 2020 and matched peers without dermatologic conditions. Percentages represent the proportion of individuals with a new anxiety diagnosis within each follow-up interval, excluding those with anxiety prior to the index date.



Time-to-Event Analysis

Time-to-event analyses using Kaplan-Meier methods further supported these findings. Individuals with dermatologic conditions exhibited a modestly increased hazard of incident anxiety across follow-up periods.

The hazard ratio for incident anxiety was 1.135 (95% CI 1.088 - 1.185) at 1 year, 1.121 (95% CI 1.094 - 1.149) at 3 years, and 1.096 (95% CI 1.074-1.119) at 5 years, consistent with the cumulative incidence findings in Table 2; log-rank tests were significant across all analyses (all $P<.001$).

Despite statistical significance, Kaplan-Meier curves showed substantial overlap between cohorts, reflecting the modest magnitude of the association and reinforcing the importance of interpreting relative measures alongside absolute risk differences.

Discussion

In this large, propensity score–matched cohort study of young adults, we found that individuals with common dermatologic conditions experienced a consistently higher incidence of anxiety compared with matched peers without dermatologic conditions. Although the magnitude of the association was modest, the excess risk was observed across several follow-up intervals and persisted over time. These findings suggest that dermatologic conditions in early adulthood are associated with a small but sustained increase in the likelihood of subsequent anxiety diagnoses. The consistency of findings across cumulative incidence, relative risk, and time-to-event analyses supports the robustness of the observed association.

The absolute differences in anxiety incidence were relatively small, ranging from approximately one-half of a percentage point at 1 year to just over 1 percentage point at 5 years. However, given the high prevalence of dermatologic conditions among young adults, even modest increases in risk may have meaningful implications at the population level. Importantly, the consistency of the association across 1-, 3-, and 5-year follow-up intervals strengthens confidence that the observed relationship is not driven by short-term diagnostic clustering or surveillance bias alone.

The relative effect size attenuated with longer follow-up, while the absolute difference in cumulative incidence increased over time. This pattern is expected in large observational cohorts and reflects the increasing background incidence of anxiety in young adulthood. From a clinical perspective, the absolute risk difference may be more informative than relative measures, as it better contextualizes the magnitude of excess risk attributable to dermatologic conditions. Our findings underscore the importance of interpreting statistically significant associations in the context of absolute effect sizes.

Several mechanisms may plausibly explain the observed association. Dermatologic conditions, such as acne, psoriasis, and dermatitis, often emerge or worsen during adolescence and early adulthood, a period characterized by heightened psychosocial vulnerability. Visible skin disease may contribute to anxiety through effects on self-image, social interactions, and perceived stigma. In addition, chronic or recurrent

XSL•FO
RenderX

symptoms, such as pruritus, pain, or flares, may increase psychological distress over time. While our study was not designed to identify causal pathways, the persistence of the association across multiple years is consistent with a cumulative psychosocial burden rather than a transient response to diagnosis alone.

Our findings are consistent with prior literature reporting associations between dermatologic disease and mental health outcomes, while extending this work in several important ways. First, we focused specifically on young adults, a population in which both dermatologic conditions and anxiety commonly emerge but remain underrepresented in large-scale longitudinal studies. Second, by using propensity score matching and excluding individuals with prior anxiety diagnoses, we sought to minimize confounding and focus on incident outcomes. Third, the use of multiple follow-up intervals allowed us to characterize the temporal evolution of risk rather than relying on a single time point estimate.

This study has several limitations. As an observational analysis based on EHR data, it is subject to residual confounding despite propensity score matching. The study period included the onset of the COVID-19 pandemic, which may have influenced background rates of anxiety diagnoses as well as patterns of health care use. Although both cohorts were drawn from the same calendar period and matched on baseline characteristics and ambulatory visit history, residual confounding related to pandemic-era behavioral and health system changes remains possible. However, because both cohorts were derived from the same period, such effects are likely to be nondifferential and unlikely to materially affect the internal validity of the comparison. Anxiety diagnoses may be underascertained or misclassified, and diagnostic practices may vary across health care organizations. Additionally, patients with dermatologic conditions may have more frequent health care encounters, potentially increasing opportunities for anxiety diagnosis. Although matching on ambulatory visit history partially addresses this concern, differential health care use cannot be fully excluded. Finally, the TriNetX platform does not allow for detailed assessment of disease severity, symptom burden, or patient-reported outcomes, which may further modify the relationship between dermatologic conditions and anxiety.

Despite these limitations, the strengths of this study include its large, geographically diverse population, rigorous matching strategy, and consistent findings across multiple analytic approaches. The concordance between cumulative incidence estimates and time-to-event analyses supports the robustness of the observed association.

In conclusion, young adults with common dermatologic conditions experience a small but persistent increase in the incidence of anxiety compared with matched peers without dermatologic conditions. These findings highlight the importance of holistic care approaches that consider both physical and psychological well-being in patients with dermatologic disease. Future studies incorporating patient-reported outcomes and disease severity measures may help clarify which subgroups are at greatest risk and inform targeted screening or intervention strategies.

Acknowledgments

The authors acknowledge the TriNetX Research Network for providing access to deidentified electronic health record data used in this study.

Funding

This study did not receive any external funding.

Data Availability

The data used in this study were obtained from the TriNetX Research Network. These data are available only under license for this study and are not publicly available. Qualified researchers may access data through the TriNetX platform, subject to institutional agreements, data use approvals, and compliance with applicable privacy and security requirements. Aggregate results supporting the findings of this study are included within the paper.

Authors' Contributions

Conceptualization: IZ
Data curation: IZ
Formal analysis: IZ
Methodology: AZ
Project administration: AZ
Supervision: AZ
Validation: AZ
Writing—original draft: IZ
Writing—review and editing: AZ
Both authors approved the final version of the manuscript for submission and agree to be accountable for all aspects of the work.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Ho T, Taylor MT, Marathe KS, Lucky AW, Barbieri JS. Most common pediatric skin conditions managed in outpatient dermatology clinics in the United States stratified by race and ethnicity. Pediatr Dermatol 2021 Nov;38 Suppl 2:129-131. [doi: 10.1111/pde.14693] [Medline: 34339074]
2. Chang J, Nock MR, Cohen JM, Bunick CG. Acne accounts for an almost 2.5-fold higher proportion of dermatology visits among adult females compared to adult males in the United States: a study of the national ambulatory medical care survey from 2002-2016. PLoS One 2023;18(9):e0290763. [doi: 10.1371/journal.pone.0290763] [Medline: 37733776]
3. Potterton R, Austin A, Robinson L, Webb H, Allen KL, Schmidt U. Identity development and social-emotional disorders during adolescence and emerging adulthood: a systematic review and meta-analysis. J Youth Adolesc 2022 Jan;51(1):16-29. [doi: 10.1007/s10964-021-01536-7] [Medline: 34783954]
4. Branje S, de Moor EL, Spitzer J, Becht AI. Dynamics of identity development in adolescence: a decade in review. J Res Adolesc 2021 Dec;31(4):908-927. [doi: 10.1111/jora.12678]
5. De Vere Hunt IJ, Howard E, McPherson T. The impact of chronic skin disease in adolescence and the need for specialist adolescent services. Clin Exp Dermatol 2020 Jan;45(1):5-9. [doi: 10.1111/ced.14021] [Medline: 31236992]
6. Fournier H, Calcagni N, Morice-Picard F, Quintard B. Psychosocial implications of rare genetic skin diseases affecting appearance on daily life experiences, emotional state, self-perception and quality of life in adults: a systematic review. Orphanet J Rare Dis 2023 Feb 23;18(1):39. [doi: 10.1186/s13023-023-02629-1] [Medline: 36823650]
7. Salari N, Heidarian P, Hosseinian-Far A, Babajani F, Mohammadi M. Global prevalence of anxiety, depression, and stress among patients with skin diseases: a systematic review and meta-analysis. J Prev (2022) 2024 Aug;45(4):611-649. [doi: 10.1007/s10935-024-00784-0]
8. Balieva F, Abebe DS, Dalgard FJ, Lien L. Risk of developing psychiatric disease among adult patients with skin disease: a 9-year national register follow-up study in Norway. Skin Health Dis 2023 Dec;3(6):e294. [doi: 10.1002/ski2.294] [Medline: 38047256]
9. Cortés H, Rojas-Márquez M, Del Prado-Audelo ML, Reyes-Hernández OD, González-Del Carmen M, Leyva-Gómez G. Alterations in mental health and quality of life in patients with skin disorders: a narrative review. Int J Dermatol 2022 Jul;61(7):783-791. [doi: 10.1111/ijd.15852] [Medline: 34403497]
10. Hughes O, Bewley A. Is it really ever “just acne”? Considering the psychodermatology of acne. Br J Dermatol 2023 Oct 30;189(Suppl 1):i11-i16. [doi: 10.1093/bjd/ljad251] [Medline: 37903071]
11. Sood S, Jafferany M, Vinaya Kumar S. Depression, psychiatric comorbidities, and psychosocial implications associated with acne vulgaris. J Cosmet Dermatol 2020 Dec;19(12):3177-3182. [doi: 10.1111/jocd.13753] [Medline: 33006820]
12. Baker N, Billick SB. Psychiatric consequences of skin conditions: multiple case study analysis with literature review. Psychiatr Q 2022 Sep;93(3):841-847. [doi: 10.1007/s11126-022-09991-6] [Medline: 35771406]
13. Wu JJ, Feldman SR, Koo J, Marangell LB. Epidemiology of mental health comorbidity in psoriasis. J Dermatolog Treat 2018 Aug;29(5):487-495. [doi: 10.1080/09546634.2017.1395800] [Medline: 29057684]
14. Storer B, Kershaw KA, Braund TA, et al. The prevalence of anxiety disorders in dermatology outpatients: a systematic review and meta-analysis. J Eur Acad Dermatol Venereol 2023 Apr 28. [doi: 10.1111/jdv.19151] [Medline: 37118899]
15. Dalgard FJ, Gieler U, Tomas-Aragones L, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015 Apr;135(4):984-991. [doi: 10.1038/jid.2014.530] [Medline: 25521458]
16. Dalgard FJ, Svensson Å, Halvorsen JA, et al. Itch and mental health in dermatological patients across Europe: a cross-sectional study in 13 countries. J Invest Dermatol 2020 Mar;140(3):568-573. [doi: 10.1016/j.jid.2019.05.034] [Medline: 31491369]
17. Schuster B, Peifer C, Ziehfreund S, et al. Happiness and depression in psoriasis: a cross-sectional study in Germany. Qual Life Res 2022 Jun;31(6):1761-1773. [doi: 10.1007/s11136-021-02991-2] [Medline: 34532835]
18. Pärna E, Aluoja A, Kingo K. Quality of life and emotional state in chronic skin disease. Acta Derm Venereol 2015 Mar;95(3):312-316. [doi: 10.2340/00015555-1920] [Medline: 24978135]
19. Picardi A, Abeni D, Pasquini P. Assessing psychological distress in patients with skin diseases: reliability, validity and factor structure of the GHQ-12. J Eur Acad Dermatol Venereol 2001 Sep;15(5):410-417. [doi: 10.1046/j.1468-3083.2001.00336.x] [Medline: 11763380]
20. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 2020 Jul;158(1S):S65-S71. [doi: 10.1016/j.chest.2020.03.012]
21. Grujičić S, Nikolić A. Cross-section studies: advantages and disadvantages. Zdravstvena Zaštita 2021;50(4):43-54. [doi: 10.5937/zdravzast50-35574]
22. Zhou H, Roberts P, Dhaliwal S, Della P. Transitioning adolescent and young adults with chronic disease and/or disabilities from paediatric to adult care services - an integrative review. J Clin Nurs 2016 Nov;25(21-22):3113-3130. [doi: 10.1111/jocn.13326] [Medline: 27145890]

XSL•FO
RenderX

23. Hendrickx G, De Roeck V, Maras A, et al. Challenges during the transition from child and adolescent mental health services to adult mental health services. BJPsych Bull 2020 Aug;44(4):163-168. [doi: 10.1192/bjb.2019.85] [Medline: 31931898]
24. Marchini S, Laroche MA, Nemorin H, et al. From adolescence to adulthood: understanding care trajectories in an early detection and intervention centre in France. Early Interv Psychiatry 2025 Jan;19(1):e13605. [doi: 10.1111/eip.13605] [Medline: 39223792]
25. De Vere Hunt I, Chapman K, Wali G, et al. Establishing and developing a Teenage and Young Adult dermatology clinic with embedded specialist psychological support. Clin Exp Dermatol 2019 Dec;44(8):893-896. [doi: 10.1111/ced.13950] [Medline: 30784103]
26. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health 2016;37:61-81. [doi: 10.1146/annurev-publhealth-032315-021353] [Medline: 26667605]
27. Huguet N, Kaufmann J, O'Malley J, et al. Using electronic health records in longitudinal studies: estimating patient attrition. Med Care 2020 Jun;58:S46-S52. [doi: 10.1097/MLR.0000000000001298] [Medline: 32412953]
28. Li Q, Patrick MT, Sreeskandarajan S, et al. Large-scale epidemiological analysis of common skin diseases to identify shared and unique comorbidities and demographic factors. Front Immunol 2024;14:1309549. [doi: 10.3389/fimmu.2023.1309549]
29. Manemann SM, St Sauver JL, Liu H, et al. Longitudinal cohorts for harnessing the electronic health record for disease prediction in a US population. BMJ Open 2021 Jun 8;11(6):e044353. [doi: 10.1136/bmjopen-2020-044353] [Medline: 34103314]
30. Madhavan S, Bastarache L, Brown JS, et al. Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. J Am Med Inform Assoc 2021 Feb 15;28(2):393-401. [doi: 10.1093/jamia/ocaa287] [Medline: 33260207]
31. Simon GJ, Peterson KA, Castro MR, Steinbach MS, Kumar V, Caraballo PJ. Predicting diabetes clinical outcomes using longitudinal risk factor trajectories. BMC Med Inform Decis Mak 2020 Jan 8;20(1):6. [doi: 10.1186/s12911-019-1009-3] [Medline: 31914992]
32. TriNetX. URL: https://trinetx.com [accessed 2026-05-27]
33. eCFR: 45 CFR 46.104 -- exempt research. Office for Human Research Protections, Department of Health and Human Services. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.104 [accessed 2026-06-01]
34. eCFR: 45 CFR 164.514 other requirements relating to uses and disclosures of protected health information. Office for Human Research Protections, Department of Health and Human Services. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-C/part-164/subpart-E/section-164.514 [accessed 2026-06-01]

Abbreviations

EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
***ICD-10*:** *International Classification of Diseases, Tenth Revision*

Edited by S Subhadarshani; submitted 04.Jan.2026; peer-reviewed by S Gulliver; revised version received 19.Apr.2026; accepted 20.May.2026; published 03.Jun.2026.

Please cite as:
Zai I, Zai A
Dermatologic Conditions and Incident Anxiety in Young Adults: Propensity Score–Matched Retrospective Cohort Study
JMIR Dermatol 2026;9:e90820
URL: https://derma.jmir.org/2026/1/e90820
doi:10.2196/90820

© Isabella Zai, Adrian Zai. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging AI Large Language Models for Writing Clinical Trial Proposals in Dermatology: Instrument Validation Study

Megan Hauptman[1], MD; Daniel Copley[2], BSc; Kelly Young[1], MD, PhD; Tran Do[1], MD, PhD; Joseph S Durgin[1], MD; Albert Yang[1], MD; Jungsoo Chang[1], MD; Allison Billi[1], MD, PhD; Mio Nakamura[1], MD; Trilokraj Tejasvi[1], MD

[1]Department of Dermatology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI, United States

[2]Anduril, Costa Mesa, CA, United States

Corresponding Author:
Trilokraj Tejasvi, MD
Department of Dermatology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI, United States

Abstract

Background: Large language models (LLMs) are becoming increasingly popular in clinical trial design but have been underused in research proposal development.

Objective: This study compared the performance of commonly used open access LLMs versus human proposal composition and review.

Methods: A total of 10 LLMs were prompted to write a research proposal. Six physicians and each of the LLMs assessed 11 blinded proposals for capabilities and limitations in accuracy and comprehensiveness.

Results: ChatGPT-o1 and Llama 3.1 were rated the most and least accurate, respectively, by human scorers. LLM scorers rated ChatGPT-o1 and DeepSeek R1 as the most accurate. ChatGPT-o1 and Llama 3.1 were rated as the most and least comprehensive, respectively, by human and LLM scorers. LLMs performed poorly on scoring proposals and, on average, rated proposals 1.9 points higher than humans for both accuracy and comprehensiveness.

Conclusions: Paid versions of ChatGPT remain the highest-quality and most versatile option of the available LLMs. These tools cannot replace expert input but serve as powerful assistants, streamlining the development process and enhancing productivity.

(JMIR Dermatol 2026;9:e76674) doi:10.2196/76674

KEYWORDS

artificial intelligence; AI; large language model; research proposal; clinical research; clinical trials; deep learning; machine learning; research design

Introduction

Advancements in artificial intelligence (AI) have led to the development of large language models (LLMs) using algorithms that learn from data and recognize patterns to make decisions based on all available data within a training set [1]. However, AI is limited by the data it is trained on and an inability to account for the nuanced contexts of individual research studies [2]. Researchers are increasingly using LLMs in clinical trial design to improve patient selection, cohort composition, and recruitment [3]. In contrast, the use of LLMs in research proposal development is largely unexplored, and thus, they are perhaps underused. This study aimed to address this gap by comparing the performance of LLMs versus the current gold standard of human proposal composition and review. Our goals were 3-fold: to rate LLMs in composing clinical trial proposals, assess LLMs in scoring clinical trial proposals, and evaluate the ease of using LLMs (including usability and efficiency).

Methods

Overview

Commonly used open access AI platforms (DeepSeek R1, ChatGPT-o3-mini [OpenAI], ChatGPT-o1 [OpenAI], ChatGPT-4o [OpenAI], Claude Sonnet [Anthropic], Claude Opus [Anthropic], OpenEvidence, Grok 2 [xAI], Gemini Advanced [Google], and Llama 3.1 [Meta AI]) were evaluated for use in research proposal drafting. We requested each of the models to do the following:

> *Write a research proposal for a study looking at the use of narrowband-ultraviolet B phototherapy for psoriasis treatment for psoriasis patients of varying skin pigmentation with 3 aims: 1. To understand the factors that affect the response of NB-UVB in psoriasis patients of varying skin pigmentation. 2. Evaluate adverse effects of NB-UVB and their impact on psoriasis patients of varying skin pigmentation. 3. Compare the acute immunologic response to NB-UVB*

XSL•FO
RenderX

in psoriasis patients of varying skin pigmentation using bulk and single-cell RNA sequencing. Include the following sections: 1 page 'Specific Aims' with details on each of the 3 aims, 1/2 page background and significance of the topic, 1 page of 'preliminary data/studies' relevant to the study, 1 page 'experimental design' (include summary of study, inclusion and exclusion criteria, study visits and procedures with an associated table describing specifics of study visits), 1/2 page of 'statistical methods, power calculations and bioinformatic analyses' specific for each aim, 1/4 page of 'potential problems and alternative strategies.' Please have approximately 30 references from reputable sources. Make the proposal a total of 7 pages long in paragraph form, in formal scientific language and at a graduate level.

To assess the outputs, each of the 11 blinded proposals (n=10, 90.9% LLM generated and n=1, 9.1% human written) was systematically reviewed and scored by 6 independent physician evaluators, all with strong research backgrounds. Each evaluator used a standardized Likert scale ranging from 1 to 5 (1="strongly disagree"; 5="strongly agree") to rate each proposal for capabilities and limitations in the LLMs' accuracy and comprehensiveness (Table 1).

Table . Criteria for assessing the accuracy, usability, comprehensiveness, and efficiency of large language models (LLMs).

Domain	Assessment criteria	Scoring methodology
Accuracy	Raters systematically fact-checked all proposal content. Only proposals with fully correct and verified factual information (including cited data, statistics, and conclusions) were rated highly. All references were checked for verifiability, relevance, and reputable source quality.	Rated independently by each evaluator on a Likert scale from 1 to 5 (1="strongly disagree: not accurate"; 5="strongly agree: fully accurate"). Scores were aggregated by calculating the mean of all raters' scores for each proposal.
Comprehensiveness	Assessed by evaluating inclusion and completeness of required proposal sections: specific aims, background and significance, preliminary data and studies, experimental design with inclusion and exclusion criteria and study visits and procedures, statistical methods, power calculations and bioinformatic analyses, and potential problems and alternative strategies. Proposals were further checked to meet format requirements: approximately 7 pages in length and 30 reputable references.	Rated independently on a Likert scale from 1 to 5. The mean score was calculated for all evaluators per proposal.
Usability	Assessed qualitatively based on researchers' (MH and DC) experience using each LLM. Criteria included intuitiveness of the interface, clarity of documentation, and ease of generating proposals without technical guidance.	Rated by 2 nontechnical investigators on a Likert scale from 1 to 5; scores were descriptively summarized.
Efficiency	The time from user input to final output was measured in minutes. Minimal delays and rapid response were rated favorably.	The time (minutes and seconds) for the LLM to complete the query was recorded.

For each domain assessed by human reviewers, individual scores were first tabulated. Scores from the 6 evaluators for each proposal were then aggregated by calculating the mean domain score, yielding an overall mean score per domain for each proposal. These aggregated scores provided a quantitative measure of each proposal's performance relative to evaluator consensus. No additional weighting was applied; each evaluator's score carried equal weight in the final aggregation.

In addition to scientific content review, LLM usability and efficiency, including description of pros and cons, were evaluated by 2 investigators. These qualitative evaluations were collected separately and did not contribute to the aggregated proposal scores.

Ethical Considerations

The authors have adhered to local, national, regional, and international law and regulations regarding protection of personal information, privacy, and human rights. This study did not involve human participants, identifiable private information, or interactions requiring human subjects protections. Accordingly, formal human ethics review approval was not required, and informed consent was not necessary. All data used in this study were deidentified prior to analysis to ensure participant confidentiality. No compensation was provided for participation in this study. These determinations are in accordance with University of Michigan policies and federal regulations (45 CFR 46) governing human research [4]. The research was conducted in compliance with the University of Michigan's guidelines on research ethics.

Results

LLMs Composing Proposals

The human-written proposal obtained a score of 5 for accuracy and comprehensiveness across all human scorers and remained the gold standard (Table 2). Human scorers rated ChatGPT-o1 as the most accurate and Llama 3.1 as the least accurate. When

XSL•FO
RenderX

assessed in scoring LLM-derived clinical trial proposals, LLM scorers rated ChatGPT-o1 and DeepSeek R1 as the most accurate (Multimedia Appendix 1). ChatGPT-o1 and Llama 3.1 were found to be the most and least comprehensive, respectively, by both human and LLM scorers.

Table . Full scores by evaluation criterion for each proposal and model.

Proposal and model	Accuracy (1-5), mean (SD)	Comprehensiveness (1-5), mean (SD)	Usability (1-5), mean (SD)	Efficiency
ChatGPT-4o	2.2 (1.2)	1.8 (1.4)	5.0 (0.0)	1 min, 37 s
Claude Opus	3.3 (1.4)	2.7 (0.6)	5.0 (0.0)	1 min, 30 s
ChatGPT-o1	3.5 (1.6)	4.3 (0.5)	3.5 (0.7)	1 min
ChatGPT-o3-mini	2.8 (1.7)	4.0 (0.6)	4.0 (0.0)	30 s
Claude Sonnet	2.0 (1.3)	1.8 (0.8)	4.0 (0.0)	28 s
DeepSeek R1	3.2 (1.5)	3.3 (1.4)	4.0 (0.0)	1 min, 23 s
OpenEvidence	2.3 (1.5)	1.3 (0.5)	3.5 (0.7)	45 s
Grok 2	3.2 (1.5)	3.0 (0.6)	4.0 (0.0)	1 min, 15 s
Gemini Advanced	2.5 (1.0)	1.5 (0.5)	4.5 (0.7)	37 s
Llama 3.1	1.7 (1.0)	1.5 (0.8)	4.5 (0.7)	20 s
Human proposal	5.0 (0.0)	5.0 (0.0)	N/A[a]	N/A (>10 working d)

[a]N/A: not applicable.

Mean and SD scores per criterion are reported for each proposal and model as assessed by 6 independent physician raters (except for usability, which was rated by 2 nontechnical investigators). Efficiency is reported as actual proposal generation time.

All raw scores are available in Multimedia Appendix 1.

LLMs Scoring Proposals

Overall, LLMs performed poorly on scoring proposals and, on average, rated proposals 1.9 points higher than humans for both accuracy (range 1.3-2.8) and comprehensiveness (range 0.7-3). The Claude Sonnet proposal showed the largest discrepancy between human and LLM scoring, with an average difference of 2.8 (SD 3.4) points for accuracy and 3 (SD 4.2) points for comprehensiveness. Interestingly, the ChatGPT-o1 and DeepSeek proposals both received top scores of 5 for both accuracy and comprehensiveness from all LLMs versus human averages of 4.3 (SD 2.2) and 3.3 (SD 1.9), respectively. The absence of variance at the top of the range (and wide variance in the middle of the range) suggests that the discriminatory power of the LLMs plateaued at the top LLM quality.

Ease of Using LLMs

All open access LLMs were highly efficient and ran in a matter of seconds to minutes (minimum of 20 seconds for Llama 3.1 and maximum of 1 minute and 37 seconds for ChatGPT-4o). When assessed for ease of use, ChatGPT-4o and Claude Opus offered the most intuitive interfaces and were highly usable for researchers (DC and MH) without computer science backgrounds.

Discussion

Principal Findings

LLMs offer powerful tools to assist humans in clinical trial proposal creation. LLMs take only minutes to generate proposals, whereas prior investigations into time commitment for generation of proposals by humans have reported estimates of 116 principal investigator hours, 55 coinvestigator hours, and 38 working days [5,6]. Therefore, judicious use of LLMs in proposal development allows researchers to save significant time in organizing sections, formatting, and ensuring coherence.

To provide guidance for readers, we performed a direct comparison of the tested LLMs, highlighting meaningful differences in performance, usability, and application. Table 3 summarizes these findings, with clear delineation of unique strengths and limitations for each model.

XSL•FO
RenderX

Table . Pros and cons of open access large language models (LLMs).

LLM (AI platform)	Pros	Cons
Overall	• Generally reliable, very user-friendly, and highly comprehensive and efficient	• Occasional factual inaccuracies and hallucinations (eg, fabricated references) • Lack of access to the most recent studies due to their training data cutoffs[a]
ChatGPT	• Most advanced and versatile option of the available LLMs • GPT-4o is the lowest-latency[b] and cheapest model	• Offers more advanced, paid "reasoning" models (GPT-o1 and GPT-o3), but they are computationally expensive and slower
Claude	• Designed with emphasis on alignment with human values • Tends to be more cautious about controversial or sensitive topics	• Models less tailored to clinical contexts compared to ChatGPT
DeepSeek	• Fully open source, promoting transparency and community contributions • Does not have associated license fees	• Struggles with fine-tuning on dialogue • Large models (eg, DeepSeek-Coder-33B) require large amounts of GPU[c] memory
Gemini	• Gemini 1.5 Pro boasts the largest context window[d] as a part of Google's ecosystem • Gemini 1.5 Flash is one of the fastest models	• Struggles to produce quality responses without significant prompt engineering • Concerns about data privacy and use with integration into various Google services
Grok 2	• Integration into X's (formerly known as Twitter) ecosystem allows Grok to stay up-to-date with current events and trends • Offers conversational capabilities tailored for social interaction	• Remains suboptimal compared to Claude 3.5 or GPT-4o • As a result of being directly linked to X, a platform with frequent user-generated content, Grok struggles to moderate sensitive or controversial interactions
Llama 3.1	• Llama 3.2 is one of the fastest models (along with Gemini 1.5) • Optimized for efficiency with lower computational requirements compared to other models	• Technical expertise required for it to run properly • Less user-friendly for researchers without technical support
OpenEvidence	• Offers access to the most recently curated medical research • Most robust and relevant citations	• Weaker reasoning capabilities than those of leading frontier models

[a]LLM training data cutoffs: October 2023 for ChatGPT, April 2024 for Claude Sonnet and July 2024 for Claude Haiku, December 2023 for Llama 3.1, May 2024 for Gemini, and unknown for OpenEvidence and Grok.

[b]Time to first token of tokens received, in seconds, after the application programming interface request is sent.

[c]GPU: graphics processing unit.

[d]Maximum number of combined input and output tokens.

ChatGPT-o1 and ChatGPT-o3-mini demonstrated the highest overall accuracy and comprehensiveness, delivering well-structured proposals with robust citations and high scientific rigor. Llama 3.1 and Gemini Advanced were notably efficient, reliably delivering full proposals with rapid turnaround times, but occasionally produced less nuanced sections in preliminary data or limited discussion. Regarding ease of use, ChatGPT-4o and Claude Opus feature intuitive interfaces and require minimal learning curves, making them ideal for researchers new to AI-powered tools. In contrast, Llama 3.1 and OpenEvidence ranked the lowest in usability as their technical requirements and specialized interfaces can be challenging for new users.

All open access LLMs can aid in initial outlining and creation of research proposals. They can assist in initial brainstorming of a clear researchable question and generating hypotheses based on existing literature. LLMs are useful in literature review and can summarize existing studies related to the proposal topic and identify gaps in current knowledge. Furthermore, all open access LLMs can propose data collection methods, define eligibility criteria based on study objectives, recommend appropriate statistical tests based on study design, and help draft proposal sections. They also allow for iterative refinements, enabling tailored outputs to meet specific requirements or needs. While human verification is always required, LLMs can greatly improve time spent on initial proposal drafting and aid in

XSL•FO
RenderX

mundane tasks associated with proposal writing, including proofreading and revisions, writing administrative sections, and optimizing citations.

Limitations to Consider

All LLMs operate similarly to traditional autocomplete and work by using available contextual clues and a statistical model to predict the most likely next "token" or word. Due to the training data cutoffs of AI models, researchers must manually incorporate the latest literature findings. AI researchers are working on incorporating more access to real-time data, for example, generative pretrained transformer actions [6], but these solutions come with their own trade-offs. Another limitation is that users must verify citations as the model may "hallucinate" or fabricate realistic-sounding but false information. Finally, although AI models such as DALL-E (or others) can create images, they are less effective at producing accurate, clinical-grade figures.

Additionally, current LLMs were largely unable to score proposals and should not replace human review for quality control. The high scores from the LLM raters indicate that the LLMs were unable to detect entire missed protocol sections. Other than Gemini Advanced (who self-scored its written proposal with 3 for accuracy and comprehensiveness), Claude Sonnet, and Llama 3.1, all the LLMs self-scored their own proposals with 5 for both accuracy and comprehensiveness, suggesting overlapping "blind spots" in LLM proposal generation and evaluation.

One limitation of this study is that the order in which the proposals were sent for respondents to review was not randomized. Additionally, the "gold standard" (human proposal) was last, and question order likely played a role, with kinder grading of the LLM-derived proposals before reviewing the human-written proposal. Had the human proposal been first, this would have highlighted missing components of LLM-derived proposals and likely led to harsher human grading of the latter.

Another important limitation is the rapid and frequent versioning of LLM platforms, which poses challenges for scientific reproducibility. As models are updated, their performance and outputs can meaningfully change over time, making it difficult to reproduce results or maintain consistency in studies that rely on AI-generated content. Researchers should document model versions and use dates to mitigate this issue and ensure transparency.

Conclusions

The future of AI in clinical research is expected to be transformative and far-reaching. As AI algorithms continue to evolve, they are likely to become more accurate, comprehensive, efficient, and interpretable, enabling researchers to leverage AI-driven insights for personalized medicine, disease prevention, and improved patient outcomes. In the coming years, AI is anticipated to play a crucial role in optimizing clinical trial design and accelerating drug discovery [7]. The integration of AI with other emerging technologies, such as blockchain and the Internet of Medical Things, could further revolutionize clinical research by improving data security, privacy, and real-time patient monitoring [8]. As these advancements continue to unfold, AI has the potential to democratize access to novel therapies, reduce health care costs, and, ultimately, usher in an era of precision medicine [9].

LLMs offer a transformative approach to drafting research proposals [10]. Paid versions of ChatGPT (ChatGPT-o3-mini and ChatGPT-o1) currently remain the highest-quality (as determined by the Artificial Analysis Quality Index) and most versatile option of the available LLMs, balancing usability, speed, accuracy, and customization [11]. While these tools cannot entirely replace expert input, they serve as powerful assistants, streamlining the development process and enhancing productivity. For optimal results, researchers should combine AI-generated content with their expertise, ensuring precision and adherence to the latest research standards.

Funding

The authors received no specific funding for this work.

Data Availability

The data used in this study can be made available upon request to the corresponding author.

Authors' Contributions

Study conception and design were completed by MH and TT. Material preparation and data collection were performed by MH, DC, KY, TD, JSD, AY, JC, AB, and MN. Data analysis was performed by MH. The first draft of the manuscript was written by MH and DC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Human and large language model (LLM) scoring of LLM performance on accuracy and comprehensiveness.
[DOCX File, 21 KB - derma_v9i1e76674_app1.docx]

XSL•FO
RenderX

References

1. Fliorent R, Fardman B, Podwojniak A, et al. Artificial intelligence in dermatology: advancements and challenges in skin of color. Int J Dermatol 2024 Apr;63(4):455-461. [doi: 10.1111/ijd.17076] [Medline: 38444331]
2. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial intelligence applied to clinical trials: opportunities and challenges. Health Technol (Berl) 2023;13(2):203-213. [doi: 10.1007/s12553-023-00738-2] [Medline: 36923325]
3. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. Trends Pharmacol Sci 2019 Aug;40(8):577-591. [doi: 10.1016/j.tips.2019.05.005] [Medline: 31326235]
4. eCFR :: 45 CFR Part 46 -- Protection of Human Subjects. Electronic Code of Federal Regulations (eCFR). URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46 [accessed 2026-01-05]
5. von Hippel T, von Hippel C. To apply or not to apply: a survey analysis of grant writing costs and benefits. PLoS ONE 2015;10(3):e0118494. [doi: 10.1371/journal.pone.0118494] [Medline: 25738742]
6. Herbert DL, Barnett AG, Clarke P, Graves N. On the time spent preparing grant proposals: an observational study of Australian researchers. BMJ Open 2013 May 28;3(5):e002800. [doi: 10.1136/bmjopen-2013-002800] [Medline: 23793700]
7. GPT actions: customize ChatGPT with GPT Actions and API integration. OpenAI Platform. URL: https://platform.openai.com/docs/actions/introduction [accessed 2025-01-07]
8. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
9. AI: the future of clinical research. ICON. URL: https://careers.iconplc.com/blogs/2024-2/ai-the-future-of-clinical-research [accessed 2025-03-19]
10. Markey N, El-Mansouri I, Rensonnet G, van Langen C, Meier C. From RAGs to riches: utilizing large language models to write documents for clinical trials. Clin Trials 2025 Oct;22(5):626-631. [doi: 10.1177/17407745251320806] [Medline: 40013826]
11. Independent analysis of AI. Artificial Analysis. URL: https://artificialanalysis.ai [accessed 2025-01-15]

Abbreviations

AI: artificial intelligence
LLM: large language model

Edited by R Alhusayen; submitted 28.Aug.2025; peer-reviewed by RQ Wang, T Karen; revised version received 07.Nov.2025; accepted 06.Dec.2025; published 12.Jan.2026.

Please cite as:
Hauptman M, Copley D, Young K, Do T, Durgin JS, Yang A, Chang J, Billi A, Nakamura M, Tejasvi T
Leveraging AI Large Language Models for Writing Clinical Trial Proposals in Dermatology: Instrument Validation Study
JMIR Dermatol 2026;9:e76674
URL: https://derma.jmir.org/2026/1/e76674
doi:10.2196/76674

© Megan Hauptman, Daniel Copley, Kelly Young, Tran Do, Joseph S Durgin, Albert Yang, Jungsoo Chang, Allison Billi, Mio Nakamura, Trilokraj Tejasvi. Originally published in JMIR Dermatology (http://derma.jmir.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of TikTok in Education on Hidradenitis Suppurativa in Skin of Color: Cross-Sectional Analysis

Arsema K Zadu*, MS; Jordan Young*, BS; Janyla A Seltzer*, MD; Angel S Byrd*, MD, PhD; Cheri Frey*, MD

Department of Dermatology, College of Medicine, Howard University, 2041 Georgia Ave NW, Washington, DC, United States

*all authors contributed equally

Corresponding Author:
Arsema K Zadu, MS
Department of Dermatology, College of Medicine, Howard University, 2041 Georgia Ave NW, Washington, DC, United States

Abstract

Abstract: This study analyzed 50 TikTok videos returned by a search for "hidradenitis suppurativa in Black skin," revealing that nearly half were patient-created, few had physician involvement (n=10, 20% dermatologists; n=7, 14% plastic surgeons), and few had commission-based (n=7, 14%) or sponsored content (n=2, 4%); they were predominantly patient testimonials on various treatments, highlighting the need for greater physician engagement to address patient needs, hidradenitis suppurativa product safety, and efficacy.

(JMIR Dermatol 2026;9:e71566) doi:10.2196/71566

KEYWORDS

TikTok; skin of color; hidradenitis suppurativa; patient education; social media

Introduction

TikTok has become a prominent platform and search engine for dermatologic information [1]. A majority of videos on the top 20 most commonly diagnosed skin conditions were created by patients [2]. This study aimed to evaluate the primary sources of education on hidradenitis suppurativa (HS) in Black skin on TikTok.

HS is a chronic inflammatory condition marked by sinus tracts, nodules, and abscesses in intertriginous areas [3]. It often affects African American women and is frequently misdiagnosed, leading to delayed identification, accelerated progression, and treatment challenges [4]. Black patients experience a 1.5-year longer diagnostic delay than White patients, and after diagnosis, they typically wait 5 years to see a dermatologist compared to 3 years for White patients [5]. Early intervention is critical to prevent advanced-stage manifestations, which are often refractory to standard therapeutic approaches [6].

A study on dermatologist visibility on TikTok found that most dermatologic education videos are made by individuals without formal medical training [7]. While these platforms allow patients to connect and share experiences with skin conditions, the risk of spreading misinformation is significant, potentially worsening disease severity and reducing treatment effectiveness. This study aimed to describe who produces HS TikTok content featuring Black skin and to quantify the treatments, products, and themes presented.

Methods

TikTok was selected because users often turn to the platform as a general search tool, including for health topics related to HS [1]. The app was used to search for "hidradenitis suppurativa in black skin" in the search bar. A total of 50 videos were viewed one by one by a single reviewer to assess the creator's role, brands or products discussed, key themes, and whether the products were part of paid sponsorships or if the creator received commission from sales.

Ethical Considerations

All data were deidentified prior to analysis. User IDs, screenshots, images, and quoted content contained no personally identifiable information. Data were analyzed and reported in anonymized form only.

Results

Of the 50 analyzed videos, 24 (48%) were created by patients. Board-certified dermatologists produced 10 (20%) videos, while board-certified plastic surgeons produced 7 (14%) videos. One (2%) video was created by a nurse practitioner. Beauty service providers accounted for 2 (4%) videos. The "other" category, which included dietitians, social workers, product creators, and creators with unclear professional titles, accounted for 6 (12%) videos. Board certification status was verified using creator profile biographies and linked professional websites.

Seven videos featured products associated with sales commissions, and 2 videos involved paid sponsorships.

XSL•FO
RenderX

Treatment content represented the largest category, comprising 35 (70%) videos (Table 1). Fourteen (28%) of the 50 videos focused on explaining HS. Of these, 9 (18%) were created by health care professionals, 4 (12%) by patients, and 1 by a product creator. The remaining videos focused on living with HS and dietary approaches to symptom management.

Table . Video themes stratified by creator type across 50 analyzed videos.

	Video type, n				
Creator type	Education	Medical treatment	Over-the-counter treatment	Lifestyle	Mental health
Patient	4	4	14	4	3
Board-certified dermatologist	7	3	5	0	0
Board-certified plastic surgeon	1	2	3	0	0
Nurse practitioner	1	0	0	0	0
Beauty service provider	0	1	1	0	0
Other	1	0	3	0	0

Products appeared across multiple categories, including face masks and exfoliants, body scrubs and cleansers, acne treatments, topical oils, salves, body butters, creams, antiseptics, spot treatments, systemic medications, and hair or scalp care, as shown in Table 2. PanOxyl and Hibiclens washes were the most frequently recommended products, appearing in 5 (10%) and 6 (12%) videos, respectively. Magic Healer body butter was mentioned in 4 (8%) videos, and Humira was mentioned in 3 (6%) videos. CeraVe acne foaming cream cleanser, a turmeric kojic acid soap, and Tend Skin Solution were all separately mentioned in 2 (4%) videos. All other products were mentioned in one video each.

Table . Categories of products mentioned across 50 analyzed hidradenitis suppurativa–related videos.

Category	Products
Face masks and exfoliants	Skinfix Glycolic Renewing Mask, Cocokind Turmeric Mask
Body scrubs and cleansers	Skinfix body scrub, Olay body wash, Naturium Vitamin C body wash, lemon turmeric and kojic acid soap, Dr. Bronner's Soap, turmeric kojic acid soap, Dial bar soap
Acne cleansers and treatments	PanOxyl, PanOxyl acne foaming wash, Inkey List 5 % Benzoyl Peroxide Cleanser, CeraVe acne foaming cream cleanser, Zapzyt, Tend Skin Solution
Topical oils and salves	Relief Natural Company, GuruNanda Tea Tree, Zunda Turmeric, black seed oil, vitamin E, clove water, Magic Healer body butter, Palmer's Body Oil
Body butters, creams, and moisturizers	HS Body Butter, Fenty Skin Cream, Healing Ocean Cream
Antiseptics and antimicrobials	Hibiclens, Magic Healer Product
Specialty treatments and patches	Mighty Patch
Medications	Humira
Hair and scalp care	Head & Shoulders

Discussion

Patients produced the majority of lifestyle-based and experiential content, while physicians concentrated on treatment education and procedural intervention. Commercial messaging remained isolated to product owners and patient testimonials, with minimal clinical creator involvement. This distribution highlights a content gap between medically accurate education and the lived experience narrative dominating public HS discourse on TikTok. Research shows that only about 20% of skin of color videos are created by board-certified dermatologists [8]. These findings highlight an opportunity for dermatologists on TikTok to engage with patients and address the safety and efficacy of these products. Dermatologists, who already have a strong social media presence, can foster patient trust by encouraging open discussions about non–medical-grade treatments. By acknowledging the value of herbal and alternative remedies that patients find helpful, they can assess their safety and efficacy while supporting their continued use when appropriate. Meanwhile, they can provide evidence-based guidance to minimize the risks of harmful or ineffective treatments.

XSL•FO
RenderX

Acknowledgments

AKZ is a recipient of the 2024-2025 Women of Dermatology Society and La-Roche Posey Diversity in Dermatology Fellowship at Howard University.

Conflicts of Interest

ASB is the inaugural recipient of the Skin of Color Society Career Development Award as well as the Society for Investigative Dermatology Freinkel Diversity Fellowship Award, and a recipient of the Robert A. Winn Excellence in Clinical Trials Career Development Award (Winn CDA) funded by Bristol Myers Squibb Foundation (BMSF); consultant for SENTÉ, Inc. and Sonoma Biotherapeutics; Advisory boards for Merck and Novartis. CF is a Primary Investigator for Novartis. All other authors declare no conflicts of interest.

References

1. Zadu AK, Adibe I, Welch D, et al. Social media as a catalyst for health education and early diagnosis of hidradenitis suppurativa. J Drugs Dermatol 2025 Jun 1;24(6):e45. [Medline: 40465496]
2. Villa-Ruiz C, Kassamali B, Mazori DR, Min M, Cobos G, LaChance A. Overview of TikTok's most viewed dermatologic content and assessment of its reliability. J Am Acad Dermatol 2021 Jul;85(1):273-274. [doi: 10.1016/j.jaad.2020.12.028] [Medline: 33359080]
3. van Straalen KR, Prens EP, Gudjonsson JE. Insights into hidradenitis suppurativa. J Allergy Clin Immunol 2022 Apr;149(4):1150-1161. [doi: 10.1016/j.jaci.2022.02.003] [Medline: 35189127]
4. Zouboulis CC, Goyal M, Byrd AS. Hidradenitis suppurativa in skin of colour. Exp Dermatol 2021 Jun;30 Suppl 1(S1):27-30. [doi: 10.1111/exd.14341] [Medline: 34085333]
5. Serrano L, Ulschmid C, Szabo A, Roth G, Sokumbi O. Racial disparities of delay in diagnosis and dermatologic care for hidradenitis suppurativa. J Natl Med Assoc 2022 Dec;114(6):613-616. [doi: 10.1016/j.jnma.2022.08.002] [Medline: 36511276]
6. Andersen RK, Jemec GBE. Treatments for hidradenitis suppurativa. Clin Dermatol 2017;35(2):218-224. [doi: 10.1016/j.clindermatol.2016.10.018] [Medline: 28274363]
7. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
8. Campbell JR, Woolery-Lloyd H, Williams K. TikTok and black skin: is this a missed opportunity for dermatologists? J Drugs Dermatol 2023 Jul 1;22(7):698-700. [doi: 10.36849/JDD.7061] [Medline: 37410044]

Abbreviations

HS: hidradenitis suppurativa

Edited by J Meisenheimer; submitted 26.Jan.2025; peer-reviewed by ED Quincey, G Padilla-Castillo, S Bestavros; revised version received 15.Dec.2025; accepted 15.Dec.2025; published 27.Feb.2026.

Please cite as:
Zadu AK, Young J, Seltzer JA, Byrd AS, Frey C
The Role of TikTok in Education on Hidradenitis Suppurativa in Skin of Color: Cross-Sectional Analysis
JMIR Dermatol 2026;9:e71566
URL: https://derma.jmir.org/2026/1/e71566
doi:10.2196/71566

© Arsema K Zadu, Jordan Young, Janyla A Seltzer, Angel S Byrd, Cheri Frey. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Harmonized Dual Deep Learning Architectures for Image-Based Diagnostics of Skin Neglected Tropical Diseases: Benchmark Study via Novel Funnel Framework

Yohannes Minyilu[1*], MSc; Mohammed Abebe Yimer[1*], PhD; Million Meshesha[2*], PhD

[1]Faculty of Computing and Software Engineering, Institute of Technology, Arba Minch University, SE, Arba Minch, Ethiopia

[2]School of Information Science, College of Natural and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia

[*]all authors contributed equally

Corresponding Author:
Yohannes Minyilu, MSc
Faculty of Computing and Software Engineering, Institute of Technology, Arba Minch University, SE, Arba Minch, Ethiopia

Abstract

Background: While deep learning–based methods are the potential technological solutions for the diagnosis of skin Neglected Tropical Diseases (skin NTDs), limited efforts were seen toward the use of such tools in Ethiopia. Data scarcity, methods, and models selection issues created further challenges in an attempt to close the previous gap.

Objective: This study attempts to design a benchmark image-based diagnostic model for skin NTDs through a synergistic combination of feature extraction pretrained models, a custom-designed convolutional neural network (CNN) model trained on the extracted features, and an integrated data augmentation method applied dynamically.

Methods: For this study, a new skin images dataset is created using skin photographs collected by a team of researchers from the NTDs research center of Arba Minch University Medical College. The new dataset contains 1495 images in 3 classes having severe class imbalance. Extensive experiments were conducted to find the optimal deep learning approach by designing a new CNN model, applying transfer learning, and designing the 2-stage approach that uses pretrained models for feature extraction and trains the new CNN model using the extracted features from the pretrained models and applying data augmentation based on the integrated 2-stage approach. For model selection, the study proposed a novel approach, the funnel framework with cascaded selection of methods and models.

Results: After hyperparameter tuning, the model trained using DenseNet121 feature extractor scored the highest accuracy of 96.6%, F_1-score of 95%, and sensitivity of 95%, while the MNv2-based model scored comparable results of 95.6% accuracy, 90% F_1-score, and 90% sensitivity. This study finally selected the DenseNet121 and MNv2 models for feature extraction to build the final model for skin NTDs classification.

Conclusions: The 2-stage approach significantly boosted the models' performance compared with other methods, while the data augmentation method further enhanced the performance of the selected models. Finally, this study suggests further studies using advanced class-balancing methods with more data and a possible integration of other clinical data types.

