doi.org/10.58931/cdt.2026.72164

ABOUT THE AUTHOR

Sheila Wang, MD, PhD, FRCPC

Dr. Sheila Wang is an Assistant Professor of Dermatology in the Department of Medicine at the University of Toronto, a Staff Dermatologist and Clinician-Investigator at Women's College Hospital, and the co-founder of Swift Medical, a leader in AI-powered imaging and digital dermatology. Her work bridges dermatology, wound care, and data science, advancing how skin and wound conditions are assessed, monitored, and implemented in real-world practice, particularly for patients with skin of colour and those facing barriers to care. She leads a digital health research program focused on AI and advanced imaging in wound care and inflammatory skin disease, with over 11 million dollars in research funding. She is the co-founder of Swift Medical, an AI-powered imaging platform deployed in more than 5,200 facilities and used to monitor over one million patients annually. Her contributions to innovation and health equity have been recognized with honours including the Governor General's Innovation Award, the Canadian Medical Association Young Leader Award, the Joule Innovation Award, and the American Academy of Dermatology Quality Improvement Award.

Affiliations: University of Toronto, Toronto, ON
Women's College Hospital, Toronto, ON

Seeing More Than Meets the Eye: Artificial Intelligence-Based Imaging in Dermatology and the Future of Equitable Care

Sheila Wang, MD, PhD, FRCPC

Introduction

Artificial intelligence (AI)-based imaging is rapidly reshaping how skin disease is documented, triaged, and monitored, yet its clinical adoption still lags far behind its technological promise.[1-10] This gap is especially pressing for patients with skin of colour (SOC) and others who already experience barriers to dermatologic care.

What AI Actually Does in Dermatology

AI refers to computer programs designed to perform tasks that simulate human intelligence, including learning from data, recognizing patterns, and making decisions. Within health care, AI systems are increasingly used to support clinicians by enabling more consistent interpretation of complex information at scale and is intended to augment, rather than replace, clinical judgment.[1,2]

Figure 1. Smartphone-based AI wound imaging at the point of care; A clinician uses a smartphone application to capture a standardized wound image with automated boundary detection and area calculation, illustrating how AI-enabled imaging can support objective measurement and longitudinal monitoring in everyday clinical settings, including primary care and community practice; Smartphone-based wound imaging was performed using the Swift Ray 1 platform; *courtesy of Swift Medical Inc., Toronto, Canada.*

In dermatology, most current tools fall under visual AI, or computer vision, which focuses on interpreting images and videos to support clinical tasks. Common imaging tasks include classification (assigning a diagnosis), detection (localizing a lesion), and segmentation (outlining a lesion), each supporting different aspects of care such as triage, biopsy decision-making, and longitudinal follow-up. Deep learning approaches, particularly convolutional neural networks, can learn subtle image features directly from pixel data and have demonstrated dermatologist-level performance for certain skin cancer classification tasks in research settings.[3,11]

Because skin disease can manifest across multiple visual scales, AI methods have been applied to clinical photography, dermoscopy, dermatopathology, and teledermatology imaging. In practice, many systems are deployed as clinical decision-support tools that suggest likely diagnoses, highlight regions of interest, or provide standardized measurements over time rather than offering stand-alone diagnostic decisions (**Figure 1**). More recently, vision-language and other multimodal architectures have been proposed that integrate both image and textual information, creating opportunities for richer, context-aware decision support.[3,12]

Despite rapid growth in the research literature, regulatory approval of dermatology AI tools remains limited. A recent review identified only a small number of AI-based dermatology devices with regulatory clearance worldwide, including just a few systems authorized by the United States Food and Drug Administration, in sharp contrast to the hundreds of AI-enabled devices in radiology. Lengthy medical device approval timelines and the relative novelty of adaptive algorithms within existing regulatory frameworks contribute to this slower uptake.

Why Performance in Practice Often Falls Short

Many AI systems demonstrate impressive accuracy in curated research datasets but perform less reliably in day-to-day practice. This performance gap reflects how dermatologic data are collected, labelled, and evaluated.[8,10,13-15]

Most training datasets comprise images captured under relatively controlled lighting and camera conditions and are often drawn from specialist centres rather than primary care clinics or patient homes. In these settings, models may inadvertently learn spurious correlations, such as device-specific artifacts, background patterns, or image quality, rather than lesion morphology itself, leading to brittle performance when deployed across different environments.

Establishing robust ground truth represents another challenge. Many datasets rely on single-clinician assessments without histopathologic confirmation or longitudinal follow-up, and details of labelling workflows are often underreported.[5,8,10,13-15] Well-documented inter-rater variability in diagnosis and severity scoring, especially for inflammatory skin diseases, further undermines label reliability and limits its generalizability. Together, these issues contribute to discrepancies between algorithm performance reported in development studies and effectiveness in routine care.

External validation is frequently limited, and relatively few studies evaluate AI performance prospectively in real-world workflows. A systematic review comparing AI systems with clinicians for skin cancer diagnosis found that most studies relied on retrospective data rather than prospective trials. In studies that have included randomized or prospective designs, AI assistance tends to function best as an adjunct to clinician judgment, helping to improve diagnostic accuracy particularly for generalist physicians, rather than as an independent diagnostic system.[2]

The Blind Spot: Skin of Colour and Dataset Bias

Underrepresentation of SOC is one of the most critical limitations of current dermatologic AI systems. Many training datasets disproportionately contain images of lighter skin tones and classic "textbook" disease presentations, whereas darker skin, early or subtle disease, and treatment-modified lesions are often underrepresented. These imbalances mirror long-standing inequities in dermatology education, research, and workforce representation and risk perpetuating the same disparities within AI tools.[5,7,16,17]

Dermatology also lacks a robust, standardized approach to measuring skin tone for AI development and evaluation. Much of the literature has used the Fitzpatrick skin type, a classification system designed to capture sun reactivity rather than actual skin pigmentation and performs poorly in those with SOC.[16,17] This limitation can distort both malignancy risk assessment and evaluations of model performance across skin tones. Without careful skin tone characterization and stratified reporting, apparent "overall" accuracy can mask substantially worse performance in darker skin.[5,7,16,17]

Alternative frameworks, such as the Monk Skin Tone Scale, were specifically developed to better represent a continuum of pigmentation across diverse populations and may offer a more appropriate foundation for skin tone-stratified evaluation in dermatologic AI.[18] Regardless of the specific scale, intentional sampling strategies, transparent reporting of skin tone distribution, and skin tone-stratified performance metrics should be adopted as standard practice.

Evidence from curated image datasets and systematic reviews demonstrates that AI models may show lower sensitivity for melanoma in patients with SOC compared with pooled performance across all skin types, raising concern that uncritical deployment could inadvertently exacerbate existing outcome gaps. At the same time, much of the oncology-focused AI literature has been generated in predominantly lighter-skinned populations, highlighting a clear need for more inclusive data and evaluation.[5,7,16,17]

Beyond the Image: Multimodal and Context-Aware Models

Most dermatology AI tools are image-only models that do not incorporate clinical history, risk factors, or treatment context. In everyday practice, morphology is only one part of the diagnostic puzzle, particularly for inflammatory dermatoses, drug eruptions, pigmentary disorders, and complex presentations where pattern recognition alone is insufficient.[4,5,19]

Multimodal approaches that combine imaging with structured metadata, such as age, sex, cancer history, or lesion evolution, have already demonstrated superior performance to image-only models for common malignant and benign lesions.[5] For example, models that combine smartphone photographs with patient metadata have improved overall diagnostic accuracy and discrimination for distinguishing between malignancies and benign lesions in clinically relevant classes (**Figure 2**). Vision-language systems and transformer-based architectures are also emerging, offering the potential to jointly process clinical notes, patient-reported symptoms, and images in a single pipeline.[3]

From a clinician's perspective, AI is most often valuable not as a source of definitive diagnoses but as a tool for risk stratification, helping determine which lesions can be monitored, which warrant biopsy, and which require urgent referral.[4-7,11,13] In this context, calibrated risk outputs, clear uncertainty estimates, and intuitive visual explanations may be more useful than rigid categorical labels.