(JMIR Dermatol 2026;9:e91544) doi:10.2196/91544

KEYWORDS

skin NTDs classification; 2-stage approach; feature extraction; funnel framework; hyperparameter optimization

Introduction

Background

Neglected Tropical Diseases (NTDs) represent 21 different diseases, including podoconiosis, scabies, and tungiasis, affecting more than 1 billion people globally among underserved communities in the tropical areas [1]. As a tropical country, NTDs are highly prevalent in Ethiopia, with the majority of NTDs identified by the World Health Organization present except for Chagas disease and yaws [2], particularly in the remote areas of the country [3,4]. As estimated by the federal ministry of health [5], more than 75 million people are at risk of contracting at least 1 NTD. Of the 21 NTDs recognized by the World Health Organization, more than half (about 18 of them) have skin manifestations and are called skin NTDs [6-8]. Accordingly, the diagnosis of skin NTDs primarily involves examination of patient skin, presenting opportunities for integrated diagnosis [8], involving the use of artificial intelligence–based skin NTDs diagnostic tools [9]. Although it is not fully explored, previous studies [10-12] have shown that deep learning (DL) methods can be used for the diagnosis of skin NTDs.

XSL•FO
RenderX

This study proposes a DL-based diagnostic model for skin NTDs based on skin images of patients using a new skin images dataset we created for this study. However, apart from the limited previous efforts in DL-based skin NTDs diagnostics, dataset-related issues created challenges in building our proposed model, as the dataset used is characterized by small-sized image samples with a severe class imbalance. Generally, as a computer vision task, the development of a DL-based diagnostic model using skin images requires a large-scale higher-quality skin images dataset [11]. In the case of skin NTDs, data scarcity and dataset-related issues are the major challenges in building DL-based intelligent diagnostic tools for skin NTDs [13,14]. These challenges mainly arise from several factors that include poor record keeping, management, and reporting practices [13]. Additionally, model characteristics regarding usability and efficiency issues are also major challenges in building diagnostic models for skin NTDs. Beyond prediction accuracy, efficiency parameters, such as model complexity, inference speed, response time, and deployment platform options (web-based and mobile-based), are also major factors that determine the selection of DL tools, techniques, and models.

Accordingly, given the data-related issues of the dataset we used for this study and the expected operational platforms, which DL method is appropriate to develop an image-based diagnostic model for skin NTDs? While answering this question requires properly devised strategies based on carefully designed experiments, analysis, and interpretation of results, it establishes a foundational benchmarking effort that helps in identifying suitable DL methods to overcome the mentioned challenges. Therefore, this study conducts extensive experimentations to find the optimal DL solution based on the following guiding questions: (1) Which DL approach (baseline model design, transfer learning, or hybrid) would be a feasible strategy to develop the proposed model? (2) Which DL model architecture would help in developing a high-performance diagnostic model for skin NTDs, given the nature of the dataset used? (3) Which approach helps in creating the robust model development pipeline based on the experimental screening of both DL methods and architectures that collectively address high predictive performance with lower architectural and computational complexity?

While addressing these questions, this study develops a benchmark image-based diagnostic DL model for skin NTDs based on experimentally identified suitable methods and approaches that involve designing a new convolutional neural network (CNN) model and applying transfer learning. Evidently, the data scarcity created higher difficulty in capturing relevant features from input images using baseline models, including the custom-designed model, given the diversity and nature of the manifestations of skin NTDs that include “mossy” limbs in podoconiosis [14]. While the use of transfer learning is ultimately the recommended DL strategy, several factors related to pretrained models create challenges that include huge data requirements, domain incompatibility, tendency of capturing noise features, and lack of a standardized robust diagnostic model development pipeline including skin NTDs. To address such challenges, we designed and implemented the 2-stage approach that presents a robust and extensible architectural pipeline integrating the feature mapping (extraction) models and applying domain adaptation. Furthermore, as the 2-stage approach integrates the 10-layer classification head with different regularization methods, it provides deeper feature filtering architectures.

Overall, this study presents several contributions to the problem domain (skin NTDs diagnostics) and to the field through multiple achievements, which include identification of optimal DL methods for skin NTDs that have higher data scarcity problems; establishment of a robust DL model development pipeline, which incorporates designing the 2-stage approach; methodological rigor that includes robust experimental setup and systematic architectural screening by adopting the funnel framework; and, ultimately, development of DL diagnostic models for skin NTDs, which can serve as an architectural benchmark for skin NTDs (skin-related diseases in general).

Related Works

Previous studies showed the potential of the DL-based methods for skin NTDs. Accordingly, Steyve et al [10] proposed an optimized real-time diagnostic approach for 3 skin NTDs (Buruli, leishmaniasis, and leprosy) using a support vector machine classifier optimized by a black hole optimization algorithm. Yotsu et al [15] also presented DL methods for using major CNN architectures (ResNet50 and VGG16 models) for the diagnosis of 5 skin NTDs (Buruli ulcer, leprosy, mycetoma, scabies, and yaws). Another study by Pattnayak et al [16] proposed a DL method for 5 skin NTDs (Buruli ulcer, leprosy, mycetoma, scabies, and yaws). Beesetty et al [17] applied a Siamese-based Few Shot Learning model, trained it on an extremely small dataset with fewer disease classes (368 clinically diagnosed leprosy and 28 nonleprosy skin lesions), and reported higher accuracy.

Challenges Toward Applying DL Methods for Skin NTDs

Multiple factors created challenges toward digitizing the skin NTDs diagnostic procedures using intelligent digital diagnostic tools. Some of the challenges are insufficient infrastructure, data security issues, and limited efforts toward the integration of digital diagnostic tools [18], including the general DL model development challenge, data scarcity, and class imbalance. Multiple large-scale skin image datasets, such as the HAM10000 (Human Against Machine with 10,000 training images) [19] and ISIC (International Skin Imaging Collaboration) [20], are available to train DL models for non-NTD skin diseases. However, it is difficult to find such massive skin image datasets that are publicly available to train DL models for skin NTDs. The other major issues related to data scarcity are the completeness and class distribution imbalance.

Potential DL Solutions for Intelligent Skin NTDs Diagnosis

Data augmentation, mostly for image classification tasks, is the most widely used machine learning operation to address the problem of data scarcity and distribution imbalance [21-23] through artificially generating images. There are 2 major approaches of data augmentation. The traditional augmentation

XSL•FO
RenderX

method uses basic general geometric transformations [21,24], such as cropping, padding, flipping (horizontal or vertical), rotations, permutations, scaling, translations, and addition of noise [23,25]. The cropping, flipping, rotations, and scaling transformations are applied to simulate imaginary variations that might occur in reality. The other data augmentation approach is the class-based conditional augmentation, which is conducted based on predefined conditions either by applying basic geometric transformations or by using generative adversarial (GAN) models [23,26,27].

Gaps Identified

Overall, the vast literature exploration confirmed that very few efforts were seen toward the use of DL-based diagnostic tools for skin NTDs, specifically in the Ethiopian context. Additionally, the data scarcity issues are the subsequent challenges creating the other major research gaps. These major research gaps clearly suggest that further efforts in the area are mandatory, with a clear indication of having an initially established DL-based diagnostic framework that can serve as a benchmark for current and future studies. Therefore, we conducted this foundational study to develop an image-based skin NTDs diagnostic model using a novel skin NTDs image dataset created by using skin photographs of patients collected from one of the remote and highly affected areas in Ethiopia. We conducted this benchmarking study to identify optimal methods and DL model architectures based on properly designed experimental settings, given the nature of the dataset used for the study. Regarding the dataset-related problems, the study demonstrates the dynamic (online) data augmentation method based on the standard general geometric transformations to initially address the data scarcity problem.

Methods

Ethical Considerations

For this study, we created a new dataset using skin photographs of patients with skin NTDs obtained during clinical data collection from a remote affected area in the southwest of Ethiopia. While the data were initially collected by trained health care professionals based on strict ethical procedures using an ethical clearance letter obtained from the institutional review board (IRB) in Arba Minch University (AMU), we acquired the data through institutional research collaboration ("Acknowledgments" section) that warrants full access and use of the firsthand data. To acquire and use the skin images data for this study, proper ethical approval procedures were followed, starting with the acquisition of the ethical clearance letter from the IRB in AMU. Hence, for this study, the authors obtained a specific ethical clearance letter from the concerned institutional review board in AMU (approval protocol number YM23161).

The data collection process was conducted with the full respect of participants' privacy, where data were acquired from each individual participant based on their will confirmed by a written consent. This was achieved using a dedicated checklist prepared and used to systematically monitor the data collection, ensuring that they were collected based on the free will of each participant. This, specifically, involved a mandatory checklist to ensure that each participant has read and signed the consent form using targeted questions, which include the following: "Does the participant read information statement and willing to sign the consent form?" and the consideration of underaged participants—"if the age is below 18, does the guardian sign the consent form?" Overall, all involved participants provided written consent for their participation and authorized the use of their clinical data for academic purposes, including publications.

As the data collection was performed by a team of trained health care professionals, including public health officers, the entire data collection process was carried out in a professional and responsible manner, authorized by the IRB. Regarding this study, we have been authorized by the same institutional ethical review committee before acquiring and using the data based on a series of verification procedures that include careful analysis of the required data for our study, analysis of participant privacy vulnerability issues as a result of using the data, anonymizing all participant-level data in the images to remove all information in the images that could be used to identify patients (study participants), and strictly complying with ethical requirements for using human patient data—which was confirmed by the ethical clearance letter from the IRB. Therefore, before using the images for this study, all images have been anonymized to remove any personal data to ensure that all participant-identifiable features in any images of the manuscript or supplementary material are not visible.

Finally, as the data were collected from one of the highly affected areas during a mass drug administration (MDA) campaign, the ultimate goal of the data collection was to assess the burden of the skin NTDs that will be used for immediate public health decisions. In this process, the patients living in the affected remote community primarily benefited from the MDA-based data collection, with the patients being diagnosed at the MDA site. However, no special compensations were implemented for the participants as a result of using the data for further study. Additionally, the acquired data are used to build a diagnostic model that intends to serve for diagnosing skin NTDs in the same resource-limited areas.

Data Collection and Dataset Description

In this study, we used a new handcrafted dataset containing skin photographs of patients with skin NTDs that were captured to show skin areas affected by the skin NTDs. Initially, the data were collected by a team of researchers from the Collaborative Research and Training Center for Neglected Tropical Disease, College of Medicine Health Sciences of AMU. Data were collected in a project-based research for the assessment of skin NTDs burden through community screening during the scabies MDA campaign from Gacho Baba District, Gamo Zone, southwest of Ethiopia. The dataset contains skin photographs (images) of 3 different skin NTDs, namely, podoconiosis, scabies, and tungiasis. These 3 diseases were included in the dataset since they are identified as the most prevalent skin NTDs identified in the specified affected area. The entire data collection process was conducted in a professional and ethical manner, where the whole process was initiated after all legal and ethical issues were addressed, and an ethical clearance was obtained to collect the data. For this study, we acquired the collected data through institutional research collaboration

XSL•FO
RenderX

between the NTDs research center of the medical college and computing faculty of the technology institute of AMU. Using the acquired data, we created a new skin NTDs image dataset and used it for this study to develop the proposed DL-based skin NTDs diagnostic model.

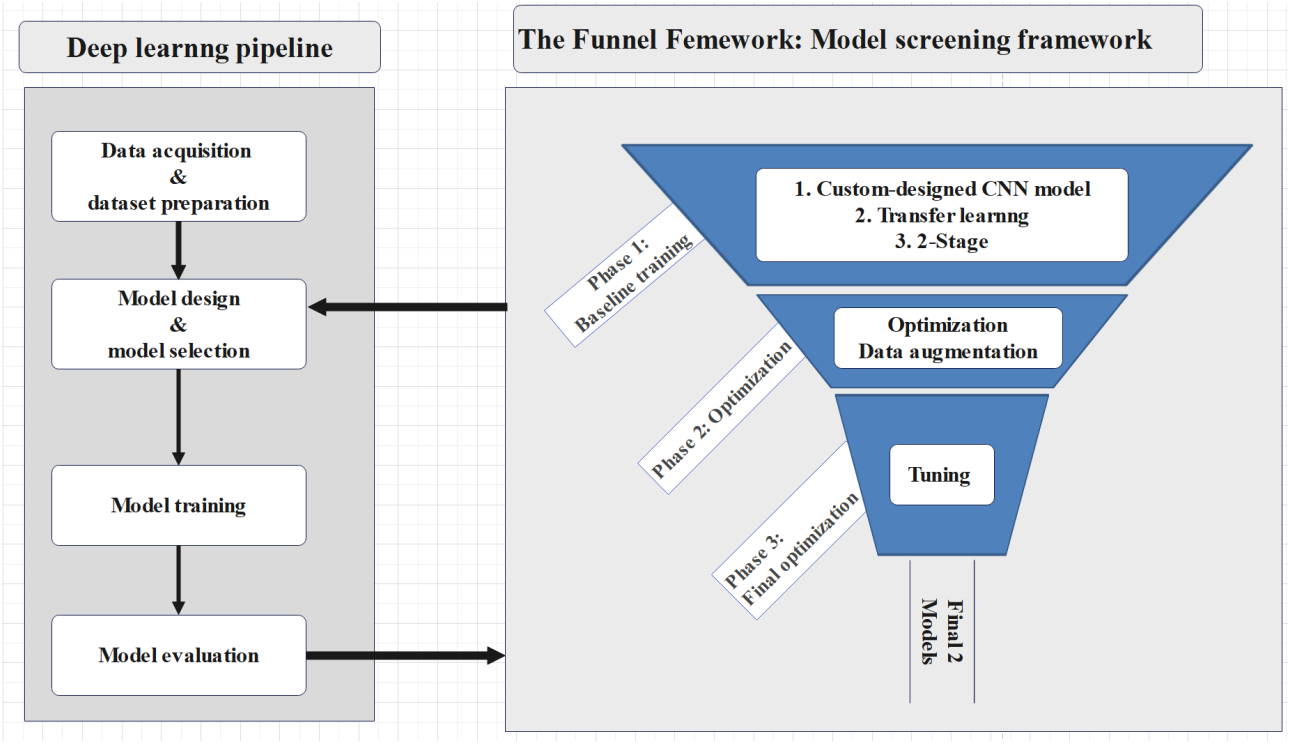
Exploratory Data Analysis

The images were obtained for dermatologist verification and present skin lesions, scratches, excoriations, and other infestations that are typical clinical signs of podoconiosis, scabies, and tungiasis. After acquiring the data, we created 3 separate initial datasets containing the skin images, the unique image IDs, and their corresponding labels for each disease. The final dataset contains 1495 images, as shown in Multimedia Appendix 1. As shown in the statistical distribution, scabies has the largest proportion among the 3 disease classes with a total of 955 instances (955/1495, 63.88%), while tungiasis represents the second largest size with 474 instances (474/1495, 31.71%) and podoconiosis having only 66 instances (66/1495, 4.41%).

Study Design

This study intends to conduct a DL architectural benchmarking research that requires a systematic approach to select optimal DL methods and algorithms, and we proposed a mixed research strategy based on the newly proposed funnel framework we adopted for this study, as summarized in Figure 1. Overall, our study needs a systematic approach to select optimal DL methods and models through multistaged experimental filtering. Specifically, the selection of DL pretrained models requires multiple experimentations with systematically devised screening and filtering methods based on comparative analysis of model performance results. Initially, we propose and experiment with 3 different DL methods: first, train a new custom-designed CNN model; second, transfer learning using the 21 selected pretrained DL models; and third, demonstrate the 2-stage approach.

Figure 1. Architecture of the proposed funnel framework based on the cascaded model selection mechanism, CNN: convolutional neural network.



After the completion of these 3 experiments, the best approach and top 5 pretrained DL models will be selected to apply further enhancements to finally select the best 2 models. All these tasks required a systematically designed approach that can be used as a framework to guide the training, screening, and analysis processes. Therefore, we propose a new approach, the funnel framework with cascaded (phased) selection of models and DL methods, as shown in Figure 1. The funnel framework, adapted from the business-related fields, is used to screen out top-performing models and methods initially identified based on a comparative and phased approach. Accordingly, the experiments are conducted in different training settings to select the optimal DL approach that produces comparatively maximum model performance and screen out the high-performing pretrained models in transfer learning and feature extraction.

Dataset Preparation and Preprocessing

The entire data-splitting process is conducted using the stratified splitting approach with a ratio of 80:20 train-test split, followed by data preprocessing operations. Accordingly, for image resizing, we applied the standard image resolution to all images in the dataset, which required resizing the images to 224 × 224 × 3 pixels for 17 models, while 240 × 240 × 3 pixels and above were used to resize the images for higher EfficientNet models (B1, B3, B5, and V2S). As a next task in data preprocessing, image normalization was programmatically applied to all images.

Model Development: Model Design and Selection

In this study, we use 3 strategies regarding the development of the proposed DL model, which include (1) training a new custom-designed CNN model, (2) applying transfer learning

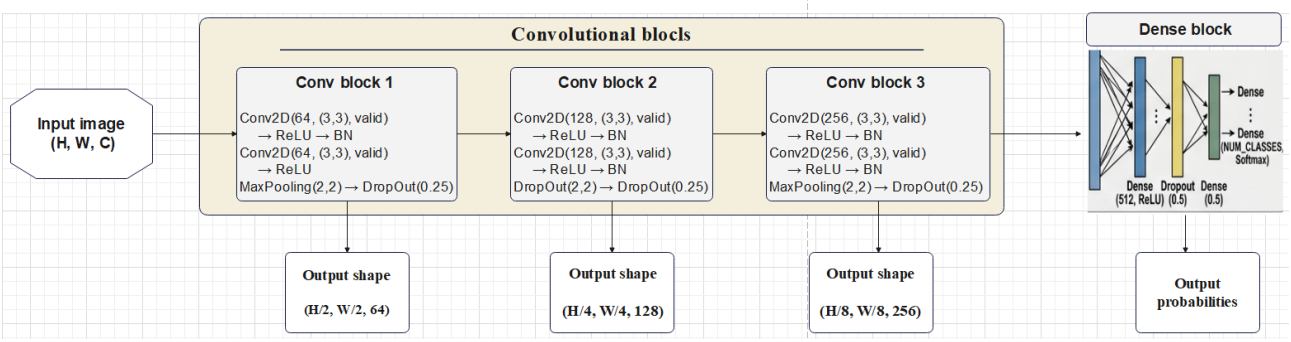
XSL•FO
RenderX

using selected pretrained DL model architectures, and (3) applying the proposed 2-stage approach—a hybrid approach that uses pretrained DL architectures for feature extraction and training the new CNN model for classification. We applied this set of strategies to experimentally demonstrate possible methods to select the optimal strategy resulting in overall higher performance of models.

The New CNN Model

As a first study, we begin our experiments by designing a new custom-designed CNN model that will be used for baseline training and evaluation as well as for the classification of the skin NTDs in the 2-stage approach. Accordingly, we designed a new CNN model consisting of 3 major components (blocks) that represent the 3 different stages of the entire pipeline: feature extraction, the dense layers (including the flattening layer), and the classification head. Based on this general architectural layout, the new model is designed to have 30 layers, containing 8 weight-bearing layers from 6 convolutional and 2 dense layers, 11 regularization layers properly applied across all blocks, 7 activation layers, and 4 spatial refinement layers—3 pooling and 1 flatten layer. Figure 2 visualizes the overall architecture of our custom-designed CNN model.

Figure 2. Architecture of the proposed custom-designed convolutional neural network model.



The model is designed in a consistent manner, implementing a hierarchical structure to capture complex and relevant visual patterns from each input sample image using 3 repeated convolutional blocks having variable filter sizes. Accordingly, the output depth of each block (number of filters) is sequentially doubled across the 3 feature extraction blocks (64 filters in block 1 to 256 filters in block 3). Internally, each feature extraction block is designed to have 2 convolutional layers (Conv2D with 3 × 3 kernels), including properly applied activation, regularization, and dimensionality reduction techniques. Accordingly, the model uses the “relu” activation function, along with robustly implemented regularization methods, which include (1) normalization (BatchNormalization)—applied across all blocks including the dense layers, (2) dropout—applied both in the feature extraction blocks (using Dropout [0.25]) and in the dense block (using Dropout [0.5]), and (3) pooling layers that apply the maximum pooling method—serving dual purposes including regularization, while the maximum pooling method is primarily used for dimensionality reduction [28]. This architectural setup is used to ensure that each feature extraction block extracts and refines features and applies normalization, all before dimensionality reduction (Dropout), to ensure that relevant features and patterns are captured. The overall feature dimension transformations and output shapes of each extraction block are presented in Table 1.

Table . Architectural summary of the new custom-designed CNN[a] model.

Model block	Input shape	Convolutional layers	After pooling (2 × 2)	Output
Input layer	224 × 224 × 3	—[b]	—[b]	3
Block 1 (convolutional block 1)	224 × 224	220 × 220	110 × 110	64
Block 2 (convolutional block 2)	110 × 110	106 × 106	53 × 53	128
Block 3 (convolutional block 3)	53 × 53	49 × 49	24 × 24	256
Dense block	Flattened vector	—[b]	512 units	3: number of classes

[a]CNN: convolutional neural network.

[b]Not available.

On the final block (classification head), the model applies feature map transformation and final skin NTDs classification using 6 layers: flatten—the layer used to transform the final spatial feature map into a 1D feature vector, fully connected (Dense)—the largest layer in the model having 512 units, regularization—2 independent layers applying BatchNormalization and heavier dropout (Dropout [0.5]), activation layer using “relu,” and output layer—the final layer that predicts the probability distribution among the 3 disease classes (podoconiosis, scabies, and tungiasis) using the SoftMax

XSL•FO
RenderX

activation function. Overall, the Adam optimizer and "categorical cross entropy" loss function are used for the final model compilation. All these strategies are properly applied along with a synchronized implementation of early stopping, all to prevent overfitting.

Given the nature of our new dataset, having only 1495 images in the dataset, training a DL model from scratch appeared to be a bit challenging, as the limited size of the feature maps would potentially force the models to learn all the details (including noise pixels) resulting in difficulty to generalize well on new skin NTDs images due to overfitting [28,29]. To overcome this challenge, we conducted further experimental inquiries to identify and use the optimal DL method based on our newly designed CNN model, demonstrating the transfer learning method entirely based on pretrained DL models, followed by the 2-stage approach.

Transfer Learning

To improve the performance of our baseline CNN model and demonstrate the other potential DL methods, we deployed the transfer learning method by using a diverse set of pretrained DL architectures (CNN and transformer-based). To achieve this goal, we applied a systematic DL architectural selection procedure using a predefined set of model selection parameters to validate, comparatively analyze, and finally select the best model and method for the proposed skin NTDs diagnostic model. Hence, we identified 21 pretrained model architectures, selected based on 4 major selection parameters, which include architectural distribution—defining model family and operational principles (CNN and transformer-based models), model complexity—including both architectural (model size) and computational (efficiency) complexity, and novelty (recency) of models. Table 2 presents a summary of the 21 selected pretrained models.

XSL•FO
RenderX

Table . Summary of DL[a] model architectures considered during initial screening.

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
ResNet50	Residual Networks	Skip connections [30]	25.6 M	Large	4.1	Heavy
ResNet18	Residual Networks	Skip connections [30]	11.3 M	Moderate	1.8	Moderate
ConvNext-Small	Modern CNN[c] Architectures	Transformer-like CNN components using standard ConvNet modules [31]	49.7 M	Large	4.5	Heavy
ConvNext-Tiny	Modern CNN[c] Architectures	Transformer-like CNN components using standard ConvNet modules [31]	28M	Moderate	4.5	Heavy
CovNeXtv2-Tiny	Modern Pure CNN Architectures	Fully convolutional masked autoencoder framework with a global response normalization layer [32]	28.1M	Large	4.47	Moderate
CovNeXtv2-Atto	Modern Pure CNN Architectures	Fully convolutional masked autoencoder framework with a global response normalization layer [32]	3.7M	Lightweight	0.55	Lightweight
DenseNet121	Densely Connected Networks	Dense CNN (DenseNet) block architecture with feature reuse [33]	7.3M	Moderate	2.9	Moderate
Xception	Depth-wise Separable CNNs	Depth-wise separable convolutions [34]	21.4	Large	8.4	Heavy
EfficientNetB5	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	29M	Large	9.9	Heavy
EfficientNetB3	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	11.2M	Moderate	1.8	Moderate
EfficientNetB1	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	6.9M	Moderate	0.7	Moderate

XSL•FO
RenderX

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
EfficientNetB0	Compound-Scaled CNNs	Compound scaling—uniform scale of all dimensions (depth, width, and resolution), built on BMConv[d] blocks [35]	4.4M	Lightweight	0.39	Lightweight
MobileNetV2	Mobile CNNs	Inverted residual blocks with linear bottlenecks [36]	2.6M	Lightweight	0.3	Lightweight
MobileNetV3-Large	Mobile CNNs	Hardware-aware NAS[e] along with the NetAdapt algorithm (platform-aware adaptation [37,38]	3.2M	Lightweight	0.22	Lightweight
MobileNetV3-Small	Mobile CNNs	Hardware-aware NAS[e] along with the NetAdapt algorithm (platform-aware adaptation [37,38]	1.1M	Lightweight	0.06	Lightweight
EfficientNetV2B0	Advanced compound-scaled CNNs	Training-aware NAS and scaling—joint optimization of training speed and parameter efficiency [39]	6.2M	Moderate	0.72	Moderate
EfficientNetV2S	Advanced compound-scaled CNNs	Training-aware NAS and scaling—joint optimization of training speed and parameter efficiency [39]	20.7M	Moderate	2.9	Moderate
Recent CNN and transformer-based models						
RepViT	ViT[f]-Inspired Pure lightweight CNN	Reparameterization convolutions in ViT-like Meta-Former structure [40]	5.1M	Moderate	0.80	Lightweight
FasterViT-0	Hybrid (CNN + ViT)	HAT[g] using window-based self-attention, carrier tokens for local-global representation learning [41]	31.4M	Large	3.34	Moderate
FastViT	Hybrid (CNN + ViT)	RepMixer (structural reparameterization for token mixing) and skip-connection elimination [42]	3.6M	Lightweight	0.70	Lightweight

XSL•FO
RenderX

Model	Architectural family	Core architectural principle(s)	Model complexity		Model efficiency	
			Total parameters	Model size	GFLOPs[b]	Efficiency class
EfficientViTB0	Hybrid (CNN + ViT)	Lightweight multi-scale attention (for context extraction) and MBConv (for local information extraction) [43]	0.7M	Lightweight	0.07	Lightweight

[a]DL: deep learning.

[b]GFLOPs: Giga floating point operations.

[c]CNN: convolutional neural network.

[d]BMConv: mobile-inverted bottleneck convolution.

[e]NAS: neural architecture search.

[f]ViT: vision transformer.

[g]HAT: hierarchical attention.

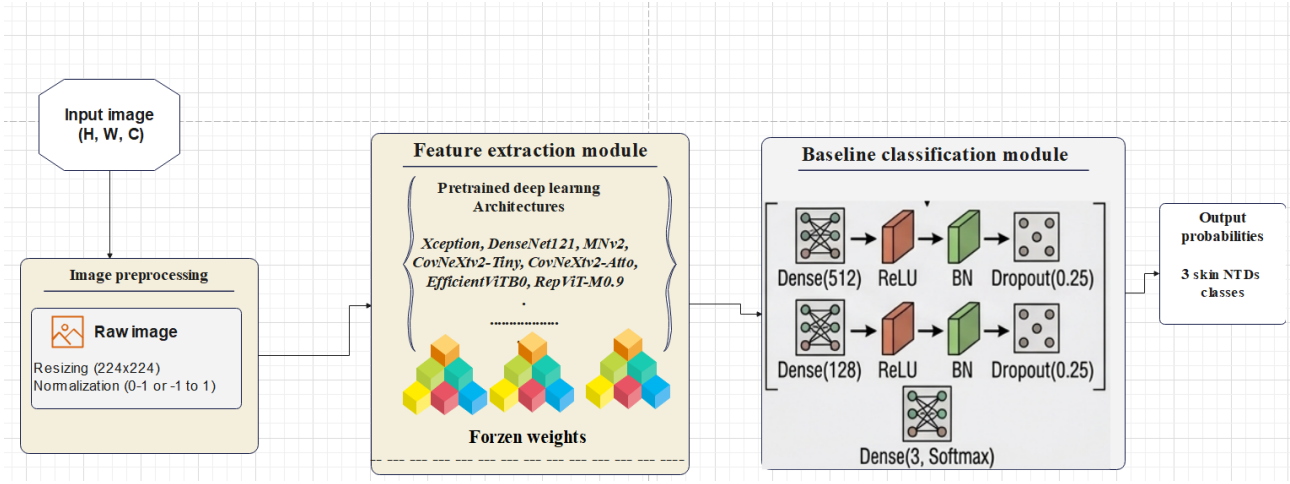
All the 21 selected pretrained models were initially trained on ImageNet-1K dataset, initialized with weights from this standard, large, and diversified dataset having around 3.2 million images [44], helping to create general feature extraction baselines. Furthermore, as shown in Table 1, maintaining architectural distribution, we selected the 21 representative models from nine architectural families, which include (1) models of residual networks (ResNet50 and ResNet18) that apply residual learning (or skip connections) method, (2) modern CNN architectures (ConvNextV1 Small/Tiny and ConvNextV2 Tiny/Atto) that have transformer-like CNN components, (3) model with densely connected CNNs (DenseNet121), (4) model with depth-wise separable CNNs (Xception), (5) the EfficientNet family (B0-B5), (6) mobile CNNs MobileNet (V2 and V3), (7) models that apply advanced compound scaling method (EfficientNetV2 B0 and S), (8) recent state-of-the-art lightweight CNN architecture (RepViT), and (9) hybrid transformer-based architectures (FastViT, FasterViT-0, and EfficientViTB0). Regarding architectural and computational complexities, 6 of these models are heavier models having parameters between 30 and 50 million (M) (30 M < parameters < 50 M), while having computational complexities that range between 4.0 and 9.9 Giga floating point operations (GFLOPs). The other 8 models have moderate levels of complexities (6.2 M < parameters < 28.1 M, and 0.7 billion [B] < FLOP < 4.0 B), while 7 of the 21 models are lightweight models (0.7 M < parameters < 5.1 M, and 0.06 B < FLOP < 0.55 B).

The 2-Stage Approach: Feature Extraction With Integrated CNN Model

On the third experimental setting, our proposed 2-staged approach is demonstrated. In this approach, we crafted a robust hybrid model development pipeline that incorporates 2 different DL model architectures, the selected pretrained and our custom-designed CNN models, integrated to the utility modules (data loading, preprocessing, and evaluation). These 2 groups of models are used independently in 2 phases (stages) subsequently operating one after the other to achieve 2 exclusive DL operations, feature extraction (feature mapping) and disease classification (inference), representing the 2 fundamentally isolated but highly interdependent modules in the pipeline. The selected pretrained models are used only for the purpose of feature extraction. Given these 2 modules are highly fundamental that operate subsequently, the disease classification model operates using the output of the feature extraction model, we named the overall hybrid pipeline as the 2-stage approach. Figure 3 presents the overall architecture of our 2-stage approach, depicting the 2 major stages as modules in the DL architecture.

XSL•FO
RenderX

Figure 3. Architecture of the new 2-stage approach. BN: batch normalization; DL: deep learning; skin NTDs: skin Neglected Tropical Diseases.



In the first stage (feature extractor module), the selected 21 DL architectures are solely used for feature extraction, which we named them as extractor(s) or feature map(s), and we created 21 specific extractors (maps) by freezing all the trainable weights in each of these 21 models. We use this method to effectively extract features using the mapping logic of each pretrained model and derive feature representations within the high-dimensional feature space (R^{2048}) for each input image in our dataset. This method helped us in creating the full feature representations by avoiding pooling layers, as in the case of our custom-designed CNN model, helping us in preserving important spatial representations of features that would potentially be dropped as a result of using pooling method (MaxPooling). After mapping all the input images using the pretrained models, feature matrices are created for each of the separate train and test sets and are prepared to be used by our custom-designed model. This completes the first phase where the feature extraction models completed their only purposes of feature extraction, which are no longer used.

In the second stage (baseline classification module; Figure 3), the feature matrices created using the previously extracted features by each extractor are used to train the new CNN model. However, since our new model was initially designed as a full-fledged CNN architecture, its use in this phase required restructuring. Hence, we restructured our newly designed CNN model by excluding the extraneous feature extraction layers, as these operations are performed by different models in the separate previous module. Guided by our initially designed 30-layer deep and highly regularized CNN model, we specifically redesigned the final 10-layer architecture to enhance the robustness of the classification head, improve classification performance, and add architectural novelty. This approach totally eliminates the 3 feature extraction blocks of the new CNN model, leading to the exclusion of 20 layers. Specifically, this approach directly eliminates the 6 convolutional layers and their corresponding 10 layers (7 activation and 3 maximum pooling), while minimizing the normalization layers to only 3. This creates a final classification model having 10 layers with 3 dense blocks and a final output layer, where each dense block in turn includes regularization methods (normalization and dropout).

Overall, unlike most standard transfer learning pipelines where the pretrained models serve complex tasks with only 1 or 2 classification layers finally added [45], the classification head in our 10-layer customized (redesigned) CNN model has deeper architectural layers. Through its 9 regularization layers, the new CNN head also serves a separate feature filtering purpose, where the final classification layer makes the classification decision using highly relevant skin lesion features.

Model Training and Evaluation

We used a 5-fold cross-validation method to train and evaluate the models. The metrics used to evaluate the models are selected to assist performance analyses from different perspectives. The macro F_1-score and the class-specific metrics (using sensitivity and specificity) are highly used, as accuracy was found to be a misleading metric due to the highly skewed nature of our new dataset used. Hence, we prioritize the macro F_1-score of models on both the train and test sets to effectively evaluate the models' generalizability and learning ability. The overall sensitivity (macro recall), the area under the precision-recall curve (AUPRC), and the area under the receiver operating characteristic (AUROC) scores are also highly used to measure how well each model performs in identifying positive disease classes. Furthermore, visual tools using tables, confusion matrices, ROC curves, including performance plots such as the slope charts and radar plots, are also used to analyze models' performance.

Final Model Selection

Ultimately, this study selects the final best-performing skin NTDs classification models with the highest classification performance. To achieve this, we applied systematic model selection procedures in 2 phases: first-level selection—applied for the initial feature extraction model screening based on baseline performance score (end of phase 1), and final model selection—applied at the end of model screening experiments (end of phase 2). Overall, the selection of top-performing models involved analysis of performance scores that include macro F_1-score, sensitivity, AUPRC, and 4 class-specific performance metrics (podo-recall, tungiasis-recall, scabies-recall, and podo-F_1-score). Hence, at the end of phase 1, using these extended screening parameters, top-performing models that

XSL•FO
RenderX

achieve stable and outperforming scores across the 2 experiments are selected. Accordingly, selected models are used in the next experimental training that applies performance optimization using the dynamic (online) data augmentation method.

Finally, at the end of phase 2, the final 2 best-performing models are selected based on the results achieved during the experiment with the optimization method, and we applied a robust selection procedure based on weighted comparison of performance scores achieved during this last experiment. Accordingly, 6 evaluation metrics (F_1-score [macro], podo-F_1, macro recall [mean], podo-recall, inference speed [samples per second (sps)], and number of parameters) are used, and we applied the weighted scores comparison method with each metric given different weight. To achieve this strategic comparison, we conducted 4 procedures. First, the metrics are categorized as performance (F_1-score [macro], podo-F_1, macro recall [mean], and podo-recall) and efficiency metrics (inference speed and number of parameters). Out of all these metrics, the model complexity (number of parameters) is a metric that is mostly desired to be lower ("lower is good"), while the other 5 metrics represent best model performance when their values are higher ("higher is better"), having opposite directional symmetry with model size. Second, we normalized all values of the 6 comparison metrics using the minimum-maximum normalization method to ensure that all values fall between 0 and 1 and facilitate the combined scores comparison, which includes normalizing the values that already have values between 0 and 1, mainly for (1) avoiding range dilution—minor differences in performance mostly lead to larger ranges that determine comparison; (2) baseline value definition, combined scores mostly perform well with the least score defined as "0.0"; and (3) simplifying directional symmetry. Therefore, to address all these, we applied the minimum-maximum normalization method by defining 2 normalization formulas. First, we normalized the "higher is better" metrics (the metrics that represent best values when their values are higher) using the formula:

$$\text{xnormalized} = \text{x} - \text{xminxmax} - \text{xmin}$$

where, $x_{\text{normalized}}$ is a value in any 1 of the 5 metrics (eg, macro F_1-score) representing a single value for a specific model that is being normalized; x_{min} is the least value in that specific metric; and x_{max} is the maximum score in the same group of metric.

Next, for the "lower is good" metric (model complexity), we applied the normalization using the formula:

$$\text{xnormalize} = \text{xmax} - \text{xminxmax} - \text{xmin}$$

Third, we assigned weights *w* for all evaluation metrics, and as a diagnostic model, disease prediction performance scores are the primary requirements. Hence, we assigned higher weights (0.2) for performance metrics and a relatively lower weight (0.1) for the 2 efficiency metrics (inference speed and number of parameters). Finally, we define a unified objective function that computes the weighted sum of score (WSS) for each model using the formula:

$$\text{WSSm} = \sum \text{i}=16\text{wi} * \text{xi}$$

where WSS_m is the weighted sum of score of any given model evaluated, x_i is a specific evaluation metric, and w_i is the weight of a given metric. After computing the weighted scores, models are ranked accordingly to identify the top 2 models on the overall performance. Accordingly, the first selected model would be a model with higher complexity, with a consideration of using the model as a back-end (web-based) classification model, while the second selected model would be a lightweight model with the potential of being embedded in edge (mobile) devices for the actual and real-time diagnosis of skin NTDs.

Experimental Setup

Overview

As the study is guided by the funnel framework based on extended experiments, the whole training experiments are conducted in a phased approach having 3 phases. The first phase deals with the establishment of baseline skin NTDs classification models—using the new CNN model, transfer learning, and the 2-stage approach; the second phase applies performance optimization (data augmentation); and the third phase applies the final performance optimization through hyperparameter tuning.

Phase 1: Baseline Model Training With Cascaded Model Selection

In this first phase, only baseline skin NTDs classification models are trained in 3 separate experimental setups: the custom-designed CNN model, transfer learning, and the 2-staged approach. This phase is intended for the overall evaluation of baseline models' performance, which includes comparative analysis and first-level model screening.

Custom-Designed CNN Model

The first experimental setting involves training the custom-designed 30-layered CNN model, where the new model is trained and evaluated under 2 different methods. The first training involves the baseline training, where the model is barely trained without applying any advanced machine learning methods for tweaking performance. On the next experiment, the same model is trained by applying data augmentation techniques. These experiments are conducted to evaluate and analyze the performance of this newly crafted model on our new small-sized skin NTDs dataset, as the results from these experiments determine subsequent strategies.

Transfer Learning: Baseline Models Using Pretrained Models

On the second experimental setting, we demonstrated the transfer learning method using the selected pretrained DL architectures to build baseline models. Hence, 21 baseline DL models are fully trained on our new skin NTDs image dataset, with proper evaluation of each model on the classification of the skin NTDs, where the results are used for the overall analysis of models' performance.

Two-Stage Approach: Feature Extraction With Integrated CNN Model

On the third experimental setting, the 2-staged approach is demonstrated. To apply this robust training pipeline, we used

XSL•FO
RenderX

the 21 systematically selected DL architectures for feature extraction and the restructured new 10-layered CNN model architecture. Hence, 21 different skin NTDs classification models are trained using this approach, where each model is evaluated using the predefined appropriate evaluation methods.

Phase 2: Performance Optimization

Data Augmentation

Evidently, DL models are highly data-intensive and require a large amount of data to achieve excellent performance [46]. Hence, using a relatively small-sized dataset, as the case of this study with only 1495 samples, developing a DL-based diagnostic model with higher classification accuracy is a real challenge; it might not be beneficial at all, compared with the use of large-sized datasets [47]. Methodologically, several DL-based data augmentation methods are available, including the advanced standard and conditional augmentation using GAN models, which are also highly suitable for class imbalance handling. However, their execution requires further experiments, analysis, and selection, including higher computational requirements, which in turn extends this study by deviating from the intended benchmarking objective. Therefore, as a benchmarking study, the dynamic or online (on-the-fly) data augmentation approach [29] is used by applying transformations on the training set to increase the number of training samples by a factor of 5 (including the original training images) and create a total of 4785 training images. This method is primarily used to alleviate the data scarcity issue, analyze performance changes, and increase both the size and the diversity of the training set without the permanent creation of the images. Accordingly, we applied selected geometric transformations such as rotation (0.2), scaling (0.2), and horizontal flipping during model training on the input images at the time of model training to mathematically simulate the real-world diagnostics of skin NTDs. Therefore, using this method, 5 different models are trained based on the 2-stage approach using the final 5 selected extractors.

Phase 3: The Final Optimization Methods and Hyperparameter Tuning

As the last experimental training, the hyperparameter tuning operation is applied to the final 2 selected models. To achieve this, the "Hyperband" algorithm was used, which is a faster and resource-efficient hyperparameter optimization algorithm than other hyperparameter searching algorithms. Hyperband is an efficient bandit-based Keras algorithm for hyperparameter optimization that uses early stopping with a successive halving algorithm to quickly find good configurations for models [48,49]. Therefore, the hyperparameter tuning operation is applied by running the Hyperband algorithm to the maximum of 30 epochs, with the maximum number of trials being 60. The optimization was carried out based on the validation accuracy, which was set to be the objective metric. Finally, the final best hyperparameters were saved to use for the final optimized training of the skin NTDs classification model.

Results

This section presents model evaluation results of the 5 experimental settings (baseline training, transfer learning, and the 2-stage approach, including the training with data augmentation and hyperparameter tuning).

Phase 1: Baseline Model Training With Cascaded Screening

The New CNN Model

The new CNN baseline model is trained on the new skin NTDs dataset in 2 different experimental settings, the baseline training that applies no advanced DL method and applying the data augmentation method, as summarized by the overall results in Table 3. Accordingly, on the baseline training, the model achieved an accuracy of 0.674 and F_1-score of 0.42 (with AUROC=0.676 and AUPRC=0.444), with a very high loss (0.978). However, during the second training with data augmentation, the new CNN model achieved improved performance with an F_1-score of 0.446 and an increased loss of 1.458.

However, class-wise, the new model showed the worst sensitivity for podoconiosis (scoring all 0.0 in precision, recall, and F_1-scores), with a macro recall of 0.43 (having a recall of 0.457 and 0.827 for tungiasis and scabies, respectively), as shown in Table 4.

Table . Overall performance scores of the new CNN[a] model across the 2 experimental settings.

Experiment	Model/method	Accuracy	Loss	F_1-score (macro)	AUROC[b] (mean)	AUPRC[c] (mean)
First	Baseline	0.674	0.978	0.422	0.676	0.444
Second	Dynamic data augmentation	0.691	1.458	0.446	0.790	0.608

[a]CNN: convolutional neural network.

[b]AUROC: area under the receiver operating characteristic curve.

[c]AUPRC: area under the precision-recall curve.

XSL•FO
RenderX

Table . Class-specific performance of the new CNN[a] model across the 2 experiments.

Experiment	Model/Method	Recall				F_1-score			
		Podoconiosis	Tungiasis	Scabies	Macro recall	Podoconiosis	Tungiasis	Scabies	Macro F_1-score
First	Baseline	0.0	0.457	0.827	0.428	0.0	0.489	0.776	0.422
Second	Dynamic data augmentation	0.154	0.181	0.979	0.438	0.235	0.296	0.806	0.446

[a]CNN: convolutional neural network.

With data augmentation, however, the model achieved a macro recall of 0.438 and macro F_1-score of 0.446, with improved class-specific scores in recall (podoconiosis=0.154, tungiasis=0.181, and scabies=0.979) and F_1-score (podoconiosis=0.235, tungiasis=0.296, and scabies=0.806). As confirmed by the results, the use of the standard data augmentation method significantly improved the sensitivity and macro F_1-score of the model for the podoconiosis class (SD +0.109) and (SD +0.017).

Transfer Learning: Baseline Performance

On the second and third training settings of the first phase, baseline models are trained using similar pretrained DL model architectures demonstrating (1) the transfer learning method, and (2) the 2-stage approach. Table 5 presents the overall performance of models scored during these 2 experimental settings.

XSL•FO
RenderX

Table . Performance of the models across the 2 experiments of phase 1 model screening experiments[a].

Model	Experiment 1: transfer learning (baseline models)					Experiment 2: 2-stage approach (baseline models)				
	Accuracy	Log-loss	F_1-score (macro)	AUPRC[b] (mean)	Sensitivity (mean)	Accuracy	Log-loss	F_1-score (macro)	AUPRC (mean)	Sensitivity (mean)
ResNet50	0.695	0.733	0.37	0.516	0.39	0.641	1.027	0.26	0.430	0.33
ConvNext-Small	0.641	0.751	0.26	0.552	0.33	0.805	0.532	0.67	0.747	0.62
Xception	*0.973*	*0.106*	*0.94*	*0.975*	*0.91*	*0.94*	*0.219*	*0.9*	*0.937*	*0.88*
EfficientNetB5	0.641	0.785	0.26	0.378	0.33	0.638	0.86	0.34	0.368	0.36
ConvNext-Tiny	0.671	0.71	0.33	0.573	0.37	0.718	0.693	0.43	0.568	0.43
DenseNet121	*0.95*	*0.119*	*0.89*	*0.973*	*0.84*	*0.96*	*0.133*	*0.91*	*0.974*	*0.9*
EfficientNetB3	0.641	0.782	0.26	0.370	0.33	0.668	0.767	0.37	0.429	0.39
MNv3-Large	0.641	0.77	0.26	0.494	0.33	0.681	0.995	0.59	0.625	0.57
MNv3-Small	0.641	0.777	0.26	0.388	0.33	0.708	0.687	0.4	0.640	0.41
EfficientNetB0	0.641	0.786	0.26	0.360	0.33	0.641	1.012	0.26	0.396	0.33
EfficientNetB1	0.641	0.785	0.26	0.363	0.33	0.628	0.828	0.35	0.437	0.37
MNv2	*0.956*	*0.109*	*0.91*	*0.973*	*0.87*	*0.943*	*0.181*	*0.88*	*0.915*	*0.86*
ResNet18	0.641	0.805	0.26	0.352	0.33	0.735	0.737	0.48	0.545	0.49
EfficientNetV2B0	0.641	0.786	0.26	0.349	0.33	0.641	0.793	0.26	0.393	0.33
EfficientNetV2S	0.674	0.694	0.39	0.501	0.4	0.654	0.867	0.44	0.487	0.45
CovNeXtv2-Tiny	*0.95*	*0.124*	*0.919*	*0.983*	*0.886*	*0.94*	*0.222*	*0.876*	*0.935*	*0.862*
CovNeXtv2-Atto	*0.94*	*0.226*	*0.878*	*0.921*	*0.837*	*0.936*	*0.25*	*0.882*	*0.909*	*0.857*
EfficientViTB0	*0.963*	*0.145*	*0.93*	*0.967*	*0.898*	*0.963*	*0.124*	*0.903*	*0.967*	*0.876*
FasterViT0-T8	0.936	0.165	0.883	0.969	0.855	0.671	0.814	0.382	0.469	0.392
FastViT-T8	0.896	0.37	0.602	0.903	0.61	0.93	0.223	0.862	0.939	0.833
RepViT-M0.9	*0.943*	*0.188*	*0.915*	*0.959*	*0.888*	*0.956*	*0.166*	*0.925*	*0.959*	*0.893*

[a]Values presented in italics represent high-performance

[b]AUPRC: area under the precision-recall curve.