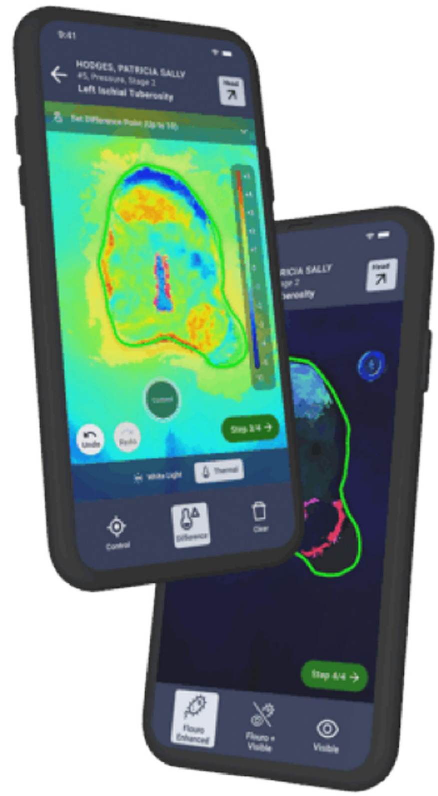

Figure 2. Thermal and segmentation-based assessment of cutaneous inflammation; Example screenshots from an AI-enabled mobile platform showing thermal maps and automated lesion segmentation, demonstrating how multimodal imaging can highlight areas of subclinical inflammation and provide consistent, quantitative metrics for disease activity over time, particularly useful for poorly demarcated inflammatory dermatoses in people with skin of colour; *courtesy of Swift Medical Inc., Toronto, Canada.*

Health Equity: Risks and Opportunities

The same technologies that risk amplifying disparities can, if deliberately designed, help reduce them.[4,5,8,16,17,18,20] AI-augmented teledermatology has expanded access to specialist expertise through both real-time video and asynchronous "store-and-forward" models. For patients in remote, rural, or otherwise underserved settings, AI-enabled mobile applications can provide preliminary lesion assessment, triage recommendations, or monitoring support that might otherwise be unavailable.

Patients from marginalized groups may view algorithmic decision-making as one potential way to reduce exposure to implicit bias encountered in traditional care, although this depends heavily on the transparency and validation of the tools. Equity-centred design therefore requires diverse training data, rigorous bias assessment, transparent reporting, and meaningful engagement with the communities most likely to be affected.[5,15-17]

Large language models (LLMs) can also help address communication barriers in dermatologic care. Early evidence suggests that LLMs can simplify patient-education materials, improve readability, and generate multilingual content tailored to different literacy levels, which may be particularly valuable for newcomers and patients with limited health literacy.[4,5,8,16,17,20] Wearable and sensor-based technologies, including devices that monitor scratching or skin temperature, offer additional opportunities for low-cost, remote monitoring of chronic skin disease, particularly in settings with limited specialist access.

Building Trustworthy, Clinically Useful AI

Moving from proof-of-concept models to trustworthy, widely adopted tools in dermatology will require sustained attention to data quality, transparency, and evaluation.[4,5,8,10,13-15]

Several priorities stand out:

- Representative training and validation datasets that intentionally include diverse skin tones, care settings, body sites, ages, and disease severities, with explicit reporting of these characteristics.
- Robust ground truth anchored in histopathology or well-documented clinical follow-up wherever feasible, with clear reporting of labelling methods and inter-rater agreement when consensus labels are used.[4,5,8,10,13-15]
- Prospective, real-world studies embedded in clinical workflows to evaluate impact on diagnostic accuracy, biopsy rates, wait times, and patient outcomes.[2]
- Routine skin tone-stratified reporting of sensitivity, specificity, and calibration, accompanied by transparent discussion of any performance gaps and efforts to address them.
- Human-AI collaboration models designed to augment clinician judgment, supported by intuitive interfaces, clear communication of uncertainty, and integration into existing electronic medical record systems.
- Post-deployment monitoring frameworks to track safety, performance drift, and equity across patient groups over time.[5,8-10,14,15]

Dermatology is a highly visual, pattern-based specialty in which AI holds intuitive appeal. Yet the conditions most in need of improved access and equity, including chronic inflammatory diseases, pigmentary disorders, and malignancies in patients with SOC, are also those least likely to be well represented in existing training datasets. Thoughtfully designed, clinically integrated, and equity-focused AI has the potential to move the field beyond isolated lesion classification toward comprehensive, longitudinal, and patient-centred skin health, allowing us to see, and act on, more than meets the eye.[4-6,8,10,11,13,16-18,20]