On the training with only transfer learning (experiment 1, Table 5), only 8 models scored top results, where Xception outperformed all models with the top accuracy (97.3%), macro F_1-score (0.94), and sensitivity (0.91), followed by EfficientViTB0 (accuracy=0.963, F_1-score=0.930, and sensitivity=0.898) and MNv2 (accuracy=95.6%, F_1-score=0.91, and sensitivity=0.87). DenseNet121 and ConvNeXtV2-Tiny are the next high-performing models scoring the same accuracy (95%) and sensitivity (0.84), with ConvNeXtV2-Tiny scoring better F_1-score (0.89). The RepViT, FasterViT, and ConvNeXtV2-Atto models are the other top-performing models with macro F_1-scores of 0.92, 0.88, and 0.84. FastViT, ResNet50, EfficientNetV2S, and ConvNext-Tiny scored macro F_1-scores of >0.33 and average sensitivity >0.37. However, the

XSL•FO
RenderX

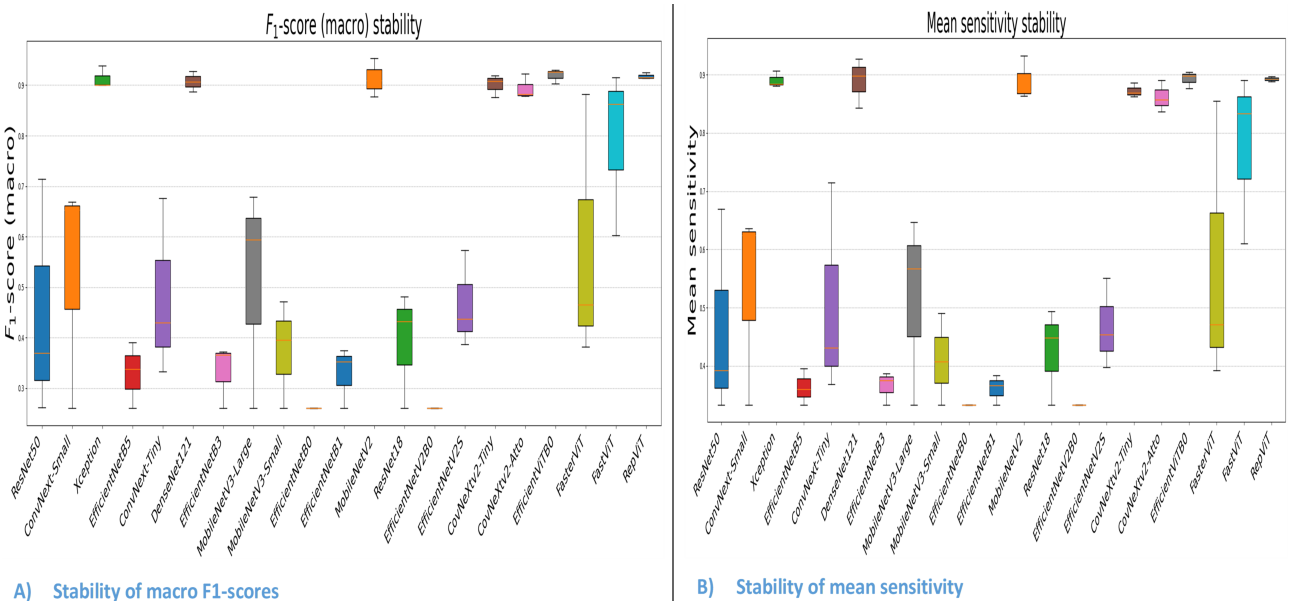
remaining 9 models similarly scored worst class-specific sensitivity (recall=0.0 for podoconiosis and tungiasis classes), leading to the least macro F_1-score (0.26) and average sensitivity (0.33). Using the 2-stage approach as shown in experiment 2 (Table 5), only 4 models trained using RepViT, DenseNet, EfficientViTB0, and Xception extractors helped their corresponding trained classification models to achieve macro F_1-scores of 0.90 and above, while 8 models scored macro F_1-scores of >086. As shown, the model trained on RepViT-extracted features scored the highest macro F_1-score (0.93, mean sensitivity=0.89), where the model using DenseNet-extracted features had the second highest F_1-score (0.91) with the maximum mean sensitivity (0.90). The model using EfficientViTB0-extracted features scored the next top F_1-score (0.90) during this experiment, followed by the models using the Xception feature extractor (F_1-score=0.90 and sensitivity=0.88), CovNeXtv2-Atto feature extractor (F_1-score=0.88 and sensitivity=0.86), MNv2 feature extractor (F_1-score=88 and sensitivity=0.86), CovNeXtv2-Tiny feature extractor (F_1-score=0.88 and sensitivity=0.86), and FastViT feature extractor (F_1-score=0.86 and sensitivity=0.83).

Method-wise, the use of the 2-stage approach improved the overall performance of the majority of the models, where 62% of the models (13 models) exhibited improved macro F_1-scores, with 2 models showing no variations, while the remaining models exhibited minor declines. Class-wise, the 2-stage approach improved the podo sensitivity (podo-recall) of 5 models (using DenseNet121, MNv3-Large, ConvNeXt-Small, ConvNeXtV2-Atto, and FastViT) by the SD of +0.109, +0.218, +0.218, +0.054, and +0.435, respectively, with 3 models (using ConvNeXtV2-Tiny, EfficientViTB0, and FasterVit) showing declining podo-recall, while the other 13 models showing no changes.

Best Models Selection

Overall, the hybrid 2-stage approach applied using the 21 pretrained models (both CNN and transformer-based) yielded superior performance compared with other methods (baseline and transfer learning). Hence, as a DL architectural benchmarking study, we selected this hybrid architecture for further experiments applying optimization methods. However, to screen out short-listed top feature mapping models, we conducted deep and multidimensional analysis that includes performance stability analyses, as shown in Figure 4A and B.

Figure 4. Analysis of models' performance stability across experiments. (A) Stability of the macro F_1-scores of each of the 21 models across the 2 experiments. (B) Stability of the mean sensitivity scores of the 21 models across the 2 experiments.



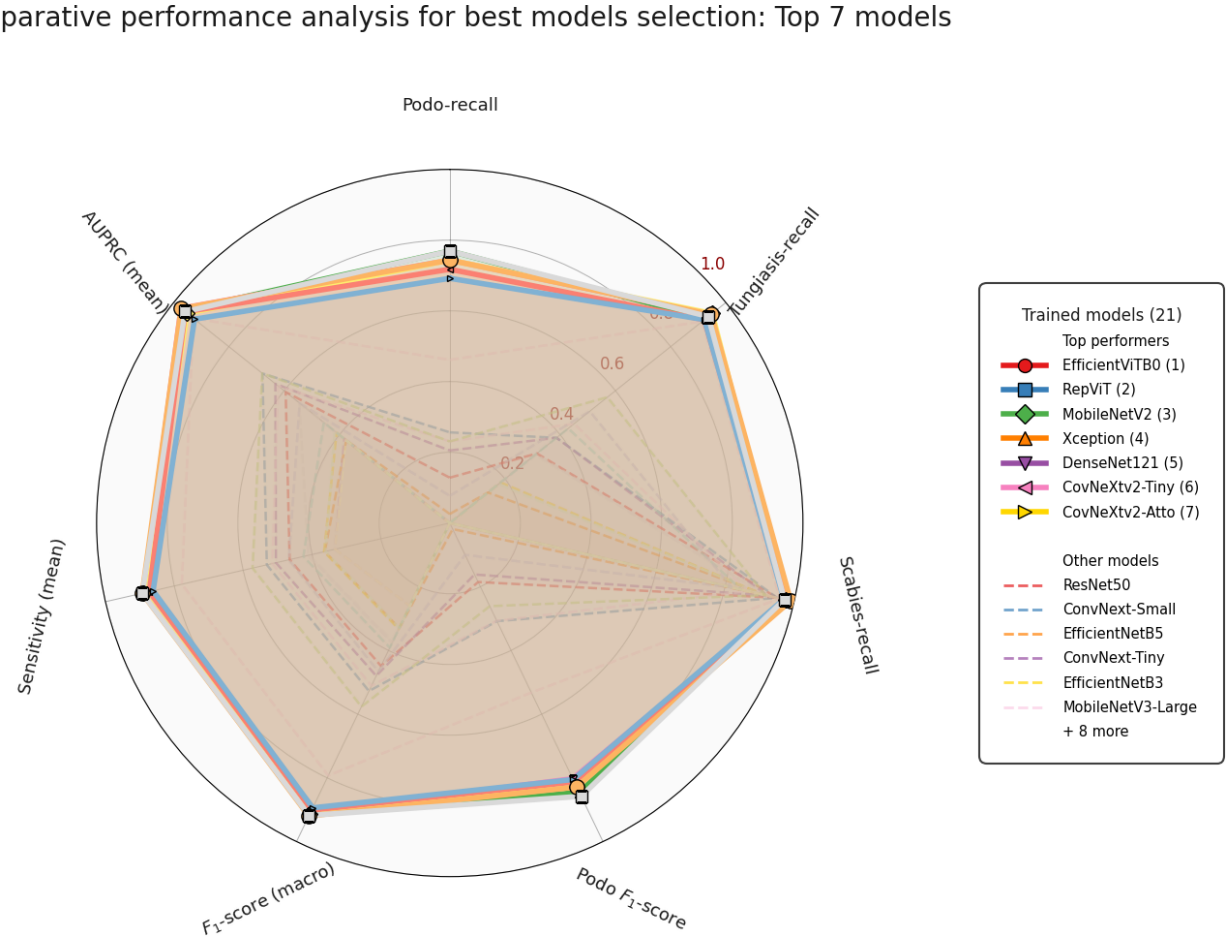
As shown, the box plots clearly depict the stability of the models' performance primarily in macro F_1-score and overall sensitivity (macro recall) across the 2 experiments. Overall, the models fall into three major categories: (1) stable and least-performing models—models trained using features extracted by EfficientNetB0 and EfficientNetV2B0, where the use of these models for feature extraction resulted in the worst macro F_1-score and sensitivity, while the other 6 feature extractors, EfficientNetB5, EfficientNetB3, MNv3-Small, EfficientNetB1, ResNet18, and EfficientNetV2S, resulted in comparatively better performance scores; (2) the unstable models—models trained on ResNet50, ConvNeXt-Small, ConvNeXt-Tiny, MNv3-Large, FasterViT0, and FastViT-T8, for feature extraction, where the models using these extractors scored highly unpredictable performance scores (macro F_1 and sensitivity), as shown by their wide-range scores (tall boxes); and (3) the stable and top-performing models—representing 7 models trained using the features extracted by Xception, DenseNet, MNv2, ConvNeXtV2-Tiny, ConvNeXtV2-Atto, EfficientViTB0, and RepViT. The models using these 7 pretrained models exhibited minimum variability in their macro F_1-score and sensitivity, while achieving the highest median macro F_1-score (0.88) and median mean sensitivity (0.86) across the 2 experiments, as shown in Figure 4A B.

The next analysis includes performance comparison using 7 parameters, as shown by the radar plot in Figure 5. As shown, only 7 models using EfficientViTB0, RepViT, MNv2, Xception,

XSL•FO
RenderX

DenseNet, ConvNeXt-Tiny, and ConvNeXt-Atto feature maps showed exceptionally outperforming performance collectively having a median F_1-score (macro) exceeding 0.95, AUPRC (macro) approaching 1.0, with a comparatively lower median podo-F_1-score approaching 0.90. However, the other 14 models (other than the 7 top-performing models) scored average macro F_1-scores <0.89, average AUPRC<0.92, while showing the worst sensitivity for the minority class with median podo-F_1-score below 0.35.

Figure 5. Overall model performance comparison for best model selection. The radar plot illustrates analysis of the models' performance based on the average scores of the 7 evaluation metrics across the 2 experimental settings, depicting the exceptionally outstanding performance of the 7 top-performing models out of the 21. AUPRC: area under the precision-recall curve.



All these results underlined that the 7 models identified as top-performing models (Figure 5) provided higher feature representation capabilities that led to better overall model performance. Hence, the EfficientViTB0, RepViT, MNv2, Xception, DenseNet, ConvNeXt-Tiny, and ConvNeXt-Atto models are selected for further experiments being used for feature mapping. However, the remaining 14 models were highly challenged to extract useful features that led to poor overall performance and are screened out from being used in further experiments.

Phase 2: Performance Optimization

On the second phase of our experimental framework, the selected 7 models are retrained by applying data augmentation method, and the results in Table 6 are achieved. In this experiment, the model trained using the MNv2 feature mapping significantly outperformed all other models with macro F_1-score of 0.95 and sensitivity of 0.93, while scoring the fourth highest AUPRC (0.97) and the fourth least errors (loss=0.137). As shown, the models using DenseNet121 and EfficientViTB0 extractors scored the next closer performance, where the model trained using DenseNet121 mapping achieved the second highest macro F_1-score and mean sensitivity of 0.928, while the model using EfficientViTB0 for feature mapping scored a macro F_1-score of 0.926 and mean sensitivity of 0.905.

XSL•FO
RenderX

Table . Performance of the 7 screened models in the second phase experiment with data augmentation[a].

Model	Accuracy	Log-loss	F_1-score (macro)	AUPRC[b] (mean)	Sensitivity (mean)
Xception	0.94	0.158	0.901	0.982	0.884
DenseNet121	0.966	0.13	0.928	0.974	0.928
MNv2	*0.973*	*0.137*	*0.953*	*0.967*	*0.933*
CovNeXtv2-Tiny	0.956	0.128	0.908	0.956	0.869
CovNeXtv2-Atto	0.953	0.209	0.922	0.944	0.891
EfficientViTB0	0.97	0.098	0.926	0.985	0.905
RepViT-M0.9	0.956	0.168	0.914	0.959	0.897

[a]Values presented in italics represent high-performance models.

[b]AUPRC: area under the precision-recall curve.

Overall, during this experiment, the 7 feature extraction pretrained models (MNv2, DenseNet121, EfficientViTB0, CovNeXtv2-Atto, RepViT, CovNeXtv2-Tiny, and Xception) helped their corresponding skin NTDs classification models to score macro F_1-scores >0.90, AUPRC >0.94, and mean sensitivity >0.86, with overall loss ≤0.21.

Final Models Selection

For the systematic selection of the final 2 best feature extraction pretrained models that provide the best feature mapping for the skin NTDs classification models, we applied the robust weighted score comparison method defined in the "Methods" section of the study. Hence, applying our predefined WSS formula, we computed the WSS for the scores achieved on the third experiment that applies data augmentation and for the overall average performance of the models across the 3 experiments. For each group of evaluation, we applied two different sets of weights for the metrics: (1) weights given only to the predictive performance metrics—equal weights of 0.25 are given to only the 4 predictive metrics (macro F_1, recall, podo-recall, and podo-F_1), while the speed and number of parameter are not considered to evaluate only predictive performance; and (2) weights given to all the 6 metrics—the predefined weights of 0.2 given for the 4 predictive performance metrics, while the efficiency metrics are given the weights of 0.1. Table 7 summarizes the overall results of the combined weighted evaluation method.

Table . Combined weighted method results for the 7 prescreened models (final model selection)[a].

Model	Performance metrics				Efficiency metrics		Weighted scores (optimization)		Weighted scores (average)	
	F_1-score (macro)	Podo-F1	Macro recall (mean)	Podo-recall	Speed (sps)	Parameters	WSS[b] (predictive)	WSS (overall)	WSS (predictive)	WSS (overall)
RepViT-M0.9	0.05	0.03	0.09	0.10	0.09	0.10	0.335	0.459	*0.974*	*0.975*
DenseNet121	*0.10*	*0.06*	*0.18*	*0.20*	*0.07*	*0.06*	*0.681*	*0.671*	0.511	0.528
EfficientViTB0	0.10	0.03	0.11	0.10	0.00	0.06	0.424	0.399	*0.784*	*0.772*
Xception	0.00	0.03	0.05	0.10	0.03	0.00	0.222	0.211	0.855	0.684
CovNeXtv2-Atto	0.08	0.10	0.07	0.10	0.10	0.10	0.442	0.554	0.009	0.204
MNv2	*0.20*	*0.20*	*0.20*	*0.20*	*0.07*	*0.04*	*1.0*	*0.916*	0.688	0.685
CovNeXtv2-Tiny	0.03	0.00	0.00	0.00	0.05	0.07	0.034	0.155	0.241	0.367

[a]Values presented in italics represent high-performance models.

[b]WSS: weighted sum of scores.

As shown, the model using the MNv2 feature mapping showed an exceptionally highest combined weighted score with an overall WSS of 0.92 on the aggregate weighted scores. This model still showed an exceptionally outperforming combined weighted score on the predictive performance metrics with a WSS of 1.0. The other model using the DenseNet121 mapping has the second highest weighted scores having an overall WSS of 0.67 and a WSS of 0.68 on the predictive performance. On

XSL•FO
RenderX

the same experiment, the models using CovNeXtv2-Atto (0.55) showed the next top overall WSS, followed by the model using RepViT-M0.9 extractors (0.46). On this last experiment, the hybrid transformer-based architecture EffcientViTB0 showed lower weighted scores (0.399 and 0.424) on the overall and predictive WSS. Conversely, using the aggregated performance derived by computing the average scores of each metric for all models across 3 experiments, the combined weighted scores present a completely different set of scores. Based on these results, the model using the RepViT-M0.9 extractors showed an exceptionally higher combined weighted scores both on the overall WSS (0.98) and the predictive WSS (0.97), followed by the model using EfficientViTB0 with an overall WSS (0.98) and predictive WSS (0.97). The models based on MNv2 and DenseNet121 showed the next higher weighted scores (overall WSS of 0.69 and 0.53, respectively) on the aggregated scores of the models across the 3 experiments.

Overall, the data augmentation method is applied to optimize the models' performance, given the study faced the data scarcity problem. However, the models using the RepViT-M0.9 and EfficientViTB0 extractors unexpectedly reacted negatively to this optimization method that expands the sample size and data variance, while the same method significantly boosted the performance of the models based on the MNv2 and DenseNet121 feature mapping. These results reveal a generalization paradox problem, showing (1) overspecialization—the models based on the RepViT-M0.9 and EfficientViTB0 extractors are experiencing performance instability and a potential overfitting due to memorization of the initial small-sized training data, and (2) latent generalization—the models using the MNv2 and DenseNet121 extractors showed their hidden generalization abilities that are exposed as the result of the data augmentation method. Therefore, based on these major driving facts, we selected the MNv2 and DenseNet121 pretrained models for feature mapping in our final skin NTDs diagnostic model.

Phase 3: The Final Optimization

On the final experiment, hyperparameter tuning is applied on the last 2 models selected, which helped the model using the DenseNet121 feature extractor achieve improved performance, as shown in Table 8.

Table . Performance scores of the final 2 selected models after hyperparameter tuning[a].

Model	Accuracy	Macro F_1-score	Loss	AUROC[b] (average)	AUPRC[c] (average)	Recall				F_1-score			
						Podoconiosis	Tungiasis	Scabies	Macro recall	Podoconiosis	Tungiasis	Scabies	Macro F_1-score
DenseNet121	*0.966*	*0.946*	*0.181*	*0.996*	*0.974*	*0.923*	*0.979*	*0.974*	*0.959*	*0.889*	*0.968*	*0.982*	*0.946*
MNv2	0.96	0.906	0.150	0.987	0.955	0.769	0.947	0.979	0.898	0.80	0.937	0.982	0.906

[a]Values presented in italics represent high-performance models.

[b]AUROC: area under the receiver operating characteristic curve.

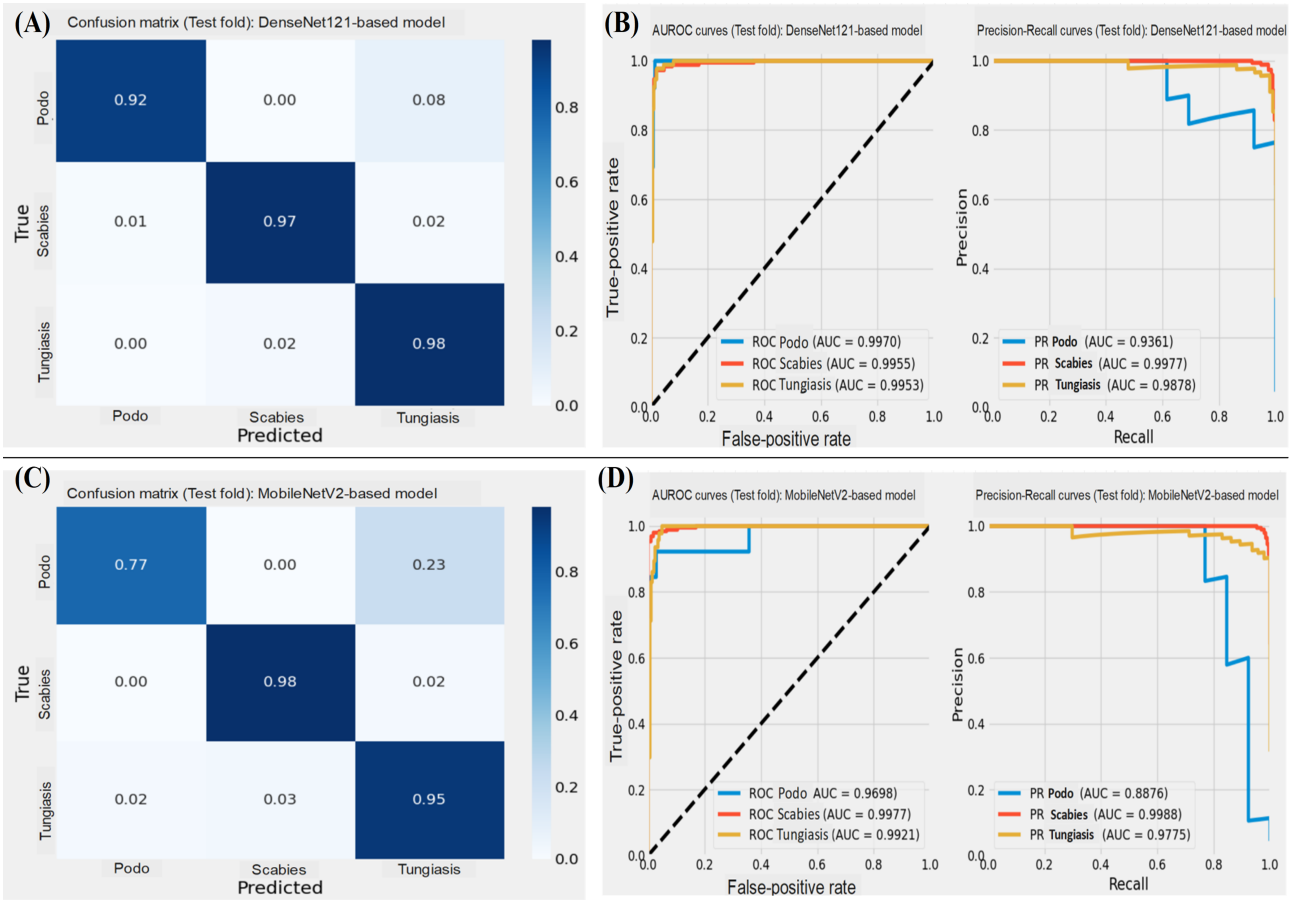
[c]AUPRC: area under the precision-recall curve.

Figures 6A-6D summarizes the overall performance of the final 2 models trained using the DenseNet121 and MNv2 feature extractors after hyperparameter tuning.

After hyperparameter tuning, the DenseNet121-based model scored improved performance, compared with the previous experiment with data augmentation, achieving an accuracy of 0.966, F_1-score of 0.95 (Δ=+0.018), and mean sensitivity of 0.96 (Δ=+0.031). However, the MNv2-based model exhibited declining performance in F_1-score (Δ=−0.047) and macro recall (Δ=−0.035), including podo-specific recall and F_1-score.

XSL•FO
RenderX

Figure 6. Comprehensive classification evaluation of the final 2 models. (A) Confusion matrix showing class-wise results for the model trained using DenseNet121 extractor, (B) AUROC and area under the precision-recall curve (AUPRC) of the model using DenseNet121 feature extractor, (C) confusion matrix showing class-wise results for the model trained using MNv2 extractor, and (D) AUROC and AUPRC of the model using MNv2 feature map. AUC: area under the curve; AUROC: area under the receiver operating characteristic curve; PR: precision-recall; ROC: receiver operating characteristic.



Discussion

Principal Findings

As a DL architectural benchmarking effort, this study developed a diagnostic model for skin NTDs based on the funnel framework with an extensive experiment-based cascaded architectural screening. A significant portion of the experimental design involved identifying robust pretrained DL architectures for feature mapping to be used with our custom-designed CNN model. On average, across the first 2 phases (5 experiments), the custom-designed CNN model and 10 other models trained using the feature mapping architectures yielded poor overall performance with macro F_1-scores below 0.50, 5 models (using MNv3-Large, ConvNext-Small, FasterViT, FastViT, and CovNeXtv2-Atto extractors) showed moderate performance with above 0.50, and 6 models using EfficientViTB0, RepViT, MNv2, Xception, DenseNet121, and ConvNeXtV2-Tiny extractors achieved exceptionally high performance with F_1-scores exceeding 0.90 and sensitivity above 0.87. This exceptionally high performance of the top 6 models is attributed to their robust feature mapping logic that allowed deriving high-dimensional representative features, given our small-sized skin NTDs dataset. Method-wise, the 2-stage approach applying data augmentation resulted in performance improvements for the majority of the models, which underscore the success of the 2-stage approach for skin NTDs classification, given the dataset problems.

The other critical finding of this study highlighted that modern and state-of-the-art DL architectures such as EfficinetNetB5 and EfficinetNetB3, even the highly anticipated lightweight models of EfficinetNetB0 variants, MNv3-Large and MNv3-Small, were unable to extract useful features from the dataset used. All these models proved to be highly data-intensive and extremely sensitive to class imbalance. Conversely, DenseNet121, MNv2, Xception, including the comparatively recent DL architectures of ConvNeXtV2 (Tiny/Atto), and RepViT provided highly robust feature mapping capability, while the transformer-based hybrid architecture EfficientViTB0 showed its potential. Ultimately, the DenseNet121 and MNv2 demonstrated their robustness across the 5 different experimental trainings, including the final optimization. Hence, these 2 models are selected due to their overall feature mapping capability to be integrated with our custom-designed CNN model. Overall, the success of the DenseNet121 model is architecturally linked to its dense connectivity patterns and inductive biases. These factors allowed the trained classifier to reuse feature maps from previous layers through dense blocks, which expands the feature maps into more representative data. Likewise, the success of the MNv2 models is attributed to its internal information bottleneck method. To combat the dataset-related problems such as high variance and possible data memorization, several

XSL•FO
RenderX

regularization strategies were synergistically applied. The strategies included data augmentation, batch normalization, dropout, and early stopping with a synchronized implementation of callbacks. While these methods successfully improved model generalization on unseen test sets, regularization alone was insufficient to fully address the overfitting issues caused by the severely skewed distribution of disease classes. Specifically, for the 12 models excluded after the first phase experiments, all these regularization methods were unable to improve performance even with the application of data augmentation. This clearly proves that the severe class imbalance, especially between scabies and podoconiosis, greatly affected the models' performance. This clearly requires experimental investigation by applying the different methods for class balancing, both at the data and algorithm levels.

The Synergistic Hybrid Approach

Overall, the most effective strategy identified was the 2-stage hybrid approach that combined the high-performing feature extractors with the properly designed CNN classification module. This final harmonized integration involved 3 key strategies: the robust feature extraction architectures, the optimal CNN classification head configuration, and the dynamic data augmentation methods. The harmonized use of all these 3 strategies, along with hyperparameter tuning, provided high-performance classification models by ensuring that each model operated on the optimized output of its preceding module. Ultimately, this study selected the final 2 skin NTD diagnosis models that achieved an F_1-score > 0.95 using MNv2, and an F_1-score > 0.93 using DenseNet121 after data augmentation. All these higher-performance scores are attributed to the combined use of all DL methods, and feature mapping models we use based on our 30-layered very deep CNN model have brought tangible disease-predictive performance improvements across all models.

The integrated 2-stage approach demonstrated its potential for skin NTDs diagnosis, given the study used a dataset with severe class imbalance and small sample images. While most of the DL methods and pretrained models used in the 2-stage approach are preexisting, the harmonization of the feature mapping with our 10-layered classification head created a different DL architecture. This architecture allowed us to establish information bottlenecks using the compact, highly focused, and properly regularized 10-layer classification head that trains on the extracted features. This new architecture added robustness to our model development pipeline, allowing us to build top-performance skin NTDs diagnostic models using the highly constrained dataset, compared with the standard transfer learning method using the same pretrained models. Accordingly, the improved sensitivity of models for the minority class (podo-recall) confirmed the achievements of the architecture, where 5 models showed their sensitivity for podoconiosis, while 62% of the models (13 models) have shown lower overall performance and no variation in podo-recall. These facts clearly underline that the 2-stage approach provided methodological solutions to the severe class imbalance problem. Furthermore, the approach also demonstrated domain adaptation through its two separate modules: (1) feature extraction module—the feature mapping logics of the models initially trained on ImageNet-1K adapted for our skin NTDs diagnostic model, where these diseases have diverse lesion types and textures such as "mossy limbs" that require symmetric analysis of limbs to detect podo based on swellings; and (2) feature refinement module—the highly regularized 10-layered classification head that applies further refinements on the feature matrices extracted by the mapping models.

Clinical Relevance and Further Considerations

This study developed benchmark image-based diagnostic model(s) for skin NTDs using the traditional augmentation method to initially alleviate data scarcity issues (limited sample images with severe class imbalance). We applied this method on our sample input image using our predefined basic geometric transformations at the time of model training to mathematically simulate the real-world diagnostics of skin NTDs. In the actual real-world scenarios, skin NTDs are highly prevalent in remote underserved areas, where the diagnostics of the diseases are mostly undertaken by middle-level health care workers under low-resource settings. Hence, our model design assumed these scenarios that our model is used by middle-level health care workers without the need for highly sophisticated devices, where moderate-quality handheld smartphones can properly serve the diagnostics. Accordingly, we added the basic transformation techniques to our augmentation pipeline that includes rotation (20 degrees), shearing (range=0.1), and zooming (range=0.1) to avoid variations in our models' predictions that can be caused by variable camera angles, inconsistent framing, and varying focal distances in photographs of the skin NTDs. We also added the horizontal flipping transformation to maintain anatomical symmetry, and random brightness adjustments are also applied to simulate the diagnostic practices that can be conducted under different lighting adjustments.

While the study achieved its objective, the data scarcity issues still remained to be the major challenges. However, the top performance scores achieved by the final 2 optimized models (the models using the MN2 and DenseNet121 feature maps) underlined that further experiments using advanced DL-based data augmentations and balancing methods have the potential of further improving the predictive performance of the models. Hence, using each of the selected benchmark models as a basis, advanced DL-based methods such as the use of GAN-generated samples and weighted balancing methods are expected to significantly boost the models' performance, as confirmed by the performance improvements achieved by applying the traditional dynamic data augmentation.

Additionally, as a clinical diagnostic model, the other important aspect expected from such systems is the treatment facility, once the diseases are correctly identified with higher diagnostic accuracy. Generally, the incorporation of treatment recommendations represents the deployment-ready (final) stage of a diagnostic tool, mostly indicating that the model passed several improvement and validation stages. It also requires further system validation and authorization. Therefore, marking the current gaps, incorporation of treatment recommendations, and performance improvements using advanced DL methods are the immediate research tasks.

XSL•FO
RenderX

Conclusions

Ultimately, this study developed a benchmark image-based diagnostic model for skin NTDs through a robust hybrid DL pipeline using a novel skin NTDs dataset collected from a remote area representing an underserved community in the Southwest of Ethiopia. The study developed a new baseline model, applied transfer learning, designed a 2-stage approach, and applied dynamic data augmentation, where all experiments were conducted based on a novel research framework proposed for this study, the funnel framework. Using the funnel framework, optimal methods and high-performing models were selected in a phased approach. With the 2-stage approach being the best model-building approach, the DenseNet121 and MobileNetV2 were top-performing feature extractors. Finally, after the last training applying hyperparameter tuning, the models trained using these extractors still showed performance improvements, DenseNet121 (accuracy=96.6% and F_1-score=0.85) and MobileNetV2 (accuracy=96.6% and F_1-score=0.85). Therefore, the study ultimately selected the DenseNet121 and MobileNetV2 models for feature mapping (extraction) for the final DL-based skin NTDs diagnostic models.

While we developed the intended model for the diagnosis of skin NTDs using the novel skin images dataset, this study also exhibits some downsides that limit the study from contributing to its potential. Data scarcity and severe class imbalance are the primary challenges. The number of diseases represented in this study is also limited to only 3 diseases while there are more than 8 skin NTDs prevalent in Ethiopia. Additionally, the study does not demonstrate data-balancing methods due to resource-related constraints and the extended experiments required. Furthermore, only image data were considered to develop the proposed diagnostic model. Therefore, we recommend further efforts to address the mentioned limitations. First, it is recommended that additional data must be collected from different affected areas (if applicable) with the data being representative of all disease classes. Additionally, DL-based class-balancing methods, such as conditional augmentation based on generative models, are recommended to be experimented. We also suggest the inclusion of multiple types of data, such as text-based patients' data, to expand the dimensionality of the dataset used.

Acknowledgments

The authors of this study gratefully acknowledge Mr Alemayehu Bekele and his team for providing the data used to complete this study, and the technical support he provided us. This study used the data collected from patients with skin Neglected Tropical Diseases (skin NTDs) living in Gacho Baba District of the Gamo Zone, Southwest Ethiopia. The data were initially collected by Mr. Alemayehu Bekele for a project-based research intended to assess the burden of skin NTDs through community screening, led by Mr Alemayehu Bekele from the Collaborative Research and Training Center for Neglected Tropical Disease, College of Medicine Health Sciences of AMU. The authors acquired the dataset for this study through collaboration after having the required ethical clearance letter. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this paper. Finally, the authors would like to declare that no generative tools and generated contents are used for the preparation of any part of this study.

Funding

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

Data Availability

Currently, the new skin NTDs dataset used for this study is not publicly available, as it requires the final verification and consensus regarding dissemination.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Distribution of data samples in the new skin Neglected Tropical Diseases image dataset. The figure illustrates data distribution portraying class imbalance among the 3 disease classes: scabies (dark blue, 63.9%), tungiasis (teal, 31.7%), and podo (light green, 4.4%).

[PNG File, 28 KB - derma_v9i1e91544_app1.png]

References

1. Yohannes Minyilu G, Yimer MA, Meshesha M. Application of deep learning-based multimodal data fusion for the diagnosis of skin neglected tropical diseases: systematic review. JMIR AI 2025 Dec 4;4:e67584. [doi: 10.2196/67584] [Medline: 41344666]

XSL•FO
RenderX

2. Deribe K, Meribo K, Gebre T, et al. The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. Parasit Vectors 2012 Oct 24;5(1):240. [doi: 10.1186/1756-3305-5-240] [Medline: 23095679]
3. The third national neglected tropical diseases strategic plan 2021-2025. : Ministry of Health; 2021 URL: https://espen.afro.who.int/sites/default/files/content/document/Third%20NTD%20national%20Strategic%20Plan%202021-2025_0.pdf [accessed 2026-06-16]
4. Ethiopia: country disease outlook. World Health Organization, Regional Office for Africa, Country Disease Outlook. 2023 Aug. URL: https://www.afro.who.int/sites/default/files/2023-08/Ethiopia.pdf [accessed 2026-06-05]
5. Elimination of neglected tropical diseases (NTDS) in Ethiopia, Woreda level coordination toolkit. : Federal Ministry of Health; 2019 URL: https://www.susana.org/_resources/documents/default/3-3709-7-1571741955.pdf [accessed 2026-06-16]
6. Skin NTDS laboratory network (Skin NTDS LABNET). World Health Organization. URL: https://www.who.int/groups/skin-ntds-laboratory-network [accessed 2026-06-05]
7. Abdela SG, Diro E, Zewdu FT, et al. Looking for NTDs in the skin; an entry door for offering patient centered holistic care. J Infect Dev Ctries 2020 Jun 29;14(6.1):16S-21S. [doi: 10.3855/jidc.11707] [Medline: 32614791]
8. Postigo JAR, Pathak P, Asiedu K. Skin health for all: update on skin neglected tropical diseases with a focus on buruli ulcer and yaws. : Mectizan Donation Program URL: https://iris.who.int/server/api/core/bitstreams/54e59c06-67c7-4c86-8ca9-f210448dab18/content [accessed 2026-06-16]
9. Quilter EEV, Butlin CR, Carrion C, Ruiz-Postigo JA. The WHO Skin NTD mobile application—a paradigm shift in leprosy diagnosis through Artificial Intelligence? Leprosy 2024 Jun;95(2):e2024030. [doi: 10.47276/lr.95.2.2024030]
10. Steyve N, Steve P, Ghislain M, Ndjakomo S, Pierre E. Optimized real-time diagnosis of neglected tropical diseases by automatic recognition of skin lesions. Inform Med Unlocked 2022;33:101078. [doi: 10.1016/j.imu.2022.101078]
11. Brigato L, Barz B, Iocchi L, Denzler J. Image classification with small datasets: overview and benchmark. IEEE Access 2022;10:49233-49250. [doi: 10.1109/ACCESS.2022.3172939]
12. Sales C, Coates SJ. Applications of artificial intelligence for high-burden, underserved skin diseases in global settings: a review. Curr Derm Rep 2025 Jun;14(1):14. [doi: 10.1007/s13671-025-00469-9]
13. Hailu A, Gebre T, Seife F, et al. Challenges and strategies for mainstreaming neglected tropical diseases campaign interventions in Ethiopia. Am J Trop Med Hyg 2025 Feb 5;112(2):467-478. [doi: 10.4269/ajtmh.24-0261] [Medline: 39591648]
14. Tekola Ayele F, Adeyemo A, Finan C, et al. HLA class II locus and susceptibility to podoconiosis. N Engl J Med 2012 Mar 29;366(13):1200-1208. [doi: 10.1056/NEJMoa1108448]
15. Yotsu RR, Ding Z, Hamm J, Blanton RE. Deep learning for AI-based diagnosis of skin-related neglected tropical diseases: a pilot study. PLoS Negl Trop Dis 2023 Aug;17(8):e0011230. [doi: 10.1371/journal.pntd.0011230] [Medline: 37578966]
16. Pattnayak P, Mohanty A, Das T, Patnaik S. Applying artificial intelligence and deep learning to identify neglected tropical skin disorders. Presented at: 2024 3rd International Conference for Innovation in Technology (INOCON); Mar 1-3, 2024; Bangalore, India p. 1-6. [doi: 10.1109/INOCON60754.2024.10511323]
17. Beesetty R, et al. Leprosy skin lesion detection: an AI approach using few shot learning in a small clinical dataset. Indian J Lepr 2023;95:89-102 [FREE Full text]
18. Barnowska EJ, Fastenau A, Penna S, Bonkass AK, Stuetzle S, Janssen R. Diagnosing skin neglected tropical diseases with the aid of digital health tools: a scoping review. PLoS Digit Health 2024 Oct;3(10):e0000629. [doi: 10.1371/journal.pdig.0000629] [Medline: 39374195]
19. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 2018 Aug;5(1):180161. [doi: 10.1038/sdata.2018.161]
20. Cassidy B, Kendrick C, Brodzicki A, Jaworek-Korjakowska J, Yap MH. Analysis of the ISIC image datasets: usage, benchmarks and recommendations. Med Image Anal 2022 Jan;75:102305. [doi: 10.1016/j.media.2021.102305]
21. Johnson JM, Khoshgoftaar TM. Survey on deep learning with class imbalance. J Big Data 2019 Dec;6(1):27. [doi: 10.1186/s40537-019-0192-5]
22. Wang Z, et al. A comprehensive survey on data augmentation. arXiv. Preprint posted online on Oct, 2025. [doi: 10.48550/arXiv.2405.09591]
23. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019 Dec;6(1):60. [doi: 10.1186/s40537-019-0197-0]
24. Kumar T, Brennan R, Mileo A, Bendechache M. Image data augmentation approaches: a comprehensive survey and future directions. IEEE Access 2024;12:187536-187571. [doi: 10.1109/ACCESS.2024.3470122]
25. Islam T, Hafiz M, Jim JR, Kabir M, Mridha MF. A systematic review of deep learning data augmentation in medical imaging: recent advances and future research directions. Healthc Analytics 2024 Jun;5:100340. [doi: 10.1016/j.health.2024.100340]
26. Ribas LC, Casaca W, Fares RT. Conditional generative adversarial networks and deep learning data augmentation: a multi-perspective data-driven survey across multiple application fields and classification architectures. AI 2025 Feb;6(2):32. [doi: 10.3390/ai6020032]
27. Wickramaratne SD, Mahmud M. Conditional-GAN based adversarial networks and deep learning data augmentation: a multi-perspective data-adversarial networks and deep learning data augmentation: a multi-perspective data-data augmentation

XSL•FO
RenderX

for deep learning task classifier improvement using fNIRS data. Front Big Data 2021;4:659146. [doi: 10.3389/fdata.2021.659146] [Medline: 34396092]
28. Wu H, Gu X. Max-pooling dropout for regularization of convolutional neural networks. In: Arik S, Huang T, Lai WK, Liu Q, editors. Neural Information Processing: Springer International Publishing; 2015, Vol. 9489:46-54. [doi: 10.1007/978-3-319-26532-2_6]
29. Cerqueira V, Santos M, Roque L, Baghoussi Y, Soares C. Online data augmentation for forecasting with deep learning. arXiv. Preprint posted online on Jan, 2025. [doi: 10.48550/arXiv.2404.16918]
30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016; Las Vegas, NV, USA p. 770-778. [doi: 10.1109/CVPR.2016.90]
31. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A convnet for the 2020s. Presented at: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-24, 2022. [doi: 10.1109/CVPR52688.2022.01167]
32. Woo S, Debnath S, Hu R, et al. ConvNeXt V2: co-designing and scaling convnets with masked autoencoders. Presented at: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 17-24, 2023; Vancouver, BC, Canada p. 16133-16142. [doi: 10.1109/CVPR52729.2023.01548]
33. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI p. 2261-2269. [doi: 10.1109/CVPR.2017.243]
34. Chollet F. Xception: deep learning with depthwise separable convolutions. Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI p. 1800-1807. [doi: 10.1109/CVPR.2017.195]
35. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. arXiv. Preprint posted online on May 28, 2019. [doi: 10.48550/arXiv.1905.11946]
36. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-23, 2018; Salt Lake City, UT p. 4510-4520. [doi: 10.1109/CVPR.2018.00474]
37. Yang TJ, et al. NetAdapt: platform-aware neural network adaptation for mobile applications. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, editors. Computer Vision ECCV: Springer International Publishing; 2018, Vol. 11214:289-304. [doi: 10.1007/978-3-030-01249-6_18]
38. Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3. Presented at: 2019 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 27 to Nov 2, 2019; Seoul, South Korea p. 1314-1324. [doi: 10.1109/ICCV.2019.00140]
39. Tan M, Le QV. EfficientNetV2: smaller models and faster training. arXiv. Preprint posted online on Apr 1, 2021. [doi: 10.48550/arXiv.2104.00298]
40. Wang A, Chen H, Lin Z, Han J, Ding G. Rep ViT: revisiting mobile CNN from ViT perspective. Presented at: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 16-22, 2024; Seattle, WA, USA p. 15909-15920. [doi: 10.1109/CVPR52733.2024.01506]
41. Hatamizadeh A, et al. FasterViT: fast vision transformers with hierarchical attention. arXiv. Preprint posted online on Jun 9, 2023. [doi: 10.48550/arXiv.2306.06189]
42. Anasosalu Vasu PK, Gabriel J, Zhu J, Tuzel O, Ranjan A. FastViT: a fast hybrid vision transformer using structural reparameterization. Presented at: 2023 IEEE/CVF International Conference on Computer Vision (ICCV); Sep 30-30, 2023; Paris, France p. 5762-5772. [doi: 10.1109/ICCV51070.2023.00532]
43. Cai H, Li J, Hu M, Gan C, Han S. EfficientViT: lightweight multi-scale attention for high-resolution dense prediction. Presented at: 2023 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 1-6, 2023; Paris, France p. 17256-17267. [doi: 10.1109/ICCV51070.2023.01587]
44. Deng J, Dong W, Socher R, Li LJ. ImageNet: a large-scale hierarchical image database. Presented at: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops); Jun 20-25, 2009. [doi: 10.1109/CVPR.2009.5206848]
45. Sahito A, Frank E, Pfahringer B. Transfer of pretrained model weights substantially improves semi-supervised image classification. In: Gallagher M, Moustafa N, Lakshika E, editors. AI 2020: Advances in Artificial Intelligence: Springer International Publishing; 2020, Vol. 12576:433-444. [doi: 10.1007/978-3-030-64984-5_34]
46. Alzubaidi L, Bai J, Al-Sabaawi A, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. J Big Data 2023 Apr;10(1):46. [doi: 10.1186/s40537-023-00727-2]
47. Rukundo O. Effects of image size on deep learning. Electronics (Basel) 2023 Feb;12(4):985. [doi: 10.3390/electronics12040985]
48. Li L, Jamieson K, DeSalvo G, Rostamizadeh A, Talwalkar A. Hyperband: a novel bandit-based approach to hyperparameter optimization. arXiv. Preprint posted online on Jun, 2018. [doi: 10.48550/arXiv.1603.06560]
49. Song Y, Chiangpradit M, Busababodhin P. Hyperband-optimized CNN-BiLSTM with attention mechanism for corporate financial distress prediction. Appl Sci (Basel) 2025 May;15(11):5934. [doi: 10.3390/app15115934]

XSL•FO
RenderX

Abbreviations

AMU: Arba Minch University
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
CNN: convolutional neural network
DL: deep learning
GAN: generative adversarial
GFLOP: Giga floating point operation
IRB: institutional review board
ISIC: International Skin Imaging Collaboration
MDA: mass drug administration
NTD: Neglected Tropical Disease
skin NTD: skin Neglected Tropical Disease
WSS: weighted sum of scores

Edited by S Subhadarshani; submitted 15.Jan.2026; peer-reviewed by R Wu, T Banerjee; revised version received 11.May.2026; accepted 20.May.2026; published 23.Jun.2026.

Please cite as:
Minyilu Y, Yimer MA, Meshesha M
Harmonized Dual Deep Learning Architectures for Image-Based Diagnostics of Skin Neglected Tropical Diseases: Benchmark Study via Novel Funnel Framework
JMIR Dermatol 2026;9:e91544
URL: https://derma.jmir.org/2026/1/e91544
doi:10.2196/91544

© Yohannes Minyilu, Mohammed Abebe Yimer, Million Meshesha. Originally published in JMIR Dermatology (http://derma.jmir.org), 23.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Navigating the Intersection of Radiofrequency Microneedling and Surgical Facelifts: Scoping Review

Mia Panlilio[1], BA; Rebecca Bolen[1], BA; Olnita Martini[1], MS; Alexa Bonk[1], BS; John Tedesco[2], DO

[1]Department of Biomedical Sciences, College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States
[2]College of Osteopathic Medicine, Oklahoma State University, Tulsa, OK, United States

Corresponding Author:
Mia Panlilio, BA
Department of Biomedical Sciences, College of Osteopathic Medicine, Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

Abstract

Background: Optimal management of facial skin laxity requires a nuanced approach by health care providers working in aesthetics. Radiofrequency microneedling (RFMN) devices have emerged as a popular noninvasive treatment for facial rejuvenation and improving skin laxity. While RFMN has demonstrated efficacy in enhancing skin tightening and complementing aesthetic procedures, its long-term impact on subsequent surgical facelifts remains uncertain.

Objective: The objective of this scoping review is to explore the interplay between RFMN and surgical facelift outcomes, with a focus on potential complications such as excessive skin tightening, dermal scarring, and altered tissue planes that may pose surgical challenges.

Methods: A search using PubMed and Google Scholar was conducted, and articles were selected from peer-reviewed journals based on specific inclusion and exclusion criteria. Only articles available in English were selected. In total, 21 articles were included in this scoping review.

Results: Papers included in this review discussed the mechanisms of action involved with RFMN, RFMN-related tissue changes, and how these changes could impact future facelift procedures. Most of the papers found that RFMN may drastically alter multiple tissue planes involved in facelift procedures due to collagen deposition through multiple tissue layers and increased tissue fibrosis. Patient factors influencing the effectiveness of RFMN and its role in facial rejuvenation were also examined, emphasizing the importance of navigating patient-specific demographics as a future consideration when creating an individualized treatment plan for each patient.

Conclusions: Patients should be informed that RFMN may lead to dermal fibrosis, tissue adhesions, and altered superficial musculoaponeurotic system composition, which could interfere with future facelift procedures and the patient's desired treatment goals. This emphasizes the importance of detailed discussion between the patient and health care provider to improve pretreatment consultation, increase patient education, and set realistic expectations. Further research is needed to determine optimal timing and treatment strategies for patients considering both RFMN and surgical facelifts to achieve the best aesthetic outcomes.

(JMIR Dermatol 2026;9:e78385) doi:10.2196/78385

KEYWORDS

radiofrequency; microneedling; facial; laxity; facelift

Introduction

Radiofrequency (RFMN) microneedling is a minimally invasive procedure that combines mechanical injury and thermal stimulation via tiny microneedles with radiofrequency (RF) energy to induce collagen remodeling [1]. The microneedling component creates controlled microtraumas, triggering a postinflammatory cascade that promotes neocollagenesis, elastin production, and angiogenesis. RF energy delivered through the microneedles then generates thermal coagulation within the dermis and hypodermis to induce collagen denaturation and subsequent contraction of tissue for skin tightening benefits [2]. This fractional approach allows for targeted treatment while preserving surrounding tissue, reducing recovery time [3]. Compared to traditional microneedling, evidence suggests that RFMN brings about greater improvements in aged skin, likely by eliminating senescent fibroblasts and increasing the number of nonsenescent fibroblasts [4].

A systematic review done in 2021 by Tan et al [5] analyzed 42 studies evaluating RFMN use across various conditions, with the largest evidence base for skin rejuvenation, followed by acne scars, acne vulgaris, striae, and axillary hyperhidrosis. A smaller number of studies were available supporting RFMN use for melasma, rosacea, cellulite, and androgenetic alopecia.

XSL•FO
RenderX

Based on the large and growing body of evidence for skin rejuvenation, RFMN devices have gained immense popularity for addressing skin laxity in patients seeking noninvasive alternatives to surgical facelifts. These devices offer treatment options for individuals outside the average age range for a facelift, those who have previously undergone a facelift, or patients desiring minimally invasive interventions [5]. Advances in RFMN technology, such as interchangeable tips with various microneedle pin configurations and dual treatment modes, allow for targeted treatments in delicate anatomical areas like the periorbital region [5]. While these technological refinements enhance customization, the process may impact deeper dermal structures critical to surgical outcomes, setting the stage for potential interference with future facelift procedures.

Long-term effects of RFMN before surgical facelifts remain unclear, raising concerns about potential complications. RFMN treatment prior to an elective facelift may have the potential to interfere with optimal facelift results due to excessive skin tightening, scarring, and damage to the dermis. This emphasizes the importance of pretreatment discussion about expectations and the adverse effects of RFMN if it is being used with a patient who is considering a facelift in the future. Given the interplay between RFMN and surgical facelifts, what are the long-term implications of RFMN for patients who may eventually pursue surgical facelift procedures? Could the very technology we're using to delay surgery unintentionally complicate it later? This scoping review explores the anatomical and clinical intersections between RFMN and surgical facelifts for patients realistically considering either a facelift or RFMN. It draws on current literature, evolving device technology, and real-world considerations to guide thoughtful treatment planning for optimal patient outcomes.

Methods

Searches done on PubMed and Google Scholar using the terms "skin laxity and radiofrequency microneedling," "skin laxity and microneedling," "skin laxity and facelift," "facelift and radiofrequency microneedling," and "facelift and microneedling" were conducted on January 4, 2024, and again on June 20, 2025, to account for any newly published or updated literature since the original search. The second search did not yield any new articles. Articles from peer-reviewed journals were included if they provided information on the mechanism of action of RFMN, described the techniques involved when performing a facelift procedure, or examined the effects of RFMN and/or facelift procedures on the skin. Only articles available in English were selected, and articles were excluded if they provided information on the use of RFMN on parts of the body other than the face and neck, or if they described RFMN treatment or surgical intervention unrelated to improving facial skin laxity. A summary of the inclusion and exclusion criteria are highlighted in Textbox 1. The review was conducted based on the 2005 methodology of Arksey and O'Malley [6]. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) reporting guidelines were followed, and the completed PRISMA-ScR checklist for this review can be found as Checklist 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- From peer-reviewed journals
- Content of article: mechanism of action of radiofrequency microneedling (RFMN); facelift procedures; effects of RFMN and/or facelift procedures on the skin
- Available in English

Exclusion criteria

- Content of article: RFMN being used on parts of the body other than the face and neck; RFMN treatment or surgical intervention unrelated to improving facial skin laxity
- Not available in English or an English translation

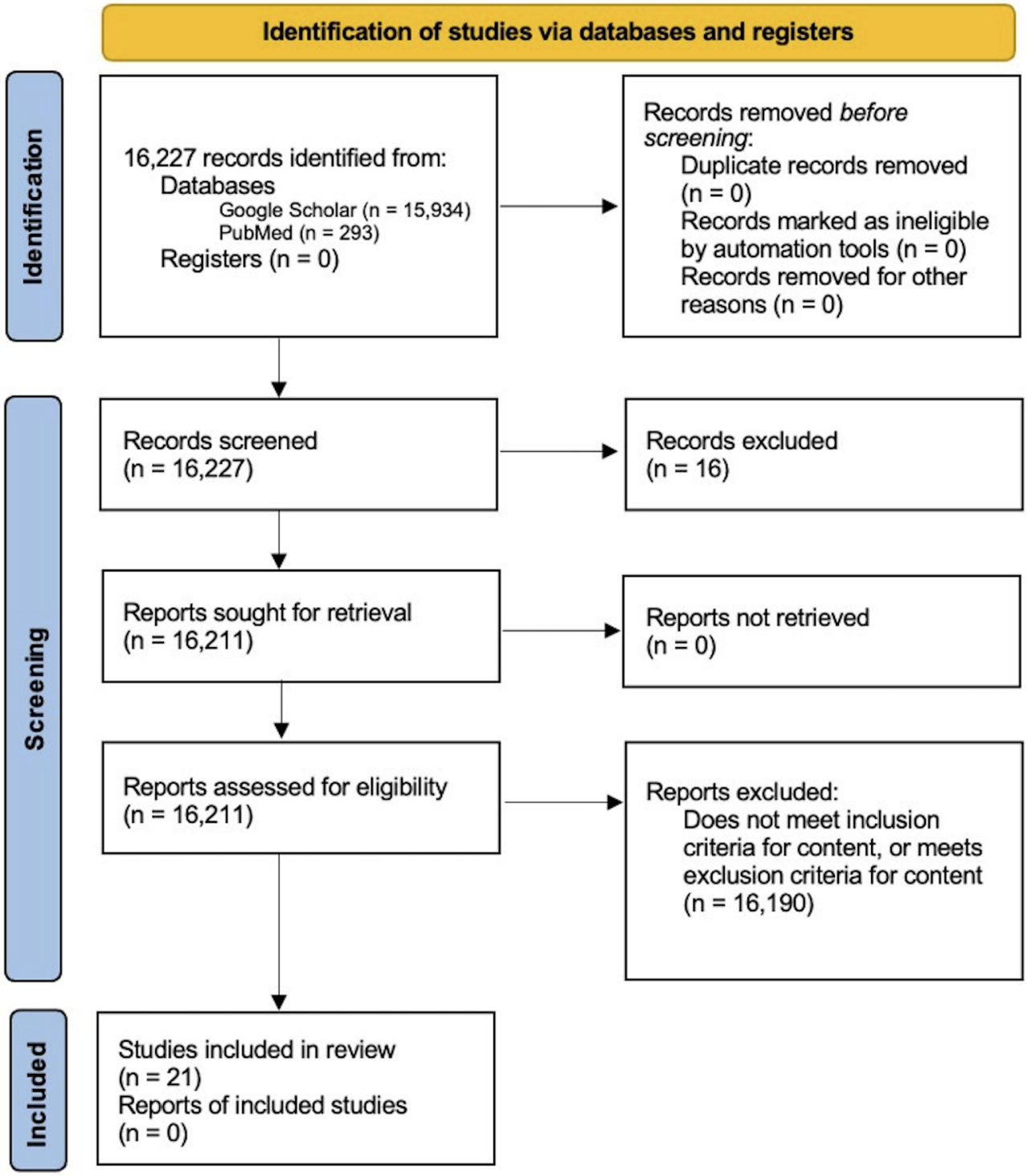
Results

Article Selection

From the initial search using the selected search terms, 15,934 and 293 articles were respectively identified from Google Scholar and PubMed, for a total of 16,277 articles (Figure 1). Out of the 16,277 articles screened, 16 PubMed articles were excluded because they were not available in English. Google Scholar does not have a language screening filter, so all 15,934 articles from the initial search were still considered. After all authors screened the remaining articles for content based on the inclusion and exclusion criteria, 21 articles in total were selected for this review. A PRISMA diagram of the article selection is available below in Figure 1.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of article selection.



Summary of Included Articles

After article screening was completed and the 21 articles were selected, 12 articles were identified as original research, 5 as narrative literature reviews, 2 as systematic reviews, and 1 as an educational reference article. One article, by Arksey and O'Malley [6], was used as a framework to guide writing of this review, and did not necessarily fit the inclusion and exclusion criteria for articles specifically related to the topic of this scoping review. Table 1 displays a summary of the articles mentioned, as well as their article types and study designs.

XSL•FO
RenderX

Table . Summary of included articles.

Authors (year)	Journal	Article type	Study design
Devgan et al (2019) [1]	*Otolaryngology Clinics of North America*	Literature review	Narrative review
Spataro et al (2022) [2]	*Facial Plastic Surgery Clinics of North America*	Literature review	Narrative review
Hendricks and Farhang (2022) [3]	*Journal of Cosmetic Dermatology*	Literature review	Narrative review
Hwang et al (2025) [4]	*Scientific Reports*	Original research	Split-face comparative clinical trial
Tan et al (2021) [5]	*Dermatologic Surgery*	Literature review	Narrative review
Arksey and O'Malley (2005) [6]	*International Journal of Social Research Methodology*	Original research	Not applicable
Dayan et al (2020) [7]	*Plastic and Reconstructive Surgery—Global Open*	Original research	Prospective case series (single-arm clinical study)
Arnaoutakis et al (2022) [8]	*Facial Plastic Surgery & Aesthetic Medicine*	Literature review	Narrative review
Ramaut et al (2018) [9]	*Journal of Plastic, Reconstructive & Aesthetic Surgery*	Systematic review	Systematic review
Huang et al (2014) [10]	*Biochemistry*	Basic science research	In vitro experimental study
Nguyen et al (2025) [11]	*Lasers in Medical Science*	Original article	Clinical and histologic study (prospective cohort)
Xu et al (2025) [12]	*Lasers in Surgery and Medicine*	Original research	Animal study (porcine model)
Zheng et al (2014) [13]	*Dermatologic Surgery*	Original research	Experimental histologic study
Wang et al (2025) [14]	*Lasers in Medical Science*	Original research	Pilot clinical study
Wang et al (2024) [15]	*Lasers in Surgery and Medicine*	Original research	Animal study (porcine model)
Cho et al (2024) [16]	*Skin Research & Technology*	Original research	Animal study (minipig model)
Hohman et al (2023) [17]	*StatPearls*	Reference article	Narrative review (educational)
Ghassemi et al (2003) [18]	*Aesthetic Plastic Surgery*	Original research	Anatomical cadaveric study
Demesh et al (2021) [19]	*Journal of Cosmetic Dermatology*	Original research	Clinical case series
Seo et al (2012) [20]	*Lasers in Surgery and Medicine*	Original research	Clinical and histologic study (prospective cohort)
Austin et al (2022) [21]	*Lasers in Surgery and Medicine*	Systematic review	Systematic review

RFMN-Related Tissue Changes

While RFMN effectively improves skin laxity and wrinkle reduction, its impact on future facelift procedures remains uncertain. An article published in 2022 found that a single session of noninvasive fractional bipolar RFMN achieved approximately 37% of the skin laxity improvement seen with a surgical facelift, suggesting multiple treatments may be required for significant results [8]. However, a 2018 systematic review from the *Journal of Plastic, Reconstructive & Aesthetic Surgery* states that repeated sessions risked dermal fibrosis, particularly in the papillary dermis, potentially complicating future surgical interventions [9].

As RFMN devices introduce repeated and organized microtraumas into the skin, the depth of penetration directly impacts targeted tissue layers. Microneedling from such devices introduces a targeted mechanism of repair that excludes highly inflammatory cellular cascades, such as transforming growth factor β-1 (TGF-β1) and transforming growth factor β-2 (TGF-β2), and instead is driven through a less-inflammatory cascade via transforming growth factor β-3 (TGF-β3), a protein known to lead to fibroblast migration and collagen matrix remodeling [10]. Platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), and epidermal growth factor (EGF) are all released locally in response to the microtrauma, allowing natural skin tightening via angiogenesis and collagen deposition [2]. Repeated treatments thus increase collagen deposition, which increases the risk for the development of dermal fibrosis.

A better understanding of the timeline of collagenesis following RFMN is critical when considering the interplay with surgical facelifts. Acute inflammation and early collagen deposition dominate the first week after treatment, followed by organized collagen remodeling and maturation over the ensuing 1 to 3 months [2,10]. Persistent changes in dermal structure, including fibrosis or altered tensile strength, may interfere with surgical flap elevation, tissue pliability, and healing after a facelift.

RFMN energy settings can be optimized to balance skin tightening with control of fibrosis by titrating energy per needle,

XSL•FO
RenderX

pulse duration, and depth to achieve sufficient dermal coagulation for neocollagenesis and elastogenesis while avoiding excessive thermal injury that may promote fibrotic remodeling. Data supports targeting moderate energy settings (eg, energy per needle 20 - 60 mJ, pulse durations 100-300 ms) and limiting the number of passes to induce controlled dermal coagulation, maximizing skin tightening while minimizing the risk of fibrosis [11,12]. Adjusting needle depth to target the reticular dermis and using insulated needles can further localize RFMN thermal effects, thus reducing epidermal damage and unwanted fibrosis [13,14]. Sequential or pulsed energy delivery, as well as energy feedback systems, can help regulate tissue response and prevent overtreatment [15,16].

Effects on Future Elective Facelift Procedures

Surgical facelift procedures rely on the manipulation of the superficial musculoaponeurotic system (SMAS), a fibrofatty connective tissue layer continuous with the superficial cervical fascia, connected to the platysma muscle inferiorly and the galea superiorly [17]. It plays an integral role in the anatomic relation of the superficial dermis to the underlying facial muscles. There are two distinct SMAS compositions given anatomic regions, and the abrupt junction of differing compositions resides at the nasolabial fold region, where medially, there are fewer fat lobules and a more direct connection of the SMAS to the superficial dermis as muscle fibers are seen to extend superficially into the dermis [17]. However, the other regions of the SMAS lateral to the nasolabial fold still carry the same properties of communication of facial muscle to skin by muscle tendon fibers connecting both regions via the SMAS [18]. Beneath this layer, the SMAS has a complex relation with deep ligaments and connections that limits the mobility of superficial structures. These connections are crucial to release to generate the most optimal movement for desired facelift outcomes [17].

Due to the sophisticated relationship of neighboring structures, RFMN, particularly at greater depths, may alter these structural relationships. Traditional surgical facelifts target a single plane of tissue in a primary horizontal plane [17]. On the contrary, RFMN targets a small treatment area in a vertical configuration through multiple planes of tissue, potentially leading to increased tissue adhesions, difficult surgical dissection, impaired flap mobility, and suboptimal facelift outcomes [10].

Although undergoing RFMN treatments prior to a facelift could potentially induce a level of fibrosis that may help delay the timeline when a patient would be a candidate for a facelift, evidence suggests RFMN may be more beneficial postoperatively. Following a facelift, RFMN could enhance skin tightening and improve aesthetic outcomes by stimulating additional collagen production either as an immediate adjunctive therapy or as a method to combat recurrent long-term skin laxity [19]. Careful planning is necessary to determine the appropriate time frame between treatments to avoid excessive fibrosis and impaired wound healing.

Patient Considerations for RFMN Procedures

When considering therapy using RFMN, specific patient populations should be considered when discussing treatment options, as certain age groups, as well as patients with jowl laxity, have been shown to experience better outcomes with facial surgery [19]. Taking into consideration that each treatment has varying mechanisms and different anatomical targets, a comparative study found that surgical facelifts improved skin laxity by 46% relative to baseline, whereas RFMN alone achieved only a 16% improvement [20]. These findings underscore the importance of setting realistic patient expectations regarding treatment efficacy. Additionally, patient age should be considered, as older individuals (≥55 years) experience more pronounced skin tightening with RFMN compared to younger patients [2]. Younger patients typically have a higher collagen content in their skin compared to older patients, who undergo collagen loss due to age; thus, the relative resulting decrease in skin laxity is much more noticeable in older patients versus younger patients. This knowledge prompts early discussion of RFMN treatment to address its potential effectiveness or lack thereof, especially at a younger age.

In addition to considering a patient's age, premature neck and jowl laxity are common concerns among patients seeking skin-tightening treatments. However, RFMN does not effectively target subplatysmal fat, necessitating careful patient selection—individuals with significant subplatysmal fat may achieve superior results with surgical interventions such as liposuction [19]. For patients with pronounced skin laxity, RFMN alone may be insufficient and could exacerbate sagging if incidental heat-induced fat loss occurs without concurrent skin excision [19]. Other complications of RFMN reported include hyper- or hypopigmentation of treated skin, thermal burns, blistering, and scarring; these can often be mitigated with proper technique and equipment settings [8].

Discussion

Main Findings

The nuances of RFMN in facial rejuvenation necessitate a deeper understanding of its implications for future surgical facelifts. This calls for detailed discussion between the patient and health care provider to improve pretreatment consultation, patient education, and results. Patients should be informed that RFMN may lead to dermal fibrosis, tissue adhesions, subcutaneous adipose denaturation, and altered SMAS composition, which could complicate facelift procedures and their desired outcomes. Additionally, providers should address the limitations of RF to effectively target jowl laxity and set realistic expectations regarding RF results in patients younger than 55 years. Understanding these points would allow for individualized treatment planning, ensuring patients receive the most appropriate interventions based on their anatomical considerations and aesthetic goals. Furthermore, current evidence suggests RFMN may be better positioned as a postoperative adjunct rather than a presurgical intervention, especially in patients known to be surgical candidates in the future.

It is also important to note that RFMN has primarily been used for skin rejuvenation, mild laxity, and conditions such as acne scars, rather than as a substitute for surgical facelift procedures [5]. Additional systematic reviews and clinical trials, such as those by Austin et al [21] and Nguyen et al [11], further support

XSL•FO
RenderX

that early RFMN and radio frequency protocols are designed for modest rejuvenation, targeting mild-to-moderate laxity and stimulating collagen production. Among these studies, reported patient satisfaction was highest among those seeking subtle improvements rather than facelift-level results.

Limitations

This scoping review is limited by the minimal availability of long-term studies specifically examining the effects of RFMN on subsequent surgical facelift procedures. The available literature mainly consists of case reports, small-scale studies, and expert opinion, which restricts the generalizability of conclusions. In addition, the large amount of variation between RFMN device settings, treatment protocols, and patient demographics across studies further limits the ability to standardize findings or establish definitive treatment guidelines.

Conclusion

Cosmetic surgical providers trained in RFMN and/or facelift procedures should give careful consideration to this new technology when discussing different options for facial rejuvenation. Factors to weigh in these considerations should include, but are not limited to, age-related expectations, area of treatment, and the potential impact of subsequent facelifts. Rather than viewing RFMN and surgical facelifts as isolated interventions, providers should consider how early noninvasive treatments may influence future surgical options. This has important implications for clinical decision-making, patient education, and informed consent.

Ultimately, optimizing aesthetic outcomes will require a more integrated strategy that aligns patient goals with both immediate and long-term treatment trajectories. Future studies are needed to evaluate the cumulative impact of RFMN on facial anatomy and to guide safe, evidence-based treatment planning for patients considering both noninvasive and surgical facial rejuvenation options. Additional research should also focus on establishing guidelines for the optimal timing of RFMN relative to surgical facelifts and identifying strategies to minimize adverse effects of RFMN. As RFMN technology evolves, ongoing studies will be critical in refining its role in facial rejuvenation and improving patient outcomes.

Acknowledgments

All authors declare no use of generative artificial intelligence in drafting any portion of the manuscript (this includes any and all tools, including ChatGPT, Gemini, Claude, or any other comparable tool not explicitly listed).

Funding

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

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 19 KB - derma_v9i1e78385_app1.docx]

References

1. Devgan L, Singh P, Durairaj K. Minimally invasive facial cosmetic procedures. Otolaryngol Clin North Am 2019 Jun;52(3):443-459. [doi: 10.1016/j.otc.2019.02.013] [Medline: 30954270]
2. Spataro EA, Dierks K, Carniol PJ. Microneedling-associated procedures to enhance facial rejuvenation. Facial Plast Surg Clin North Am 2022 Aug;30(3):389-397. [doi: 10.1016/j.fsc.2022.03.012] [Medline: 35934440]
3. Hendricks AJ, Farhang SZ. Dermatologic facial applications of Morpheus8 fractional radiofrequency microneedling. J Cosmet Dermatol 2022 Oct;21 Suppl 1:S11-S19. [doi: 10.1111/jocd.15231] [Medline: 35916259]
4. Hwang JM, Lee SH, Baek EJ, et al. Comparison of the effects of fractional microneedle radiofrequency and microneedling on modulating the senescent fibroblast milieu in aged skin. Sci Rep 2025 May 26;15(1):18296. [doi: 10.1038/s41598-025-02545-3] [Medline: 40419589]
5. Tan MG, Jo CE, Chapas A, Khetarpal S, Dover JS. Radiofrequency microneedling: a comprehensive and critical review. Dermatol Surg 2021 Jun 1;47(6):755-761. [doi: 10.1097/DSS.0000000000002972] [Medline: 33577211]
6. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
7. Dayan E, Rovatti P, Aston S, Chia CT, Rohrich R, Theodorou S. Multimodal radiofrequency application for lower face and neck laxity. Plast Reconstr Surg Glob Open 2020 Aug;8(8):e2862. [doi: 10.1097/GOX.0000000000002862] [Medline: 32983756]
8. Arnaoutakis D, Kochhar A, Magro I, Karimi K. The role of subcutaneous radiofrequency-assisted liposculpture in the facial plastic surgeon's practice. Facial Plast Surg Aesthet Med 2022;24(S1):S11-S16. [doi: 10.1089/fpsam.2022.0225] [Medline: 36103269]

XSL•FO
RenderX

9. Ramaut L, Hoeksema H, Pirayesh A, Stillaert F, Monstrey S. Microneedling: where do we stand now? A systematic review of the literature. J Plast Reconstr Aesthet Surg 2018 Jan;71(1):1-14. [doi: 10.1016/j.bjps.2017.06.006] [Medline: 28690124]
10. Huang T, Schor SL, Hinck AP. Biological activity differences between TGF-β1 and TGF-β3 correlate with differences in the rigidity and arrangement of their component monomers. Biochemistry 2014 Sep 16;53(36):5737-5749. [doi: 10.1021/bi500647d] [Medline: 25153513]
11. Nguyen L, Bartholomeusz J, Schneider SW, Herberger K. Histological and clinical dose-response analysis of radiofrequency microneedling treatment for skin rejuvenation. Lasers Med Sci 2025 Feb 7;40(1):75. [doi: 10.1007/s10103-025-04335-9] [Medline: 39915343]
12. Xu Y, Zhang Y, Wang H, Tang H, Zeng W, Wen X. Microneedle radiofrequency induces extracellular matrix remodeling through fibroblast activation: a histological study in a porcine model. Lasers Surg Med 2025 Aug;57(6):528-543. [doi: 10.1002/lsm.70033] [Medline: 40444432]
13. Zheng Z, Goo B, Kim DY, Kang JS, Cho SB. Histometric analysis of skin-radiofrequency interaction using a fractionated microneedle delivery system. Dermatol Surg 2014 Feb;40(2):134-141. [doi: 10.1111/dsu.12411] [Medline: 24373135]
14. Wang M, Li Y, Lai X, Shi C, Yan Y. The clinical effects and skin histological changes induced by a novel insulated radiofrequency microneedle: a pilot study. Lasers Med Sci 2025 Feb 17;40(1):98. [doi: 10.1007/s10103-025-04354-6] [Medline: 39960528]
15. Wang H, Hamblin MR, Zhang Y, Xu Y, Wen X. Histological evaluation of monopolar and bipolar radiofrequency microneedling treatment in a porcine model. Lasers Surg Med 2024 Mar;56(3):288-297. [doi: 10.1002/lsm.23768] [Medline: 38334177]
16. Cho SB, Kang SY, Lee YJ, Choi M, Kim B, Ahn JC. Effect of sequential delivery of 1- and 2-MHz bipolar microneedling radiofrequency energy on thermal tissue reactions in a minipig model. Skin Res Technol 2024 Sep;30(9):e13898. [doi: 10.1111/srt.13898] [Medline: 39213160]
17. Hohman MH, Raggio BS, Patel BC. Deep plane facelift. In: StatPearls: StatPearls Publishing; 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK545277/ [accessed 2024-01-22]
18. Ghassemi A, Prescher A, Riediger D, Axer H. Anatomy of the SMAS revisited. Aesthetic Plast Surg 2003;27(4):258-264. [doi: 10.1007/s00266-003-3065-3] [Medline: 15058546]
19. Demesh D, Cristel RT, Gandhi ND, Kola E, Dayan SH. The use of radiofrequency-assisted lipolysis with radiofrequency microneedling in premature jowl and neck laxity following facialplasty. J Cosmet Dermatol 2021 Jan;20(1):93-98. [doi: 10.1111/jocd.13824] [Medline: 33128284]
20. Seo KY, Yoon MS, Kim DH, Lee HJ. Skin rejuvenation by microneedle fractional radiofrequency treatment in Asian skin; clinical and histological analysis. Lasers Surg Med 2012 Oct;44(8):631-636. [doi: 10.1002/lsm.22071] [Medline: 22936274]
21. Austin GK, Struble SL, Quatela VC. Evaluating the effectiveness and safety of radiofrequency for face and neck rejuvenation: a systematic review. Lasers Surg Med 2022 Jan;54(1):27-45. [doi: 10.1002/lsm.23506] [Medline: 34923652]

Abbreviations

EGF: epidermal growth factor
FGF: fibroblast growth factor
PDGF: platelet-derived growth factor
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RF: radio frequency
RFMN: radio frequency microneedling
SMAS: superficial musculoaponeurotic system
TGF-β1: transforming growth factor β-1
TGF-β2: transforming growth factor β-2
TGF-β3: transforming growth factor β-3

Edited by R Dellavalle; submitted 01.Jun.2025; peer-reviewed by A Edalatpour, P Huang; revised version received 08.Dec.2025; accepted 06.Jan.2026; published 27.Apr.2026.

Please cite as:
Panlilio M, Bolen R, Martini O, Bonk A, Tedesco J
Navigating the Intersection of Radiofrequency Microneedling and Surgical Facelifts: Scoping Review
JMIR Dermatol 2026;9:e78385
URL: https://derma.jmir.org/2026/1/e78385
doi:10.2196/78385

XSL•FO
RenderX

© Mia Panlilio, Rebecca Bolen, Olnita Martini, Alexa Bonk, John Tedesco. Originally published in JMIR Dermatology (http://derma.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Informatics-Based Psychotherapeutic and Psychiatric Interventions in Dermatology: Scoping Review of Impacts on Skin Disease Severity and Mental Health Outcomes

Caroline Lamarre[1], BSc; Jeffrey Chivinski[2], BPharm, MD; Alexandre Hudon[3,4,5,6,7], BEng, MSc(Ed), MD, PhD

[1]Department of Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada

[2]Beacon Dermatology, Calgary, AB, Canada

[3]Department of Psychiatry and Addictology, Faculty of Medicine, Université de Montréal, 2900 Boulevard Édouard-Montpetit, Montréal, QC, Canada

[4]Department of Psychiatry, Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada

[5]Department of Psychiatry, Institut national de psychiatrie légale Philippe-Pinel, Montréal, QC, Canada

[6]Centre de recherche de l'Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada

[7]Centre de recherche en pédagogie de la santé, Montréal, QC, Canada

Corresponding Author:
Alexandre Hudon, BEng, MSc(Ed), MD, PhD
Department of Psychiatry and Addictology, Faculty of Medicine, Université de Montréal, 2900 Boulevard Édouard-Montpetit, Montréal, QC, Canada

Abstract

Background: Chronic dermatologic conditions such as psoriasis, atopic dermatitis, and hidradenitis suppurativa are associated with a high burden of psychiatric comorbidities, including depression, anxiety, and suicidality. Despite growing awareness of the psychosocial impact of skin diseases, mental health needs remain underaddressed in dermatologic care. Digital technologies (including teledermatology, mobile health apps, and internet-delivered psychotherapies) offer promising avenues for integrating psychotherapeutic and psychiatric interventions into dermatology. However, the scope, effectiveness, and implementation of such informatics-based approaches remain poorly mapped in the literature.

Objective: This scoping review aimed to systematically identify, categorize, and synthesize studies on digital psychotherapeutic and psychiatric interventions targeting patients with dermatological conditions, with a focus on clinical, mental health, and implementation outcomes.

Methods: Following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines, we conducted a comprehensive search across 5 databases (MEDLINE, Embase, Web of Science, PsycINFO, and Google Scholar) for articles published up to March 2025. Studies were included if they involved patients with dermatologic conditions and assessed interventions that combined a digital informatics component (eg, telehealth, apps, artificial intelligence, virtual platforms) with a psychotherapeutic or psychiatric element (eg, cognitive behavioral therapy [CBT], mindfulness, consult-liaison psychiatry). Eligible study designs included clinical trials, observational studies, and mixed methods research. Data were extracted systematically, and methodological quality was assessed using JBI tools.

Results: Out of 15,176 records identified, 11 studies met the inclusion criteria. Most interventions targeted psoriasis (9/11) and used asynchronous digital platforms such as internet-based CBT and mobile apps. Across studies, dropout rates ranged from 10% to 76%. Improvements in dermatologic quality of life were reported in 6 of 11 studies, with statistically significant reductions in depression and anxiety observed in multiple trials (eg, internet-based CBT and mindfulness-based interventions), alongside reductions in psoriasis severity (Psoriasis Area and Severity Index) and itch intensity in randomized controlled trials. Intervention duration ranged from single-session virtual reality exposure to 8 - to 12-week structured programs. However, long-term outcomes beyond 3 to 12 months were rarely assessed, and reporting of sociodemographic variables and equity-related factors was limited.

Conclusions: Informatics-based psychotherapeutic and psychiatric interventions represent a promising frontier in psychodermatology, with early evidence suggesting feasibility and potential clinical benefit. Digital platforms may expand access to mental health support and improve holistic care for patients with dermatologic conditions. However, significant gaps remain in terms of equity, long-term effectiveness, integration into clinical workflows, and adaptation for diverse populations. Future research should focus on rigorous, inclusive trials and the development of hybrid models that blend digital and face-to-face care to ensure sustainable and equitable impact.

(JMIR Dermatol 2026;9:e82096) doi:10.2196/82096

XSL•FO
RenderX

KEYWORDS

psychodermatology; digital health; cognitive behavioral therapy; teledermatology; mental health; dermatologic conditions; mobile health apps; digital therapeutics; psychiatric comorbidity; scoping review

Introduction

Dermatologic Issues and Mental Health

Chronic dermatologic conditions, such as psoriasis, atopic dermatitis (AD), acne, vitiligo, and hidradenitis suppurativa (HS), are among the most psychologically burdensome medical conditions and impose a substantial, multifaceted impact on mental health. Not only do these cutaneous diseases cause visible disfigurement, discomfort, and functional and social impairment, but they are also associated with a high prevalence of various psychiatric comorbidities. Chronic symptoms like pain and pruritus associated with the relapsing nature of many skin conditions often mediate these psychological effects [1]. Large-scale studies have found that around 10% of patients with dermatologic conditions meet the criteria for clinical depression and around 17% for clinical anxiety, with about 12% to 13% reporting suicidal ideation [1]. These rates are roughly 2 to 3 times higher than in the general population, underscoring how common serious psychological comorbidities (eg, depression, anxiety, social anxiety or phobia, and suicidal thoughts) significantly add to patients' overall disease burden [1]. For instance, psoriasis and AD are each associated with notably elevated risks of new-onset depression and anxiety. A recent UK cohort study reported that adults diagnosed with psoriasis or AD had a 10% to 20% greater hazard of developing depression or anxiety within the first year, compared to matched individuals without skin disease [2]. In particular, patients with more severe skin involvement (such as extensive psoriasis) show even higher relative risks of depression and other psychiatric sequelae [2].

Psychiatric morbidity appears especially pronounced in HS, which research indicates is accompanied by a high prevalence of depression and anxiety, as well as increased rates of substance use disorders and even serious mental illnesses such as psychosis or bipolar disorder [3]. This higher prevalence is due to a combination of biopsychosocial factors, such as chronic pain and physical discomfort secondary to malodor and constant drainage. Many patients develop social withdrawal and avoidance behaviors out of fear of stigmatization. As the disease often affects intertriginous and genital areas, patients experience an impact on their sexual health and relationship difficulties. A lack of cure or repeated ineffective treatments leads to hopelessness and despair over time. Most alarmingly, patients with HS face a dramatically elevated risk of suicidality. For example, epidemiologic studies have found that suicide is roughly twice as likely in HS sufferers compared to those without HS [3]. Such findings illustrate that chronic inflammatory skin diseases carry not only physical burdens but also profound psychological and social challenges for patients. This underlines the need for integrated, multidisciplinary approaches to care for these patients.

Despite this heavy psychosocial burden, patients' mental health needs are often overlooked in dermatologic care. In 1 large survey of over 7200 patients with inflammatory dermatoses, two-thirds reported high stress levels, yet fewer than 15% had ever been offered any psychological support [4]. Dermatology clinics traditionally focus on cutaneous findings, and many patients never receive screening or referral for underlying depression, anxiety, or other psychological issues related to their skin condition [1]. These unmet needs can be due to under-recognition by care providers, limited access to mental health care like psychology or psychiatry, or simply the stigma associated with mental illness. Dermatologists often lack the time or referral pathways to manage the psychologic comorbidities. This gap in care means that a significant portion of patients with dermatologic disorders endure unaddressed emotional distress, which can, in turn, exacerbate disease outcomes. Furthermore, this can lead to poor adherence to medical treatments and a worsening quality of life. Consequently, there is a growing consensus that integrating mental health screening and psychotherapeutic support into routine dermatologic practice is critical for the optimal management of chronic skin diseases. Experts advocate for the regular use of mental health questionnaires (eg, Patient Health Questionnaire-2 for depression and Generalized Anxiety Disorder-7 for anxiety) in dermatology settings and prompt referral to psychiatry or psychology when needed [5]. Studies suggest that comprehensive care, combining dermatologic treatment with interventions, such as cognitive behavioral therapy (CBT), stress management techniques, or support groups, can improve patients' quality of life and potentially even their skin disease severity [6]. By recognizing and addressing the psychological comorbidities of chronic dermatoses, clinicians can substantially reduce the overall disease burden and improve long-term outcomes for these patients.

Psychotherapeutic Interventions and the Emerging Role of Digital Technologies

Digital platforms (including teledermatology, mobile health [mHealth] apps, artificial intelligence [AI]–supported therapeutics, and virtual assistants) are reshaping dermatologic care and have emerged as innovative ways to provide and deliver therapeutic interventions to patients with chronic medical conditions. Teledermatology (via store-and-forward images or live video) enables remote diagnosis and management of chronic inflammatory skin diseases with high reliability. It can significantly shorten waiting times for patients while maintaining diagnostic accuracy comparable to in-person visits [7]. For example, 1 study on patients with chronic skin conditions found that live-video teledermatology achieved 96% diagnostic concordance with face-to-face consultations and reduced consultation turnaround time by 90% [8]. This translates to faster treatment and greatly improved access for patients who might otherwise face long delays or travel barriers due to living in remote areas. In addition to bypassing geographic barriers, teledermatology can help reduce stigma or anxiety associated with in-person therapy and is more flexible. mHealth apps offer

XSL•FO
RenderX

accessible tools for patient self-monitoring, education, and behavior change. Dermatology apps can help patients track symptoms (for instance, logging eczema or psoriasis flares), adhere to treatment plans, and learn about skincare, thereby empowering them in day-to-day disease management [9]. Multiple scoring systems self-reported by patients can be integrated into these apps, for example: the Numeric Rating Scale for evaluating itch, or Skindex, which measures the emotional, symptomatic, and functional burden of skin diseases. Recent evidence confirms that such mHealth interventions improve patient outcomes: a 2024 meta-analysis showed that smartphone apps and telemonitoring significantly enhanced patients' quality of life and self-management of AD, while also increasing patient satisfaction and engagement in care [10]. Meanwhile, AI-driven digital therapeutics and virtual assistants are being integrated into dermatology workflows to support both providers and patients. In a teledermatology program for psoriasis, adding a virtual assistant (chatbot) improved medication adherence and gave patients greater peace of mind by connecting them to their dermatologist; notably, participants' Dermatology Life Quality Index (DLQI) scores improved with the chatbot intervention [11]. Such digital assistants can streamline triage, answer patient questions, and automate follow-ups, augmenting the efficiency of teledermatology services [11].

Importantly, these technologies do not just address skin health in isolation. They also impact mental health outcomes. Many chronic dermatoses carry a significant psychological burden (stress, low self-esteem, anxiety), so bridging dermatologic care with psychosocial support is crucial. Digital platforms are beginning to fill this need. For instance, an internet-based cognitive behavioral therapy (iCBT) program for patients with atopic eczema significantly reduced itching and insomnia while improving quality of life, with outcomes comparable to traditional therapist-led care [12]. This kind of online psychodermatologic intervention is valuable because access to in-person psychological services for patients with skin diseases is often limited [12]. In sum, teledermatology and related digital innovations are enhancing chronic skin disease management by improving convenience and access, supporting patient self-management, and addressing the psychosocial dimensions of dermatologic care in ways that were not previously possible. These interventions enhance the resilience and quality of life of patients with chronic skin diseases.

Gaps in the Literature

Despite the growing promise of digital platforms in dermatologic and psychodermatologic care, their real-world implementation remains fraught with challenges. The presence of digital psychodermatology tools continues to be fragmented and sparse in clinical practice. Clinicians cite significant barriers related to workflow integration, including the time required for image capture and documentation; limited interoperability with electronic medical records; and medicolegal concerns surrounding data privacy, reimbursement models, and clinical liability [13,14]. From the patient perspective, digital literacy and technology acceptance remain key hurdles (particularly among older adults), whose comfort with telehealth interfaces is shaped by varying levels of support, familiarity, and personal values [15]. While digital tools such as teledermatology and online psychotherapeutic interventions have demonstrated improved access and high patient satisfaction, their clinical effectiveness remains inconsistent. Some models yield optimal outcomes only when used in conjunction with traditional in-person care [13,16]. Furthermore, many digital mental health interventions in dermatology, such as app-based CBT or stress management platforms, face high attrition rates and low long-term engagement, limiting their sustained impact on mental health outcomes [17]. These implementation and engagement challenges highlight the need for better-designed, clinically integrated, and user-centered digital solutions in psychodermatology.

Objectives and Hypotheses

This scoping review aims to systematically map, categorize, and synthesize the scientific literature on informatics-enabled psychotherapeutic and psychiatric interventions designed to improve both dermatologic conditions and their associated mental health outcomes. It also seeks to address key gaps identified in the existing literature. Specifically, the objectives are to: (1) chart the landscape of digital platforms (such as teledermatology, mHealth apps, AI-based decision-support tools, iCBT modules, mindfulness-based digital programs, and e-psychiatry consult-liaison services) used therapeutically in dermatologic care; (2) profile the dermatologic diagnoses most frequently targeted and characterize the demographic and clinical features of the patient populations involved; (3) synthesize reported clinical, psychological, and quality-of-life outcomes, including dermatologic severity indices, itch and pain scores, validated mental health scales, and patient-reported outcomes; (4) analyze implementation models, including modes of delivery, duration, provider roles, integration into clinical care pathways, and equity-related considerations (eg, digital literacy, socioeconomic factors, and geographic access); (5) assess the methodological quality of included studies using validated critical appraisal tools to identify strengths, limitations, and risks of bias; and (6) identify knowledge gaps and research priorities to guide the future development, evaluation, and equitable implementation of digital psychodermatologic interventions.

Methods

Search Strategies

To comprehensively capture the current state of research on informatics-based psychotherapeutic and psychiatric interventions in dermatology, we conducted a structured scoping review of studies published up to March 2025. Searches were performed across 5 major electronic databases: MEDLINE (via PubMed), Web of Science, Embase, PsycINFO (via Ovid), and Google Scholar. The review followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Checklist 1), which provides methodological guidance for mapping emerging fields with conceptual or methodological diversity [18].

The search strategy was developed iteratively, combining free-text keywords and controlled vocabulary (Medical Subject

XSL•FO
RenderX

Headings and Emtree terms) to ensure broad yet relevant coverage. Terms related to dermatologic diagnoses (eg, "psoriasis," "acne," "atopic dermatitis," "hidradenitis suppurativa") were combined with those describing digital modalities (eg, "teledermatology," "mobile health," "artificial intelligence," "digital therapeutics," "virtual assistants") and psychosocial or psychiatric constructs (eg, "depression," "anxiety," "psychotherapy," "CBT," and "mental health"). The objective was to capture literature at the intersection of digital tools and mental health outcomes in populations with dermatologic disorders. Search strategies were peer-reviewed internally and piloted for sensitivity before full execution. The scoping review was not registered.

All searches were conducted by the lead reviewer (CL), with results independently verified by a second reviewer (AH) to ensure consistency, accuracy, and completeness. No restrictions on study geography, clinical setting, or publication language were applied at the search stage, although inclusion criteria were applied during screening. Complete search algorithms for each database are provided in Multimedia Appendix 1.

Study Eligibility

Studies were eligible for inclusion if they met all of the following criteria: (1) the population under investigation comprised children, adolescents, or adults with a clinically diagnosed chronic dermatologic disorder (eg, psoriasis, AD, acne, HS, vitiligo, alopecia areata, chronic urticaria, or other specified inflammatory or pigmentary skin diseases); (2) the study evaluated a therapeutic approach that incorporated both an informatics component (eg, teledermatology, mobile or e-health apps, web-based platforms, AI-driven tools, virtual reality [VR], chatbot technologies) and a psychotherapeutic or psychiatric element aimed at improving mental health or dermatologic outcomes. Acceptable psychological components included, but were not limited to, CBT, mindfulness-based interventions, psychoeducation, liaison psychiatry, or psychotropic medication management. Studies were eligible if they reported at least 1 skin-related outcome (eg, disease severity scores, itch, pain, or DLQI) or 1 mental health outcome (eg, anxiety, depression, psychological distress, or health-related quality of life). Eligible study designs encompassed randomized and nonrandomized controlled trials, quasi-experimental studies, cohort or cross-sectional analyses, mixed methods evaluations, and implementation studies with quantitative or qualitative outcome reporting.

Studies were excluded if they focused exclusively on diagnostic tools, image recognition systems, or data-capture technologies without a therapeutic or psychiatric objective. Similarly, studies assessing only pharmacological, surgical, or cosmetic interventions without an integrated digital or psychological component were not considered. Single-case reports with fewer than 5 participants, editorials, opinion articles, conference abstracts without full text, narrative reviews, and protocol papers lacking original data were also excluded. Additionally, studies involving only healthy volunteers, individuals without a dermatologic diagnosis (eg, cosmetic users), or interventions targeting sun protection, skin cancer screening, or esthetic counseling were not included. Only articles published in English or French were considered for review; gray literature and unpublished manuscripts were excluded.

Data Extraction

A structured data extraction process was conducted using a predesigned coding framework in Microsoft Excel to ensure systematic and reproducible capture of study details. One reviewer (CL) performed the initial data extraction, and a second reviewer (AH) independently reviewed all entries to verify accuracy and consistency across studies. Discrepancies in interpretation or classification were resolved through discussion until consensus was reached.

Variables extracted encompassed bibliographic details (eg, author names, publication year, country of origin, and publication type), participant characteristics (eg, dermatologic diagnosis, sample size, age distribution, and gender), and intervention attributes (eg, program title, type of informatics platform, psychotherapeutic or psychiatric components, mode of delivery, and intervention duration). Study design parameters were also recorded, including methodological approach, setting, comparator groups, and factors related to health equity. Outcome measures included both clinical (eg, skin severity scores, itch or pain levels, DLQI) and psychological indicators (eg, symptoms of depression, anxiety, stress, or overall mental well-being), as well as implementation-related variables, such as adherence, usability, accessibility, and equity considerations. Where applicable, methodological quality and potential sources of bias were assessed using validated critical appraisal tools.

To enhance interpretability and align with digital health implementation frameworks, interventions were further categorized according to interaction modality and level of human involvement into three predefined groups: (1) *synchronous interventions*, defined as real-time interactions between patients and providers (eg, telepsychiatry, video consultations, live e-consults); (2) *asynchronous interventions*, defined as time-independent, user-initiated, or platform-delivered content without real-time clinician interaction (eg, mobile apps, web-based iCBT modules, educational platforms); and (3) *automated interventions*, defined as systems delivering therapeutic or support functions with minimal or no human involvement, often leveraging algorithmic or AI components (eg, chatbots, AI-assisted triage tools, VR environments). Classification was performed independently by 2 reviewers, and discrepancies were resolved by consensus.

Quality Assessment

The methodological quality of all included studies was assessed using design-specific critical appraisal tools developed by the Joanna Briggs Institute (JBI) [19]. Two reviewers (CL and AH) independently conducted the quality assessment, with any discrepancies resolved through consensus.

The appropriate JBI checklist was applied according to the study design: randomized controlled trials, quasi-experimental studies, cohort studies, cross-sectional analyses, case series (with ≥5 participants), and qualitative research were all appraised using the corresponding JBI instruments. For mixed methods studies, both quantitative and qualitative components were assessed separately using the relevant JBI tools.

XSL•FO
RenderX

Each checklist evaluated key domains of methodological rigor, including selection bias, measurement validity, confounding, and clarity of outcome reporting. Studies were not excluded based solely on quality ratings; however, appraisal results were documented to guide the interpretation of the strength and credibility of the evidence. Findings from the quality assessments are presented in tabular format in Multimedia Appendix 2 and summarized narratively to highlight common methodological strengths and limitations across the included literature.

Results

Description of Studies

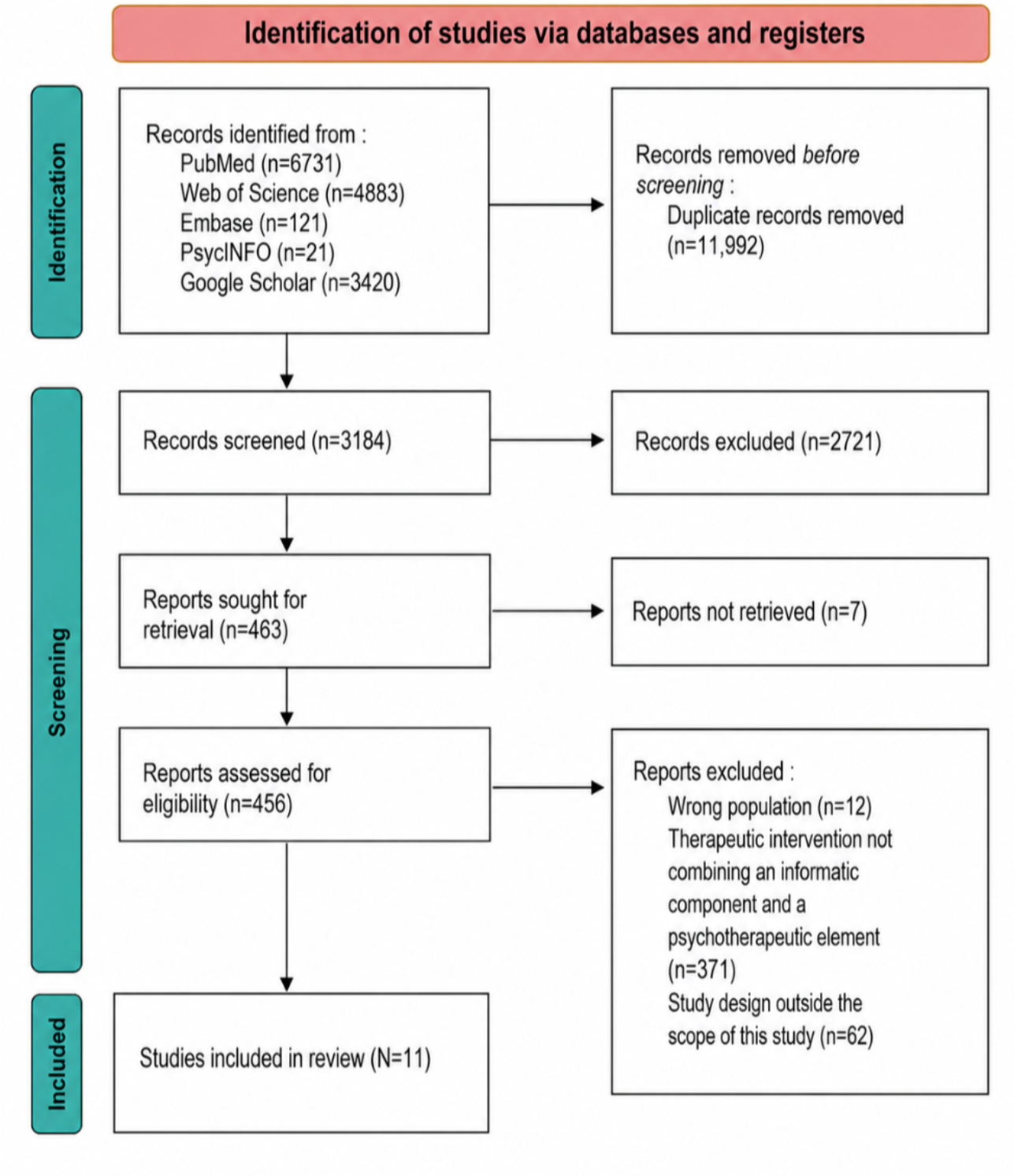
A total of 15,176 records were identified through electronic database searches. After removing 11,992 duplicates, 3184 unique titles and abstracts were screened for relevance. Of these, 2721 were excluded based on eligibility criteria during the initial title and abstract review. Seven additional full-text reports could not be retrieved, leaving 456 articles for full-text review.