Correspondence

Sheila Wang, MD, PhD, FRCPC
Email: sheila@swiftmedical.com

Financial Disclosures

S.W.: Co-founder: Swift Medical

References

1. Amisha A, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. J Family Med Prim Care. 2019;8(7):2328-2331. doi: 10.4103/jfmpc.jfmpc_440_19
2. D'Adderio L, Bates DW. Transforming diagnosis through artificial intelligence. NPJ Digit Med. 2025;8(1):54. doi: 10.1038/s41746-025-01460-1
3. Sarker IH. AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. SN Comput Sci. 2022;3(2):158. doi: 10.1007/s42979-022-01043-x
4. Li Z, Koban KC, Schenck TL, Giunta RE, Li Q, Sun Y. Artificial intelligence in dermatology image analysis: current developments and future trends. J Clin Med. 2022;11(22):6826. doi:10.3390/jcm11226826
5. Omiye JA, Gui H, Daneshjou R, Cai ZR, Muralidharan V. Principles, applications, and future of artificial intelligence in dermatology. Front Med (Lausanne). 2023;10:1278232. doi:10.3389/fmed.2023.1278232
6. Nahm WJ, Sohail N, Burshtein J, Goldust M, Tsoukas M. Artificial intelligence in dermatology: a comprehensive review of approved applications, clinical implementation, and future directions. Int J Dermatol. 2025;64(9):1568-1583. doi:10.1111/ijd.17847
7. Salinas MP, Sepúlveda J, Hidalgo L, Peirano D, Morel M, Uribe P, et al. Artificial intelligence versus clinicians for skin cancer diagnosis: a systematic review and meta-analysis. NPJ Digit Med. 2024;7(1):125. doi:10.1038/s41746-024-01103-x
8. Grzybowski A, Jin K, Wu H. Challenges of artificial intelligence in medicine and dermatology. Clin Dermatol. 2024;42(3):210-215. doi:10.1016/j.clindermatol.2023.12.013
9. Van Norman GA. Drugs, devices, and the FDA: Part 2. an overview of approval processes: FDA approval of medical devices. JACC Basic Transl Sci. 2016;1(4):277-287. doi:10.1016/j.jacbts.2016.03.009
10. Thomas L, Hyde C, Mullarkey D, Greenhalgh J, Kalsi D, Ko J. Real-world post-deployment performance of a machine learning-based digital health technology for skin lesion assessment and suggestions for post-market surveillance. Front Med (Lausanne). 2023;10:1264846. doi:10.3389/fmed.2023.1264846
11. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118. doi:10.1038/nature21056
12. Lin B, Xu Y, Bao X, Zhao Z, Wang Z, Yin J. SkinGEN: an explainable dermatology diagnosis-to-generation framework with interactive vision-language models. arXiv. 2024.14755v2. https://doi.org/10.48550/arXiv.2404.14755
13. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. J Med Syst. 2018;42(11):226. doi:10.1007/s10916-018-1088-1
14. Tschandl P. Risk of bias and error from data sets used for dermatologic artificial intelligence. JAMA Dermatol. 2021;157(11):1271-1273. doi:10.1001/jamadermatol.2021.3128
15. Daneshjou R, Smith MP, Sun MD, Rotemberg V, Zou J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. JAMA Dermatol. 2021;157(11):1362-1369. doi:10.1001/jamadermatol.2021.3129
16. Daneshjou R, Vodrahalli K, Novoa RA, Jenkins M, Liang W, Rotemberg V, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv. 2022;8(32):eabq6147. doi:10.1126/sciadv.abq6147
17. Narla S, Heath CR, Alexis A, Silverberg JI. Racial disparities in dermatology. Arch Dermatol Res. 2023;315(5):1215-1223. doi:10.1007/s00403-022-02507-z
18. Monk E. The Monk Skin Tone Scale. SocArXiv. 2023 May 5. https://doi.org/10.31235/osf.io/pdf4c
19. Du-Harpur X, Watt FM, Luscombe NM, Lynch MD. What is AI? Applications of artificial intelligence to dermatology. Br J Dermatol. 2020;183(3):423-430. doi:10.1111/bjd.18880
20. Giansanti D. The artificial intelligence in teledermatology: a narrative review on opportunities, perspectives, and bottlenecks. Int J Environ Res Public Health. 2023;20(10):5810. doi:10.3390/ijerph20105810