During this stage, 445 reports were excluded: 12 due to an ineligible population, 371 for lacking a combined informatics and psychotherapeutic intervention, and 62 for using study designs outside the scope of the review.

Ultimately, 11 studies met all inclusion criteria and were included in the final synthesis. A detailed flowchart of the screening and selection process is presented in Figure 1, and key study characteristics are summarized in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart illustrating the inclusion of studies.



Main Outcomes

This scoping review synthesized findings from 11 peer-reviewed studies evaluating digital interventions with psychotherapeutic or psychiatric components aimed at improving both dermatologic and mental health outcomes. Most studies focused on chronic inflammatory skin diseases, especially psoriasis (n=9), with fewer addressing AD, HS, or vitiligo.

Digital platforms included iCBT, mindfulness-based self-help modules, mobile apps (eg, *DermaScope*, *Allay*), VR interfaces, and AI-assisted screening tools. Psychotherapeutic elements comprised guided iCBT, compassion-focused therapy, acceptance-based therapy, and self-guided mindfulness recordings, while psychiatric components involved consult-liaison psychiatry or psychotropic medication management within digitally integrated pathways.

Many studies reported statistically significant improvements in both dermatologic severity (eg, Psoriasis Area and Severity Index [PASI] scores, itch scales, DLQI) and mental health outcomes (eg, depression, anxiety, illness perceptions), though

XSL•FO
RenderX

effect sizes and adherence rates varied. Dropout rates ranged from 10% to 46%, with generally high acceptability among users who completed the interventions. Most interventions were delivered as adjuncts to standard dermatologic care, but only a minority were integrated within structured care pathways.

Risk of bias was moderate in most studies, with common limitations including small sample sizes, nonblinded designs, and short follow-up periods. Despite these constraints, the evidence suggests that digital psychodermatologic interventions are both feasible and potentially beneficial, warranting further development and refinement. The main outcomes are summarized in Table 1.

Table . Main outcomes across the 11 identified peer-reviewed studies.

Study, year	Population with condition	Intervention summary	Key outcomes	Interaction modality
van Beugen et al, 2016 [20]	Adults with psoriasis	Tailored therapist-guided (iCBT[a]) targeting coping and disease adjustment	Improved physical functioning and daily life impact; mixed psychological effects	Hybrid (asynchronous + synchronous)
Hedman-Lagerlöf et al, 2021 [21]	Adults with AD[b]	Therapist-guided iCBT (12 wk)	Reduced AD severity, itch, stress, depression; sustained at 12 months	Hybrid (asynchronous + synchronous)
Kern et al, 2023 [22]	Adults with AD	Self-guided digital CBT[c]-based intervention (8 wk)	Improvements in QoL[d], itch, depression, stress; feasibility demonstrated	Asynchronous
Kern et al, 2025 [12]	Adults with AD	Self-guided vs clinician-guided online CBT (RCT[e])	Self-guided noninferior to clinician-guided; reduced symptoms (POEM[f])	Hybrid (asynchronous + synchronous)
Fortune et al, 2022 [23]	Adults with psoriasis	Mobile app (*Allay*) delivering digital therapeutic content (CBT-informed)	Improved QoL, resilience, depression, illness perceptions	Asynchronous
Domogalla et al, 2021 [24]	Adults with psoriasis	Smartphone app + educational program integrated into care	Reduced depressive symptoms and QoL improvements (long-term)	Hybrid (asynchronous + synchronous)
Muftin et al, 2022 [25]	Adults with psoriasis	Online compassion-based and mindfulness-based self-help (RCT)	Reduced shame; improved QoL	Asynchronous
Clarke et al, 2024 [26]	Adults with skin conditions and depression	Online compassion-based guided self-help intervention	Improved depression, QoL, self-compassion; feasibility established	Hybrid (asynchronous + synchronous)
Sengupta and Wagani, 2025 [27]	Adults with various skin conditions	Online mindful self-compassion intervention (structured sessions)	Reduced depression, anxiety, stress; improved QoL and well-being	Synchronous (scheduled sessions)
Zhou et al, 2024 [28]	Adults with psoriasis	Online MBCT[g] + usual care	Improved PASI[h], depression, QoL, itch	Asynchronous
Leibovici et al, 2009 [29]	Adults with pruritus (AD/psoriasis)	VR[i] immersion and audiovisual distraction	Reduced pruritus intensity (short-term)	Automated

[a]iCBT: internet-based cognitive behavioral therapy.

[b]AD: atopic dermatitis.

[c]CBT: cognitive behavioral therapy.

[d]QoL: quality of life.

[e]RCT: randomized controlled trial.

[f]POEM: Patient-Oriented Eczema Measure.

[g]MBCT: mindfulness-based cognitive therapy.

[h]PASI: Psoriasis Area and Severity Index.

[i]VR: virtual reality.

Landscape of Digital Interventions

The digital interventions identified in this review encompassed a broad range of informatics modalities designed to support psychotherapeutic or psychiatric care in populations with dermatologic disorders.

Web-based iCBT was the most extensively studied approach. In the trial by van Beugen et al [20], a 12-week secure online platform delivered guided modules targeting stress,

XSL•FO
RenderX

dysfunctional beliefs, and disease coping in patients with psoriasis, producing moderate improvements in quality of life and psychological flexibility compared to care-as-usual. Similarly, Fortune et al [23] embedded a therapist-supported iCBT program within the *Allay* mobile app, providing structured psychoeducation and skill-building for emotional distress in psoriasis, with early evidence of improved illness perceptions and emotional adjustment.

Compassion and mindfulness-based audio programs were examined in 2 studies. They were delivered via MP3 or digital links. Muftin et al [25] compared self-directed compassion-focused therapy with mindfulness-based therapy for psoriasis, with both interventions showing moderate reductions in distress and improved skin-related attitudes. Clarke et al [26] evaluated a guided audio intervention using kindness and acceptance-based recordings for patients with AD and HS, noting favorable user engagement and emotional improvements.

mHealth apps such as *DermaScope* and *Allay* were also common platforms [23,24]. *DermaScope* offered educational content and emotional support prompts, achieving high user satisfaction and reducing depressive symptoms in a German psoriasis cohort. *Allay*, implemented in the United States and Ireland, enabled daily symptom monitoring, psychoeducational content, and therapist communication, with most users requesting continued access after the study.

AI-driven tools were more exploratory. Zhou et al [28] developed a machine learning model to predict psychiatric comorbidity risk in HS, demonstrating the feasibility of AI-assisted mental health triage within dermatology.

Similarly, VR was used in Leibovici et al [29] as a distraction technique for patients experiencing severe itching, with immersive 3D environments reducing perceived discomfort during acute episodes.

Emerging modalities included Sengupta et al [27] conversational chatbot integrated with skin-tracking functions to provide CBT-based support to adolescents with vitiligo, which showed positive user attitudes and symptom control over time. Kern et al [12,22] piloted a novel e-psychiatry consult-liaison model for dermatology referrals, enabling synchronous psychiatric evaluation via teleplatforms, thereby improving continuity of care for unmet psychiatric needs.

Target Populations and Diagnoses

Psoriasis was the primary focus in most included studies, with 9 out of 11 trials targeting individuals with moderate-to-severe plaque psoriasis. Participants were typically adults aged from 35 to 53 years, with a predominance of female respondents and relatively high levels of education. Several protocols assumed or prescreened for digital literacy, reflecting a bias toward digitally engaged populations. Few studies reported detailed breakdowns of socioeconomic status, race, or rural versus urban residence, limiting the generalizability to marginalized groups.

HS was the sole focus in Zhou et al [28], which applied AI-based stratification to improve psychiatric detection and triage. Sengupta et al [27] included a broader diagnostic mix, engaging adolescents and young adults with vitiligo, AD, and other pigmentary disorders via a chatbot-enabled support system. AD and chronic itch were secondarily mentioned in Clarke et al [26] and Leibovici et al [29], respectively, though without significant subgroup analyses.

No study targeted pediatric or geriatric patients exclusively, and few interventions were adapted for linguistic or cognitive accessibility. Clinical samples were largely restricted to mild-to-moderate symptom severity, and most trials excluded patients with active suicidality, substance use disorders, or psychosis. This suggests a significant gap in psychodermatologic research involving more complex or underserved patient populations.

Outcomes

All included studies assessed at least 1 validated mental health outcome, most commonly using the Hospital Anxiety and Depression Scale, Patient Health Questionnaire-9, Brief Illness Perception Questionnaire, and visual analog scales for stress and emotional burden. Dermatologic outcomes were also consistently reported, with 8 studies using standardized measures such as the PASI, the Dermatology Life Quality Index (DLQI), and Numeric Rating Scale for itch and pain.

Several studies have demonstrated statistically significant, though modest, reductions in emotional distress. Domogalla et al [24] reported a mean 3.2-point decrease on the Patient Health Questionnaire-9 following a mobile app intervention, while van Beugen et al [20] found reduced self-reported helplessness and avoidance behavior using a digital CBT protocol. Fortune et al [23] observed improvements in illness beliefs (particularly in perceived control and emotional representation) following therapist-supported iCBT via the *Allay* app.

Multicomponent programs incorporating mindfulness or compassion elements yielded improvements not only in depression and anxiety scores but also in broader functional outcomes, such as sleep quality, social participation, and perceived stigmatization [25,26]. DLQI scores improved in 6 studies, with reductions of 3 to 7 points from baseline, indicating clinically meaningful gains in daily functioning and disease coping.

Long-term effects were less certain, as few studies included follow-up beyond 8 to 12 weeks. Notable exceptions included Hedman-Lagerlöf et al [21] who conducted a randomized controlled trial of exposure-based iCBT for AD, demonstrating sustained reductions in skin-related avoidance, anxiety, and depressive symptoms at a 12-month follow-up.

The study by Zhou et al [28] was the only one to focus on predicting psychiatric morbidity algorithmically, rather than evaluating therapeutic outcomes directly, underscoring a gap between predictive and interventional research.

Overall, while effect sizes varied, the evidence consistently pointed to positive impacts on both dermatologic and psychological well-being when digital tools incorporated structured therapeutic content. A summary of the findings is presented in Table 2.

XSL•FO
RenderX

Table . Summary of the findings.

Category	Summary
Most studied condition	Psoriasis (n=9 studies), with some inclusion of HS[a], AD[b], and vitiligo
Age range of participants	Adults aged 18 - 65 y; mean range across studies: 35 - 53 y
Gender representation	Majority female, digitally literate, higher education level
Digital modalities used	Web platforms, mobile apps, chatbots, VR[c], AI[d] triage tools, e-consults
Therapeutic approaches	CBT[e], ACT[f], mindfulness, compassion-based therapy, liaison psychiatry
Mental health outcomes measured	HADS[g], PHQ-9[h], B-IPQ[i], stress and mood scales, QoL[j] indicators
Dermatologic outcomes measured	PASI[k], DLQI[l], itch or pain scales, visual symptom tracking
Key mental health results	Small-to-moderate reductions in anxiety, depression, and distress; improved illness beliefs
Key dermatologic results	DLQI improved in 6 studies; itch and pain modestly reduced in short-term
Duration of interventions	From 10-min VR to 8- to 12-wk programs; mostly 4 - 8 wk
Provider involvement	Ranged from self-guided applications to full psychiatric consults; minimal hybrid models
Barriers identified	Digital fatigue, limited literacy, emotional burden of disease reminders
Facilitators identified	Autonomy, ease of access, therapist support, personalization of content
Equity considerations	Minimal focus on older adults, underserved groups, or multilingual adaptation
Gaps in research	Lack of long-term outcomes, economic data, implementation strategies, equity, and digital alliance measures

[a]HS: hidradenitis suppurativa.
[b]AD: atopic dermatitis.
[c]VR: virtual reality.
[d]AI: artificial intelligence.
[e]CBT: cognitive behavioral therapy.
[f]ACT: acceptance and commitment therapy.
[g]HADS: Hospital Anxiety and Depression Scale.
[h]PHQ-9: Patient Health Questionnaire-9.
[i]B-IPQ: Brief Illness Perception Questionnaire.
[j]QoL: quality of life.
[k]PASI: Psoriasis Area and Severity Index.
[l]DLQI: Dermatology Life Quality Index.

Implementation and Sustainability

The majority of interventions were designed as adjunctive supports rather than standalone treatments, aiming to complement rather than replace dermatologic care. Delivery was predominantly asynchronous, allowing users to access modules, audio content, or app-based features on their own schedules without real-time clinician involvement. Intervention duration varied widely—from a single 10-minute VR session for itch relief to 8- to 12-week therapist-supported iCBT or self-guided compassion-based therapy protocols [20,25,26,29]. Clinical involvement ranged from fully automated, self-guided platforms (eg, Sengupta 2025 chatbot, Domogalla mobile app) to integrated psychiatry consultations delivered via telehealth, as seen in Kern et al [12], where dermatologists referred patients to e-psychiatry services for diagnostic clarification and psychotropic management [22].

Reported barriers to implementation included digital fatigue (with daily app reminders or disease-focused content perceived as burdensome) [23,25], technological literacy challenges, and data privacy concerns, particularly among older or digitally underserved populations [28]. Dropout rates ranged from 10% to 46%, with reasons for attrition often unreported or multifactorial. Conversely, facilitators of engagement included content personalization, 24/7 accessibility, and a sense of autonomy and empowerment, especially in self-help platforms [24,26].

Despite these insights, few studies examined long-term sustainability, cost-effectiveness, or integration into routine dermatology workflows. Only Kern et al [12,22] assessed provider workflow impact, and no study conducted formal equity analyses by race, income, rurality, or language. Overall, while the evidence supports feasibility, substantial gaps remain in

XSL•FO
RenderX

implementation science and health equity evaluation for digital psychodermatology.

Guidance Level and Adherence

Across the included studies, interventions varied substantially in the degree of human support provided, ranging from fully self-guided (unguided) digital interventions to clinician- or therapist-supported (guided) models, with several hybrid approaches combining asynchronous content with limited human interaction. Guided interventions, including therapist-supported iCBT and structured clinician feedback embedded within digital platforms, were reported in studies such as Hedman-Lagerlöf et al [21] and van Beugen et al [20], where participants received periodic support, feedback, or monitoring alongside digital modules. In contrast, unguided interventions—including self-help mobile apps and fully automated online programs (eg, Fortune et al [23], Muftin et al [25], and Kern et al [12] self-guided arm)—relied primarily on user-driven engagement without real-time or personalized clinician input [20,21]. Hybrid models, such as those incorporating minimal email support or periodic check-ins (eg, [24,26]), occupied an intermediate position in terms of resource intensity. When examined in relation to adherence, a pattern emerged whereby studies reporting higher levels of human support tended to demonstrate lower attrition and more sustained engagement, whereas fully unguided interventions were more frequently associated with higher dropout rates [24,26]. For example, therapist-guided iCBT trials reported relatively stable completion rates despite minimal clinician time investment, while several self-guided interventions noted usability challenges or attrition over time. Overall dropout rates varied across studies with variability partly attributable to differences in intervention intensity, duration, and user support. Although the heterogeneity of study designs and reporting precludes formal quantitative comparison, these findings suggest a potential association between the degree of human support and intervention adherence, with guided or hybrid models appearing to mitigate disengagement relative to fully self-directed approaches.

Quality Appraisal

Using JBI appraisal tools, most studies demonstrated a moderate risk of bias. Randomized controlled designs were generally of higher methodological quality but remained susceptible to open-label effects and attrition [20,25]. Quasi-experimental and observational studies provided useful insights but were limited by small sample sizes and short follow-up periods [24,27]. Overall, quality ratings across studies ranged from moderate to high based on the JBI checklists. Detailed assessments are presented in Multimedia Appendix 2.

Knowledge Gaps

A number of critical knowledge gaps were identified across the reviewed literature. First, the demographic reach of interventions was limited: most samples comprised middle-aged, digitally literate adults, with minimal inclusion of older adults, individuals with low health or digital literacy, or patients from rural or socioeconomically disadvantaged backgrounds. No study provided stratified analyses by race, ethnicity, or income, and none were explicitly designed to address the needs of underserved populations.

While several interventions demonstrated short-term clinical or psychological benefits, none assessed long-term psychiatric outcomes, such as suicidal ideation, substance use disorders, or psychiatric hospitalization, despite the elevated prevalence of these conditions in chronic dermatologic illnesses. Furthermore, sustainability and cost-effectiveness were rarely addressed. Only Kern et al [12] investigated system-level integration into dermatologic workflows, and no study conducted a formal economic evaluation or budget impact analysis.

There was also a notable absence of key implementation metrics, such as intervention fidelity, training burden, and reimbursement feasibility, limiting the applicability of findings to real-world clinical settings. Although several interventions were delivered via mobile apps or web-based portals, blended care models that combine digital and in-person therapeutic components remain largely unexplored.

Additionally, the digital therapeutic alliance, a key factor in mental health treatment efficacy, was not systematically assessed in any study. Equity-focused implementation strategies, including culturally adapted content or multilingual delivery formats, were entirely absent.

Addressing these gaps is essential to ensure that psychodermatological digital tools are scalable, inclusive, and clinically meaningful across diverse patient populations and care systems.

Discussion

Principal Results

This scoping review synthesized findings from 11 studies evaluating informatics-enabled psychotherapeutic or psychiatric interventions in dermatology. Most interventions were delivered via digital platforms, including web-based cognitive CBT, mHealth apps, mindfulness audio modules, VR tools, and AI-supported systems. The observed effects of digital psychodermatologic interventions can be interpreted through the lens of the brain-skin axis, a bidirectional neuroimmunological pathway linking psychological stress and cutaneous inflammation. By reducing stress and improving coping and behavioral responses (eg, scratching), interventions such as iCBT may attenuate activation of the hypothalamic-pituitary-adrenal axis and downstream inflammatory processes, contributing to improvements in both mental health and skin severity (eg, PASI). This mechanistic framework supports the integration of digital mental health tools as adjunctive strategies in the management of chronic inflammatory skin diseases.

Psoriasis emerged as the most frequently studied dermatologic condition, while fewer studies addressed AD, HS, or mixed skin disorders. All studies assessed mental health outcomes, most commonly targeting depression, anxiety, stress, or illness perception. Several interventions demonstrated moderate improvements in emotional distress, self-efficacy, and

XSL•FO
RenderX

dermatologic quality of life, particularly those integrating CBT or self-compassion frameworks.

However, digital equity considerations, long-term adherence, and integration into routine dermatologic care were rarely assessed. Methodological quality was generally moderate to high, but many studies were underpowered, lacked control groups, or experienced high attrition.

Comparison With Prior Work

Psychological therapies have demonstrated measurable benefits in chronic skin conditions such as psoriasis [30]. A recent systematic review of randomized trials found that interventions such as cognitive behavioral therapy (CBT), mindfulness-based interventions, and structured patient education can improve both mental health outcomes and dermatologic severity in psoriasis [31]. Similarly, Qureshi et al [30] reported that adjunctive psychological programs (including stress management and biofeedback) were associated with better skin clearance and quality-of-life improvements in psoriatic disease. However, most interventions in these studies were delivered in person, with only a minority using digital or telehealth formats [31]. This gap highlights an underexplored opportunity in dermatology, given the logistical advantages of digital platforms for chronic care.

Our review highlights the unique value of informatics platforms (such as web portals and mobile apps) in psychodermatology. Digital tools can reduce access barriers and extend care beyond clinic visits through on-demand support and self-management resources. Unlike weekly face-to-face sessions, online programs can provide asynchronous, low-cost coaching that patients can access between dermatology appointments. Evidence from general mental health research suggests this convenience does not necessarily compromise efficacy: multiple meta-analyses have shown that guided iCBT produces outcomes equivalent to those of in-person therapy for anxiety and depression [32]. Andersson and colleagues [32] further demonstrated that across conditions, iCBT achieves symptom reductions comparable to face-to-face care, although direct comparisons remain limited.

Despite these findings, dermatology still lacks head-to-head trials explicitly comparing digital and traditional psychotherapeutic delivery. This research underscores the importance of formally evaluating whether internet-delivered psychotherapies can match the clinical outcomes of office-based care in populations with dermatologic disorders.

Digital CBT has also proven effective in other chronic medical conditions, such as improving depression in diabetes and reducing pain catastrophizing in arthritis [33]. However, adaptation of these interventions to dermatology-specific needs remains limited. Content must be tailored to address skin-related stressors, such as visible lesions, itch-scratch cycles, and social stigma. Yet, personalized psychodermatology programs remain rare, with few adapted to specific clinical phenotypes (eg, itch-dominant psoriasis or high-stigma AD), which reinforces the call for more personalized digital therapeutics. As Kazdin [34] has argued in his framework for mental health innovation, digital platforms offer a unique opportunity to personalize therapy content to individual patient profiles at scale, which is an important direction for future research.

A striking finding of this review is the disproportionate focus on psoriasis, which accounted for 9 of the 11 included studies, with comparatively minimal representation of other high-burden dermatologic conditions such as AD and HS. This imbalance likely reflects several structural drivers within the field. First, psoriasis benefits from a long-standing research infrastructure, including the widespread use of standardized and validated outcome measures such as the PASI and the DLQI, which facilitate trial design, comparability, and regulatory alignment. In contrast, while AD has measures such as the Patient-Oriented Eczema Measure and Eczema Area and Severity Index, and HS has emerging indices such as Hidradenitis Suppurativa Clinical Response, these tools have been less consistently integrated into psychodermatology and digital intervention research, potentially limiting study development and cross-study synthesis. Second, psoriasis has been a primary target of substantial pharmaceutical investment, particularly with the expansion of the biologics market, which has historically driven broader innovation ecosystems, including adjunctive behavioral and digital interventions. This contrasts with conditions such as HS, which, despite significant psychosocial burden and high rates of psychiatric comorbidity, has received comparatively less research funding and fewer large-scale intervention studies. Third, psoriasis populations are often easier to recruit within structured clinical pathways and registries, enabling digital trial implementation, whereas patients with other conditions may be more fragmented across care settings. The consequence is a literature that may overrepresent conditions with strong biomedical and commercial ecosystems while underrepresenting those with equally or greater psychosocial need. This raises important concerns regarding the generalizability of current findings and highlights a critical gap in the development of digital psychodermatologic interventions for underserved patient populations. Future research should prioritize conditions such as HS and AD not only to ensure equity in innovation but also to better align digital intervention development with the true distribution of psychological burden across dermatologic diseases.

The present findings should be interpreted in light of the mixed methods systematic review by Hewitt et al [35], which represents the most closely related synthesis of digital psychological interventions in dermatology and identified 23 studies, including 15 randomized controlled trials, concluding that such interventions may improve psychological outcomes such as mood, quality of life, and illness-related knowledge despite substantial heterogeneity limiting definitive conclusions. While both reviews examine the intersection of digital technologies and psychological support in populations with dermatologic disorders, this study extends and refines this evidence base through several key distinctions. Our inclusion criteria were deliberately more restrictive, requiring interventions to integrate both an informatics component and a structured psychotherapeutic or psychiatric element, and to report dermatologic and/or mental health outcomes; this approach excludes purely educational or communication-based tools without explicit therapeutic intent, thereby enhancing

XSL•FO
RenderX

clinical interpretability despite yielding a smaller number of included studies (n=11). The review also extends the literature temporally and conceptually by incorporating studies published up to March 2025, capturing the emergence of newer modalities such as AI-assisted tools, virtual assistants, and hybrid care models that were largely absent from earlier syntheses and signal a shift toward more interactive, personalized, and scalable psychodermatologic interventions. In contrast to prior work, this synthesis explicitly treats dermatologic and psychological outcomes as coprimary domains, integrating clinical severity indices (eg, PASI, DLQI) with validated mental health measures to adopt a more comprehensive biopsychosocial perspective that reflects the bidirectional relationship between skin disease and psychological distress. Additionally, greater emphasis is placed on implementation considerations, including adherence, integration into clinical workflows, and equity-related barriers; although prior work identified issues of acceptability and usability, this analysis further examines how such interventions are operationalized in practice and highlights persistent gaps, particularly regarding long-term engagement and access among underserved populations. Taken together, this review does not contradict prior findings but rather refines and extends them through a more clinically focused definition of intervention, the inclusion of emerging technologies, and the positioning of outcomes within an integrated dermatologic-psychiatric framework, thereby contributing to the advancement of scalable and clinically embedded models of psychodermatologic care.

Limitations

This scoping review has several limitations. Despite a comprehensive multidatabase search strategy, relevant studies may have been missed, particularly those published in nonindexed journals, in languages other than English or French, or under alternative terminologies not captured by our search terms.

The heterogeneity in study designs, populations, interventions, and outcomes limited our ability to perform quantitative synthesis or direct cross-study comparisons. Many included studies had small sample sizes, high attrition rates, and limited follow-up, reducing the generalizability and robustness of their findings. Most trials also lacked standardized measures of adherence, usability, or cost-effectiveness, which are essential for assessing scalability. Also, a key finding of this review is the widespread lack of reporting on sociodemographic variables such as race or ethnicity, education, and socioeconomic status, despite frequent references to barriers like digital literacy. This gap limits the interpretability and generalizability of findings and raises concerns that current evidence may disproportionately reflect more digitally literate and advantaged populations, underscoring the need for standardized reporting and more inclusive study designs.

This review did not systematically include gray literature, which may have excluded unpublished or program-level innovations from clinical settings. Finally, although we aimed to capture both dermatological and mental health outcomes, few studies reported comprehensively on both, and even fewer examined long-term effects or implementation into routine dermatological care.

Conclusions

This scoping review provides a comprehensive synthesis of the emerging field of informatics-based psychotherapeutic and psychiatric interventions in dermatology. The reviewed studies highlight the feasibility and preliminary efficacy of digital tools in addressing both psychological distress and dermatologic symptom burden across a range of chronic skin conditions. These interventions offer accessible and scalable, patient-centered models of care that can complement traditional dermatology services, particularly in resource-limited or remote settings.

However, this evidence base is constrained by methodological heterogeneity, small sample sizes, high attrition, and insufficient reporting on equity, usability, and long-term impact. The lack of integration into dermatologic workflows and the absence of head-to-head comparative studies with face-to-face psychotherapy also represent additional knowledge gaps.

Future research should prioritize designing trials that evaluate digital interventions in diverse populations, incorporate equity-informed frameworks, and test hybrid models of care that blend digital tools with human support. As the fields of psychodermatology and digital health continue to converge, the development of personalized, adaptable, and clinically integrated digital mental health interventions will be essential to improving patient outcomes and reducing disparities in dermatologic care.

Funding

This study received indirect funded from the La Fondation de l’Institut universitaire en santé mentale de Montréal and operating funds from IVADO for AH. We did not use generative AI tools for this manuscript.

Authors' Contributions

Conceptualization: CL, AH.
Formal analysis: CL, AH.
Methodology and validation: CL, AH.
Funding acquisition: AH.
Investigation: AH.
Project administration: AH.
Supervision: AH, JC.

XSL•FO
RenderX

Resources: AH.
Data curation: CL, AH, JC.
Writing — original draft: CL, AH, JC.
Writing — review and editing: CL, AH, JC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Electronic search strategy for the scoping review conducted.
[DOCX File, 17 KB - derma_v9i1e82096_app1.docx]

Multimedia Appendix 2
Systematic review study selection detailed results.
[DOCX File, 33 KB - derma_v9i1e82096_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 111 KB - derma_v9i1e82096_app3.docx]

References

1. Dalgard FJ, Gieler U, Tomas-Aragones L, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol 2015 Apr;135(4):984-991. [doi: 10.1038/jid.2014.530] [Medline: 25521458]
2. Henderson AD, Adesanya E, Mulick A, et al. Common mental health disorders in adults with inflammatory skin conditions: nationwide population–based matched cohort studies in the UK. BMC Med 2023 Aug 4;21(1):285. [doi: 10.1186/s12916-023-02948-x] [Medline: 37542272]
3. Caccavale S, Tancredi V, Boccellino MP, Babino G, Fulgione E, Argenziano G. Hidradenitis suppurativa burdens on mental health: a literature review of associated psychiatric disorders and their pathogenesis. Life (Basel) 2023 Jan 9;13(1):189. [doi: 10.3390/life13010189] [Medline: 36676137]
4. Misery L, Chesnais M, Merhand S, et al. Perceived stress in four inflammatory skin diseases: an analysis of data taken from 7273 adult subjects with acne, atopic dermatitis, psoriasis or hidradenitis suppurativa. J Eur Acad Dermatol Venereol 2022 Aug;36(8):e623-e626. [doi: 10.1111/jdv.18016] [Medline: 35181931]
5. Mostaghimi L, Noughani H. “Psychodermatology”: the present standing and a path forward. J Eur Acad Dermatol Venereol 2022 Aug;36(8):1162-1163. [doi: 10.1111/jdv.18278] [Medline: 35848144]
6. Christensen RE, Jafferany M. Psychiatric and psychologic aspects of chronic skin diseases. Clin Dermatol 2023;41(1):75-81. [doi: 10.1016/j.clindermatol.2023.03.006] [Medline: 36878453]
7. Chow A, Smith HE, Car LT, et al. Teledermatology: an evidence map of systematic reviews. Syst Rev 2024 Oct 12;13(1):258. [doi: 10.1186/s13643-024-02655-5] [Medline: 39396040]
8. Seghers AC, Seng KH, Chio MTW, Chia E, Ng SK, Tang MBY. A prospective study on the use of teledermatology in psychiatric patients with chronic skin diseases. Australas J Dermatol 2015 Aug;56(3):170-174. [doi: 10.1111/ajd.12297] [Medline: 25754857]
9. Kaliyadan F, Ashique KT. Use of mobile applications in dermatology. Indian J Dermatol 2020;65(5):371-376. [doi: 10.4103/ijd.IJD_422_20] [Medline: 33165393]
10. Cherrez-Ojeda I, Robles-Velasco K, Osorio MF, et al. A systematic review and meta-analysis of mobile health applications and telemonitoring in atopic dermatitis self-management. Dermatol Ther (Heidelb) 2024 Jul;14(7):1787-1798. [doi: 10.1007/s13555-024-01213-0] [Medline: 38909171]
11. Roca S, Almenara M, Gilaberte Y, et al. When virtual assistants meet teledermatology: validation of a virtual assistant to improve the quality of life of psoriatic patients. Int J Environ Res Public Health 2022 Nov 5;19(21):14527. [doi: 10.3390/ijerph192114527] [Medline: 36361408]
12. Kern D, Ljótsson B, Lönndahl L, et al. Self-guided vs clinician-guided online cognitive behavioral therapy for atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2025 Feb 1;161(2):183-190. [doi: 10.1001/jamadermatol.2024.5044] [Medline: 39693097]
13. Peracca SB, Fonseca A, Hines A, et al. Implementation of mobile teledermatology: challenges and opportunities. Telemed J E Health 2021 Dec;27(12):1416-1422. [doi: 10.1089/tmj.2020.0500] [Medline: 33691074]
14. Dovigi E, Kwok EYL, English 3rd JC. A framework-driven systematic review of the barriers and facilitators to teledermatology implementation. Curr Dermatol Rep 2020;9(4):353-361. [doi: 10.1007/s13671-020-00323-0] [Medline: 33200042]

XSL•FO
RenderX

15. Johnson A, Shukla N, Halley M, et al. Barriers and facilitators to mobile health and active surveillance use among older adults with skin disease. Health Expect 2021 Oct;24(5):1582-1592. [doi: 10.1111/hex.13229] [Medline: 34190397]
16. Maddukuri S, Patel J, Lipoff JB. Teledermatology addressing disparities in health care access: a review. Curr Dermatol Rep 2021;10(2):40-47. [doi: 10.1007/s13671-021-00329-2] [Medline: 33747638]
17. Ghani M, Adler C, Yeung H. Patient factors associated with interest in teledermatology: cross-sectional survey. JMIR Dermatol 2021 May 10;4(1):e21555. [doi: 10.2196/21555] [Medline: 37625162]
18. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
19. Munn Z, Barker TH, Moola S, et al. Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. JBI Evid Synth 2020 Oct;18(10):2127-2133. [doi: 10.11124/JBISRIR-D-19-00099] [Medline: 33038125]
20. van Beugen S, Ferwerda M, Spillekom-van Koulil S, et al. Tailored therapist-guided internet-based cognitive behavioral treatment for psoriasis: a randomized controlled trial. Psychother Psychosom 2016;85(5):297-307. [doi: 10.1159/000447267] [Medline: 27508937]
21. Hedman-Lagerlöf E, Fust J, Axelsson E, et al. Internet-delivered cognitive behavior therapy for atopic dermatitis: a randomized clinical trial. JAMA Dermatol 2021 Jul 1;157(7):796-804. [doi: 10.1001/jamadermatol.2021.1450] [Medline: 34009282]
22. Kern D, Ljótsson B, Lönndahl L, et al. A digital self-help intervention for atopic dermatitis: analysis of secondary outcomes from a feasibility study. JMIR Dermatol 2023 Mar 20;6:e42360. [doi: 10.2196/42360] [Medline: 37632924]
23. Fortune DG, Ravnkilde V, Ryan S, Ramsay B, Clough S, Richards HL. A digital therapeutic for management of psychosocial aspects of psoriasis: a pre-post proof of concept study. Skin Health Dis 2022 Jun;2(2):e103. [doi: 10.1002/ski2.103] [Medline: 35677915]
24. Domogalla L, Beck A, Schulze-Hagen T, Herr R, Benecke J, Schmieder A. Impact of an eHealth smartphone app on the mental health of patients with psoriasis: prospective randomized controlled intervention study. JMIR mHealth uHealth 2021 Oct 25;9(10):e28149. [doi: 10.2196/28149] [Medline: 34431478]
25. Muftin Z, Gilbert P, Thompson AR. A randomized controlled feasibility trial of online compassion-focused self-help for psoriasis. Br J Dermatol 2022 Jun;186(6):955-962. [doi: 10.1111/bjd.21020] [Medline: 35041766]
26. Clarke EN, Norman P, Thompson AR. Online compassion-based self-help for depression in people with skin conditions: a feasibility study. Pilot Feasibility Stud 2024 Apr 16;10(1):63. [doi: 10.1186/s40814-024-01486-4] [Medline: 38627850]
27. Sengupta A, Wagani R. Mindful self-compassion for psychological distress associated with skin conditions: an online intervention study. Indian J Dermatol Venereol Leprol 2025;91(1):71-75. [doi: 10.25259/IJDVL_451_2023] [Medline: 38314971]
28. Zhou Y, Zhao Y, Li Y, et al. Online mindfulness-based cognitive therapy as a treatment for patients with psoriasis: a randomized controlled trial. SSRN. Preprint posted online on Aug 30, 2024. [doi: 10.2139/ssrn.4940246]
29. Leibovici V, Magora F, Cohen S, Ingber A. Effects of virtual reality immersion and audiovisual distraction techniques for patients with pruritus. Pain Res Manag 2009;14(4):283-286. [doi: 10.1155/2009/178751] [Medline: 19714267]
30. Qureshi AA, Awosika O, Baruffi F, Rengifo-Pardo M, Ehrlich A. Psychological therapies in management of psoriatic skin disease: a systematic review. Am J Clin Dermatol 2019 Oct;20(5):607-624. [doi: 10.1007/s40257-019-00437-7] [Medline: 30937679]
31. Xiao Y, Zhang X, Luo D, et al. The efficacy of psychological interventions on psoriasis treatment: a systematic review and meta-analysis of randomized controlled trials. Psychol Res Behav Manag 2019;12:97-106. [doi: 10.2147/PRBM.S195181] [Medline: 30799963]
32. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295. [doi: 10.1002/wps.20151] [Medline: 25273302]
33. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
34. Kazdin AE. Annual research review: expanding mental health services through novel models of intervention delivery. J Child Psychol Psychiatry 2019 Apr;60(4):455-472. [doi: 10.1111/jcpp.12937] [Medline: 29900543]
35. Hewitt RM, Ploszajski M, Purcell C, et al. A mixed methods systematic review of digital interventions to support the psychological health and well-being of people living with dermatological conditions. Front Med (Lausanne) 2022;9:1024879. [doi: 10.3389/fmed.2022.1024879] [Medline: 36405626]

Abbreviations

AD: atopic dermatitis
AI: artificial intelligence
CBT: cognitive behavioral therapy
DLQI: Dermatology Life Quality Index
HS: hidradenitis suppurativa

XSL•FO
RenderX

iCBT: internet-based cognitive behavioral therapy
JBI: Joanna Briggs Institute
PASI: Psoriasis Area and Severity Index
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
VR: virtual reality
mHealth: mobile health

Edited by R Dellavalle; submitted 08.Aug.2025; peer-reviewed by H Sadeghsalehi; revised version received 22.May.2026; accepted 25.May.2026; published 10.Jun.2026.

Please cite as:
Lamarre C, Chivinski J, Hudon A
Informatics-Based Psychotherapeutic and Psychiatric Interventions in Dermatology: Scoping Review of Impacts on Skin Disease Severity and Mental Health Outcomes
JMIR Dermatol 2026;9:e82096
URL: https://derma.jmir.org/2026/1/e82096
doi:10.2196/82096

© Caroline Lamarre, Jeffrey Chivinski, Alexandre Hudon. Originally published in JMIR Dermatology (http://derma.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Interest in Janus Kinase (JAK) Inhibitors for Alopecia Areata: A Google Trend Analysis

Jade Howard[1], MD; Nourine A H Kamili[2], PhD; Hala Idris[2], BS; Loren D Krueger[2], MD

[1]Transitional Year Residency Program and BayCare Health System (Winter Haven), College of Medicine, Florida State University, 1201 1st St S, Suite 100A, Winter Haven, FL, United States

[2]Department of Dermatology, School of Medicine, Emory University, Atlanta, GA, United States

Corresponding Author:
Jade Howard, MD
Transitional Year Residency Program and BayCare Health System (Winter Haven), College of Medicine, Florida State University, 1201 1st St S, Suite 100A, Winter Haven, FL, United States

Abstract

Public interest in Janus kinase (JAK) inhibitors for alopecia areata increased following media coverage and the 2022 US Food and Drug Administration (FDA) approval of baricitinib, highlighting the need for patient education and physician guidance on appropriate indications and treatment selection for hair loss disorders.

(JMIR Dermatol 2026;9:e75119) doi:10.2196/75119

KEYWORDS

alopecia areata; hair loss; Janus kinase inhibitors; Google Trends; alopecia areata treatment; hair loss treatment

Introduction

Janus kinase (JAK) inhibitors are a class of medications that work by targeting and inhibiting JAK enzymes, which play a significant and diverse role in the immune system. The JAK system has been implicated in a variety of immune pathways including inflammation and autoimmunity [1]. On June 13, 2022, the US Food and Drug Administration (FDA) approved the use of a JAK inhibitor, baricitinib, commercially known as Olumiant, for adults with severe alopecia areata, an autoimmune form of hair loss [2]. Additionally, this medication was also approved for the use of severe alopecia areata in children aged 12 years and over on June 26, 2023. The objective of this study was to assess the impact of the FDA approval of the JAK inhibitor, baricitinib (Olumiant), for severe alopecia areata on public interest in JAK inhibitors for alopecia areata.

Methods

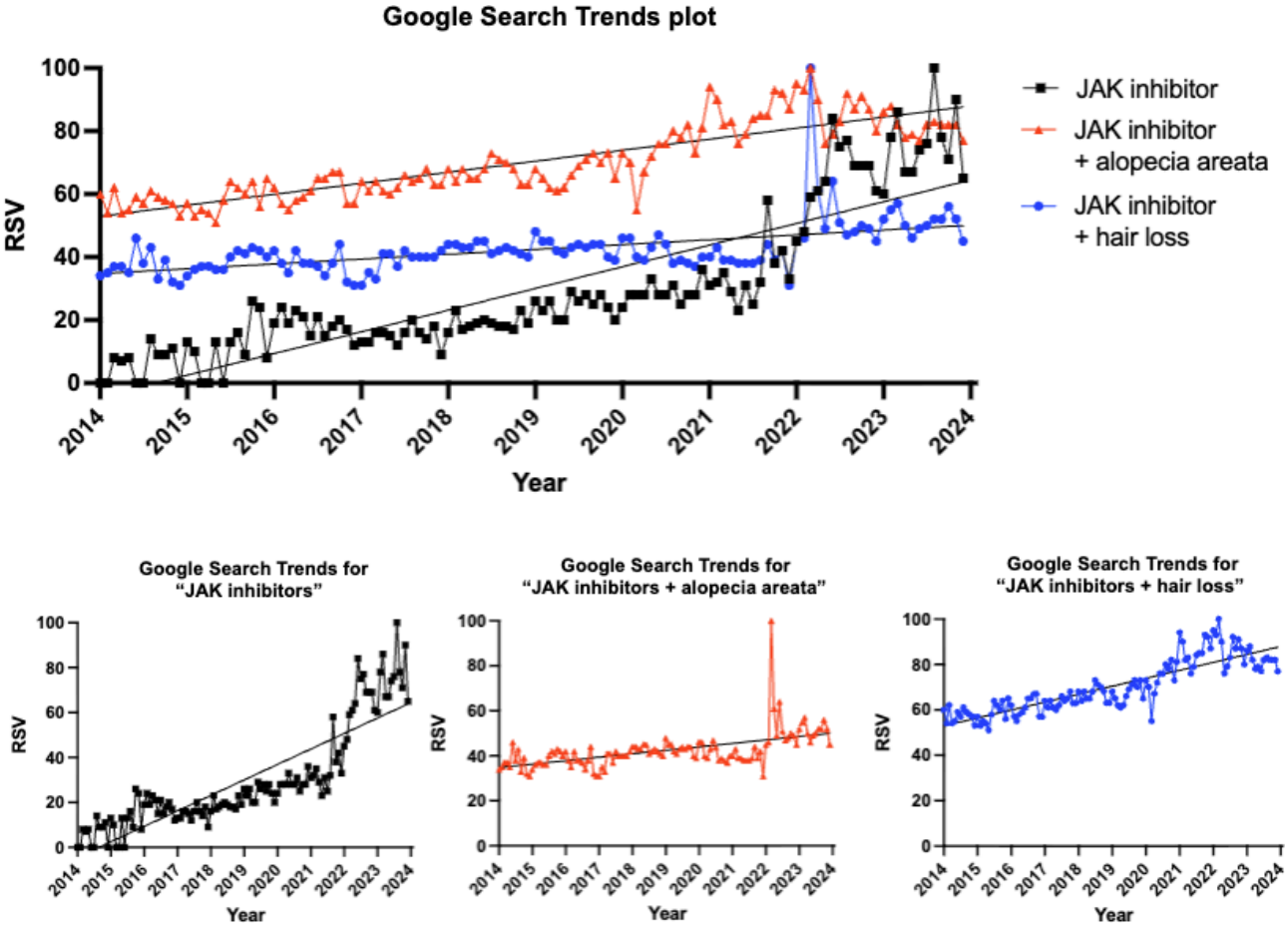
To evaluate public interest, we used Google Trends, a free analysis tool that provides insight into public interest for search terms. Google Trends calculates a relative search volume (RSV), ranging from 0 to 100. An RSV value of 0 indicates minimal interest in a keyword and a value of 100 represents the peak interest within a given period [3]. For our study, we searched the following keywords from January 2014 to January 2024: “JAK inhibitor,” “JAK inhibitor + hair loss,” and “JAK inhibitor + alopecia areata.” These keywords were chosen to encompass generic and specific terms related to the topic of JAK inhibitors and hair loss.

Results

A time series analysis was used to assess changes in RSV between January 2014 and December 2023. Linear regression demonstrated significant upward trends in RSV for all 3 search terms (Figure 1). Search interest for “JAK inhibitor” increased at a rate of 6.89 (95% CI 6.16 - 7.62; P<.001), RSV units per month with a strong model fit (R^2=0.748). Searches for “JAK inhibitor+ hair loss” also increased significantly, at a rate of 3.50 (95% CI 3.13 - 3.87; P<.001; R^2=0.749) RSV units per month. Similarly, “JAK inhibitor+ alopecia areata” demonstrated a statistically significant upward trend, though with a smaller magnitude of increase (1.54, 95% CI 1.12 - 1.96 RSV units per month; P<.001) and a more modest model fit (R^2=0.306).

XSL•FO
RenderX

Figure 1. Google Trends results for the search items "JAK inhibitor," "JAK inhibitor + alopecia areata," and "JAK inhibitor + hair loss" (data accessed April 30, 2024). Data are presented as relative search volume (RSV), where RSV of 100 represents peak search activity in a time period. The plot represents the time series analysis of Google Trends RSV per month between January 2014 and December 2023 for the terms "JAK inhibitor," "JAK inhibitor + alopecia areata," and "JAK inhibitor + hair loss." Search interest for "JAK inhibitor" increased at a rate of 6.89 (95% CI 6.16 - 7.62; P<.001; R^2=0.748) RSV units per month. Searches for "JAK inhibitor + hair loss" and "JAK inhibitor + alopecia areata" increased at a rate of 3.50 (95% CI 3.13 - 3.87; P<.001; R^2=0.749) RSV units per month and 1.54 (95% CI 1.12 - 1.96; P<.001; R^2=0.306) RSV units per month, respectively. JAK: Janus kinase.



Discussion

All 3 search terms ("JAK inhibitor," "JAK inhibitor + hair loss," "JAK inhibitor + alopecia areata") demonstrated significant increases over time (P<.001), indicating rising public interest from January 2014 to December 2023. Searches for "JAK inhibitor + hair loss" showed a strong and consistent upward trend (R^2≈0.75), while general searches for "JAK inhibitor" increased at the fastest rate, suggesting expanding overall awareness of the drug class. The increasing trend is not perfectly linear; rather, it reflects periods of spikes, drops, and plateaus. This variability is possibly influenced by social factors such as periods of media coverage, FDA approvals, etc. For example, there is an appreciable spike in all 3 search terms between 2020 - 2023, which may indicate increased public interest following increased media coverage about the FDA approval of JAK inhibitors for severe alopecia reported by news outlets as early as 2019 [4]. Baricitinib became the first FDA-approved JAK inhibitor for alopecia areata in 2022. This was followed by the approval of ritlecitinib in 2023 and deuruxolitinib in 2024 [5]. Additionally, social media applications like TikTok could be possible contributors to the increase in public interest, due to user-friendly content, the ability of content creators to freely express their journey with hair loss and JAK inhibitors, and users being able to engage in conversations with their peers anonymously. Although detailed TikTok data trends from 2014 - 2023 were not readily attainable, the TikTok Creator Search Insights tool, which provides trends for the past 6 months, indicates that more than 1 million searches were made on TikTok for JAK inhibitors and alopecia from September 2025 to February 2026. These searches yield a variety of videos, including content from patients describing their experience with JAK inhibitors. A more in-depth exploration of search trends and content on social media platforms would provide further valuable insight into the trends in public interest on this topic.

Given the increased public interest in JAK inhibitor treatment as indicated by Google Search Trends in this study, it is important that the public receives proper education regarding the implications of taking JAK inhibitors and the knowledge that JAK inhibitors are treatment for severe alopecia areata, which is an autoimmune condition, and not currently efficacious for other hair loss disorders. It is important for physicians to educate their patients on hair loss disorder treatments and increase patient awareness of certain hair loss treatments for the patient's form of hair loss.

XSL•FO
RenderX

Authors' Contributions

JH, NAHK, HI performed data analysis;
JH, NAHK, HI, LDK all contributed to manuscript preparation;
JH, NAHK, HI and LDK revised the manuscript, approved the final version, and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Ryguła I, Pikiewicz W, Kaminiów K. Novel Janus kinase inhibitors in the treatment of dermatologic conditions. Molecules 2023 Dec 13;28(24):8064. [doi: 10.3390/molecules28248064] [Medline: 38138551]
2. Freitas E, Guttman-Yassky E, Torres T. Baricitinib for the treatment of alopecia areata. Drugs (Abingdon Engl) 2023 Jun;83(9):761-770. [doi: 10.1007/s40265-023-01873-w] [Medline: 37195491]
3. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill 2019 May 29;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]
4. Abdelmalek M. New drug for alopecia shows promise: what you need to know. ABC News. 2019. URL: https://abcnews.com/GMA/Wellness/drug-alopecia-shows-promise/story?id=60128818&utm_source=chatgpt.com [accessed 2024-05-11]
5. Sanchez K, Englander H, Salloum L, et al. Evaluating current and emergent JAK inhibitors for alopecia areata: a narrative review. Dermatol Ther (Heidelb) 2025 Oct;15(10):2749-2764. [doi: 10.1007/s13555-025-01517-9] [Medline: 40794245]

Abbreviations

FDA: US Food and Drug Administration
JAK: Janus kinase
RSV: relative search volume

Edited by S Subhadarshani; submitted 01.Apr.2025; peer-reviewed by B Southwell; revised version received 02.Mar.2026; accepted 14.Mar.2026; published 24.Apr.2026.

Please cite as:
Howard J, Kamili NAH, Idris H, Krueger LD
Public Interest in Janus Kinase (JAK) Inhibitors for Alopecia Areata: A Google Trend Analysis
JMIR Dermatol 2026;9:e75119
URL: https://derma.jmir.org/2026/1/e75119
doi:10.2196/75119

© Jade Howard, Nourine A H Kamili, Hala Idris, Loren D Krueger. Originally published in JMIR Dermatology (http://derma.jmir.org), 24.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Differences in Electronic Consultation Conversion Rates Between Advanced Practice Providers and Board-Certified Dermatologists

Dakota Hitchcock[1*], MD; Sabrina Newman[2*], MD

[1]Department of Internal Medicine, Saint Joseph Hospital, 1960 Ogden Street Suite 400, Denver, CO, United States

[2]Department of Dermatology, University of Colorado School of Medicine, Department of Dermatology, Aurora, CO, United States

[*]all authors contributed equally

Corresponding Author:
Dakota Hitchcock, MD
Department of Internal Medicine, Saint Joseph Hospital, 1960 Ogden Street Suite 400, Denver, CO, United States

Abstract

In this analysis of dermatology e-consults at a large academic health system, advanced practice providers had nearly threefold higher conversion rates to in-person visits compared to board-certified dermatologists, with potential implications for access and resource utilization.

(JMIR Dermatol 2026;9:e83922) doi:10.2196/83922

KEYWORDS

e-consult; telehealth; dermatology; advanced practice provider; teledermatology

Introduction

Electronic consultations (e-consults) have become an increasingly valuable tool in improving access to specialty care, reducing unnecessary in-person referrals, and supporting timely management of patients by primary care providers [1,2]. By allowing clinicians to consult with specialists asynchronously through the electronic health record, e-consults can help streamline workflows, decrease wait times, and conserve specialist resources [2,3]. Dermatology services receive a high number of e-consult requests, likely due to the visual diagnostic nature of the specialty [3,4]. As the use of e-consults expands across health care systems, understanding how different provider types use this tool, particularly in high-demand specialties such as dermatology, is critical to optimizing efficiency and effectiveness. Furthermore, identifying whether variations in conversion patterns reflect provider-level practice differences or system-level routing processes is essential for ensuring that e-consults function as intended.

Methods

We conducted a retrospective analysis to evaluate whether e-consult conversion rates differed by provider type, specifically comparing advanced practice providers (APPs), including nurse practitioners and physician assistants, to board-certified dermatologists. e-consult data specific to dermatology were extracted from the University of Colorado Hospital electronic health record system for the period of January 2020 to April 2025. An e-consult was considered “converted” if it resulted in a subsequent in-person specialist visit or full referral, rather than being resolved entirely through asynchronous communication.

In this system, e-consults are routed to APPs versus dermatologists primarily based on provider availability rather than consult content or patient acuity. As a result, patients evaluated by APPs and physicians likely represent comparable clinical populations, reducing the likelihood that differences in conversion rates were driven by systematic triage of more complex cases to one provider group.

Results

A total of 2572 dermatology e-consults were submitted during the study period. Of these, 1205 were addressed by APPs, with 321 (26.6%) resulting in conversion to an in-person visit (Table 1). In contrast, only 125 of the 1367 e-consults addressed by physicians (9.1%) were converted (Table 2). e-consults managed by APPs were nearly three times more likely to lead to an in-person referral compared to those managed by physicians.

XSL•FO
RenderX

Table . Total number and percentage of e-consults converted from advanced practice professionals.

e-consult converted	N (%)
No	884 (73.4)
Yes	321 (26.6)
Total	1205 (100.0)

Table . Total number and percentage of e-consults converted from dermatologists.

e-consult converted	N (%)
No	1242 (90.9)
Yes	125 (9.1)
Total	1367 (100.0)

Discussion

This analysis reveals a notable difference in e-consult conversion rates between APPs and physicians. This disparity suggests potential differences in how each provider group approaches triage and decision-making in specialty care. If APP-handled e-consults were converted at the same rate as physician-handled e-consults, over 200 additional dermatology clinic appointments during the study period may have been available for patients with higher-acuity needs. Despite this variation in appointment conversion, it is important to note that the majority of e-consults from both groups were resolved without the need for in-person follow-up. This reinforces the broader value of e-consults in improving efficiency and reducing unnecessary specialist visits and aligns with current literature [2,3].

The higher conversion rate observed among APPs may reflect a range of underlying factors. One possibility is that APPs may be more likely to convert e-consults conservatively due to comparatively less specialty-specific training or comfort managing complex cases. Importantly, in our system, APPs and dermatologists receive e-consults based largely on provider availability rather than clinical complexity. This reduces the likelihood that differences in patient or case characteristics explain the observed variation. Existing literature on provider-level differences in e-consult use and impact have shown mixed results. For example, one study comparing e-consults submitted by nurse practitioners and family physicians found that nurse practitioners were more likely to report that the consultation led to new clinical guidance and less likely to say it avoided an in-person referral [5]. In contrast, a systematic review of referral practices found no significant difference in overall referral rates between nurse practitioners and family physicians [6]. However, these studies largely examine differences among referring providers rather than responding providers. Because our study investigates variation among the providers performing the e-consults themselves, it represents a novel contribution to the literature. To our knowledge, no published studies have specifically examined provider-level variation in dermatology e-consult outcomes from the specialist side, underscoring the importance of our findings.

While our findings shed light on differences in provider behavior, they also raise questions about the clinical appropriateness of these conversions. Without detailed outcome data, it remains unclear whether the higher conversion rate among APPs were clinically necessary or reflective of a lower threshold for referral. Future research should explore the clinical drivers and downstream outcomes of converted e-consults, considering patient complexity, consult content, and specialty-specific considerations.

In addition to clinical impact, the higher conversion rate among APPs may have broader implications for system efficiency and resource use. Given the higher conversion rate, APP-managed e-consults may increase health care utilization, with potential cost implications for patients and health systems. Assuming a standard new patient visit billed at a level 3 or level 4 (estimated reimbursement US $120–$180 per visit), the additional ~200 appointments potentially consumed due to higher APP conversion rates translates to an estimated US $24,000–$36,000 in additional health care costs during the study period. Future work could further investigate whether these conversions lead to improved outcomes or represent avoidable costs.

This study contributes to the growing body of literature on e-consult optimization and provider practice variation. As health systems increasingly adopt team-based models of care and integrate APPs more fully into specialty workflows, ensuring consistent and effective use of e-consults across provider types will be essential. Implementing structured guidance, standardized triage protocols, and targeted training modules, particularly for APPs, may help promote more consistent decision-making and appropriate referral thresholds. Additionally, health systems may consider establishing limitations or clinical guidelines regarding the types of dermatologic conditions appropriate for independent APP e-consult management to ensure high-quality care, reduce unnecessary referrals, and minimize avoidable health care costs. By equipping all members of the care team with the tools and guidance needed to manage e-consults effectively, we can improve access, preserve specialist capacity, and enhance the overall efficiency of care delivery.

XSL•FO
RenderX

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

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

Authors' Contributions

Conceptualization: DH, SN
Data curation: SN
Formal analysis: DH, SN
Investigation: DH
Methodology: DH, SN
Supervision: SN
Validation: DH, SN
Writing – original draft: DH, SN
Writing – review & editing: DH, SN

Conflicts of Interest

None declared.

References

1. Vimalananda VG, Orlander JD, Afable MK, et al. Electronic consultations (E-consults) and their outcomes: a systematic review. J Am Med Inform Assoc 2020 Mar 1;27(3):471-479. [doi: 10.1093/jamia/ocz185] [Medline: 31621847]
2. Vimalananda VG, Gupte G, Seraj SM, et al. Electronic consultations (e-consults) to improve access to specialty care: a systematic review and narrative synthesis. J Telemed Telecare 2015 Sep;21(6):323-330. [doi: 10.1177/1357633X15582108] [Medline: 25995331]
3. Keely E, Liddy C, Afkham A. Utilization, benefits, and impact of an e-consultation service across diverse specialties and primary care providers. Telemed J E Health 2013 Oct;19(10):733-738. [doi: 10.1089/tmj.2013.0007] [Medline: 23980939]
4. Bradley C, Smith L, Youens K, White BAA, Couchman G. Formalizing the curbside: digitally enhancing access to specialty care. Proc (Bayl Univ Med Cent) 2023;36(6):716-720. [doi: 10.1080/08998280.2023.2240364] [Medline: 37829223]
5. Liddy C, Deri Armstrong C, McKellips F, Keely E. A comparison of referral patterns to a multispecialty eConsultation service between nurse practitioners and family physicians: The case for eConsult. J Am Assoc Nurse Pract 2016 Mar;28(3):144-150. [doi: 10.1002/2327-6924.12266] [Medline: 25965249]
6. Laurant M, Reeves D, Hermens R, Braspenning J, Grol R, Sibbald B. Substitution of doctors by nurses in primary care. Cochrane Database Syst Rev 2005 Apr 18(2):CD001271. [doi: 10.1002/14651858.CD001271.pub2] [Medline: 15846614]

Abbreviations

APP: advanced practice provider
E-consults: electronic consultations

Edited by R Dellavalle; submitted 10.Sep.2025; peer-reviewed by ES Temin, P Aggarwal; revised version received 30.Dec.2025; accepted 06.Jan.2026; published 30.Jan.2026.

Please cite as:
Hitchcock D, Newman S
Differences in Electronic Consultation Conversion Rates Between Advanced Practice Providers and Board-Certified Dermatologists
JMIR Dermatol 2026;9:e83922
URL: https://derma.jmir.org/2026/1/e83922
doi:10.2196/83922

© Dakota Hitchcock, Sabrina Newman. Originally published in JMIR Dermatology (http://derma.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

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

XSL•FO
RenderX

Patient Perceptions of Climate Change Impacts on Atopic Dermatitis: Cross-Sectional Survey Study

Gunnar Mattson[1], MPH, MD; Sarah Coates[2], MD; Amanda R Twigg[2], MD

[1]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]Department of Dermatology, University of California, San Francisco, 1701 Divisadero Street, Third Floor, San Francisco, CA, United States

Corresponding Author:
Amanda R Twigg, MD
Department of Dermatology, University of California, San Francisco, 1701 Divisadero Street, Third Floor, San Francisco, CA, United States

Abstract

This cross-sectional survey study (63.5% response rate) characterized how patients with atopic dermatitis (AD) perceive and experience the effects of climate change on their AD. Most participants reported that environmental factors such as heat and air pollution worsened their AD and expressed a desire for climate-health education, yet few had discussed these concerns with their dermatologist. These findings reveal a gap in patient-centered dermatologic care and support the development of tools to integrate environmental health into atopic dermatitis management.

(JMIR Dermatol 2026;9:e80679) doi:10.2196/80679

KEYWORDS

atopic dermatitis; eczema; climate change dermatology; climate change; environmental health; climate health impacts; patient experience; survey study; health communication; patient education

Introduction

Climate change is recognized as the foremost global health threat of the 21st century [1]. Environmental shifts (rising temperatures, air pollution, and extreme weather) can impair the skin barrier, alter the microbiome, and induce inflammation, increasing the prevalence and severity of atopic dermatitis (AD) among other skin conditions [2,3]. Among dermatologists, 79.6% agree it affects their patients [4]. Yet, few routinely discuss this with patients, and limited research explores how patients perceive and experience these impacts. To address these gaps, this cross-sectional study surveyed patients with AD to assess how they perceive climate change's impact on their condition and whether these concerns are addressed in dermatologic care.

Methods

Survey Instrument Development

The survey was informed by climate-health literature, dermatologic environmental impacts, and health communication frameworks (eg, message framing, perceived susceptibility, and severity from the Health Belief Model) [5]. Five UCSF (University of California, San Francisco) dermatologists reviewed the instrument for clinical relevance and clarity. Ten adult AD patients piloted it, and feedback informed wording and usability.

Study Population & Recruitment

Eligible participants were English-speaking adults with AD seen at UCSF dermatology clinics between August 2023 and August 2024. A total of 2164 patients were identified by the electronic health record (EHR) query. To reduce selection bias, patients were contacted via EHR messaging or mailed letters to account for differences in digital health access; 326 patients expressed interest and became the study population. These patients were sent the study description and a secure Qualtrics link to the online survey.

Statistical Analysis

Descriptive statistics using Microsoft Excel were used to summarize participant demographics and survey responses. Frequencies were calculated for categorical variables. No inferential or hypothesis testing was conducted, as the study aimed to characterize trends and patient-reported experiences rather than test associations or determine causality.

Ethical Considerations

This study received exempt certification from the UCSF medical ethical review committee (IRB 21 - 33538). All participants provided consent to participate in the study, and their responses were deidentified.

Results

Of 326 individuals, 207 completed the survey (63.5% response rate). A majority of individuals (n=166/207, 80.2%, 95% CI

XSL•FO
RenderX

74.8% - 85.6%) reported that environmental-climate factors impact their AD, particularly extreme heat (n=157, 75.8%, 95% CI 70.0% - 81.7%) and poor air quality (n=81, 39.1%, 95% CI 32.5% - 45.8%). Commonly reported effects included increased medication use (n=168, 81.2%, 95% CI 75.8% - 86.5%), more symptomatic flares (n=167, 80.7%, 95% CI 75.3% - 86.1%), more skin affected (n=139, 67.1%; 95% CI 60.8% - 73.5%), and changes to daily behaviors (n=130, 62.8%; 95% CI 56.2% - 69.4%). Most participants (n=179, 86.5%; 95% CI 81.8% - 91.1%) expressed interest in understanding how environmental-climate factors affect their AD, yet only 76 participants (36.7%; 95% CI 30.1% - 43.3%) said their dermatologist addressed these concerns. The most valued strategies for addressing climate-health impacts included more information (n=164, 79.2%; 95% CI 73.7% - 84.8%), dedicated time during visits to plan for exposures (n=105, 50.7%; 95% CI 43.9% - 57.5%), and more in-person visits (n=101, 48.8%; 95% CI 42.0% - 55.6%). Table 1 shows participant characteristics, and Table 2 shows survey response data.

Table . Participant demographics and background information.

Demographics	Participants (N=207)
Age in years (mean, SD)	46.4 (18.6)
Sex, n (%)	
Male	75 (36.2)
Female	129 (62.3)
Nonbinary	3 (1.4)
Race/Ethnicity, n (%)	
American Indian or Alaskan Native	2 (1.0)
Asian or Asian American	82 (39.6)
Black or African American	12 (5.8)
Hispanic or Latino	12 (5.8)
Native Hawaiian or Pacific Islander	1 (0.5)
White	107 (51.7)
Other	5 (2.4)
Years living with atopic dermatitis (mean, SD)	21.6 (18.3)
Treatments used for atopic dermatitis, n (%)	
Topical steroid	193 (93.7)
Topical medication other than a steroid	145 (70.4)
Topical over the counter product (does not require a prescription)	139 (67.4)
Pill medication (eg, methotrexate, cellcept, tofacitinib, upadacitinib)	47 (22.8)
Injection medication (eg, dupilumab, tralokinumab)	94 (45.6)
Phototherapy	41 (19.9)

XSL•FO
RenderX

Table . Responses to survey questions using the 5-point Likert scale, where 1 indicates “Strongly disagree,” 2 “Somewhat disagree,” 3 “Neutral,” 4 “Somewhat agree,” and 5 “Strongly agree.” A reported mean greater than 3 indicates agreement and less than 3 indicates disagreement.

Statement, agreement ranked using the 5-point Likert scale	Score, mean (SD)
Climate and environmental factors have impacted your experience with eczema	4.2 (1.0)
The following factor has impacted your experience with eczema:	
Extreme Heat	4.2 (1.1)
Wildfires	3.3 (1.1)
Poor Air Quality	3.4 (1.1)
Drought	3.2 (1.1)
Extreme Rainfall	3.0 (1.3)
Sea Level Rise	2.4 (1.0)
Flooding	2.6 (1.1)
Climate and environmental factors’ impact on your eczema include:	
More symptomatic with exacerbations or flares	4.2 (1.0)
More skin affected	3.9 (1.2)
Need for extra appointments with healthcare team	3.1 (1.2)
Sending additional messages to dermatologist or calling their office	3.0 (1.2)
Using medication more often	4.1 (1.0)
Change to your medication	3.2 (1.3)
Change to lifestyle or daily behaviors	3.8 (1.1)
You want to know how the climate and environment impact your eczema	4.2 (1.0)
Your dermatologist has talked about how the climate and environment affect your eczema	2.9 (1.3)
This strategy would be helpful in managing changes to your eczema from the climate and environment:	
More visits in person	3.4 (1.1)
More telehealth visits	3.2 (1.1)
Time during visits to make plans for climate or environmental problems	3.5 (1.1)
More information on the topic	4.1 (0.9)
Support groups	2.9 (1.1)

Discussion

Principal Findings

While this study does not evaluate clinical causality, it provides novel insight into how patients perceive and experience the effects of environmental-climate factors on their AD. Most participants perceived climate-related changes in their AD and desired clinical guidance, yet few reported receiving it. These findings suggest that dermatologists should initiate brief conversations about common triggers, particularly heat and air pollution, and provide anticipatory guidance and resources. This insight underscores previously reported low self-efficacy among dermatologists in discussing climate change with patients [4]. Understanding these patient insights is vital to providing patient-centered care and forming effective partnerships with patients about their skin health. These efforts align with the American Academy of Dermatology’s commitment to “educate our patients about the effects of climate change on the health of their skin.” [6]

Limitations and Future Direction

Limitations include a single-center design limiting generalizability, reliance on self-reported data with potential recall bias, and possible self-selection bias, as patients more affected by climate change may have been more likely to participate. Future research should validate these findings in broader populations, explore climate-health experiences in other skin conditions, and develop educational and clinical strategies to help navigate these climate-health conversations with patients. Even in short visits, dermatologists can explore patient experiences with climate change using supportive prompts (eg, “Would it be helpful to discuss how environmental factors might relate to your flares?”) to validate patient concerns and provide opportunities for personalized climate-health conversations to be continued in subsequent visits.

Conclusions

This study highlights a disconnect between how patients with AD experience climate-related triggers and how often these

XSL•FO
RenderX

concerns are addressed in clinical care. Findings underscore the need for tools and strategies to support climate-health conversations in dermatology. Integrating environmental health into AD management can enhance patient-centered care, improve outcomes, and reinforce dermatology's role at the intersection of clinical care, public health, and patient advocacy.

Data Availability

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

Conflicts of Interest

None declared.

References

1. Watts N, Amann M, Arnell N, et al. The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. Lancet 2021 Jan 9;397(10269):129-170. [doi: 10.1016/S0140-6736(20)32290-X] [Medline: 33278353]
2. Mieczkowska K, Stringer T, Barbieri JS, Williams M, Rosenbach M. Surveying the attitudes of dermatologists regarding climate change. Br J Dermatol 2022 Apr;186(4):748-750. [doi: 10.1111/bjd.20900] [Medline: 34787308]
3. Ahn K. The role of air pollutants in atopic dermatitis. J Allergy Clin Immunol 2014 Nov;134(5):993-999. [doi: 10.1016/j.jaci.2014.09.023] [Medline: 25439225]
4. Celebi Sozener Z, Ozdel Ozturk B, Cerci P, et al. Epithelial barrier hypothesis: effect of the external exposome on the microbiome and epithelial barriers in allergic disease. Allergy 2022 May;77(5):1418-1449. [doi: 10.1111/all.15240] [Medline: 35108405]
5. Alyafei A, Easton-Carr R. The health belief model of behavior change. In: StatPearls [Internet]: StatPearls Publishing; 2025. URL: https://www.ncbi.nlm.nih.gov/books/NBK606120 [accessed 2025-03-30]
6. Position Statement on Climate and Health. American Academy of Dermatology. 2018. URL: https://server.aad.org/forms/policies/Uploads/PS/PS%20-%20Climate%20and%20Health.pdf [accessed 2024-03-31]

Abbreviations

AD: atopic dermatitis
EHR: electronic health record
UCSF: University of California, San Francisco

Edited by E Parker; submitted 15.Jul.2025; peer-reviewed by C Nnajiofor, J Volkman; revised version received 01.Oct.2025; accepted 01.Jan.2026; published 20.Jan.2026.

Please cite as:
Mattson G, Coates S, Twigg AR
Patient Perceptions of Climate Change Impacts on Atopic Dermatitis: Cross-Sectional Survey Study
JMIR Dermatol 2026;9:e80679
URL: https://derma.jmir.org/2026/1/e80679
doi:10.2196/80679

© Gunnar Mattson, Sarah Coates, Amanda R Twigg. Originally published in JMIR Dermatology (http://derma.jmir.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Chronic Facial Abscess Mimicking Cervicofacial Actinomyces From Dermal Filler Migration: Case Report

Monika Ziogaite[1], BS; Sarah Mannlein[2], BS; Nicole Bender[3], MD; Scott J Mahlberg[2], DO

[1]College of Osteopathic Medicine, Kansas City University, 1750 Independence Ave, Kansas City, MO, United States

[2]Colorado Center for Dermatology & Skin Surgery, Longmont, CO, United States

[3]Sagis Dermatopathology, Houston, TX, United States

Corresponding Author:
Monika Ziogaite, BS
College of Osteopathic Medicine, Kansas City University, 1750 Independence Ave, Kansas City, MO, United States

Abstract

Dermal fillers are commonly used for facial augmentation, but delayed complications such as granulomatous inflammation and filler migration can mimic chronic bacterial infections, such as cervicofacial actinomycosis, and lead to diagnostic misdirection. We present the case of a woman aged 56 years with a chronic, draining abscess on the right cheek that persisted for 3 years and was initially suspected to represent cervicofacial actinomycosis. Tissue cultures were negative, and histopathologic analysis following excisional biopsy revealed polymethyl methacrylate microspheres and hyaluronic acid surrounded by granulomatous inflammation and reactive lymphoid aggregates, consistent with a foreign body reaction to dermal filler. The patient experienced complete resolution after surgical excision. This case underscores the diagnostic challenges posed by delayed filler complications and highlights the importance of considering prior cosmetic procedures in patients with chronic facial abscesses.

(JMIR Dermatol 2026;9:e80278) doi:10.2196/80278

KEYWORDS

dermal filler; foreign body reaction; facial abscess; granulomatous inflammation; filler complications; cosmetic dermatology

Introduction

The use of dermal fillers for facial augmentation has increased significantly, with both temporary (hyaluronic acid) and semipermanent or permanent fillers (polymethyl methacrylate [PMMA], calcium hydroxylapatite, silicone) [1]. While most complications occur immediately or within weeks, delayed reactions, including granuloma formation and filler migration, can present months to years after injection [2].

Foreign body granulomas are a known complication of PMMA-based fillers (Bellafill/Artefill; Suneva Medical Inc) and result from a chronic inflammatory response to nondegradable microspheres [3]. These reactions may be triggered by delayed hypersensitivity, biofilm formation, or immune dysregulation and can resemble infectious or inflammatory processes. They can also closely mimic chronic infectious processes such as cervicofacial actinomycosis, characterized by draining sinuses and subcutaneous abscesses, often prompting an extensive infectious workup before the true etiology is recognized [4]. Diagnosis relies on histopathologic evaluation, which typically reveals multinucleated giant cells, lymphoid aggregates, and fibrosis surrounding filler particles [3,4].

This case highlights the importance of early recognition of iatrogenic causes in the differential diagnosis of chronic facial abscesses and underscores the long-term risks associated with semipermanent fillers, particularly PMMA-based products. Given the varied histopathologic presentations of different filler materials, distinguishing PMMA from other injectables is crucial for accurate diagnosis and management.

Ethical Considerations

The authors obtained written consent from the patient for their photographs and medical information to be published in print and online, with the understanding that this information may be publicly available. Patient consent forms were not provided to the journal but are retained by the authors.

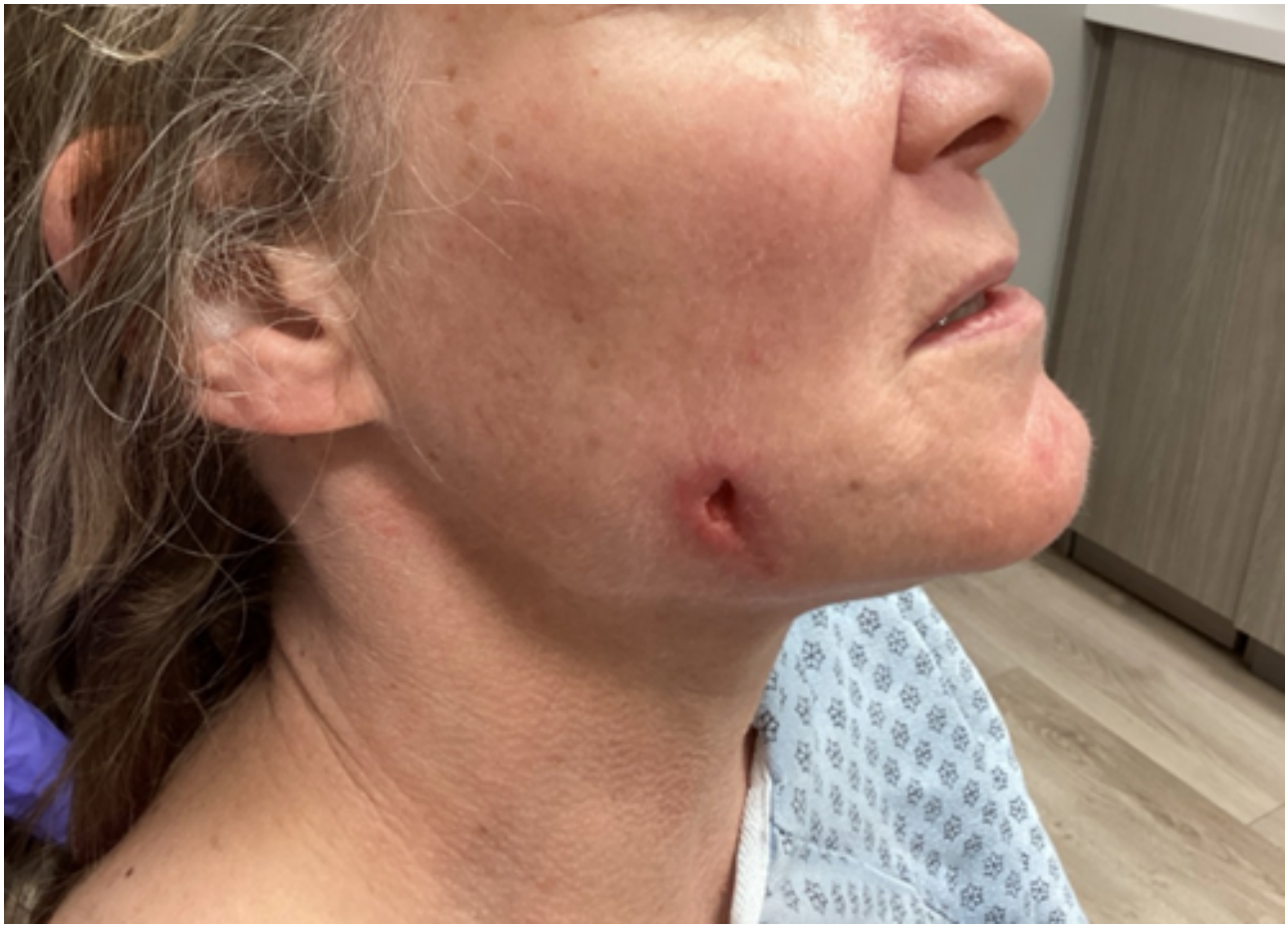
Case Report

A woman aged 56 years presented with a chronic, nonhealing, draining abscess localized to the right cheek. Characterized by intermittent drainage, localized tenderness, and surrounding erythema, the nodule persisted for approximately 3 years, during which time the patient sought care from specialists on at least one occasion. The patient denied systemic symptoms such as fever, chills, dental caries, oral drainage, pain with salivation, or malaise. Past medical history was noncontributory, and the patient had no known history of immunosuppression, diabetes, or recurrent skin infections.

XSL•FO
RenderX

On physical examination, the deep nodule was ulcerated, with erythematous borders localized to the right inferior central malar cheek. The ulcer base exhibited crusting and purulent material, with 3 cm of surrounding induration (Figure 1). No regional lymphadenopathy was present.

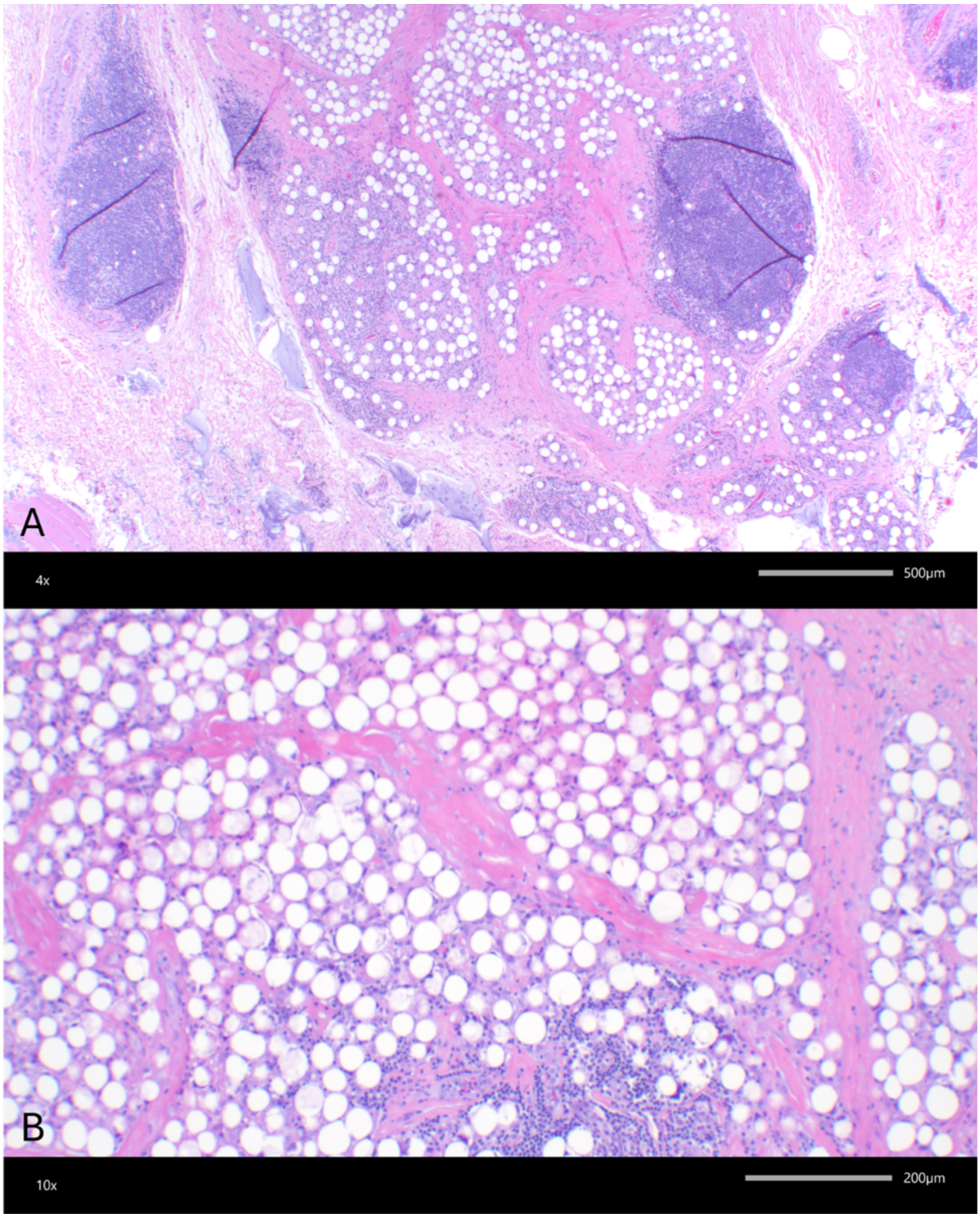
Figure 1. Nonhealing ulcer with surrounding erythema and induration on the right cheek.



Given the persistent nature of the abscess, a tissue culture was obtained via punch biopsy, which was negative for bacterial, fungal, and atypical mycobacterial growth. Due to the size and depth of the lesion, an excisional biopsy was performed to identify potential inflammatory or neoplastic pathology. Histopathologic analysis (Figure 2) revealed foreign body granulomas with abundant reactive lymphoid tissue, along with an accumulation of PMMA microspheres and hyaluronic acid, consistent with semipermanent dermal fillers. Surrounding granulomatous inflammation was also noted. Special stains, including Grocott methenamine silver, Fite, and Gram stains, were negative for fungi, mycobacteria, and bacteria, ruling out infectious etiologies.

XSL•FO
RenderX

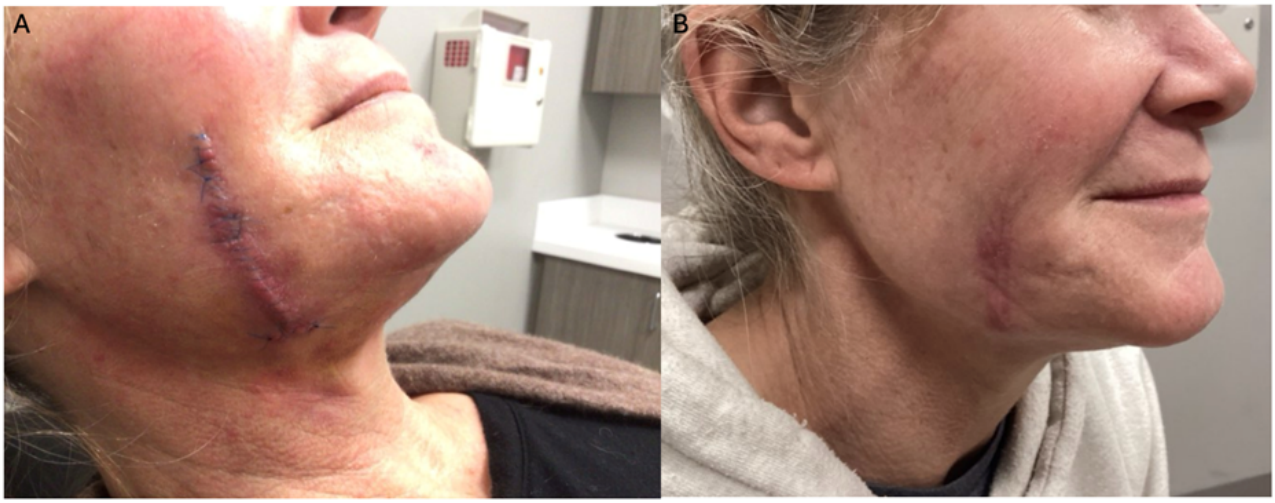
Figure 2. Histologic features consistent with a foreign body reaction to polymethylmethacrylate (PMMA) filler. (A) Incisional biopsy with numerous rounded vacuolated spaces, 30 - 50 µm in size, consistent with PMMA spherules and surrounding granulomatous inflammation and reactive lymphoid aggregates. Adjacent homogeneous basophilic material consistent with hyaluronic acid is also present (hematoxylin and eosin stain; 4x). (B) Higher magnification of PMMA spherules (hematoxylin and eosin stain; 10×).



The patient was seen for suture removal at 1 week (Figure 3) with an uneventful postoperative course. At 4 weeks postoperatively, the surgical site was well healed, with no signs of infection or recurrence (Figure 3). Although initially unable to recall prior dermal filler use, review of the pathology report prompted the patient to remember having received Bellafill (Suneva Medical Inc) approximately 6 years prior and Restylane (Galderma Laboratories) approximately 10 years earlier, both for treatment of acne scarring on the cheeks.

XSL•FO
RenderX

Figure 3. (A) Postoperative appearance at 1 week showing early healing. (B) Postoperative appearance at 4 weeks with complete resolution and no recurrence.



Discussion

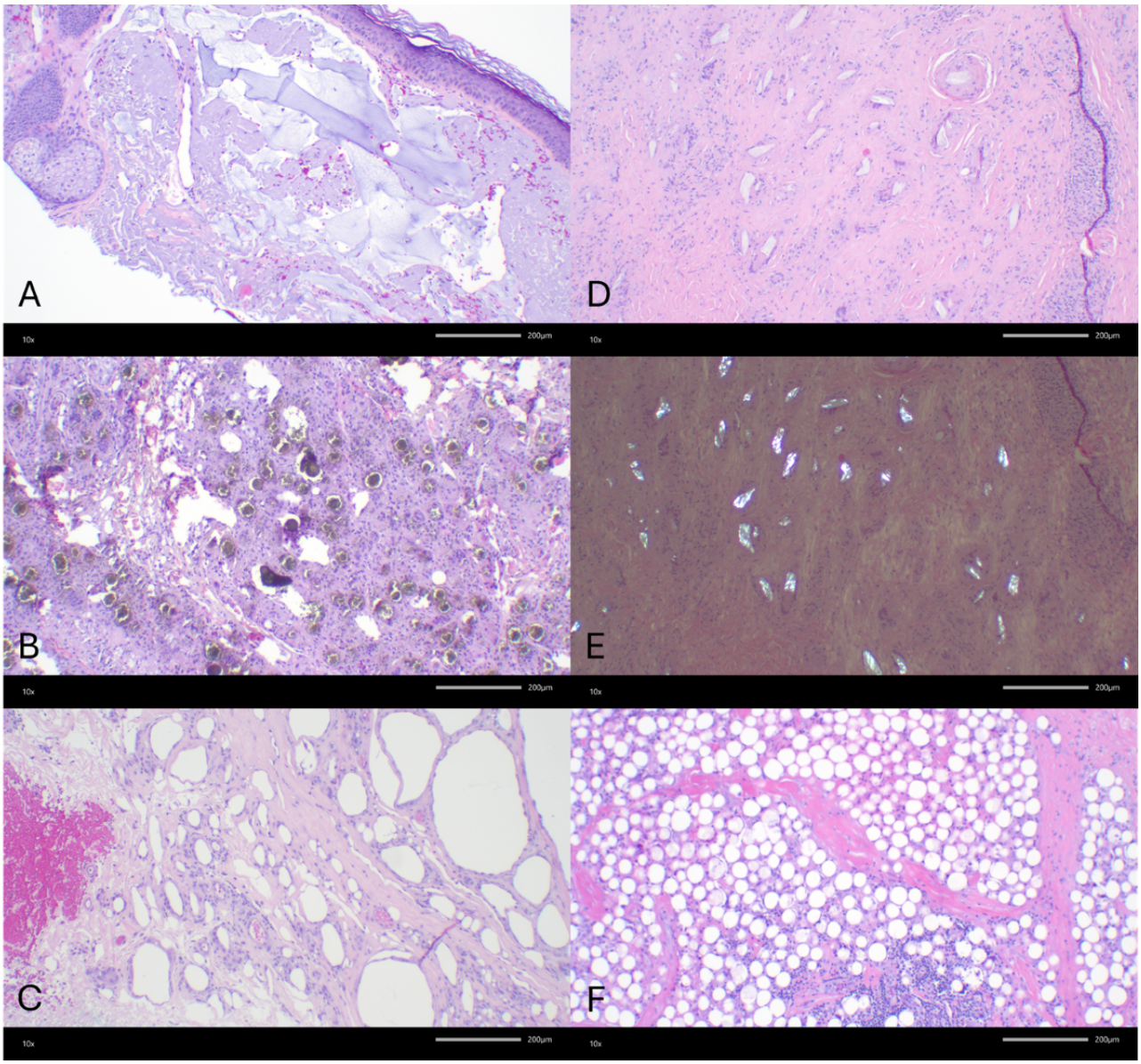
While dermal fillers are widely used for aesthetic enhancement, their delayed complications remain an evolving area of clinical concern. Although most adverse reactions occur shortly after injection, late-onset complications can develop months to years later, often leading to diagnostic uncertainty [5]. Unlike temporary hyaluronic acid–based fillers, PMMA and other nondegradable materials persist within tissues long-term, increasing the risk of prolonged inflammatory responses [6]. The delayed presentation of PMMA-related granulomas frequently results in misdiagnosis as infection or inflammatory dermatoses, delaying appropriate intervention [7].

While complications such as filler migration and granulomatous reactions are well documented, the development of a chronic filler reaction mimicking a cervicofacial actinomycetoma is rare. Actinomycetomas are chronic, subcutaneous infections caused by filamentous bacteria, characterized by abscesses, draining sinuses, and granule production [8]. The striking clinical resemblance between the foreign body granuloma in this case and a deep-seated actinomycotic infection underscores the diagnostic challenges posed by delayed filler reactions. This case highlights the need for broad infectious and histopathologic workups in atypical, chronic soft tissue infections to prevent unnecessary antibiotic treatment and delayed surgical intervention.

Histopathologic evaluation is essential for diagnosing PMMA-related granulomas, which are characterized by multinucleated giant cells, chronic lymphohistiocytic infiltrates, and fibrosis surrounding filler particles [9]. In this case, a negative infectious workup and biopsy findings of PMMA microspheres with reactive lymphoid tissue confirmed the diagnosis and guided treatment. Management remains challenging, as PMMA-based fillers lack a reversal agent comparable to that of hyaluronic acid fillers [10]. While intralesional corticosteroid injections may offer partial improvement, surgical excision is often required for definitive diagnosis and treatment, as in this case [4]. Given the histologic variability among filler types, distinguishing PMMA granulomas from reactions to calcium hydroxylapatite, poly-L-lactic acid, and silicone is critical for guiding optimal management strategies (Figure 4).

XSL•FO
RenderX

Figure 4. Histologic features consistent with foreign body reactions to soft tissue augmentation materials. (A) Pools of wispy homogeneous basophilic material consistent with hyaluronic acid (hematoxylin and eosin; 10×). (B) Gray-green granular nonrefractile microspheres consistent with calcium hydroxylapatite with surrounding granulomatous inflammation (hematoxylin and eosin; 10×). (C) Variably sized empty lipidlike vacuoles within histiocytes consistent with silicone granuloma (hematoxylin and eosin; 10×). (D) Oval, rhomboidal, and rice-shaped clear refractile and polarizable structures consistent with poly-L-lactic acid (PLLA) particles within histiocytes (hematoxylin and eosin; 10×). (E) Polarization of PLLA fragments (hematoxylin and eosin; 10×). (F) Fairly uniform 30 - 50-µm rounded vacuolated spaces consistent with polymethylmethacrylate spherules (hematoxylin and eosin; 10×).



While PMMA fillers are used less frequently than hyaluronic acid–based products, their potential for chronic inflammatory complications requires heightened clinical awareness and a detailed risk-benefit discussion prior to injection. Semipermanent fillers pose unique challenges due to their prolonged tissue retention and risk of delayed reactions. Clinicians should maintain a high index of suspicion for foreign body granulomas and probe for a history of prior filler use in cases of chronic, nonhealing facial abscesses, particularly when standard antimicrobial therapy fails or imaging reveals localized nodularity. Following excision, patients should be informed of and monitored for delayed recurrence as well as contralateral lesions, as these may occur months to years later. This case highlights the value of histopathologic microbiologic evaluation in diagnosing facial abscesses, the limitations of nonsurgical management for PMMA-induced granulomas, and the need for increased awareness of iatrogenic factors in chronic soft tissue reactions.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Jones DH. Semipermanent and permanent injectable fillers. Dermatol Clin 2009 Oct;27(4):433-444. [doi: 10.1016/j.det.2009.08.003] [Medline: 19850193]
2. Hong GW, Hu H, Chang K, et al. Review of the adverse effects associated with dermal filler treatments: part I nodules, granuloma, and migration. Diagnostics (Basel) 2024 Jul 30;14(15):1640. [doi: 10.3390/diagnostics14151640] [Medline: 39125515]
3. Sivam S, Mackay C, Humphrey C, Kriet JD. Giant PMMA foreign body granulomas with imaging. J Cutan Aesthet Surg 2023;16(3):256-258. [doi: 10.4103/JCAS.JCAS_194_20] [Medline: 38189071]
4. Lee JM, Kim YJ. Foreign body granulomas after the use of dermal fillers: pathophysiology, clinical appearance, histologic features, and treatment. Arch Plast Surg 2015 Mar;42(2):232-239. [doi: 10.5999/aps.2015.42.2.232] [Medline: 25798398]
5. Woodward J, Khan T, Martin J. Facial filler complications. Facial Plast Surg Clin North Am 2015 Nov;23(4):447-458. [doi: 10.1016/j.fsc.2015.07.006] [Medline: 26505541]
6. Funt D, Pavicic T. Dermal fillers in aesthetics: an overview of adverse events and treatment approaches. Clin Cosmet Investig Dermatol 2013 Dec 12;6:295-316. [doi: 10.2147/CCID.S50546] [Medline: 24363560]
7. Limongi RM, Tao J, Borba A, et al. Complications and management of polymethylmethacrylate (PMMA) injections to the midface. Aesthet Surg J 2016 Feb;36(2):132-135. [doi: 10.1093/asj/sjv195] [Medline: 26446059]
8. Welsh O, Vera-Cabrera L, Welsh E, Salinas MC. Actinomycetoma and advances in its treatment. Clin Dermatol 2012;30(4):372-381. [doi: 10.1016/j.clindermatol.2011.06.027] [Medline: 22682184]
9. Lemperle G, Gauthier-Hazan N, Wolters M, Eisemann-Klein M, Zimmermann U, Duffy DM. Foreign body granulomas after all injectable dermal fillers: part 1. Possible causes. Plast Reconstr Surg 2009 Jun;123(6):1842-1863. [doi: 10.1097/PRS.0b013e31818236d7] [Medline: 19483587]
10. Rao V, Chi S, Woodward J. Reversing facial fillers: interactions between hyaluronidase and commercially available hyaluronic-acid based fillers. J Drugs Dermatol 2014 Sep;13(9):1053-1056. [Medline: 25226005]

Abbreviations

PMMA: polymethyl methacrylate

Edited by J Solomon; submitted 07.Jul.2025; peer-reviewed by EA Sanchez-Romero, NC Balica; revised version received 16.Oct.2025; accepted 25.Dec.2025; published 03.Feb.2026.

Please cite as:
Ziogaite M, Mannlein S, Bender N, Mahlberg SJ
Chronic Facial Abscess Mimicking Cervicofacial Actinomyces From Dermal Filler Migration: Case Report
JMIR Dermatol 2026;9:e80278
URL: https://derma.jmir.org/2026/1/e80278
doi:10.2196/80278

© Monika Ziogaite, Sarah Mannlein, Nicole Bender, Scott J Mahlberg. Originally published in JMIR Dermatology (http://derma.jmir.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Informed Consent Practices for Publication of Patient Images in Dermatology Journals

Toluwani Taiwo[1*], BA; Bianca Obiakor[1*], BA; Sarah McClung[2], MIS; Kanade Shinkai[3], MD, PhD

[1]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]University of California, San Francisco, San Francisco, CA, United States

[3]Department of Dermatology, University of California, San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Kanade Shinkai, MD, PhD
Department of Dermatology
University of California, San Francisco
1701 Divisadero St
San Francisco, CA, 94115
United States
Phone: 1 415 353 7800
Email: kanade.shinkai@ucsf.edu

Related Article:

Correction of: https://derma.jmir.org/2025/1/e60795

(JMIR Dermatol 2026;9:e94194) doi:10.2196/94194

In "Informed Consent Practices for Publication of Patient Images in Dermatology Journals" [1], the authors made one correction.

In Multimedia Appendix 1, we included a Dryad link to our raw dataset, which includes Clarivate data. We have recently learned that Clarivate's terms of use and Dryad's data reuse policy are incompatible, and we will not be able to deposit the raw data in Dryad. The raw dataset can be made available by contacting the corresponding author.

In Multimedia Appendix 1, the following text has been revised:

> *The raw dataset for this study is available for review via the Dryad platform and accessible via the following unique link:http://datadryad.org/share/k6Eiw-OfOGf0vJrxD6sOS8GC59H5SFuhtudjzHR1EH8.*

The text now reads:

> *The raw dataset for this study is available by contacting the corresponding author.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1
List of the top 50 dermatology journals ranked by the 2023 Clarivate Journal Citation Report and a link to the publicly available raw dataset used in the study.
[DOCX File , 17 KB - derma_v9i1e94194_app1.docx]

Reference

1. Taiwo T, Obiakor B, McClung S, Shinkai K. Informed Consent Practices for Publication of Patient Images in Dermatology Journals. JMIR Dermatol 2025 Apr 18;8:e60795 [FREE Full text] [doi: 10.2196/60795] [Medline: 40249653]

XSL•FO
RenderX

Edited by P Kharbanda; submitted 26.Feb.2026; this is a non–peer-reviewed article;accepted 27.Feb.2026; published 13.Mar.2026.

Please cite as:
Taiwo T, Obiakor B, McClung S, Shinkai K
Correction: Informed Consent Practices for Publication of Patient Images in Dermatology Journals
JMIR Dermatol 2026;9:e94194
URL: https://derma.jmir.org/2026/1/e94194
doi:10.2196/94194
PMID:

©Toluwani Taiwo, Bianca Obiakor, Sarah McClung, Kanade Shinkai. Originally published in JMIR Dermatology (http://derma.jmir.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI and Digital Tools in Dermatology: Addressing Access and Misinformation

Dominique du Crest[1*], MBE; Monisha Madhumita[2*], MD; Wendemagegn Enbiale[3*], MD, MPH, PhD; Jose Antonio Ruiz Postigo[4*], MD, PhD; Josep Malvehy[5*], MD; Shannon Wongvibulsin[6*], MD, PhD; Somesh Gupta[7*], MD; Harald Kittler[8*], MD; Charbel Skayem[9], MD; Anjali Mahto[10*], MD; Adewole S Adamson[11*], MD, MPP; Jules B Lipoff[12*], MD; Art Papier[13*], MD; Hugues Cartier[14*], MD; Sébastien Garson[15], MD; Esther Freeman[16*], MD, PhD

[1]SkinAid SAS, 50 Rue Danton, Levallois Perret, France

[10]Self London, London, United Kingdom

[11]University of Texas at Austin Dell Medical School, Austin, TX, United States

[12]Department of Dermatology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, United States

[13]Department of Dermatology, University of Rochester College of Medicine, Rochester, NY, United States

[14]Saint-Jean Clinic of Dermatology, Arras, France

[15]SELARL Dr GARSON, Senlis, France

[16]Massachusetts General Hospital, Department of Dermatology, Boston, MA, United States

[2]Department of Dermatology, Saveetha Medical College, Chennai, India

[3]College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia

[4]World Health Organization, Geneva, Switzerland

[5]Centro de Investigación Biomédica en Red de Enfermedades Raras Instituto de Salud Carlos III, Barcelona, Spain

[6]Division of Dermatology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

[7]All India Institute of Medical Sciences, New Delhi, New Delhi, India

[8]Department of Dermatology, Medical University of Vienna, Vienna, Austria

[9]Hôpital Ambroise-Paré, Boulogne Billancourt, France

[*]these authors contributed equally

Corresponding Author:
Dominique du Crest, MBE
SkinAid SAS, 50 Rue Danton, Levallois Perret, France

Abstract

Digital dermatology, which is defined as the use of digital technologies that leverage individual- and population-level skin data to improve the diagnosis, treatment, and prevention of skin diseases, has emerged as a critical frontier for bridging persistent gaps in dermatologic care. This transformation holds particular promise for addressing long-standing inequities linked to geography, income, and skin type. According to the Global Burden of Disease 2023 study, skin and subcutaneous diseases remain among the most prevalent global health conditions, contributing substantially to disability-adjusted life years. Digital tools (including teledermatology, artificial intelligence [AI], and large language models) offer new ways to extend diagnosis, education, and patient empowerment to historically underserved populations. However, these same innovations risk amplifying disparities if they are not designed and deployed intentionally. Algorithmic bias, uneven digital access, and the absence of culturally responsive models can undermine progress. In this conceptual and narrative review, we draw on expert dialogues and illustrative literature, including multistakeholder exchanges at the *Skin and Digital Summit (2023-2025*) and related global forums, to examine how digital dermatology can promote equitable skin health. We focus on 3 interlinked priorities: expanding access through scalable digital platforms, ensuring AI fairness via comprehensive and diverse datasets, and countering dermatological misinformation. Central to the latter is a bot concept described here as a dynamic cycle that analyzes scientific literature; ranks evidence; translates complex research into clear language; and delivers trustworthy, personalized guidance to both consumers and clinicians. By embedding expert oversight and evidence prioritization, such tools can ensure that accurate, actionable information reaches users at the speed and scale of the internet. Drawing on case studies (including lessons from the World Health Organization's AI skin health app) and insights from the Skin and Digital Summit, we highlight both the transformative potential and the ethical complexities of these digital solutions. To navigate this evolving landscape, we propose the concept of radical dermatology**,** which confronts the reality that big tech is reshaping skin health whether we like it or not and insists that dermatologists and stakeholders lead the transformation through bold collaboration and unwavering clinical relevance.

XSL•FO
RenderX

(JMIR Dermatol 2026;9:e79044) doi:10.2196/79044

KEYWORDS

digital dermatology; AI in health care; global health; teledermatology; artificial intelligence; AI; large language model; LLM; health tech; skin; dermatology; digital health; misinformation; ethical AI; radical dermatology; social media

Introduction

Globally, skin and subcutaneous diseases are among the most common health conditions. In 2019 alone, there were an estimated 4.86 billion new cases globally, with fungal (34%) and bacterial (23%) infections making up the majority of cases [1]. Skin conditions accounted for nearly 43 million disability-adjusted life years, a burden that is both physically and psychosocially disabling [1].

Despite this high burden, dermatologic care remains inaccessible for many. An estimated 70% of people living with chronic skin diseases, including psoriasis and rosacea, do not receive professional medical care, largely due to limited dermatology workforce capacity, cost barriers, geographic distance, and sociocultural factors in care delivery [2]. These barriers affect both low-resource settings and underserved communities in low-, middle-, and high-income countries [3].

At the same time, rapid smartphone penetration, the maturation of generative artificial intelligence (AI) tools, and the World Health Organization's (WHO) prioritization of skin health have created an unprecedented window for digital solutions. Digital dermatology refers to the application of digital technologies and skin-related data (including telemedicine, mobile health [mHealth], AI, and digital imaging) to improve the diagnosis, treatment, and prevention of skin diseases. While it offers opportunities to bridge persistent gaps in access to care, challenges such as bias in algorithms, limited infrastructure, and variable digital literacy must be addressed through deliberate equity-focused design and governance. Unlike traditional models that rely on in-person specialist visits concentrated in urban centers, digital dermatology can decentralize expertise, shorten wait times, enable earlier diagnosis, extend specialist reach, and support patient self-management across geographies [4]. However, equity is not guaranteed. If bias goes unaddressed, if infrastructure remains weak, or if digital literacy lags, digital dermatology may deepen the very divides it seeks to close. In the following sections, we highlight strategies to counter these risks and to align innovation with equity.

This conceptual and narrative review synthesizes expert insights, stakeholder dialogue, and illustrative case examples to explore equity-centered strategies in digital dermatology. Input was drawn from international experts through 4 panels and 3 roundtables at the Skin and Digital Summit held virtually (IMCAS, December 2024) and in Paris (IMCAS, January 2025) [5]. Data were captured through structured notes and partial transcripts and then analyzed using inductive thematic coding by 2 reviewers, with disagreements resolved by consensus and findings organized through framework synthesis [6,7].

As discussions occurred in public professional forums, institutional review board approval was not required. Any direct speaker quotations were deidentified or quoted with permission, and related files were stored securely with restricted access, in compliance with data protection best practices.

To complement these expert perspectives, we conducted a targeted literature review (PubMed, Scopus, and Web of Science; English-language studies), including studies on equity, implementation, and ethics in digital dermatology and related fields, thereby reducing selection bias.

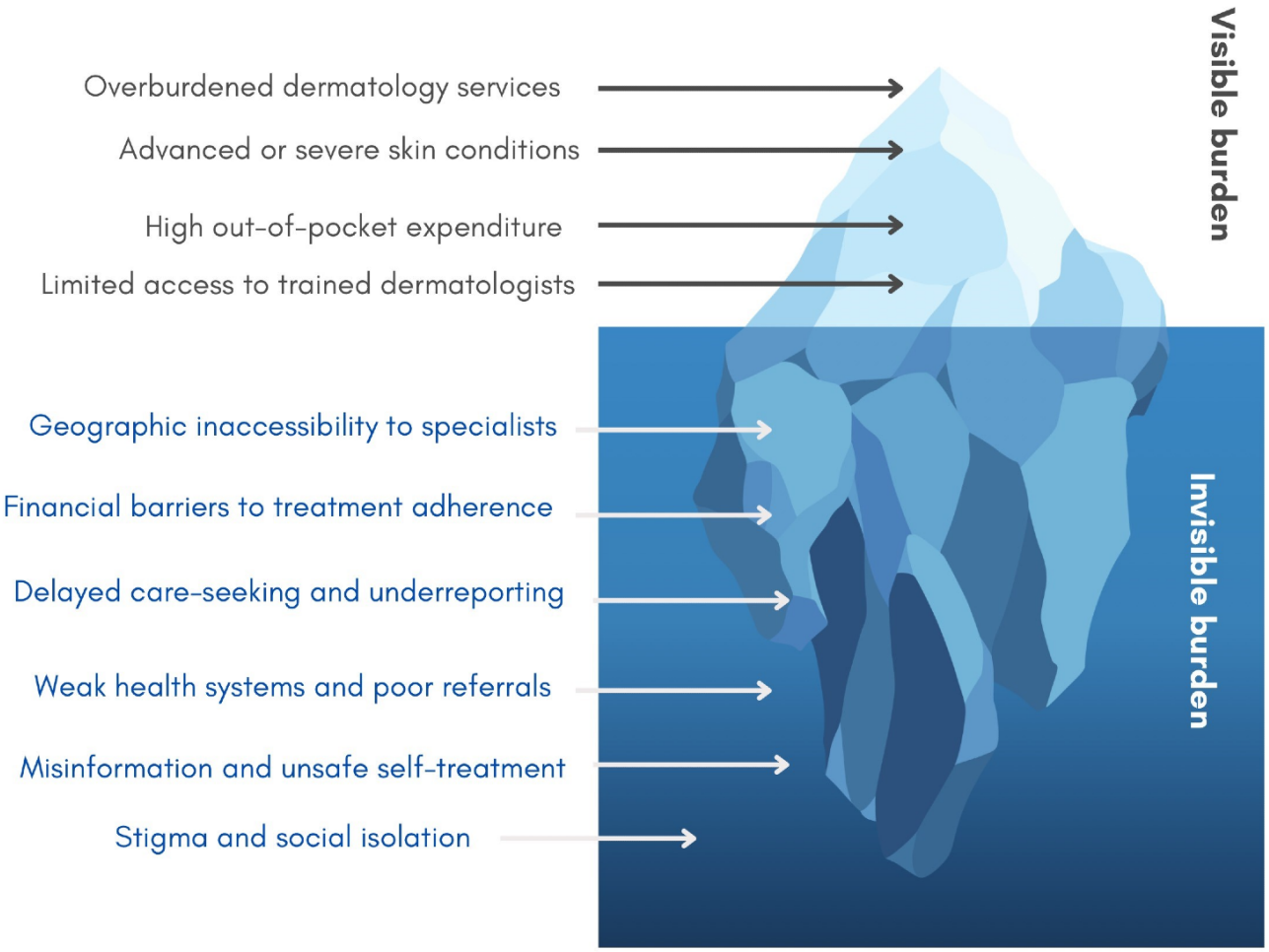
We discuss the WHO-neglected tropical disease mobile app as an example of ongoing implementation efforts and insights from global experts with references, where applicable. These examples were selected to highlight emerging real-world dynamics that are not yet captured in peer-reviewed literature. A narrative approach was used to thematically analyze insights across the digital dermatology continuum involving access to dermatologic care through digital innovation, datasets in dermatology, and combating skin-related misinformation.

Expanding Access to Dermatologic Care Through Digital Innovation

Despite the global burden of skin disease, dermatologic care remains underprioritized. Structural barriers, workforce shortages, and high out-of-pocket costs continue to limit access, especially in underserved and rural areas (Figure 1). These challenges are particularly acute for skin of color and populations in low- and middle-income countries [8,9].

XSL•FO
RenderX

Figure 1. The iceberg phenomenon of dermatologic care disparities.



Digital health innovations, including teledermatology, AI-assisted diagnostics, and mHealth apps, are increasingly recognized as promising tools to improve dermatology service delivery and care for all [10]. Given its reliance on visual diagnosis, dermatology makes it more amenable to remote care compared with many other specialties. High-quality clinical images and contextual data can enable remote triage, diagnosis, and follow-up, offering an opportunity to decentralize expertise [11,12].

One area is the proliferation of consumer-facing dermatology apps that offer mole monitoring, acne advice, and other self-assessment tools for timely triage and basic education, particularly in contexts where specialist access is limited. A 2024 review identified more than 900 such apps, including 41 with AI capabilities, and found their performance to be variable [13]. A 2021 analysis also showed sensitivity for melanoma to be highly variable across top apps, often with little dermatologist oversight or regulatory validation [14].

Alongside commercial tools, institutional efforts are advancing [15]. In 2024, the WHO piloted an AI-assisted mobile app across 5 Kenyan counties to help frontline workers identify 13 neglected tropical diseases and 24 common skin conditions. Forty Ministry of Health workers captured 605 patient images, which were reviewed by dermatologists. Although formal peer-reviewed accuracy data are pending, preliminary findings indicate more than 80% diagnostic agreement and strong usability for triage purposes rather than as a diagnostic replacement. Importantly, the app includes offline functionality and multilingual support and design features critical for resource-limited contexts [16,17]. However, several infrastructural and regulatory barriers persist as challenges, including unstable networks, device costs, limited digital literacy, and fragmented data governance frameworks that constrain uptake [18]. In some regions, legal restrictions on data sharing further complicate deployment. Addressing these challenges will require coordinated public-private partnerships and context-specific policy support.

Additionally, we must acknowledge that although digital dermatology tools can improve diagnosis and education, their impact is limited if patients cannot access the medications they are prescribed. In underserved areas, medication shortages, high costs, and weak supply chains often prevent timely treatment. Closing gaps in dermatologic care requires strategic, multisector partnerships to ensure that digital solutions are not just short-term pilot projects, but long-lasting, locally led interventions supported by active partnerships with big tech and big pharma [19]. To strengthen implementation, policy levers such as teledermatology reimbursement and data protection standards must be embedded within digital health initiatives. Such initiatives must work in tandem with public health systems, with continued monitoring of program success (such as shorter time to treatment after digital triage), through demand forecasting, community health worker training, and pharmacy

XSL•FO
RenderX

partnerships for sustained equitable access to both diagnosis and treatment [20].

Datasets: The Key to Unlocking AI Potential in Dermatology

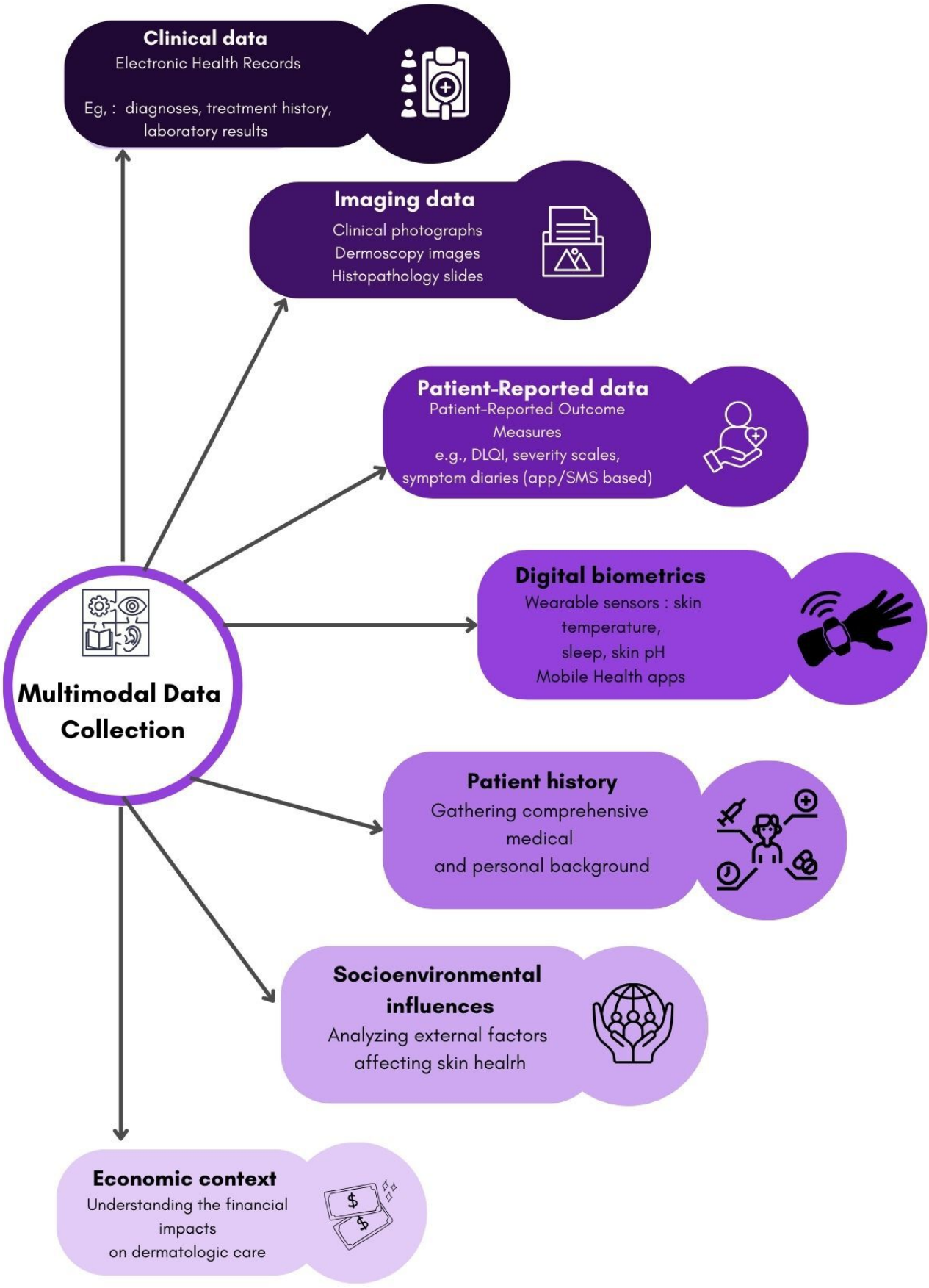
Complementing access-focused interventions, the success of AI in digital dermatology hinges on one critical factor: data. Dermatologic AI systems predominantly rely on visual data, including clinical photographs and dermatoscopy images, for training and validation [21]. However, the lack of large, diverse, and representative datasets poses a major limitation [22]. Most existing datasets overrepresent light-skinned patients from high-income settings and focus heavily on conditions such as melanoma [23]. Even newer collections, such as the Diverse Dermatology Images dataset, while addressing skin tone gaps, are limited in diagnostic breadth and geographic diversity [24,25]. To strengthen fairness and reproducibility, skin tone annotation should evolve beyond the traditional Fitzpatrick phototypes toward newer scales such as the Monk Skin Tone scale, using standardized labeling and consistent documentation protocols across datasets [26].

Without datasets that reflect all segments of the population and systems that ensure access to recommended care, AI solutions may offer precise diagnoses that are ultimately inaccessible to those who need them most. Relying on single-modality datasets, such as static images alone, restricts the ability of AI systems to capture the real-world complexity of dermatological care [27]. Dermatologic decision-making is not merely visual; it involves psychometric, environmental, and cultural factors. For instance, 2 patients with the same lesion may receive different treatment recommendations depending on their medical history, stress levels, socioeconomic context, or cultural expectations.

To address these issues, a multimodal data approach is essential (Figure 2) [28]. This includes integrating clinical photographs with patient histories, self-reported outcomes, environmental exposures, and behavioral data. For example, in conditions such as psoriasis or vitiligo, combining lesion photographs with validated severity indices (eg, Dermatology Life Quality Index and Pruritus Numerical Rating Scale), environmental parameters (UV index and humidity), comorbidities, and adherence telemetry can yield more personalized, context-aware interventions. These inputs can be captured via mHealth apps and wearables [29,30]. AI models trained on such holistic datasets, encompassing lesion images, lifestyle habits, stress levels, and treatment adherence, can deliver more accurate, individualized care. Importantly, these insights must be coupled with robust privacy safeguards, including on-device data processing and granular patient consent. Furthermore, to translate these innovations into improved outcomes in underserved areas, they must be linked to reliable access to medications through community health workers, local pharmacies, or public-sector distribution programs.

XSL•FO
RenderX

Figure 2. Expanding dermatology with multimodal data.



Enabling this shift requires interoperable and transparent data infrastructure with an emphasis on terminology mapping and handling of metadata. Platforms such as OpenMHealth, Apple HealthKit, and Fast Healthcare Interoperability Resources–compliant electronic health records can help structure and standardize multimodal data [31]. Future

XSL•FO
RenderX

dermatology-specific tools must enable seamless integration of patient-reported outcomes and subjective (eg, pruritus scores) and objective (eg, lesion photos) metrics while ensuring data privacy. Ethical development and deployment of such AI tools must follow established frameworks such as FATE (fairness, accountability, transparency, and ethics) [32] and the OECD AI Principles [33]. These frameworks call for human oversight, robust transparency about model limitations, and equitable access to benefits. Poorly curated or biased data can not only reduce diagnostic accuracy but also worsen outcomes for already marginalized groups.

While appealing, the belief in autonomous, data-driven AI-based decision-making has significant limitations; data alone are not the answer. Although increasing data volume often enhances predictive accuracy, the quality and representativeness of the datasets remain crucial [34]. Poorly curated data can reinforce existing biases and, in some cases, worsen clinical outcomes [35]. Human clinicians inherently recognize that accurate interpretation of data requires understanding its context, particularly in dermatology, where the psychosocial impact of visible skin conditions demands sensitivity beyond numerical analysis alone [36]. Effective dermatological decision-making thus integrates objective data with patient-specific considerations, including emotional, social, and cultural factors. Identical skin lesions may necessitate different diagnostic or therapeutic approaches depending on individual patient circumstances and preferences, underscoring the necessity of collaboration between data-driven insights, clinical expertise, and patient perspectives.

Combating Skin-Related Misinformation

Maximizing the impact of digital dermatology also requires confronting the parallel crisis of misinformation [37]. Fueled by algorithmic amplification and low digital health literacy, skin-related myths such as “sunscreen causes cancer” or “natural remedies are always safer” spread faster than verified medical guidance [38]. Platforms that reward engagement often privilege sensational content over accuracy, and dermatology is no exception. This dynamic is compounded by the low visibility of authoritative voices: only 4% of dermatology influencers on Instagram are board-certified dermatologists, and on TikTok, more than one-third of dermatology-related videos are created by nonprofessionals [39], frequently promoting unverified treatments such as raw potatoes for acne. On Reddit [40] and other forums, patients increasingly crowdsource diagnoses—sometimes for sensitive conditions such as sexually transmitted infections—reflecting growing distrust in formal health care and a shift toward peer-based digital health advice [41]. Studies of parenting blogs have shown that posts containing sunscreen misinformation consistently receive more engagement than scientifically accurate content [42].

Efforts to counter this phenomenon must go beyond reactive myth-busting and instead adopt structured, measurable, and trust-centered approaches and strategies grounded in successful digital literacy campaigns [43]. Public health initiatives in adjacent domains offer useful models. For example, the WHO’s “Pause Before You Share” campaign effectively encouraged users to reflect before forwarding unverified information during the COVID-19 pandemic [44]. The United Nations Children’s Fund and Gavi’s #VaccinesWork initiative combined localized messaging with influencer partnerships to rebuild vaccine confidence [45]. In sexual health, the AI-powered chatbot “Roo,” developed by Planned Parenthood, has demonstrated how conversational interfaces can deliver reliable, age-appropriate, and stigma-free health information at scale [46]. These campaigns share a common thread: they prioritize trust, accessibility, and proactive engagement, principles replicable in dermatology.

Drawing inspiration from these models, we propose an AI-powered dermatology chatbot (still in its conceptual phase) designed to address misinformation while empowering users with accurate, evidence-based content. The chatbot would operate on a retrieval-augmented generation architecture, fine-tuned on dermatology-specific content. It would draw from validated sources such as Cochrane systematic reviews, the American Academy of Dermatology (AAD), and the European Academy of Dermatology and Venereology (EADV) guidelines, patient association websites, and public health repositories from the WHO and CDC (Centers for Disease Control and Prevention). To ensure scientific rigor, the chatbot would internally rank information using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) framework. It would also include medical disclaimers, clinician escalation paths, hallucination checks, and refusal policies for diagnostic or treatment queries. Bias testing**,** data minimization**,** and content updates aligned with AAD and EADV guidelines would maintain fairness, privacy, and accuracy. Its performance should be tracked through reach, engagement, accuracy, harmful-advice rate, and user trust metrics. Dermatologists should lead content creation and partnerships to amplify credible, evidence-based information [47].

It is crucial to reiterate that dermatologists need to proactively develop accurate digital content to combat the spread of misinformation [48]. This can involve structured social media initiatives such as verified content series, myth-busting reels, or question-and-answer sessions on platforms where misinformation is most widespread. Collaborating with trusted influencers, patient advocacy groups, or public health campaigns can further expand reach and ensure that evidence-based information is both accessible and engaging [49].

Beyond Content Curation: Developing a Technological Solution to Address Skin Misinformation

The rapid, unchecked spread of skin-related misinformation is emerging as a significant public health concern—one that no dermatologist, no matter how dedicated, can tackle alone [48-51]. In this digital era, where falsehoods travel faster than facts, expert voices must be amplified by bold technological innovation. What is urgently needed is not just another fact sheet, but an intelligent, real-time solution: an AI-powered dermatology conversational agent (Figure 3) capable of delivering science-backed answers at the speed of the internet

XSL•FO
RenderX

[52]. This chatbot could serve as a critical tool, helping to dispel myths, provide accurate information, and guide individuals toward appropriate care, with empathy [53].

Figure 3. Cycle of artificial intelligence (AI)–powered dermatological chatbot. LLM: large language model; ML: machine learning; NLP: natural language processing.

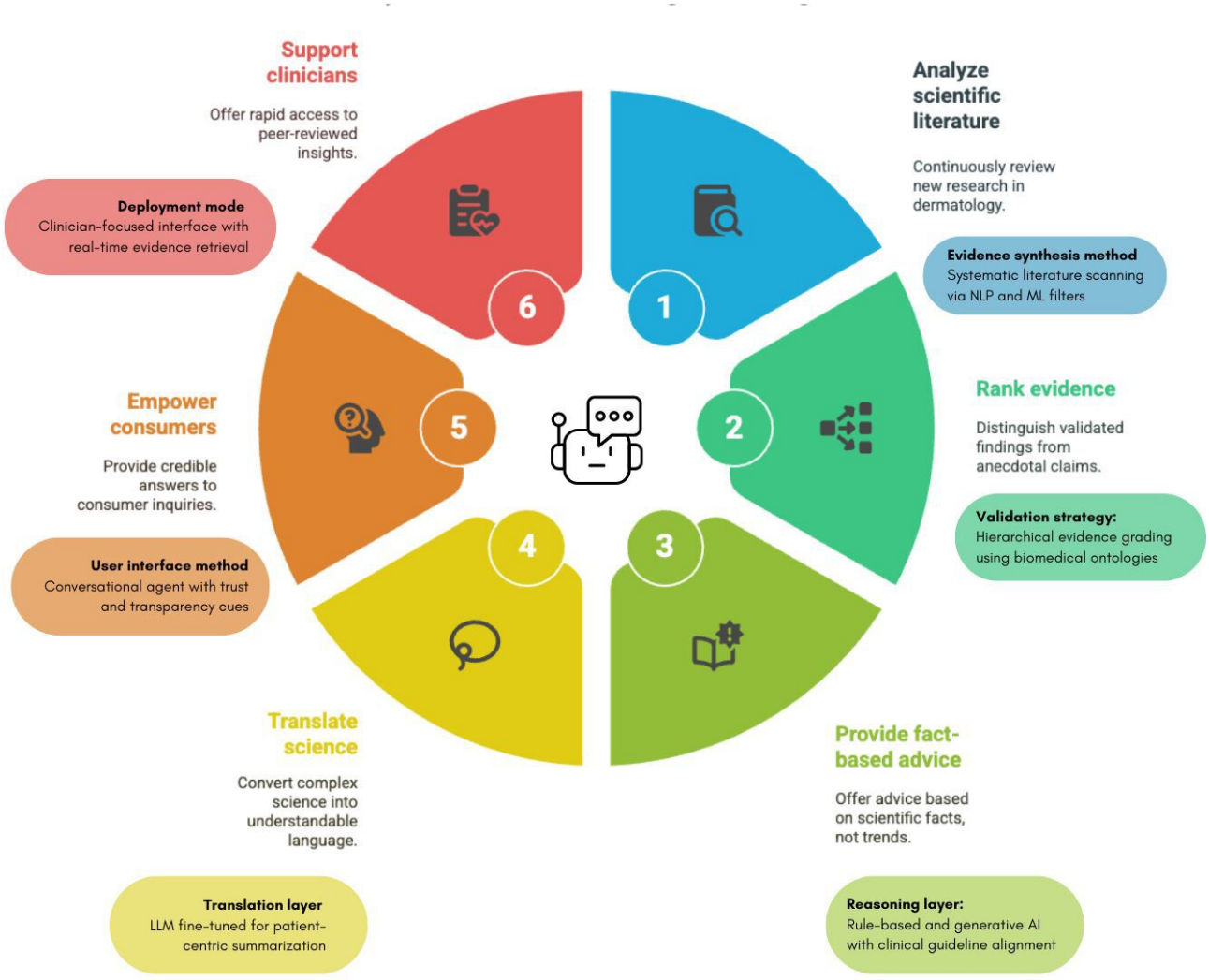


Figure 3 outlines the engine behind such a tool—a dynamic cycle that analyzes scientific literature, ranks evidence, translates complex research into clear language, and delivers trustworthy, personalized guidance to both consumers and clinicians. This is not just an upgrade; it is a paradigm shift, and it demands a multidisciplinary alliance of dermatologists, AI developers, skin scientists, tech experts, and health communicators. Only together can we build the digital frontline dermatology desperately needs.

The concept of "radical dermatology" underscores this transformation. Radical dermatology refers to a future-oriented framework that calls on dermatologists and skin health stakeholders to actively lead, rather than passively follow, the digital transformation of their field. It recognizes the growing influence of big tech in reshaping skin health and emphasizes that dermatologists must drive this change through bold collaboration, clinical relevance, and one that is ethically governed through measures such as equity dashboards and data diversity thresholds.

While conceptual in origin, radical dermatology lays the groundwork for concrete steps such as pilot studies, comparative evaluations of AI- and clinician-led care, and assessments of cost-effectiveness and equity outcomes. However, limitations of this paper include its presentation of a conceptual framework and a hypothetical chatbot model, with insights derived exclusively from expert dialogue rather than empirical clinical trials. To advance this research, future work should aim at pilot studies across diverse populations, comparative analyses of AI-supported versus clinician-led care, and comprehensive cost-effectiveness research. Dermatologists must play a central role in leading this transformative shift, guiding not only the application of digital tools but also their fundamental design, rigorous testing, and governance. This crucial role requires dermatologists to do more than just use digital tools; they must actively shape the technological, ethical, and societal landscape of digital dermatology.

Conclusions

Leveraging individual- and population-level skin data ethically through the convergence of computer vision, mHealth, and generative AI presents a pivotal opportunity for dermatology.

XSL•FO
RenderX

Conversational tools such as dialogic AI or chatbots, when well designed, can interpret validated evidence, provide real-time responses, and deliver guidance in ways that are accessible, accurate, and empathetic [52,53]. This is particularly relevant in dermatology, where care is visual, is time sensitive, and depends heavily on patient education and reassurance.

The promise of digital dermatology will only be achieved if these tools are developed for everyone, with transparency, accuracy, and trust as core principles and with an emphasis on ongoing safety auditing. People without access to dermatology care or reliable information stand to gain the most, provided their needs are addressed from the outset. Deployment should therefore prioritize underserved users through multilingual interfaces, offline functionality, and low-bandwidth design as integral components of equity-centered implementation. Dermatologists must remain central to this process, guiding design, overseeing implementation, and setting ethical standards to ensure that digital solutions are clinically sound and culturally appropriate.

The future of dermatology will be shaped not by algorithms alone but through collaboration among clinicians, technologist researchers, policymakers, and patients to ensure technology meaningfully advances skin health for all.

Data Availability

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

Conflicts of Interest

None declared.

References

1. Szeto MD, Alhanshali L, Rundle CW, et al. Dermatologic data from the global burden of disease study 2019 and the patientslikeme online support community: comparative analysis. JMIR Dermatol 2024 Dec 11;7:e50449. [doi: 10.2196/50449] [Medline: 39661989]
2. Wehausen B, Hill DE, Feldman SR. Most people with psoriasis or rosacea are not being treated: a large population study. Dermatol Online J 2016 Jul 15;22(7):13030/qt4nc3p4q2. [Medline: 27617716]
3. Yakupu A, Aimaier R, Yuan B, et al. The burden of skin and subcutaneous diseases: findings from the global burden of disease study 2019. Front Public Health 2023;11:1145513. [doi: 10.3389/fpubh.2023.1145513] [Medline: 37139398]
4. Skin diseases as a global public health priority. World Health Organization. 2025 May 27. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA78/A78_R15-en.pdf [accessed 2025-05-30]
5. Skin and Digital Summit. URL: https://www.skindigitalsummit.com/ [accessed 2026-02-13]
6. International Masters Course on Aging Science (IMCAS). URL: https://www.imcas.com/en [accessed 2925-02-20]
7. du Crest D, Madhumita M, Enbiale W, et al. Skin and digital-the 2024 narrative. Mayo Clin Proc Digit Health 2024 Sep;2(3):322-330. [doi: 10.1016/j.mcpdig.2024.05.008] [Medline: 40206126]
8. Adepoju O, Dang P, Nguyen H, Mertz J. Equity in digital health: assessing access and utilization of remote patient monitoring, medical apps, and wearables in underserved communities. Inquiry 2024;61:469580241271137. [doi: 10.1177/00469580241271137] [Medline: 39323052]
9. Gwillim EC, Azzawi S, Aigen AR. Underserved populations and health equity in dermatology: digital medicine and the role of artificial intelligence. Clin Dermatol 2024;42(5):498-506. [doi: 10.1016/j.clindermatol.2024.06.021] [Medline: 38944248]
10. Schaekermann M, Spitz T, Pyles M, et al. Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. EClinicalMedicine 2024 Apr;70:102479. [doi: 10.1016/j.eclinm.2024.102479] [Medline: 38685924]
11. Ahuja S, Briggs SM, Collier SM. Teledermatology in rural, underserved, and isolated environments: a review. Curr Dermatol Rep 2022;11(4):328-335. [doi: 10.1007/s13671-022-00377-2] [Medline: 36310767]
12. Chang AY, Kiprono SK, Maurer TA. Providing dermatological care in resource-limited settings: barriers and potential solutions. Br J Dermatol 2017 Jul;177(1):247-248. [doi: 10.1111/bjd.15372] [Medline: 28731250]
13. Wongvibulsin S, Yan MJ, Pahalyants V, Murphy W, Daneshjou R, Rotemberg V. Current state of dermatology mobile applications with artificial intelligence features. JAMA Dermatol 2024 Jun 1;160(6):646-650. [doi: 10.1001/jamadermatol.2024.1342] [Medline: 38452263]
14. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. Br J Dermatol 2022 Apr;186(4):744-746. [doi: 10.1111/bjd.20903] [Medline: 34811727]
15. Alami H, Rivard L, Lehoux P, et al. Artificial intelligence in health care: laying the foundation for responsible, sustainable, and inclusive innovation in low- and middle-income countries. Global Health 2020 Jun 24;16(1):52. [doi: 10.1186/s12992-020-00584-1] [Medline: 32580741]

XSL•FO
RenderX

16. The WHO Skin NTDs app shows encouraging results in Kenya study. World Health Organization. 2024. URL: https://www.who.int/news/item/04-12-2024-the-who-skin-ntds-app-shows-encouraging-results-in-kenya-study [accessed 2025-05-25]
17. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med 2022 Aug 18;5(1):119. [doi: 10.1038/s41746-022-00663-0] [Medline: 35982146]
18. Sylla B, Ismaila O, Diallo G. 25 years of digital health toward universal health coverage in low- and middle-income countries: rapid systematic review. J Med Internet Res 2025 May 29;27:e59042. [doi: 10.2196/59042] [Medline: 40440696]
19. Duniphin DD. Limited access to dermatology specialty care: barriers and teledermatology. Dermatol Pract Concept 2023 Jan 1;13(1):e2023031. [doi: 10.5826/dpc.1301a31] [Medline: 36892370]
20. Orton M, Agarwal S, Muhoza P, Vasudevan L, Vu A. Strengthening delivery of health services using digital devices. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S61-S71. [doi: 10.9745/GHSP-D-18-00229] [Medline: 30305340]
21. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. Bioengineering (Basel) 2023 Dec 18;10(12):1435. [doi: 10.3390/bioengineering10121435] [Medline: 38136026]
22. Daneshjou R, Vodrahalli K, Novoa RA, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv 2022 Aug 12;8(32):eabq6147. [doi: 10.1126/sciadv.abq6147] [Medline: 35960806]
23. Narla A, Kuprel B, Sarin K, Novoa R, Ko J. Automated classification of skin lesions: from pixels to practice. J Invest Dermatol 2018 Oct;138(10):2108-2110. [doi: 10.1016/j.jid.2018.06.175] [Medline: 30244720]
24. Groh M, Harris C, Daneshjou R, Badri O, Koochek A. Towards transparency in dermatology image datasets with skin tone annotations by experts, crowds, and an algorithm. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-26. [doi: 10.1145/3555634]
25. Jain T. Evaluating machine learning-based skin cancer diagnosis. arXiv. Preprint posted online on Sep 4, 2024. [doi: 10.48550/arxiv.2409.03794]
26. Arora A, Alderman JE, Palmer J, et al. The value of standards for health datasets in artificial intelligence-based applications. Nat Med 2023 Nov;29(11):2929-2938. [doi: 10.1038/s41591-023-02608-w] [Medline: 37884627]
27. Mishra S, Chaudhury S, Imaizumi H, Yamasaki T. Assessing robustness of deep learning methods in dermatological workflow. arXiv. Preprint posted online on Jan 15, 2020. [doi: 10.48550/arXiv.2001.05878]
28. Krones F, Marikkar U, Parsons G, Szmul A, Mahdi A. Review of multimodal machine learning approaches in healthcare. Information Fusion 2025 Feb;114:102690. [doi: 10.1016/j.inffus.2024.102690]
29. Hong J, Mosca M, Hadeler E, Hakimi M, Bhutani T, Liao W. The future of personalized medicine in psoriasis. Dermatological Reviews 2021 Oct;2(5):282-288. [doi: 10.1002/der2.87]
30. Ahuja R, Narayanan B, Gupta S. Precision medicine and personalized approach in vitiligo. Dermatological Reviews 2022 Oct;3(5):308-312. [doi: 10.1002/der2.159]
31. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14. [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
32. Singhal A, Neveditsin N, Tanveer H, Mago V. Toward fairness, accountability, transparency, and ethics in ai for social media and health care: scoping review. JMIR Med Inform 2024 Apr 3;12:e50048. [doi: 10.2196/50048] [Medline: 38568737]
33. Yeung K. Recommendation of the Council on Artificial Intelligence (OECD). Int leg mater 2020 Feb;59(1):27-34. [doi: 10.1017/ilm.2020.5]
34. Okolo CT. Optimizing human-centered AI for healthcare in the Global South. Patterns (N Y) 2022 Feb 11;3(2):100421. [doi: 10.1016/j.patter.2021.100421] [Medline: 35199066]
35. Green BL, Murphy A, Robinson E. Accelerating health disparities research with artificial intelligence. Front Digit Health 2024;6:1330160. [doi: 10.3389/fdgth.2024.1330160] [Medline: 38322109]
36. Breen N, Berrigan D, Jackson JS, et al. Translational health disparities research in a data-rich world. Health Equity 2019;3(1):588-600. [doi: 10.1089/heq.2019.0042] [Medline: 31720554]
37. Nelson EE, Black TA, Rousseau MA, Rashid RM. Combating misinformation in dermatology. Dermatol Pract Concept 2023 Oct 1;13(4):e2023234. [doi: 10.5826/dpc.1304a234] [Medline: 37992336]
38. Ecker U, Roozenbeek J, van der Linden S, et al. Misinformation poses a bigger threat to democracy than you might think. Nature New Biol 2024 Jun 6;630(8015):29-32. [doi: 10.1038/d41586-024-01587-3] [Medline: 38840026]
39. Ranpariya V, Chu B, Fathy R, Lipoff JB. Dermatology without dermatologists? Analyzing Instagram influencers with dermatology-related hashtags. J Am Acad Dermatol 2020 Dec;83(6):1840-1842. [doi: 10.1016/j.jaad.2020.05.039] [Medline: 32416205]
40. Chu B, Fathy R, Nobles AL, Lipoff JB. Patient crowdsourcing of dermatologic consults on a Reddit social media community. J Am Acad Dermatol 2021 Jul;85(1):226-227. [doi: 10.1016/j.jaad.2020.08.009] [Medline: 32777316]
41. Nobles AL, Leas EC, Althouse BM, et al. Requests for diagnoses of sexually transmitted diseases on a social media platform. JAMA 2019 Nov 5;322(17):1712-1713. [doi: 10.1001/jama.2019.14390] [Medline: 31688875]
42. O'Connor C, Rafferty S, Murphy M. A qualitative review of misinformation and conspiracy theories in skin cancer. Clin Exp Dermatol 2022 Oct;47(10):1848-1852. [doi: 10.1111/ced.15249] [Medline: 35514125]
43. Sood A, Sangari A, Stoff BK. Navigating internet-based misinformation with patients in the clinic. J Am Acad Dermatol 2023 Apr;88(4):961-962. [doi: 10.1016/j.jaad.2022.03.038] [Medline: 35346757]

XSL•FO
RenderX

44. Rodrigo P, Arakpogun EO, Vu MC, Olan F, Djafarova E. Can you be mindful? The effectiveness of mindfulness-driven interventions in enhancing the digital resilience to fake news on COVID-19. Inf Syst Front 2022 Mar 2:1-21. [doi: 10.1007/s10796-022-10258-5] [Medline: 35250364]
45. Germani F, Biller-Andorno N. How to counter the anti-vaccine rhetoric: filling information voids and building resilience. Hum Vaccin Immunother 2022 Nov 30;18(6):2095825. [doi: 10.1080/21645515.2022.2095825] [Medline: 35802046]
46. Mills R, Mangone ER, Lesh N, Jayal G, Mohan D, Baraitser P. Chatbots that deliver contraceptive support: systematic review. J Med Internet Res 2024 Feb 27;26:e46758. [doi: 10.2196/46758] [Medline: 38412028]
47. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490. [doi: 10.1136/bmj.328.7454.1490] [Medline: 15205295]
48. Szeto MD, Mamo A, Afrin A, Militello M, Barber C. Social media in dermatology and an overview of popular social media platforms. Curr Dermatol Rep 2021;10(4):97-104. [doi: 10.1007/s13671-021-00343-4] [Medline: 34692234]
49. El Mikati IK, Hoteit R, Harb T, et al. Defining misinformation and related terms in health-related literature: scoping review. J Med Internet Res 2023 Aug 9;25:e45731. [doi: 10.2196/45731] [Medline: 37556184]
50. Mehta-Ambalal SR, Nisarta M. Dermatology 2.0- how the internet is changing us, our patients and our practice. Indian Dermatol Online J 2021;12(4):593-596. [doi: 10.4103/idoj.IDOJ_788_20] [Medline: 34430468]
51. O'Connor C, O'Grady C, Murphy M. Spotting fake news: a qualitative review of misinformation and conspiracy theories in acne vulgaris. Clin Exp Dermatol 2022 Sep;47(9):1707-1711. [doi: 10.1111/ced.15222] [Medline: 35434841]
52. Mariani MM, Hashemi N, Wirtz J. Artificial intelligence empowered conversational agents: a systematic literature review and research agenda. J Bus Res 2023 Jun;161:113838. [doi: 10.1016/j.jbusres.2023.113838]
53. Chen D, Chauhan K, Parsa R, et al. Author correction: patient perceptions of empathy in physician and artificial intelligence chatbot responses to patient questions about cancer. NPJ Digit Med 2025 Jun 21;8(1):382. [doi: 10.1038/s41746-025-01793-x] [Medline: 40544179]

Abbreviations

AAD: American Academy of Dermatology
AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
EADV: European Academy of Dermatology and Venereology
FATE: fairness, accountability, transparency, and ethics
GRADE: Grading of Recommendations, Assessment, Development and Evaluations
mHealth: mobile health
WHO : World Health Organization

Edited by J Solomon, R Dellavalle; submitted 26.Jun.2025; peer-reviewed by H Peter Soyer, S Khosravi; revised version received 15.Dec.2025; accepted 16.Jan.2026; published 20.Feb.2026.

Please cite as:
du Crest D, Madhumita M, Enbiale W, Ruiz Postigo JA, Malvehy J, Wongvibulsin S, Gupta S, Kittler H, Skayem C, Mahto A, Adamson AS, Lipoff JB, Papier A, Cartier H, Garson S, Freeman E
AI and Digital Tools in Dermatology: Addressing Access and Misinformation
JMIR Dermatol 2026;9:e79044
URL: https://derma.jmir.org/2026/1/e79044
doi:10.2196/79044

© Dominique du Crest, Monisha Madhumita, Wendemagegn Enbiale, Jose Antonio Ruiz Postigo, Josep Malvehy, Shannon Wongvibulsin, Somesh Gupta, Harald Kittler, Charbel Skayem, Anjali Mahto, Adewole Adamson, Jules B Lipoff, Art Papier, Hugues Cartier, Sébastien Garson, Esther Freeman. Originally published in JMIR Dermatology (http://derma.jmir.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Beta-Alanine and Aquagenic Pruritus: Proposed Neuroimmune Mechanism

Natalie Piserchio[1], BS; Bailey Baratta[2], MBA, DO; Benjamin Brooks[3], PhD; Brandon Muse[4], DO; Katelin Ball[5], DO

[1]Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

[2]Internal Medicine Residency HEB/Denton, Texas Health Resources, Bedford, TX, United States

[3]Department of Biomedical Sciences, Rocky Vista University, Ivins, UT, United States

[4]Department of Obstetrics and Gynecology, Orlando Health, Orlando, FL, United States

[5]Department of Neurology, University of New Mexico, Albuquerque, NM, United States

Corresponding Author:
Natalie Piserchio, BS
Rocky Vista University, 8401 S Chambers Rd, Englewood, CO, United States

Abstract

Aquagenic pruritus (AP) is a rare itch disorder with limited effective treatments, and emerging clinical observations suggest that oral β-alanine may reduce symptoms. The purpose of this viewpoint is to propose a biologically plausible mechanism through which β-alanine may alleviate primary AP. We reviewed published case reports and patient-reported survey data describing β-alanine use in AP and integrated these clinical observations with experimental data on MAS-related G protein–coupled receptor D (MrgprD)–expressing sensory neurons and their role in mast-cell regulation. Published case reports describe marked improvement in water-induced pruritus following prophylactic oral β-alanine administration, and a recent survey of patients with idiopathic AP reported substantial symptom relief among β-alanine users. Preclinical data indicate that MrgprD-neuronal glutamate release suppresses mast cell hyperresponsiveness, suggesting a potential pathway for the observed antipruritic effect. Additional mechanisms, including β-alanine metabolism to carnosine and its potential mast cell–stabilizing effects, may also contribute. β-alanine may act through modulation of a nonhistaminergic neuroimmune circuit and represents a promising therapeutic candidate for further investigation in AP.

(JMIR Dermatol 2026;9:e90737) doi:10.2196/90737

KEYWORDS

aquagenic pruritus; beta-alanine; primary aquagenic pruritus; MrgprD; nonhistaminergic itch; mast-cell hyperresponsiveness; neuroimmune signaling; glutamatergic signaling; pruriceptive pathways; neurocutaneous mechanisms; itch modulation; MAS-related G protein–coupled receptor D

Aquagenic pruritus (AP) is a rare skin disorder characterized by intense itching, tingling, or burning after contact with water, which can significantly impact an individual's quality of life [1]. Although AP is classically associated with polycythemia vera, idiopathic (primary) AP likely represents the most common form in the general population [2]. Therapeutic approaches to alleviate this condition remain challenging and often ineffective. Existing treatments primarily focus on symptom relief through topical emollients, phototherapy, antihistamines, and avoidance of water exposure, but they provide only partial relief and do not address the underlying mechanisms triggering the pruritus [1].

The pathogenesis of idiopathic AP is incompletely understood. Several mechanisms have been proposed, including abnormal mast cell activation, dysregulated cutaneous nerve signaling, and altered neuroimmune communication within the skin [3]. A small biopsy study found increased acetylcholinesterase activity in eccrine-associated nerve fibers after water exposure in patients with AP and polycythemia vera, suggesting increased acetylcholine release and possible involvement of eccrine-associated pathways in AP pathogenesis [4]. AP has also been described in familial cases, suggesting a potential genetic predisposition [1].

Recent case reports have shed light on the role of β-alanine, a nonessential amino acid, in modulating the sensation of itch. β-alanine is a known agonist at the pruriceptive receptor MAS-related G protein–coupled receptor D (MrgprD) and induces a weak nonhistaminergic itch in humans and mice [5]. A survey-based study investigating β-alanine supplementation among 75 patients with idiopathic AP reported that 70.7% of participants who used β-alanine described substantial symptom relief, with an average relief score of 8.84 out of 10 (95% CI 8.52 - 9.16) [6].

In one case report, a 33-year-old adult male with primary AP found that prophylactic oral β-alanine taken 5 to 15 minutes before water exposure markedly reduced pruritus with sustained improvement at the 20-week follow-up [5]. In a second case, a 16-year-old male with AP refractory to traditional treatments

XSL•FO
RenderX

also found relief from oral β-alanine taken prior to water exposure [7].

In these reports, β-alanine was administered orally in powder form [5,7]. Survey data suggest that patients most commonly use doses averaging approximately 1.59 g per day during acute exacerbations. In this survey, 66% of users preferred powder formulations and 73.6% reported taking β-alanine on an as-needed basis [6].

The mechanism responsible for this improvement is unknown, and because MrgprD activation is typically pruritogenic, the antipruritic effect is unlikely to result from direct receptor agonism. Other mechanisms may also contribute to the observed therapeutic benefit. For example, β-alanine is a precursor to carnosine, a dipeptide with antioxidant and intracellular pH-buffering properties that may help stabilize mast cell responses. These effects could complement or act independently of MrgprD-mediated neural pathways [6].

Although the mechanisms underlying primary AP are not fully established, the delay between water exposure and itch onset reported in many patients may suggest a nonhistaminergic process involving abnormal neuroimmune interactions between sensory afferents and epidermal or immune cells such as keratinocytes, basophils, or mast cells [5]. A mouse study from 2021 demonstrated that MrgprD-expressing cutaneous sensory neurons help maintain skin homeostasis by releasing glutamate, which dampens mast cell hyperresponsiveness [3]. One possibility is that β-alanine may transiently enhance MrgprD-neuronal activity enough to increase local glutamate release, thereby strengthening this homeostatic inhibitory circuit. In primary AP, where mast cell hyperresponsiveness and nonhistaminergic pathways are suspected, enhanced MrgprD-mediated glutamatergic signaling could suppress mast cell activation and reduce downstream pruriceptive input. Recent work from the same research group further supports this model by demonstrating that glutamate can act directly on both mouse and human mast cells to suppress their activation, reinforcing the potential role of glutamatergic signaling in cutaneous immune homeostasis [8].

Available patient-reported data suggest that β-alanine is generally well tolerated, with transient paresthesia being the most frequently reported side effect. This is consistent with the known safety profile of β-alanine in sports nutrition studies and is typically mild and self-limited. Importantly, 90% of surveyed users reported sustained therapeutic benefit without loss of efficacy over time, suggesting minimal tachyphylaxis [6,9].

Although speculative, this mechanism aligns with both the delayed nature of AP symptoms reported by some individuals and the observed clinical benefit of prophylactic β-alanine administration. Given the debilitating nature of AP and the scarcity of effective therapies, the emerging evidence supporting β-alanine as a fast-acting, well-tolerated, inexpensive, and accessible intervention is noteworthy.

We encourage further controlled studies to characterize optimal dosing, duration of effect, long-term safety, and the underlying neurocutaneous mechanisms. If β-alanine consistently alleviates symptoms in primary AP, this observation may also support the concept that non–FcεR-mediated mast cell activation and neuroimmune dysregulation play important roles in disease pathogenesis.

β-alanine may represent a promising therapeutic avenue for a condition that currently lacks reliable treatment options. These observations highlight the need to further explore β-alanine as a targeted modulator of nonhistaminergic itch pathways in primary AP.

Funding

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

Conflicts of Interest

None declared.

References

1. Siegel FP, Tauscher J, Petrides PE. Aquagenic pruritus in polycythemia vera: characteristics and influence on quality of life in 441 patients. Am J Hematol 2013 Aug;88(8):665-669. [doi: 10.1002/ajh.23474] [Medline: 23657863]
2. Steinman HK, Greaves MW. Aquagenic pruritus. J Am Acad Dermatol 1985 Jul;13(1):91-96. [doi: 10.1016/s0190-9622(85)70149-1] [Medline: 2411768]
3. Zhang S, Edwards TN, Chaudhri VK, et al. Nonpeptidergic neurons suppress mast cells via glutamate to maintain skin homeostasis. Cell 2021 Apr 15;184(8):2151-2166. [doi: 10.1016/j.cell.2021.03.002] [Medline: 33765440]
4. Bircher AJ, Meier-Ruge W. Aquagenic pruritus: water-induced activation of acetylcholinesterase. Arch Dermatol 1988;124(1):84-89. [doi: 10.1001/archderm.1988.01670010048020] [Medline: 3337547]
5. Alinaghi F, Elberling J. Primary aquagenic pruritus successfully treated by β-alanine: a case report. Case Rep Dermatol 2024;17(1):252-254. [doi: 10.1159/000545842] [Medline: 40625687]
6. Concistrè A. Beta-alanine as a potential treatment for aquagenic pruritus: an online social media-based survey study. Acta Dermatovenerol Croat 2024 Dec;32(4):123-134. [Medline: 40657651]
7. Degala J, Degala A. The use of beta-alanine for the management of aquagenic pruritus. Ann Allergy Asthma Immunol 2025 Nov;135(5):S156-S157. [doi: 10.1016/j.anai.2025.08.469]

XSL•FO
RenderX

8. Zhang YR, Keshari S, Kurihara K, et al. Agonism of the glutamate receptor GluK2 suppresses dermal mast cell activation and cutaneous inflammation. Sci Transl Med 2024 Dec 11;16(777):eadq9133. [doi: 10.1126/scitranslmed.adq9133] [Medline: 39661706]
9. Saunders B, Elliott-Sale K, Artioli GG, et al. β-alanine supplementation to improve exercise capacity and performance: a systematic review and meta-analysis. Br J Sports Med 2017 Apr;51(8):658-669. [doi: 10.1136/bjsports-2016-096396] [Medline: 27797728]

Abbreviations

AP: aquagenic pruritus
MrgprD: MAS-related G protein–coupled receptor D

Edited by R Dellavalle; submitted 02.Jan.2026; peer-reviewed by A Concistre, D Kaplan; revised version received 10.Mar.2026; accepted 10.Mar.2026; published 25.Mar.2026.

Please cite as:
Piserchio N, Baratta B, Brooks B, Muse B, Ball K
Beta-Alanine and Aquagenic Pruritus: Proposed Neuroimmune Mechanism
JMIR Dermatol 2026;9:e90737
URL: https://derma.jmir.org/2026/1/e90737
doi:10.2196/90737

© Natalie Piserchio, Bailey Baratta, Benjamin Brooks, Brandon Muse, Katelin Ball. Originally published in JMIR Dermatology (http://derma.jmir.org), 25.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Agentic AI in Dermatology: A Call to Action

Brian Chu[1], MD; Heather Shen[2], MS; Ivy Lee[3], MD; Jules B Lipoff[4], MD

[1]Department of Dermatology, Brown University, 593 Eddy St, APC 10th Floor, Providence, RI, United States

[2]Praxis Labs, Providence, RI, United States

[3]Comprehensive Dermatology Center of Pasadena, Pasadena, CA, United States

[4]Department of Dermatology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, United States

Corresponding Author:
Brian Chu, MD
Department of Dermatology, Brown University, 593 Eddy St, APC 10th Floor, Providence, RI, United States

Abstract

Artificial intelligence (AI) tools are shifting from passive, user-initiated tools to proactive *agentic AI* systems that are capable of autonomous, multi-step actions. These agents can independently gather information, execute sequential tasks, and collaborate with humans or other agents without requiring constant prompting from humans. Early adopters in health care have demonstrated early feasibility across multiple specialties and clinical settings. Dermatology is well-positioned to benefit given its high patient volumes, administrative burdens, and clinicopathological workflows. To guide responsible adoption of agentic AI, we propose a risk-stratification framework based on clinical risk and task reversibility. Barriers to widespread adoption of agentic AI include limitations in model reliability, interoperability across health records, and unresolved questions around liability, privacy, and regulation. Dermatologists must proactively engage via professional organizations and industry partnerships to ensure that agentic AI is developed safely, equitably, and in alignment with our values.

(JMIR Dermatol 2026;9:e94909) doi:10.2196/94909

KEYWORDS

agentic AI; AI agents; AI in dermatology; artificial intelligence; technology in dermatology

From AI Tools to Agentic Teammates

Medicine is at an inflection point as artificial intelligence (AI) evolves from passive tools to proactive teammates, a concept known as *agentic AI* (throughout this article, the terms “agentic AI,” “AI agents,” and “agentic teammates” are used interchangeably) [1]. While current AI tools are powerful, they are passive: a user has to initiate every interaction, provide all context, and interpret all outputs. In comparison, autonomous AI agents act without prompting, perform sequential steps independently, gather information from multiple sources, collaborate with humans or other agents, and self-improve through feedback loops [1]. Essentially, AI agents can behave as teammates with assigned roles and responsibilities. Of note, agentic AI is distinct from traditional rule-based automation, which executes fixed logic on structured inputs. In contrast, AI agents use large language models (LLMs) to reason across unstructured data and handle novel situations, which go beyond what deterministic automation can achieve.

Large health care systems have already begun leveraging agentic AI to reduce administrative burden and improve quality of care. For example, Duke Health intends to deploy AI agents in cardiology for care coordination that autonomously handle scheduling requests, optimize resource allocation, and connect patients to clinical trials [2]. At Ochsner Health, AI agents continuously review patient panels of primary care providers, reaching out to patients for screenings, lab monitoring, and postdischarge care coordination [3]. Oxford University Hospitals is piloting AI agents in oncology that summarize charts, determine tumor staging, and draft guideline-compliant plans for multidisciplinary tumor boards [4]. These early implementations demonstrate the feasibility of agentic AI across multiple specialties and clinical settings.

Opportunities and Risks for Dermatology

Dermatology is particularly well-positioned to benefit from these emerging AI capabilities given its high patient volumes, substantial administrative burdens, and clinicopathological diagnostic approach [5]. In a future dermatology practice, AI agents could tackle high patient volumes through intelligent scheduling (eg, spacing out procedures and complex visits), previsit charting, scribing and coding, and automated patient follow-up. Voice agents can reach even more patients through phone calls, including older adults and patients with limited English proficiency. AI agents can identify and recruit patients for clinical trials, particularly to address the critical need for diverse representation. Medication access (eg, prior authorization, appeals, patient assistance programs, and laboratory monitoring) for specialty drugs like biologics is also well-suited for an agentic workflow that eases the burden of manual labor on support staff. The unique clinicopathological workflow in dermatology also lends itself to multistep agentic

XSL•FO
RenderX

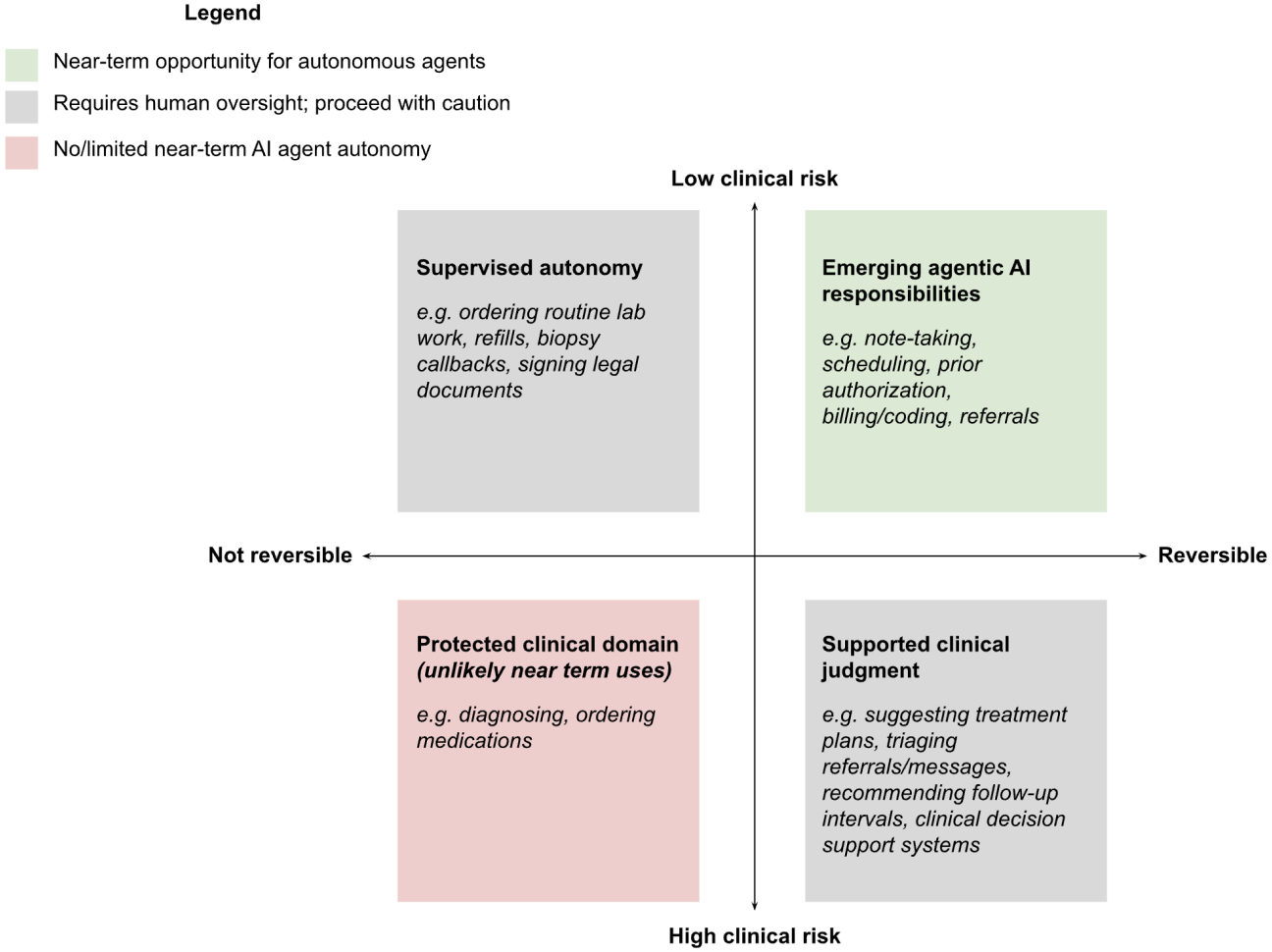
coordination: an agentic system could receive a pathology result, verify the correct patient and anatomic site against prior documentation and images, alert the dermatologist, and coordinate definitive treatment (eg, reaching out to a Mohs surgeon).

Despite the promising applications, there are limited published case studies of agentic AI in the dermatology literature. Nonetheless, there are numerous companies that have been launched in the past few years that are targeting every aspect of the dermatology workflow with an agentic approach. As health care systems and other specialties begin establishing the workflows, norms, and business models for agentic AI, dermatologists must actively educate themselves on this technology and engage to shape its real-world implementation.

Deploying agentic AI requires balancing risk; when considering pilot implementations, dermatologists should take a graduated approach for delegating tasks to autonomous agents. To assess readiness for agentic AI in dermatology practices, we propose the following framework (Figure 1): promising agentic applications ready for pilots focus on reversible, low-risk tasks. Tasks that are irreversible or carry higher clinical risk warrant stronger oversight by dermatologists or staff and therefore require additional caution before fully adopting agentic teammates. High-stakes, irreversible decisions with major clinical risk are not suitable for delegation to AI agents in near-term, real-world settings.

Significant barriers also impede the widespread deployment of AI agents in dermatology. First, current AI models are not ready for the reliability and complexity necessary in real-world clinical settings. A recent study demonstrated that the best-performing LLMs successfully completed 70% of representative clinical tasks, such as medication and test ordering, referrals, documentation, and patient communication. Models struggled most on tasks that required modifying records rather than simply retrieving information, highlighting the need for stepwise adoption [6]. Second, interoperability must be improved so that agents can gather data from siloed health records. Even with full access, agents need the right permissions to actually take action on behalf of dermatologists. Third, fundamental questions need to be addressed about liability, privacy, and regulation, as historically, regulatory agencies judge an AI-enabled medical device for a specific task, rather than a broad range of interconnected tasks [1,7].

Figure 1. Autonomous agentic AI is most ready for tasks that are reversible and of low clinical risk. When tasks are irreversible or of high clinical risk, stronger oversight with a human dermatologist is warranted. Tasks that are both irreversible and of high clinical risk (eg, diagnosing, ordering medications) are unlikely to be ready for near-term adoption and should have limited AI agent autonomy. The placement of example tasks within this framework is intended as a general guide, and exact positioning varies depending on institutional workflows, local staffing models, or electronic health record (EHR) capabilities.



XSL•FO
RenderX

Why Dermatologists Must Lead Now

Though this technology may seem distant for the practicing dermatologist, decisions about agentic AI are being made now. National and international dermatology organizations should include agentic AI on their agendas for AI use, seek partnerships with companies, and write guidelines for deploying AI agents. Unlike previous AI guidance focused on diagnostic performance, agentic AI guidelines must address multistep reasoning, minimum performance standards, mechanisms requiring human oversight, and protocols for monitoring ongoing safety. Concrete near-term actions include establishing a dedicated agentic AI task force within professional societies such as the American Academy of Dermatology, developing standardized reporting criteria for pilot programs in dermatology practices, and creating a voluntary registry of deployments to facilitate shared learning.

Research should also specifically examine agent performance across diverse patient populations to mitigate bias [7]. Although there are well-documented performance disparities of *diagnostic* AI tools on darker skin tones, it is unclear whether agentic AI systems that focus on nondiagnostic tasks will face the same challenges. However, agentic AI systems that rely on flawed components risk amplifying these disparities at scale, making them invisible in automated workflows. Safeguards should include ongoing postdeployment monitoring across demographic subgroups.

Agentic AI reflects an emerging paradigm where intelligent, autonomous agents act on behalf of physicians and patients, enhancing efficiency, continuity, and patient experience. There is no substitute for the human physician–patient relationship, and agentic AI offers dermatologists an opportunity to strengthen this relationship, not replace it. Dermatologists must not only stay informed of advances in agentic AI; we must also actively shape decisions about agentic AI to ensure appropriate, ethical, and optimized deployment to reflect our values and patients' needs.

Acknowledgments

All ideas, perspectives, and content are entirely original and written by humans. Generative AI tools were used as a collaborator to discuss strengths and weaknesses of the article, perform literature review, and suggest copyedits. Models used included OpenAI ChatGPT 5.1 and Claude Sonnet 4.5.

Conflicts of Interest

None declared.

References

1. Zou J, Topol EJ. The rise of agentic AI teammates in medicine. Lancet 2025 Feb 8;405(10477):457. [doi: 10.1016/S0140-6736(25)00202-8] [Medline: 39922663]
2. Talhelm M. DukeHealth. Duke Health, Trase Systems partner to develop AI health care tools. 2025 Sep. URL: https://corporate.dukehealth.org/news/duke-health-trase-systems-partner-develop-ai-health-care-tools [accessed 2026-05-20]
3. Madden B. AI-enabled primary care is no longer optional. Hospitalogy. 2025 Dec. URL: https://hospitalogy.com/articles/2025-12-04/ai-enabled-primary-care-is-no-longer-optional/ [accessed 2025-12-02]
4. Oxford-built multi-agent assistant for cancer care to be piloted in collaboration with Microsoft.: University of Oxford; 2025 Dec. URL: https://www.ox.ac.uk/news/2025-12-16-oxford-built-multi-agent-assistant-cancer-care-be-piloted-collaboration-microsoft [accessed 2025-12-16]
5. Omiye JA, Gui H, Daneshjou R, Cai ZR, Muralidharan V. Principles, applications, and future of artificial intelligence in dermatology. Front Med 2023;10:1278232. [doi: 10.3389/fmed.2023.1278232]
6. Jiang Y, Black KC, Geng G, et al. MedAgentBench: a virtual EHR environment to benchmark medical LLM agents. NEJM AI 2025 Aug 28;2(9). [doi: 10.1056/AIdbp2500144]
7. Arza A, Lebhar J, Lipoff JB. Applications of artificial intelligence in dermatology: ethical considerations. Dermatol Clin 2025 Oct;43(4):529-540. [doi: 10.1016/j.det.2025.05.003] [Medline: 41093475]

Abbreviations

AI: artificial intelligence
LLM: large language model

XSL•FO
RenderX

Edited by M Kirchberger; submitted 08.Mar.2026; peer-reviewed by M Hasselwander, S Drew; revised version received 19.Apr.2026; accepted 19.Apr.2026; published 05.Jun.2026.

Please cite as:
Chu B, Shen H, Lee I, Lipoff JB
Agentic AI in Dermatology: A Call to Action
JMIR Dermatol 2026;9:e94909
URL: https://derma.jmir.org/2026/1/e94909
doi:10.2196/94909

© Brian Chu, Heather Shen, Ivy Lee, Jules B Lipoff. Originally published in JMIR Dermatology (http://derma.jmir.org), 5.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Climate, Humidity, and Population-Level Interest in Dry Skin: Infodemiology Analysis Using Google Trends Across the United States

Kimiya Aframian[1*], BS; Shaya Naimi[2*], BS; Joy Xu[3], BS; Gordon Bae[1], MD

[1]Department of Dermatology, Stanford Medicine, Stanford, CA, United States

[2]Department of Sciences, CU Boulder Aerospace Engineering Sciences, Boulder, CO, United States

[3]Department of Medicine, David Geffen School of Medicine, 10833 Le Conte Ave, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Joy Xu, BS
Department of Medicine, David Geffen School of Medicine, 10833 Le Conte Ave, Los Angeles, CA, United States

Abstract

Background: Climate and weather factors of temperature and humidity are widely reported to be associated with xerosis (dry skin), a common inflammatory skin condition and frequent driver of pruritus (itchy skin) and reduced quality of life. Growing evidence supports links between environmental conditions and skin barrier function, with extreme climates associated with increased atopic dermatitis–related clinical visits. Mechanistically, temperature and humidity affect the stratum corneum, the skin's primary permeability barrier, with low humidity and high temperature increasing transepidermal water loss and promoting cutaneous inflammation.

Objective: This study examines the relationship between climate, namely temperature and humidity, and the general public's experience in dry skin and moisturizing products, throughout the United States. This study sought to address gaps in traditional epidemiologic approaches by linking climate conditions with population-level online search behavior related to dry skin and moisturizer use across the United States.

Methods: Publicly available climate data were obtained from the National Oceanic and Atmospheric Administration (NOAA), including average temperature and dew point by state over a recent nine-year period (2016 - 2025). Dew point served as a proxy for ambient humidity. Google Trends was used to assess relative search interest for five dry skin– and moisturizer-related terms by state during the same period. Search interest was normalized per million residents, and associations between climate variables and search interest were evaluated using linear regression analyses. Statistical analyses were conducted using R.

Results: Lower average temperatures and lower dew points were associated with higher dry skin–related search interest, while warmer, more humid states showed lower interest. Both temperature and dew point demonstrated significant negative associations with Google search interest. This work was not funded and data collection was performed using publicly available, free databases.

Conclusions: Population-level search behavior related to xerosis reflects national patterns of climate-associated dermatologic burden.

(JMIR Dermatol 2026;9:e93639) doi:10.2196/93639

KEYWORDS

xerosis; dry skin; humidity; dew point; temperature; google trends; infodemiology; United States

Introduction

Climate and weather factors of temperature and humidity are widely reported to be associated with xerosis (dry skin), one of the most common inflammatory skin conditions and frequent drivers of pruritus and reduced quality of life [1]. Although the mechanism is not fully understood, there is a widely growing body of evidence supporting meaningful links between environmental conditions and skin barrier function; regions characterized by extreme environmental conditions (temperature variability, precipitation patterns, sunlight exposure, and airborne pollutants) are associated with higher rates of atopic dermatitis (AD)–related clinical visits [2,3].

Mechanistically, temperature and humidity directly affect the stratum corneum, the skin's primary permeability barrier of highly organized extracellular lipids which limit transepidermal water loss (TEWL) [4]. Factors that increase evaporation from the stratum corneum, such as low humidity and high temperature, impair corneocyte cohesion and promote cutaneous inflammation. Prior studies demonstrate the impacts of dry

XSL•FO
RenderX

environment on increasing TEWL and impaired barrier repair, and high humidity on barrier restoration [5]. Such effects may be amplified in individuals predisposed to xerosis and eczematous dermatoses who already have baseline defects in barrier function [6]. Thus, “cold and dry” conditions may increase AD prevalence and flare risk, and cohort data suggests that climate factors may modulate symptom control in potentially region-specific methods [7].

Moisturizer is the mainstay management for xerosis and AD, so climate variability may plausibly be reflected in variation in skin-care needs on a regional-scale. Traditional epidemiologic analysis of the environment-skin relationship often relies on clinic visits, prescription, and registry records, however these methods do not represent subclinical symptoms and self-management behaviors outside of the formal healthcare environment [8]. Especially with the increase in accessibility to freely-available medical information −83.4% of patients in 2020 report using the Internet prior to their dermatologic appointment - it is critical to consider new metrics of analysis [9].

Google Trends is a freely accessible analytics platform which allows for assessment of Google search query volumes and has increasingly been used as an infodemiologic tool to character population-level health behaviors [10]. Search activity is reflective of self-triage and self-treatment efforts, both before and after clinical visits, offering a potential window into symptom burden captured beyond administrative data [11]. Specifically in dermatologic epidemiology, Google Trends data has been analyzed on a global scale to study temporal and geographic patterns of incidence, symptoms, and treatment of common conditions including eczema, acne, psoriasis, and skin cancer [12].

The aim of this paper is to understand the relationship between climate conditions and population-level interest in dry skin symptoms through linking of NOAA-reported temperature and humidity (dewpoint used as proxy) with Google Trends search interest across each US state for moisturizer-related and dry-skin symptom terms.

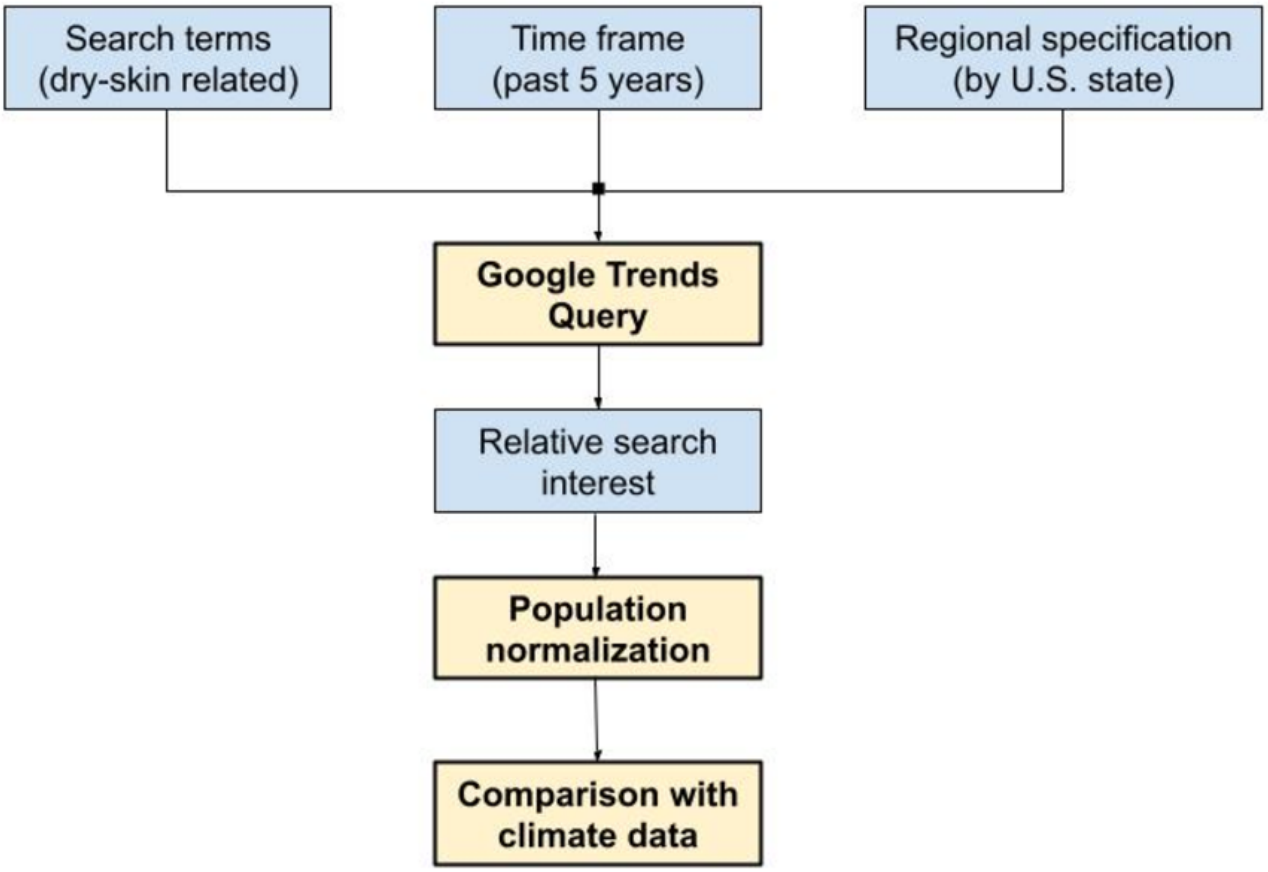
Methods

Google Trends Data

The Google Trends software allows the public to query the search interest of one or more particular search terms over a user-specified period of time and region, and allows the user to compare regions, time frames, and search terms (Figure 1). When queried, the software tool yields relative search interest between regions (eg, US states), normalized between 0 and 100 across the data set, as opposed to absolute search volume. According to this scale, a value of 0 indicates the lowest search popularity and a value of 100 indicates the highest relative search interest amongst the data set. To obtain data for this study, the Google Trends software tool was utilized to query search interest in five search terms related to dry skin and moisturizer: “eczema,” “dry skin,” “moisturizer,” “atopic dermatitis,” and “atopic eczema” over the same 9-year period specified above for climate data and by state. Relative search interest was normalized by each state’s population. Note that this additional step of population normalization is independent of the “normalized” data provided by each Google Trends query; Google Trends provides relative search interest normalized against the highest value in the dataset, which does not involve population. To normalize by state population, each state’s relative interest was divided by its population to obtain “relative search interest per million people” (Figure 1).

XSL•FO
RenderX

Figure 1. Outline of methodology. Google Trends query parameters were gathered, and the output of the query was normalized by state population prior to comparison against climate data.



Climate Data

Publicly available data were obtained from the National Oceanic and Atmospheric Administration (NOAA), a US government agency that studies the earth's oceans, atmosphere, and climate and publishes robust datasets available to the public. In particular, the NOAA offers the Climate Data Online (CDO), a web page that provides access to global historical weather and climate data, including daily, monthly, and hourly measurements of temperature, humidity, wind, and precipitation [13]. For this study, hourly measurements of average temperature and dew point by state were extracted over a recent 9-year period (2016 - 2025). Note that dew point is a widely known excellent and often preferred proxy for humidity, as it measures air moisture more directly and independently of temperature [14]. For each state, these measurements were averaged over the recent 9-year period to yield average temperature and dew point by state over the given time frame.

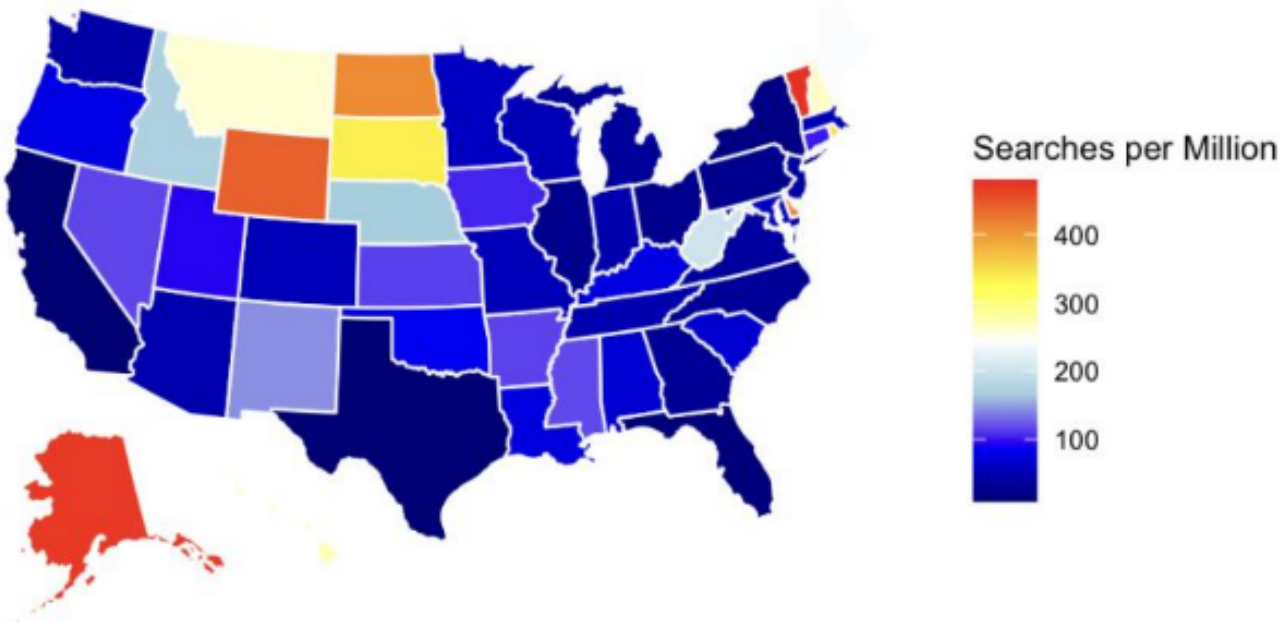
Results

Google Trends Data

Analysis of Google Trends data at the state level demonstrates substantial geographic variability in search interest for moisture-related and dry skin-associated search terms across the U.S. Higher relative search interest was observed in western and southwestern states, whereas lower search interest was observed in northeastern and midwestern states (Figure 2). Note that Google search interest was normalized per million residents, accounting for variability in population per state.

XSL•FO
RenderX

Figure 2. Google Trends Search Interest per capita by US state 2016 - 2025 period. This map was generated using data from the Google Trends software tool.



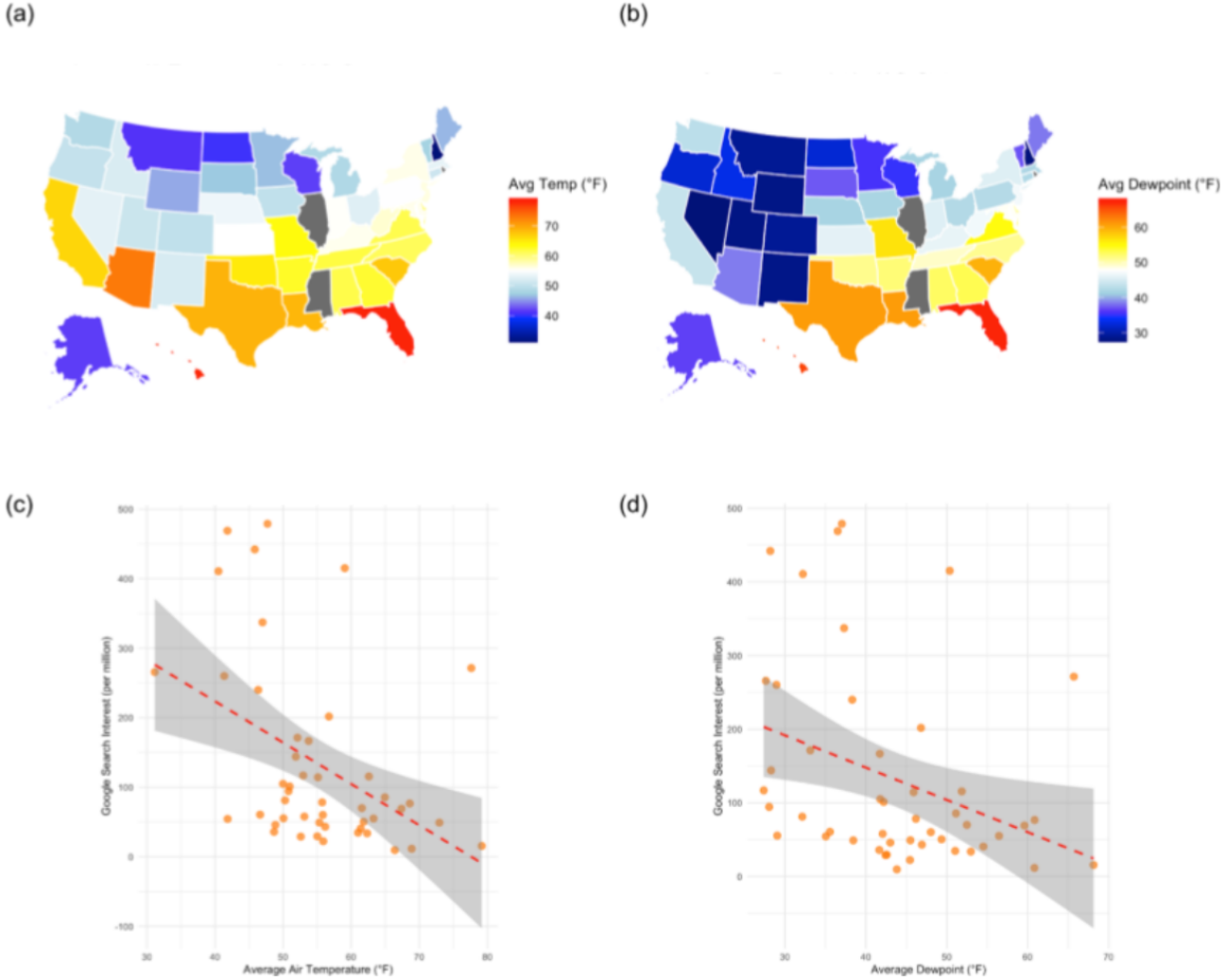
Climate Data

Comparison of the Google Trends data with NOAA-reported climate variables reveal that states such as Vermont and Wyoming, among those with the highest search interest, generally exhibited cooler average temperatures and moderate dew point values, although variation across states was observed. (Figure 3A and C). Examining temperature at the state level, Google search interest (mean 132.79, SD 132.99) was negatively correlated with Average Air Temperature (°F) (mean 55.35, SD 9.8), r(45) = −0.428, 95% CI [−0.637,−0.16], *P*=.002. Statistical analyses were conducted using R, and regression assumptions, including linearity and normality of residuals, were assessed using visual inspection of diagnostic plots.

Average dewpoint is used as a proxy for ambient humidity, showing a similar pattern as temperature on a national level. States with a lower average dewpoint and thus presumably decreased humidity conditions, especially western interior states, exhibited higher relative search interest for moisturizers and dry skin-related symptoms (Figure 3B). Contrarily, states with higher average dewpoints such as the humid southeastern states demonstrated lowest search interest in these search terms. Average dewpoints ranged from 28 degrees Fahrenheit minimum to 68 degrees Fahrenheit maximum. Similarly to temperature, at the state level, Google search interest (M=132.79, SD=132.99) was negatively correlated with Average Dewpoint (°F) (mean 43.5, SD 10.51), r(45) = −0.338, 95% CI [−0.57,−0.056], *P*=.02. (Figure 3D).

XSL•FO
RenderX

Figure 3. (A) Average air temperature by US state over 2016 - 2025 period. (B) Average dewpoint by US state over 2016 - 2025 period. (C) Plot illustrating the relationship between average dry bulb temperature (air temperature) and Google search popularity, averaged over the 2016 - 2025 period, where each point represents one US state. There is a significant downward trend with Pearson r=-0.442 and P=.002. (D) Plot illustrating the relationship between average dewpoint (proxy for humidity) and Google search popularity, averaged over the 2016 - 2025 period, where each point represents one US state. There is a significant downward trend with Pearson r=-0.348, and P=.02. All plots were generated using data from the National Oceanic and Atmospheric Administration.



Discussion

Principal Findings

This study found a significant negative correlation between average air temperature and Google search interest in dry skin-related terms per million residents, as well as a significant negative correlation between average dewpoint and Google search interest.

Regional humidity, rather than temperature alone, more closely aligns with patterns of dermatologic search interest according to geographic visualization of climate variables. Western inland states are characterized by lower average dewpoints and higher evaporation demand. These states consistently exhibit elevated Google Trends search interest in our queried terms. In contrast, coastal and southeastern states with high relative humidity and high dewpoints showed lower search interest despite the higher acreage annual temperatures. These findings suggest that atmospheric moisture content, represented by dew point, may be an important contributor to patterns of dermatologic search interest. While these findings demonstrate an association rather than causation, possible mechanisms underlying this relationship have been proposed. For example, temperature and humidity have been shown to influence skin barrier integrity, whereby low humidity may increase evaporation from the skin surface and contribute to skin dryness.

These findings cumulatively represent a robust association between climate conditions and Google search behavior related to moisturizer and dry skin symptoms across the United States. To our knowledge, this study represents an early effort to use Google Trends data as a digital epidemiologic tool to explore geographic variation in dry-skin related search activity across the United States. While search intent does not directly equate to disease prevalence, this approach provides a means of visualizing regional patterns, including clusters of elevated interest that are frequently observed in coastal areas. Google trends has also successfully predicted the outbreak of many viruses [15], indicating its validity as an epidemiological proxy.

XSL•FO
RenderX

Limitations

Limitations of this study include the difference in population per state. Normalization by population in search interest is reasonably assumed to be proportional to population, but states with extremely high or low populations may have resulted in skewed relative search interest per capita. Populations with limited internet access and healthcare utilization as well as lower socioeconomic status may be poorly represented. Demographic differences between states may also account for differences in data. Additionally, Google trends queries return relative search interest normalized between 0 - 100 across the data set instead of absolute search volumes, regardless of the data set being queried. Thus, reported findings are correlational and cannot be assumed as causation. In addition, analysis used multiyear averaged data, which may obscure important seasonal variation in skin dryness and related search behavior. These temporal patterns may not be fully captured in the present approach.

Conclusion

This study finds a strong correlation between NOAA reported temperature and humidity and state-specific Google searches, revealing that online search behavior reflects the general public's experience of the climate across the states. In other words, Americans who live in hotter and drier states are more likely to seek out moisturizer and to experience symptoms of dry skin conditions such as eczema and dermatitis. These findings reveal the potential for Google trends to monitor the general public's awareness of their dermatological response to varying climates, and can ultimately allow physicians to tailor preventative skin-care recommendations and emphasize proactive skin barrier protection in more predisposed environments. These findings highlight the potential for infodemiology tools to complement traditional dermatologic surveillance and may help guide future studies on environmental influences on skin disease.

Furthermore, physicians and medical systems at large could consider prioritization of access to moisturizers and barrier-rapid therapy in higher-risk regions with anticipated increased xerosis burden.

Funding

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

Data Availability

Data are available upon request.

Authors' Contributions

KRA & SN conducted manuscript-writing, JX provided critical revisions and edits, GHB oversaw conceptualization and guidance.

Conflicts of Interest

None declared.

References

1. Bylund S, Kobyletzki LB, Svalstedt M, Svensson Å. Prevalence and incidence of atopic dermatitis: a systematic review. Acta Derm Venereol 2020 Jun 9;100(12):adv00160. [doi: 10.2340/00015555-3510] [Medline: 32412646]
2. Hui-Beckman JW, Goleva E, Leung DYM, Kim BE. The impact of temperature on the skin barrier and atopic dermatitis. Ann Allergy Asthma Immunol 2023 Dec;131(6):713-719. [doi: 10.1016/j.anai.2023.08.007] [Medline: 37595740]
3. Park M, Mohsen ST, Katz T, et al. Climate conditions, weather changes, and air pollutants and atopic dermatitis: a meta-analysis. JAMA Dermatol 2025 Aug 1;161(8):828-839. [doi: 10.1001/jamadermatol.2025.1790] [Medline: 40560568]
4. Elias PM. Skin barrier function. Curr Allergy Asthma Rep 2008 Jul;8(4):299-305. [doi: 10.1007/s11882-008-0048-0] [Medline: 18606081]
5. Denda M, Sato J, Masuda Y, et al. Exposure to a dry environment enhances epidermal permeability barrier function. J Invest Dermatol 1998 Nov;111(5):858-863. [doi: 10.1046/j.1523-1747.1998.00333.x] [Medline: 9804350]
6. Cork MJ, Danby SG, Vasilopoulos Y, et al. Epidermal barrier dysfunction in atopic dermatitis. J Invest Dermatol 2009 Aug;129(8):1892-1908. [doi: 10.1038/jid.2009.133] [Medline: 19494826]
7. Engebretsen KA, Johansen JD, Kezic S, Linneberg A, Thyssen JP. The effect of environmental humidity and temperature on skin barrier function and dermatitis. J Eur Acad Dermatol Venereol 2016 Feb;30(2):223-249. [doi: 10.1111/jdv.13301] [Medline: 26449379]
8. Hamann CR, Andersen YMF, Engebretsen KA, et al. The effects of season and weather on healthcare utilization among patients with atopic dermatitis. J Eur Acad Dermatol Venereol 2018 Oct;32(10):1745-1753. [doi: 10.1111/jdv.15023] [Medline: 29706020]
9. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: Search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]

XSL•FO
RenderX

10. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: a scoping review. JMIR Public Health Surveill 2019;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]
11. Sivesind TE, Szeto MD, Kim W, Dellavalle RP. Google Trends in dermatology: scoping review of the literature. JMIR Dermatol 2021 May 25;4(1):e27712. [doi: 10.2196/27712] [Medline: 37632813]
12. Kamiński M, Tizek L, Zink A. Dr. Google, What Is That on My Skin?'-Internet Searches Related to Skin Problems: Google Trends Data from 2004 to 2019. Int J Environ Res Public Health 2021 Mar 4;18(5):2541. [doi: 10.3390/ijerph18052541] [Medline: 33806391]
13. NOAA. URL: https://www.ncei.noaa.gov/cdo-web/ [accessed 2026-03-25]
14. National Weather Service. Analysis information. URL: https://www.weather.gov/mfl/analysisinfo [accessed 2024-12-01]
15. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature New Biol 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]

Abbreviations

AD: atopic dermatitis
NOAA: National Oceanic and Atmospheric Administration
TEWL: transepidermal water loss

Edited by S Subhadarshani; submitted 16.Feb.2026; peer-reviewed by S Adnan; revised version received 13.Apr.2026; accepted 02.May.2026; published 25.May.2026.

Please cite as:
Aframian K, Naimi S, Xu J, Bae G
Climate, Humidity, and Population-Level Interest in Dry Skin: Infodemiology Analysis Using Google Trends Across the United States
JMIR Dermatol 2026;9:e93639
URL: https://derma.jmir.org/2026/1/e93639
doi:10.2196/93639

© Kimiya Aframian, Shaya Naimi, Joy Xu, Gordon Bae. Originally published in JMIR Dermatology (http://derma.jmir.org), 25.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Dermatology, is properly cited. The complete bibliographic information, a link to the original publication on http://derma.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX