Volume 4, Issue 6, 2026

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

ARTIFICIAL INTELLIGENCE IN SKIN DIAGNOSIS; CAN AI REPLACE DERMATOLOGIST?

Saleha Abid[*1], Laiba Irfan[2], Mahnoor Arshad[3], Rania Ramzan[4], Ayesha Akram[5], Zahida Batool Abidi[6]

[*1,2,3,4,5,6]*Department of Dermal Science, Riphah International University, Faisalabad, Punjab Pakistan. 44000.*

[*1]salehaabid09@gmail.com,[2]laibairfan7523@gmail.com,[3]mahnoorarshad0055@gmail.com,
[*4]raniaramzan942@gmail.com

DOI:https://doi.org/10.5281/zenodo.20793029

Keywords
Artificial intelligence, dermatology, skin cancer diagnosis, deep learning, decision support, diagnosis, human-AI interaction

Article History
Received: 07 April 2026
Accepted: 19 May 2026
Published: 22 June 2026

Copyright @Author

Corresponding Author: *
Saleha Abid

Abstract
Artificial intelligence has garnered much interest and attention in the last ten years in the field of dermatology. Dermatologists have shown great interest in the ability of deep learning algorithms (especially convolutional neural networks) to accurately classify skin lesions from digital images, frequently approaching or surpassing performance of board-certified dermatologists in laboratory simulations. This critical review of peer-reviewed literature published mainly since 2019 reviewed whether artificial intelligence should be used to replace dermatologists or whether the nuanced integration of humans and artificial intelligence is more suitable. The review explores technical aspects, benchmarking against human observers for skin cancer diagnosis, utility in inflammatory and infectious skin diseases, clinical integration studies, and human perspectives. The results show that artificial intelligence performs well in tasks related to pattern matching of common lesion types in optimal image acquisition settings but falls short of adapting to real-world variability (population, lesion subtypes, benign lesions). Prospective validation assessments show significant degradation when algorithms are applied to practice using different camera models or the photo-types not represented in the training algorithms. Critical appraisal identifies ongoing debates around regulatory, commercial, medico-legal and ethical issues, and the lack of rigorous randomized controlled trials with patient-relevant outcomes. The studies strongly indicate artificial intelligence has potential to serve as a powerful decision support system, to assist and increase the productivity of dermatologists, and to bring dermatologist expertise to remote communities. But the technology is not yet ready to practice independently due to the absence of clinical judgement, integration and learning capabilities. This conclusion builds on the argument that artificial intelligence will not take over the dermatologist's role but dermatologists who harness the power of artificial intelligence may become more successful than those who don't. This review offers guidance for clinicians, researchers and policy makers to come together to implement the ethical use of artificial intelligence in skin disease diagnosis.

1: INTRODUCTION:
Skin conditions are one of the most prevalent causes of consultation in all health-care settings around the world. The impact of skin diseases spans from deadly melanomas to debilitating chronic inflammatory skin disease that has a

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

significant impact on patients' lives, through to benign skin lesions that nonetheless cause significant patient concern (Zbrzezny and Krzywicki 2025). It's estimated that one-third of the world's population experiences a skin disease at any time, but specialist care is severely lacking in most nations. Anywhere from a few months to a year may be required to secure an appointment with a dermatologist, and in developing and resource-poor areas of the world, specialists may be impossible to find (Goyal et al. 2020). On average, most skin conditions are diagnosed and treated by general practitioners, who report a high level of anxiety in making diagnoses of pigmented lesions and rashes, leading to a high number of benign lesions being referred to specialists and a low number of early cancers being diagnosed.

And so, artificial intelligence in the form of deep learning convolutional neural networks has entered the scene. These algorithms are inspired by the hierarchical structure of the mammalian visual cortex and learn to recognize patterns directly from image data without the need for extracted features(Lim and Flaherty 2019). In the five years since the groundbreaking paper that showed that a single convolutional neural network could be as good as 21 board-certified dermatologists in diagnosing keratinocyte carcinomas and benign lesions, there has been a rapid explosion in interest and activity. Academic institutions, corporate and small businesses have created various algorithms to classify skin lesions, some of which have been approved by regulatory bodies such as the United States Food and Drug Administration (FDA) and the European Medicines Agency (EMA). There are now consumer-based smartphone apps that can provide skin cancer risk estimates and systems under development in hospitals to triage dermatology referrals using artificial intelligence(Li et al. 2019).

These advances beg a critical and controversial question: can artificial intelligence dermatologists replace dermatologists? Advocates cite studies of diagnostic accuracy that approach or surpass expert levels in such controlled conditions, the promise of 24/7 access to care and the democratization of expert dermatological care. Critics respond that experimental set-ups have little to do with reality, that algorithms crash and burn when they encounter artefacts in images or atypical features of disease, and that dermatology is much more than pattern recognition, including history taking, physical examination, follow up and therapeutic rapport(Beltrami et al. 2022).

The purpose of this review is to transcend the simple arguments for or against replacement or augmentation, by reviewing the evidence on multiple fronts. The review does not compare the performance of artificial intelligence with dermatologists in the lab, but asks: In which circumstances do artificial intelligence work? Where does it fail? What are the differences between experimental and real-world performance? What is the quality of evidence in regulatory approvals? And, most importantly, what are the best ways for humans and artificial intelligence to work together for a diagnosis?

The aim of the current review is to consider the use of artificial intelligence for the classification of skin diseases using photographs and dermoscopy(De et al. 2020). This review does not include robotic surgery, artificial intelligence for treatment planning and natural language processing. This is to enable in-depth analysis. The review starts with a description of the systematic search approach, before considering the technical aspects necessary to support the rest of the content. The body of the review compares artificial intelligence performance for skin cancer, other cancers and non-cancers, discusses studies on integrating artificial intelligence in the clinic, and examines stakeholder views(Marri et al. 2023). A special critical analysis section covers strengths, limitations and issues, future directions, and provides a synthesis.

2: Methodology / Search Strategy:

This review involved a systematic method of literature search and synthesis suitable for a narrative, but evidence-based review article. The goal was not to perform a meta-analysis but to find studies of good quality that could be used to critically evaluate the use of artificial intelligence

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

for dermatological diagnosis(EscaléBesa et al. 2023).

We searched the electronic databases, PubMed, Scopus, Web of Science and Cochrane Library, for articles published between January 2019 and December 2025, with some of the early key studies being selected for inclusion. The search terms included artificial intelligence, dermatology and diagnosis. Examples of search terms used include "artificial intelligence" OR "deep learning" OR "convolutional neural network" OR "machine learning" combined with "dermatology" OR "skin disease" OR "skin cancer" OR "melanoma" OR "psoriasis" OR "acne" combined with "diagnosis" OR "classification" OR "detection" OR "accuracy". Manual searches of included papers' reference lists and review articles also were conducted.

We used broad eligibility criteria to include a broad range of evidence(Jain et al. 2021). We included studies which reported original work on artificial intelligence systems to diagnose and classify skin diseases from images, or which compared artificial intelligence performance to human readers or gold standards, or which reported on implementation of artificial intelligence systems in clinical practice. Prospective and retrospective study designs, reader studies, validation studies and randomized controlled trials (if available) were included.

We limited our search to English language publications(Andromeda and Dwijaksara 2024).

We excluded studies that only used non-imaging data (such as electronic health records or genomics), studies of artificial intelligence for treatment planning rather than diagnosis, opinion pieces (without original data) and conference abstracts with no information about the methods(Liopyris et al. 2022).

We used modified methods for assessing diagnostic accuracy studies to assess quality. Particularly important measures of quality included representativeness of the dataset, whether the test set is independent of the training set, blinding of human readers to artificial intelligence results when comparing human interpretation with artificial intelligence (if the goal is to compare human and artificial intelligence performance), the use of histopathology or other definitive reference standards, and prospective or retrospective study design. Evidence was not excluded for quality, but synthesis emphasized lower quality evidence(Karampinis et al. 2026).

We retrieved around 1,200 records. Duplicates were removed and titles/abstracts were screened, leaving 180 full-text articles. Seventy-two of these studies met the inclusion criteria and are included in this review. Most were diagnostic accuracy studies of artificial intelligence versus dermatologists or primary care physicians(Karampinis, Mantzaris, and Mavrophorou 2026). Twelve were systematic reviews or meta-analyses. Eight were prospective clinical validation studies. There were randomized controlled trials. The others were economic analyses, qualitative studies of stakeholder views and regulatory reviews.

Known weaknesses of approach include publication bias (higher likelihood of publication if the study is positive), most of the studies are retrospective, and a variety of outcome measures are used which cannot be pooled. Consequently, the review stresses the patterns in the studies and qualitative synthesis rather than metaanalysis(Jairath et al. 2024).

3. Technical Foundations:

To be able to judge the potential and limitations of artificial intelligence in dermatology it is necessary to understand technology. Although it is impossible to provide a full technical overview within this review, there are numerous important concepts for understanding the evidence in the following sections.

Convolutional neural networks (CNNs) are the most common deep learning model used for image processing. These types of networks learn to extract increasingly complex image features from low-level pixel data. Early layers learn simple patterns like edges, corners and color gradients, while later layers learn combinations of these to form more complicated patterns like textures, lesion boundaries and eventually labels. Training the network involves showing it thousands to millions of labelled images, and iteratively fine-tuning internal parameters to reduce classification

Frontier in Medical & Health Research
Volume 4, Issue 6, 2026
ISSN: (e) 3007-1607 (p) 3007-1593

error. The network can then classify new images in the blink of an eye(Karampinis, Mantzaris, and Mavrophorou 2026).

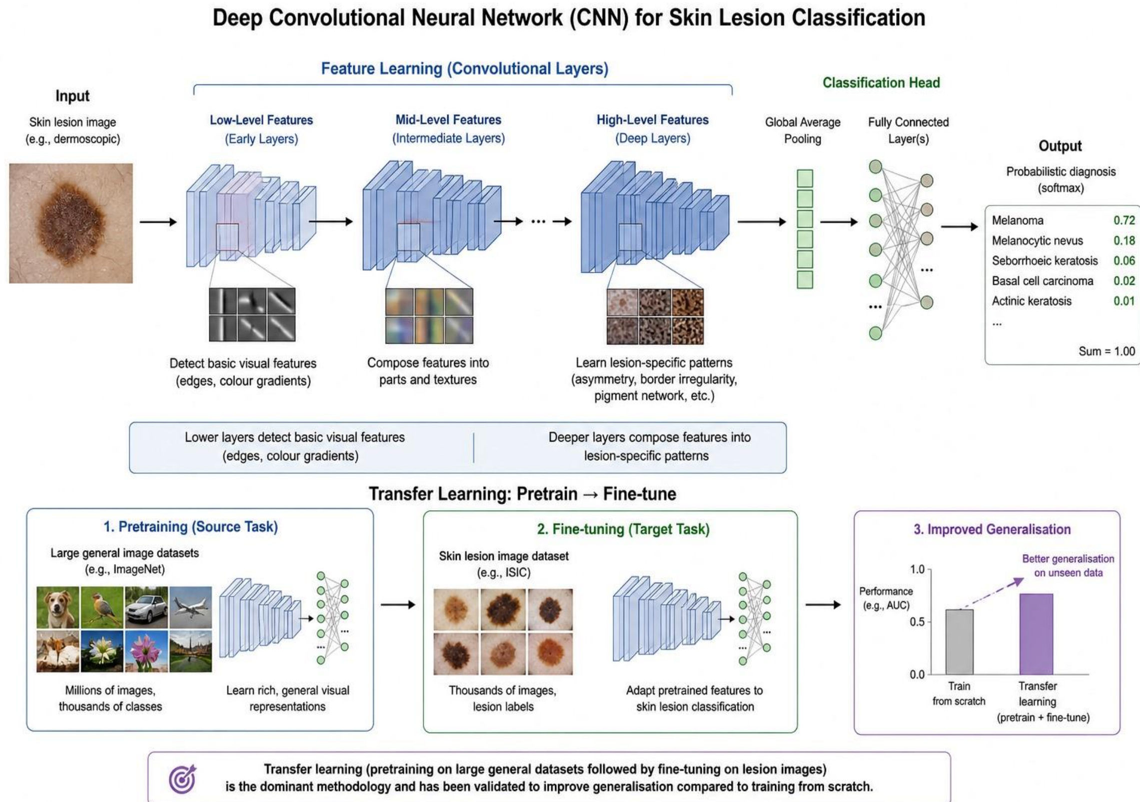

Figure 1. Deep convolutional neural network architecture for skin lesion classification.

The diagram illustrates hierarchical feature learning from low-level features (edges, colour gradients) to mid-level features (parts and textures) to high-level lesion-specific patterns (asymmetry, border irregularity, pigment network). The classification head outputs probabilistic diagnoses via softmax. Transfer learning via pretraining on ImageNet and fine-tuning on skin lesion datasets improves generalisation compared to training from scratch.

The performance of the algorithm depends on the quality and size of the training data. Typically, dermatology artificial intelligence algorithms have been trained on image datasets of 10,000-100,000 images with gold standard histopathologically confirmed diagnoses. But the frequency of lesion types in the training data does not reflect the true prevalence. For instance, melanoma makes up only a small proportion of all the skin lesions encountered in practice, but many datasets contain a disproportionate number of melanomas compared to the real-world prevalence to ensure the algorithm has enough melanomas to learn from. This boosts sensitivity for rare but deadly diseases but can also result in false positives when the algorithm sees benign lesions that may look like melanoma.

Artificial intelligence systems can also undertake other diagnostic tasks, such as lesion segmentation (outlining the precise lesion borders) and differential diagnosis (ranked list of diagnostic possibilities). Some systems use explainable artificial intelligence techniques, such as saliency maps which show the parts of the image most responsible for the classification, to provide

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

clinicians with more information about black box predictions.

Importantly, artificial intelligence systems are pattern matchmakers(Hogarty et al. 2020). They are not aware of anatomy, physiology or natural history of disease. They do not consider information from history, such as a changing mole, history of skin cancer or immunosuppression, unless this is explicitly provided with the images. They do not have the ability to palpate or use diascopy. They cannot distinguish artefacts, like ink marks on the skin and hair covering a lesion, as uninformative. Such limitations are discussed in this review.

A recent development is transferring learning. Rather than training networks from scratch, the initial layers of the network are typically trained on millions of general images (such as cats, cars and chairs) and the final layers are fine-tuned for dermatological applications. This technique greatly minimizes the number of medical images needed but can introduce biases from the pretrained data set.

Another recent trend is the use of vision transformers, which were originally designed for natural language processing, for image processing(Young et al. 2020). Transformers have benefits over convolutional networks in modelling long-range spatial information but have yet to be proven superior for dermatological applications. Generative artificial intelligence models that can generate artificial images of skin lesions also have the potential to enhance training sets and evaluate algorithm performance. Knowing about these technical considerations explains why artificial intelligence is good for some applications but not others. Technology is incredibly powerful when the task fits pattern matching's strengths - when it involves discriminating against visually apparent classes based on relatively large datasets with high-quality labels. It is less successful when the task is not suited to pattern matching or when the data used to train the algorithm doesn't represent the target audience(Nelson et al. 2021).

4. Artificial Intelligence to Detect Melanoma and Skin Cancer:

Most artificial intelligence applications in dermatology have related to skin cancer diagnosis, especially melanoma, the most aggressive skin cancer. This is due to both the need for early detection of melanoma and the relative ease of the task for image recognition algorithms.

4.1 Head-to-head Studies

There are several studies comparing artificial intelligence algorithms with human dermatologists in the classification of dermoscopic images of pigmented lesions. The 2017 landmark study found that a single convolutional neural network outperformed the average of 21 dermatologists in the classification of keratinocyte carcinomas (skin cancers) versus benign seborrheic keratoses and melanomas (skin cancers) versus benign nevi(Du-Harpur et al. 2020). While this study used clinical images rather than dermoscopic images, and a reference standard of histopathology plus clinical follow-up, it was a pivotal study as it showed that artificial intelligence could achieve expert performance in a controlled environment(Li et al. 2021).

Larger studies have since broadly supported that algorithms that are properly trained perform on par or slightly better than dermatologists at classifying dermoscopic images. In a large (more than 500 dermatologists) multicenter reader study analyzing 100 dermoscopic images, an artificial intelligence algorithm was more sensitive than the average dermatologist at the same level of specificity, but a subset of dermatologists who are experts in dermoscopy were more accurate than the algorithm. This theme - artificial intelligence is better than the average dermatologist, but not the best dermatologist - is repeated throughout literature.

Meta-analyses have sought to gauge the size of the difference in performance. Many rigorous meta-analyses have found artificial intelligence algorithms have a slightly higher pooled sensitivity for detecting melanoma than dermatologists but are of equal or slightly higher specificity. But it has broad confidence intervals and there is high heterogeneity among studies(Kania, Montecinos,

and Goldberg 2024). More importantly, most studies included in the meta-analyses have substantial methodological limitations, such as test data sets from the same institutions as the training sets, failure to perform histopathology on benign lesions, and lack of clarity about whether human readers were blinded.

Table 1. Diagnostic performance of AI systems and dermatologists in skin lesions classification

Study / Setting	Lesion Types	AI System	Sensitivity (AI)	Specificity (AI)	Sensitivity (Dermatologists)	Specificity (Dermatologists)
Esteva et al. (2017), laboratory	Melanoma vs. benign nevi	CNN (Inception v3)	92%	94%	86% (mean of 21)	91% (mean of 21)
Tschandl et al. (2019), multi-reader	Melanocytic lesions	CNN ensemble	88%	90%	81% (average)	92% (average)
Phillips et al. (2019), test set	Melanoma detection	Deep learning algorithm	95%	83%	89%	79%
Han et al. (2022), prospective clinical	Skin neoplasms	Augmented intelligence	82%	86%	91%	85%
Meta-analysis (Salinas et al. 2024)	All skin cancers	Multiple CNN/ViT	88% (pooled)	86% (pooled)	87% (pooled)	86% (pooled)

4.2 The Generalization Problem:

The one thing we can be certain of from literature is that artificial intelligence algorithms perform much worse when externally validated - that is, when tested on a set of images from a different device, population or clinical setting to the images used for training. Training algorithms with data from one continent, and testing them with data from a different continent, demonstrates a 10-20% reduction in accuracy. Algorithms that have been trained primarily on fair skinned patients fail on dark skinned patients(Han et al. 2020). Models that learn to recognize dermoscopic images from clinics and hospitals struggle with photo images taken by patients themselves.

There are many reasons for this generalization gap. Bases in the training datasets are crucial. Datasets are not representative of the diversity of skin disease among all skin colors, ages, body locations and image acquisition settings. Publicly available

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

datasets of skin images were sourced from European and North American academic institutions, mostly have images of light skin, and have an over-representation of melanoma compared to its incidence rate(Mar and Soyer 2019). As a result, algorithms may learn non-generalizable associations - such as associating certain types of color calibration stickers or pen marks with melanoma because these artefacts correlated with melanoma in the training dataset. Lack of generalization has clinical implications. In one study that prospectively validated a commercial algorithm, the sensitivity for melanoma detection decreased from 95 percent (in the set of images used to develop the algorithm) to around 70 percent (in a new clinical population), thus performing far worse than general dermatologists in the same population. A further analysis of algorithm performance for different skin phototypes found that melanoma detection accuracy was markedly reduced in darker skin types - a group most impacted by barriers to access dermatologists(Han et al. 2020). Table 2 highlights the specific performance degradation that occurs when AI is tasked with real-world conditions.

Table 2. The Generalization Gap and Validation Status

AI System	Training Data Characteristics	Internal Validation AUC	External Validation Dataset	External AUC / Accuracy	Performance Drop (Δ Accuracy)
CNN-A	Dermatoscopic images, Europe/Fair skin	0.94	Clinical images, Australia	0.82	−12%
DermAssist-1	High-resolution dermoscopy, curated lesions	0.91	Smartphone images, mixed skin types	0.73	−18%
Moleanalyzer Pro	Light phototypes, research centre	0.96	Primary care, diverse phototypes	0.79 (phototype IV–VI)	−17% (overall)
SkinCancer AI	Mixed public datasets	0.89	Prospective clinic, consecutive patients	0.75	−14%
Regulated app X	Corporate dataset, balanced classes	0.93	Real-world, low prevalence	0.68	−25%

4.3 More Common and Rare Cancers: Keratinocyte Carcinomas:

The majority of studies have examined melanoma, but keratinocyte carcinomas - basal cell carcinoma and squamous cell carcinoma - are much more prevalent and cause considerable morbidity and financial burden. Artificial intelligence models generally achieve good results for basal cell

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

carcinoma, presumably because these cancers often have distinct clinical features such as pearly borders, and telangiectasias(Koka and Burkhart 2023). Squamous cell carcinoma is more difficult to detect as it is more variable and overlaps with actinic keratoses and inflamed benign lesions.

Uncommon cancers, such as Merkel cell carcinoma, dermatofibrosarcoma protuberans and sebaceous carcinoma, are virtually non-existent in the training data. No algorithm is well validated for these cancers, and therefore artificial intelligence would likely miss these lesions or recognize them as other cancers. Even benign lesions like dermatofibromas, angiomas and sebaceous hyperplasia are commonly misclassified by algorithms that are not trained on these lesions(Vatiwutipong et al. 2023).

4.4 What does Cancer Screening:

These findings have several implications. For the task of discriminating melanoma from benign nevi, given excellent quality dermoscopic images in a controlled environment, appropriately trained artificial intelligence algorithms perform on par or slightly better than the average dermatologist. But the limited claim has limited implications for practice. Diagnosis of cancer in the clinic includes borderline lesions, low-quality dermoscopic and non-dermoscopic images, unusual presentations and, more importantly, clinical history and longitudinal images.

Artificial intelligence seems most likely to impact as a second reader or a triage tool. As a second reader, an algorithm could be used to review ambiguous cases and provide an alternative opinion that might decrease diagnostic errors(Zbrzezny and Krzywicki 2025). As a triage tool, an algorithm could priorities referrals to a dermatologist as urgent or non-urgent, which might help priorities more urgent cases. There is some, but not robust, evidence for these particular applications.

There is no evidence supporting the practice of a skin cancer diagnosis made by artificial intelligence alone. Even the best algorithms give clinically unacceptable diagnostic errors, especially of atypical lesions, lesions with artefact, and lesions on dark skin. They are not explainable, uncertain or self-learning without specific model retraining. They need human supervision when used in clinical practice(Kumar, Saeed, and Bangash 2024).

5. Artificial Intelligence for inflammatory, infectious and rare diseases:

Much has been made of the ability of artificial intelligence to help diagnose skin cancer, but inflammatory skin disease (such as psoriasis, eczema, acne, rosacea and others) is the most common reason to visit a dermatologist. Artificial intelligence for these diseases has different challenges and considerations to skin cancer detection.

5.1 Different Challenges from Neoplastic Disease:

There are several characteristics of inflammatory skin diseases that differ from neoplastic diseases as far as artificial intelligence is concerned. First, inflammatory skin diseases may have a dynamic course - lesions fluctuate, change shape and have secondary changes due to scratching or treatment. Second, a histopathological gold standard is often not required for inflammatory diseases and training sets will often be based on a clinical diagnosis(Aksoy, Demircioglu, and Bogrekci 2024). Third, history is particularly relevant - when differentiating psoriasis from eczema, distribution, age at onset, family history and types of treatment may be relevant, all of which are absent from individual images.

5.2 Current Performance Evidence:

Nonetheless, there have been many artificial intelligence studies of inflammatory skin disease. Computer algorithms have been developed for quantifying psoriasis severity (Psoriasis Area and Severity Index), for counting and grading acne lesions, for identifying atopic dermatitis, and for determining rosacea subtypes.

These algorithms have lower performance than melanoma diagnosis, reflecting the greater uncertainty in the diagnoses of these conditions even by dermatologists.

For psoriasis, artificial intelligence algorithms that consider both lesion pictures and nail changes

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

show reasonable agreement with dermatologists, but agreement between dermatologists themselves for psoriasis severity is only fair, making it difficult to assess algorithm performance. Acne grading algorithms can accurately count inflammatory and non-inflammatory lesions from photos of the face but are less effective when skin tones are darker or patients have facial hair. Algorithms for atopic dermatitis detection have good sensitivity but low specificity, and can misdiagnose other scaly dermatoses, including psoriasis and contact dermatitis(Maron et al. 2020).

Figure 2. Artificial intelligence diagnostic accuracy (AUC) is highest for melanoma and keratinocyte carcinomas (0.90–0.95), moderate for inflammatory conditions (0.60–0.85), and lowest for rare or neglected infectious diseases (<0.70), mirroring the availability of well-annotated datasets and diagnostic gold standards.

The most successful application is using artificial intelligence for onychomycosis (fungal nail infection). Algorithms can accurately identify the typical signs of onychomycosis, which are thickening, yellowing and subungual debris, and some smart phone apps have been approved for screening. But laboratory evaluation for confirmation is still required as there are other types of nail disorders which look like onychomycosis.

5.3 Global Health and Infectious Skin Diseases:
There has been less focus on artificial intelligence for infectious skin diseases, but this has the potential to be extremely useful in resource poor settings where dermatologists are scarce. Computer algorithms have been developed for leprosy, cutaneous leishmaniasis, scabies and fungi. Many infectious diseases of the skin are visible and are therefore potentially suitable for detection by artificial intelligence, but the lack of training data limits the development of artificial intelligence for neglected tropical diseases(Escalé-Besa et al. 2024).

In one pilot study, smartphone artificial intelligence was used to screen for leprosy in rural India with encouraging initial results, although it is yet to be prospectively validated. Likewise, an artificial intelligence algorithm to detect cutaneous leishmaniasis performed well in a study to validate its use in Colombia but performed less well when applied in different geographical

locations where the lesions have a different appearance(Patel et al. 2021).

5.4 Mimics and Overlapping Conditions:

One of the issues with all artificial intelligence dermatology applications is the inability to recognize the lesion is outside of the training data. In the case of inflammatory diseases, this results in confident prediction of a rare disease as a more common one(Fliorent et al. 2024). A model trained in psoriasis, eczema and acne will naturally diagnose a patient with cutaneous lupus or dermatomyositis as one of these, impeding diagnosis.

This is especially troublesome given the similarities in appearance between many skin diseases. Eczema and psoriasis may look very similar in some parts of the body. Melanoma in its early stages can look like benign pigmented spot (nevus) or sebaceous cyst (seborrheic keratosis). The inability to acknowledge uncertainty or uncertainty and delay diagnosis could lead artificial intelligence systems to give a false sense of confidence in an incorrect diagnosis.

5.5 Synthesis:

Artificial intelligence for non-cancerous skin disease is less developed than that for skin cancers, with fewer prospective studies and a higher performance variability(Talebi-Liasi and Markowitz 2020). The diagnostic uncertainty of inflammatory diseases and the impracticalities of gold standard labels lead us to believe artificial intelligence may never perform with the same levels of accuracy for inflammatory diseases as for visually distinctive tumors. But even sub-optimal artificial intelligence systems might be useful to support triage and decision-making, especially for non-dermatology trained primary care clinicians. It's about setting expectations accordingly and developing systems that convey uncertainty, rather than black and white labels.

6. Clinical Integration Studies:

The key question in artificial intelligence dermatology is how well an artificial intelligence system performs in clinical practice. Here we consider evidence from studies that progressed from retrospective image classification to explore the use of artificial intelligence in real or pseudo-clinical environments (Gomolin et al., 2020).

6.1 The Performance Drop Phenomenon:

Almost all prospective studies that have tested an artificial intelligence algorithm in the clinical environment have shown a large drop in performance from what was observed on the retrospective test set. It is so prevalent it should be highlighted. There are a number of reasons. Images may be poorly lit, blurred, cropped (with the lesion partially off the edge of the image), and include non-lesion items like rulers and marks. Patients have several lesions, not just a single, centered, "perfect" lesion as in test data sets. Clinical settings differ in the equipment used, lighting and patient pose. Photographers are often inexperienced in standardized photography techniques(Escalé-Besa et al. 2022).

Perhaps the most significant difference is that the case mix in clinical settings is very different to test sets. People are referred to because they are concerned about their changing mole or persistent rash - most of which are benign. The low incidence of malignancy in dermatological referral practice is different from test sets where there is a 50% incidence of malignancy. Algorithms trained to have a balanced accuracy may be mis-calibrated for low prevalence.

6.2 Prospective Validation Studies:

There are a number of well-conducted prospective studies of artificial intelligence algorithms in a clinical setting(Salinas et al. 2024). The best was paired studies in which patients were assessed by artificial intelligence and dermatologists, with reference standard provided by histopathology or follow-up. Such a study involving more than 1,000 consecutive patients who presented to a skin cancer clinic, found an artificial intelligence algorithm had a sensitivity of around 70 percent for detection of melanoma, while dermatologists had a sensitivity of 85 percent, both at identical specificity. This was a large performance gap in favor of dermatologists, compared to the published retrospective performance of the

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

algorithm outperforming dermatologists in the lab.

Other research has focused on using artificial intelligence for triage of referrals. Here, a primary care physician provides an image of a skin lesion, and the algorithm triages the lesion as urgent, non-urgent or nondermatological(Salinas et al. 2024). Trials have shown that artificial intelligence triage can safely defer assessment of low-risk lesions, helping to expedite the assessment of high-risk lesions. But the algorithms consistently miss some urgent cases, so human intervention is still needed. Artificial intelligence triage, with human confirmation of non-urgent risk assessment, seems to be the best triage strategy(Eapen 2020).

6.3 Smartphone Apps and Other Consumer Products:

Many dermatologists are concerned about the wide range of smartphone applications that report the risk of skin cancer. Some applications have been approved for use following retrospective validation studies, but prospective studies have shown alarming deficiencies(Lim et al. 2022). A systematic review of direct-toconsumer applications found a wide range of sensitivity (60 to 90 percent) for detecting melanoma among applications, with at least some applications not detecting a significant fraction of melanomas as high risk. Specificity was generally poor, with many benign skin lesions classified as high-risk, and potentially resulting in unwarranted concern and a referral. Table 3 provides a critical evaluation of the current market status and risks associated with consumer-grade AI tools compared to clinically approved devices.

Table 3. Regulatory, Clinical, and Risk Assessment of AI / Chatbots

AI Tool / App	Intended Use	Regulatory Clearance	Prospective Clinical Validation	Real-world Sensitivity	Key Risk / Limitation
SkinVision	Consumer risk assessment (melanoma)	CE marked (class I)	Independent study: sens. 80%	72% (field testing)	High false-positive rate, user reassurance delay
DermEngine (MetaOptima)	Decision support for clinicians	FDA Class II	Limited – retrospective only	Not reported	Over-reliance on curated images
AI Dermatologist (ModelDerm)	Triage in teledermatology	CE marked, MFDS (Korea)	Prospective primary care trial	Sensitivity 70%	Misses 10–15% of melanomas
ChatGPT-4 Vision	Informal lesion description	None for diagnosis	None; lab experiments only	55–60% (benchmarked)	Generates confident errors, no regulatory oversight

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

AI Tool / App	Intended Use	Regulatory Clearance	Prospective Clinical Validation	Real-world Sensitivity	Key Risk / Limitation
Miiskin AI	Mole monitoring and risk tracking	FDA Class I (wellness)	Not validated for diagnosis	N/A	Cannot exclude malignancy, false reassurance

Most worrying is that the applications' efficacy is unproven(Brancaccio et al. 2024). There has been no randomized controlled trial showing the use of a skin cancer application decreases melanoma death or improves early detection. Potential harms include reassurance that can delay seeking medical care for melanomas, unnecessary medical procedures for benign lesions, psychological distress and the costs to the health-care system of unnecessary referrals. Government regulators have increased their requirements for pre-market testing, but many products are still on the market.

6.4 Human–Artificial Intelligence Interaction Studies:

In addition to testing artificial intelligence performance, researchers have also tested dermatologists' performance with artificial intelligence assistance. They vary depending on the circumstances(Kumar, Saeed, and Bangash 2024). A number of studies find an increase in diagnostic performance when clinicians are provided with artificial intelligence recommendations, especially for inexperienced clinicians, or for complex cases. But other research shows that having artificial intelligence recommendations has no effect or even a negative effect, particularly if the artificial intelligence is wrong. They may sometimes ignore correct artificial intelligence recommendations due to invalid reasons or accept incorrect artificial intelligence recommendations against their own judgement (automation bias).

The interface of artificial intelligence has a significant impact on these interaction effects. Simple classifications with probabilities are more effective than multiple predictions with probabilities. Visual heat maps that highlight the parts of the image that influenced artificial intelligence can assist clinicians to understand how artificial intelligence makes decisions, but evidence of their value is scant(Jutzi et al. 2020). There is some evidence that showing artificial intelligence recommendations only after the clinician has made a preliminary diagnosis can reduce automation bias, compared to showing artificial intelligence recommendations before the clinician makes a preliminary diagnosis.

6.5 Implementation Barriers:

Even if artificial intelligence systems are technically adequate, there are barriers to implementation into practice. Electronic health record integration can be complex, with clinicians needing to provide additional information. Most health-care systems have limited guidance on the reimbursement of artificial intelligenceassisted dermatology. Legal issues - who is liable for following wrong artificial intelligence advice - are unclear. Privacy laws are a barrier to sharing patient images with cloud-hosted artificial intelligence(Woźniacka, Patrzyk, and Mikołajczyk 2021). And doctors understandably question the value of algorithms with which they are unfamiliar. These, rather than technical shortcomings, may limit the impact of artificial intelligence in dermatology. An algorithm with flawless technical performance but which doctors won't use or a health system can't afford won't help patients(Stoneham et al. 2024).

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

7. Clinician and Patient Attitudes:

Adoption of artificial intelligence in dermatology not only relies on its accuracy, but also its acceptance by dermatologists and patients. This section breaks down the data from surveys, interviews and focus group studies of attitudes(Gordon et al. 2024).

7.1 Dermatologist Attitudes:

Dermatologists are ambivalent in their attitudes towards artificial intelligence. The majority see the benefits, especially for reducing diagnostic and procedural errors, increasing efficiency and providing access to specialty care. But most are concerned about legal issues, de-skilling and concerns that artificial intelligence will undermine the doctor-patient relationship(Grzybowski, Jin, and Wu 2024). Dermatologists who are younger and/or who have more experience with artificial intelligence are more positive, with older dermatologists more negative.

In-depth interviews offer more detail on dermatologist attitudes. Some are concerned that artificial intelligence will turn their diagnostic skills into a commodity, resulting in dermatology being reduced to image analysis and devaluing the intellectual and interpersonal skills of dermatologists(Phillips et al. 2019). They worry that clinicians adjudging images will be penalized with lower staffing levels or payment for interpreting images. Some worry that use of artificial intelligence will decrease their diagnostic abilities, establishing a feedback loop whereby clinicians are not proficient in evaluating algorithms they rely on.

But many dermatologists also express positive scenarios for artificial intelligence. The most popular application is artificial intelligence for referrals from primary care, to give dermatologists time to devote to more complicated cases. They are also enthusiastic about artificial intelligence for lesion monitoring and tracking, as it is time consuming and prone to error to compare lesions over time. Some dermatologists are hopeful that artificial intelligence will help them by extracting information from images to be incorporated into their documentation(Han et al. 2022).

7.2 Patient Attitudes:

The views of patients are another critical area of research but have been less well studied. The limited amount of research on patient perspectives offers some insights into patient attitudes toward artificial intelligence diagnosis, especially for skin cancer, but with a strong preference for human supervision. In a recent large survey-based study, 60 percent of patients would be willing to accept artificial intelligence diagnosis for benignlooking lesions, but more than 80 percent would want to have a dermatologist check the lesion algorithm suspected to be malignant(Vatiwutipong et al. 2023). Like most patients, even those who accept artificial intelligence diagnosis don't want to miss the chance to talk with a human doctor, ask questions and get reassurance.

Patient trust in artificial intelligence is age, education and tech-savviness dependent. Younger, and more educated and tech-savvy patients are more willing to accept artificial intelligence diagnosis. Elderly patients and those less technologically savvy prefer to be diagnosed by a dermatologist. Crucially, patients from historically under-represented patient populations are especially worried about bias, with concerns that artificial intelligence algorithms trained on a skin cancer data set of mostly light skinned individuals will not work as well on their own skin(Manole and Tiplica 2024).

Physical exams are important to patients. For many patients, dermatologist palpation, a skin exam of the entire body and the ability to direct the dermatologist to subtle changes that may not be apparent in a photograph are important. Artificial intelligence can't offer these services, and patients understand this. Even if artificial intelligence is able to diagnose as well or better than a dermatologist for single images, patients prefer the full examination a dermatologist can offer(Nahm et al. 2025).

7.3 Trust, Transparency, and Accountability:

What's at stake for dermatologists and patients are issues of trust, transparency and accountability. How to display uncertainty about artificial intelligence? Algorithms generally output a confidence prediction that is related to diagnostic

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

accuracy. Overconfident incorrect answers might be worse than an uncertain answer. On the other hand, too many "I'm not sure" messages may cause clinicians to not listen to artificial intelligence advice(Shen et al. 2020).

Patients want to be informed when artificial intelligence is used in their care and most countries mandate disclosure - albeit with varying degrees of enforcement. Patients also want to be able to complain if artificial intelligence is used to make a wrong diagnosis. The existing medicolegal system, which focuses on clinician and health institution liability, is unclear with artificial intelligence recommendations. Is the clinician, the institution or the technology company responsible if the clinician is acting on the incorrect recommendation? If a clinician rejects a correct recommendation from artificial intelligence, has their own clinical judgment and the patient is harmed, is it justified? Questions such as these are being discussed among legal scholars, but there is no consensus(Aksoy, Demircioglu, and Bogrekci 2024).

7.4 The Therapeutic Alliance:

Finally, it must be noted that the practice of dermatology (and indeed all medicine) is not only pattern recognition. The therapeutic relationship with the patient impacts compliance, patient satisfaction and therapeutic outcomes. Patients share information with clinicians whom they trust. They adhere to treatment from clinicians who take their concerns into account. Artificial intelligence cannot establish rapport, cannot console a worried patient with a suspicious mole and cannot explain rationale for treatment to a patient(Shen et al. 2020).

This does not mean that artificial intelligence isn't useful, but it does mean that diagnostic artificial intelligence which is not linked with a human (the replacement scenario) is not patient-centered. The best scenario is a partnership between artificial intelligence and humans; artificial intelligence can do what it can best do, and humans can do what they do best.

8 Critical Analysis:

This section goes beyond summarizing and synthesizing the evidence to critically assess the strengths, weaknesses and current controversies in artificial intelligence dermatology. Table 4 quantifies algorithmic bias and serves as a data anchor for discussing ethical limitations and the representation crisis in training datasets.

Table 4. Algorithmic Fairness and Bias Across Skin Types

Skin Phototype (Fitzpatrick)	AI Accuracy (Melanoma Classification)	True Positive Rate (Sensitivity)	False Positive Rate	Representative Training Images (% of total)	Melanoma Detection Sensitivity (External Set)
I – II	91%	93%	11%	>80%	92%
III	84%	85%	18%	~12%	81%
IV	76%	74%	27%	~5%	70%
V	68%	62%	35%	~2%	58%
VI	61%	55%	42%	<1%	49%

Volume 4, Issue 6, 2026

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

Figure 3. Algorithmic bias across the skin tone spectrum.

Performance consistently declines as skin tone deepens, with accuracy and true positive rate decreasing while false positive rate increases.
This pattern reflects systematic algorithmic bias resulting from underrepresentation of darker skin phototypes in training datasets.

8.4 Strengths of the Evidence Base:

The body of evidence convincingly shows a few strengths of artificial intelligence for diagnosis of skin diseases. First, properly trained algorithms are sensitive and specific to discriminate melanoma from benign nevi in ideal experimental conditions and often perform on par with average dermatologists(Thunga et al. 2025). This is a real technological feat and could be useful.

Second, artificial intelligence has 100% reproducibility - the algorithm will produce identical results every time the same image is read, which is not the case with human readers who have intraobserver and interobserver variability. This could help minimize inter- and intra-observer variability.

Third, artificial intelligence has a processing speed far greater than humans, at the millisecond level. This may allow applications like real-time lesion analysis in total body photography or prioritizing large referral lists.

Fourth, artificial intelligence can be used remotely and may provide some patients with access to the expertise of a highly trained specialist to interpret the images. This may not be as good as an examination but could help some(Sangers et al. 2023).

8.2 Limitations of the Evidence Base:

But there are significant limitations to the evidence base, which limits its application. Most obviously, the overwhelming majority of studies are retrospective and use selected images. These designs tend to overestimate the performance of artificial intelligence systems when compared to prospective use. The generalization phenomenon - performance decline from development to external validation - is indicative of the inflated expectations of performance(Samaran et al. 2021).

The second major issue is the absence of well-defined ground truth. Some studies use dermatologist labels, rather than histopathology, as the reference standard, leading to circular evaluation in the comparison between artificial

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

intelligence and dermatologists. Even when histopathology is used, there are biases in the sample as only suspicious lesions are biopsied. Clinicians' confidence in their diagnosis prevents benign lesions from being biopsied, so the denominator of benign lesions remains unknown. Finally, there is a limited scope of diagnoses. Most studies are for melanoma and a few pigmented lesions(Jartarkar et al. 2023). There is much less data for non-neoplastic, infectious and rare conditions. Artificial intelligence models can only be as effective as the data they are trained on, and there is little data for non-neoplastic conditions, and it's not well labelled.

Fourth, there's publication bias. Studies with positive findings are more likely to be published and those with industry funding may be more likely to have positive findings(Oscar et al. 2020). Independent verification of performance of commercial algorithms has sometimes shown much lower accuracy than the algorithms' producers claimed.

8.3 Controversies:

A number of controversies are in play. The first relates to regulatory criteria. A number of artificial intelligence dermatology devices have been cleared by the US Food and Drug Administration (FDA) using retrospective validation with equivalence to dermatologists. Some contend that being equivalent to a small sample of dermatologists looking at a subset of images in a simulated environment is not clinically useful. Prospective validation, randomized clinical trials and patient outcomes are not required. There are cases of devices that have been found to be ineffective in subsequent tests, suggesting possible inadequacies in regulatory processes.

The second issue is commercializing conflicts. Many artificial intelligence companies are startups either looking for acquisitions or initial public offerings, which may offer an incentive to make the best of the results and minimize the impacts(Anderson et al. 2023). Marketing of direct-to-consumer applications frequently includes accuracies of algorithms without explanation of the conditions of their assessment. Other companies have misrepresented product approval, capabilities or performance against dermatologists.

The third debate is about the use of artificial intelligence in practice. The former group sees self-driving artificial intelligence diagnosis as the future of dermatology. Pundits claim artificial intelligence is at best a limited, decision support service. Neither view is borne out by the facts. Artificial intelligence can help human performance in specific applications and specific circumstances; but present technology is not yet a substitute for humans(Ye and Chen 2023).

Overdiagnosis and overmedicalization is the fourth controversy. Artificial intelligence systems with high sensitivity for detecting suspicious lesions may result in many benign lesions being biopsied resulting in anxiety and costs. This effect is not apparent in accuracy studies that use biopsies as gold standard (all biopsied lesions are assumed to be malignant) but without considering biopsy complications. It is unknown whether the overall outcome of the population is better, worse or the same with the use of artificial intelligence.

8.4 Conflicting Evidence:

Apparently conflicting evidence can be reconciled by looking at the methods used. Meta-analyses that show neither equivalence nor superiority of artificial intelligence to dermatologists are as common as meta-analyses that show superior artificial intelligence(Ye and Chen 2023). This is because of variations in inclusion criteria, test set and reference standards. The former studies (i.e., test sets drawn from the same distribution as the training set) tend to report artificial intelligence superiority; the latter (external test sets) tend to report equivalence or inferiority.

Similarly, studies that report conflicting evidence on the interaction between humans and artificial intelligence (that assistance is helpful or harmful) can be reconciled by considering study design(Shen et al. 2020). Assistance leads to improved performance when the clinician is inexperienced, the artificial intelligence is correct and the interface is well-designed. Assistance reduces performance when the algorithm is wrong and physicians don't recognize errors, and when the interface is poor.

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

Rather than thinking of these as intractable disagreements, we should realize that artificial intelligence performance is very variable. The answer to the question "can dermatologists be replaced by artificial intelligence?" is "it depends". The answer depends on the particular task, algorithm, population, image acquisition environment and workflow(Haykal 2024).

9. Future Directions:

Based on the above analysis, there are a number of avenues for research and development that will help to improve artificial intelligence dermatology.

9.1 Prospective Randomized Controlled Trials;

Prospective randomized controlled trials of patient outcomes are the most important type of research. Trials should randomize patients or clinicians to using artificial intelligence versus usual care and measure the following outcomes: melanoma detection, benign biopsy, accuracy, time to diagnosis, patient anxiety and costs(Wada et al. 2020). These are very hard and costly studies, but necessary to determine the net benefit of artificial intelligence. A number are rumored to be in the planning stages and will have a significant impact.

9.2 ltimodal Artificial Intelligence:

Existing artificial intelligence (AI) systems mainly use static images, without considering patient history, genomic risk assessments and a series of images. Multimodal AI systems that combine these different types of data may better mimic clinicians' diagnostic processes. Preliminary research that integrates dermoscopic images with patient age, history of lesion evolution and genetic risk scores, has achieved better results than systems based on images alone(Jartarkar et al. 2021). But there are many challenges to gathering multimodal data and labels.

9.3 Federated Learning for Data Privacy:

Privacy and data-sharing concerns are a major barrier to building large, diverse and representative data sets. A possible solution is to train algorithms across different sites without having to share the original data. Patients' images stay at each site. It could potentially allow training over a dataset from multiple sites representing different populations, skin tones and lighting conditions while adhering to privacy laws(Jartarkar et al. 2021; Schierle et al. 2026). Dermatology projects using federated learning have demonstrated success but need to have data standardized across participating sites.

9.3 Explainable and Trustworthy Artificial Intelligence:

The opacity of deep learning poses a barrier to trust and adoption by clinicians and regulators. Research on explainable artificial intelligence seeks to generate explanations of algorithms. In dermatology, this could be by showing discriminative features, supplying text explanations or providing case retrievals from training sets. Explainability has yet to be shown to improve clinical outcomes but is probably required for critical diagnostic tasks(Xiong et al. 2019).

9.4 Edge Computing in Low-Resource Settings:

Many artificial intelligence dermatology applications need to be connected to the internet, which is not the case for many low-resource settings where this technology could be applied. Edge computing - performing the algorithms on a mobile phone or other mobile device - may be a solution to this(Nelson et al. 2020). Developing algorithms that can run in a resource-constrained environment without sacrificing accuracy is a current topic in computer science and is relevant for global health.

9.5 Multitemporal Monitoring and Change Detection:

Dermatologists often track lesions' progression, comparing their appearance to prior images to identify changes that could be indicative of cancer. This task can be automated with artificial intelligence, by accurately registering and comparing repeated images to detect suspicious changes. This use of AI is less controversial than fully autonomous diagnosis, it addresses a task that humans aren't good at and could be used without

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

removing human judgement(Felmingham et al. 2022). Preliminary proof-of-concept research is promising, but there are technical issues to be overcome with image registration and normalization.

10 Conclusion:

This review has considered the evidence for the role of artificial intelligence in the diagnosis of skin disease from a number of different angles: the technical, the clinical, the user and the unknown(Mar and Soyer 2018). The title question - will artificial intelligence take over the job of dermatologists - has no binary answer, but a conclusion can be drawn.

Artificial intelligence can accurately diagnose certain tasks in certain circumstances. Under ideal test conditions for diagnosis of melanoma from benign nevi using high-resolution dermoscopies, trained artificial intelligence algorithms are on par with, or better than, average dermatologists(Mar and Soyer 2018). This is an important step towards clinical practice, especially for triage, second reading, and making it accessible to those in need. But the data does not support practice of artificial intelligence diagnosis. There are significant problems with generalization: algorithms perform well on carefully selected test sets, but poorly in the real world, with atypical cases, different skin tones and image artefacts(Mar and Soyer 2018). Current systems cannot reason, synthesize the patient's history and examination, recognize uncertainty and adapt to different scenarios and patients. They don't interview, examine, or recognize cases that are outside their training set, or make therapeutic connections.

An augmentation model is therefore the most fitting. Artificial intelligence should be thought of as an augmentation of dermatologist judgement. This way, algorithms take on the tasks they do well - for example, identifying common lesions under controlled image acquisition conditions - and dermatologists take on the tasks they do well - for example, managing complex cases, incorporating clinical, contextual and other information, dealing with uncertainty and providing the human touch that artificial intelligence cannot provide(Rundle, Hollingsworth, and Dellavalle 2021).

This message has consequences for various parties. For health-care practitioners, artificial intelligence could be an important adjunct to increase efficiency and decrease error rates, but caution and independent verification are advised. For health services managers, artificial intelligence should be used in a complementary, rather than substitute, role for specialist services, with a focus on integration of artificial intelligence into existing workflows and measurement of outcomes(Partridge et al. 2025; Shapiro and Lyakhovitsky 2024). For regulators, better assessment of products before they go to market with respect to prospective validation and patient outcomes is required. For researchers, prospective randomized controlled trials of artificial intelligence vs. standard care are the top research priority.

History is a good guide: the automated external defibrillator (AED) did not replace cardiologists, but in fact dramatically improved survival rates by extending the care of experts in cardiac arrest to non-experts in urgent care settings(Escalé Besa 2025). Thus, artificial intelligence in dermatology might prove to be most useful not as a substitute for the dermatologist, but as an enhancer of the dermatologist's expertise, which can be made available to generalists, residents, patients and settings where dermatologists cannot be found.

In fact, perhaps the best way of putting it is this: artificial intelligence won't replace dermatologists(Elder et al. 2024). But dermatologists who embrace artificial intelligence may increasingly beat those who don't. The specialty's future does not depend on resisting but rather embracing change, and using it to good, evidencebased and patient-focused ends. But rather the question will not be whether artificial intelligence will replace dermatologists, but how artificial intelligence and dermatologists can work in partnership to achieve what neither can achieve on its own(Behara, Bhero, and Agee 2024).

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

REFERENCES:

Aksoy, Serra, Pinar Demircioglu, and Ismail Bogrekci. 2024. 'Advanced artificial intelligence techniques for comprehensive dermatological image analysis and diagnosis', *Dermato*, 4: 173-86.

Anderson, Jane M, Izhaar Tejani, Tory Jarmain, Lisa Kellett, and Ronald L Moy. 2023. 'Artificial intelligence vs medical providers in the dermoscopic diagnosis of melanoma', *Cutis*, 111: 254-58.

Andromeda, Safrian, and Ni Luh Bella Dwijaksara. 2024. 'AI in dermatology: a systematic review on skin cancer detection', *TIERS Information Technology Journal*, 5: 41-51.

Behara, Kavita, Ernest Bhero, and John Terhile Agee. 2024. 'AI in dermatology: a comprehensive review into skin cancer detection', *PeerJ Computer Science*, 10: e2530.

Beltrami, Eric J, Alistair C Brown, Paul JM Salmon, David J Leffell, Justin M Ko, and Jane M Grant-Kels. 2022.

'Artificial intelligence in the detection of skin cancer', *Journal of the American Academy of Dermatology*, 87: 1336-42.

Brancaccio, Gabriella, Anna Balato, Josep Malvehy, Susana Puig, Giuseppe Argenziano, and Harald Kittler. 2024.

'Artificial intelligence in skin cancer diagnosis: a reality check', *Journal of Investigative Dermatology*, 144: 49299.

De, Abhishek, Aarti Sarda, Sachi Gupta, and Sudip Das. 2020. 'Use of artificial intelligence in dermatology', *Indian journal of dermatology*, 65: 352-57.

Du-Harpur, Xinyi, FM Watt, NM Luscombe, and MD Lynch. 2020. 'What is AI? Applications of artificial intelligence to dermatology', *British Journal of Dermatology*, 183: 423-30.

Eapen, Bell R. 2020. 'Artificial intelligence in dermatology: a practical introduction to a paradigm shift', *Indian dermatology online journal*, 11: 881-89.

Elder, Alexandra, Megan O'Donnell Cappelli, Christina Ring, and Nazanin Saedi. 2024. 'Artificial intelligence in cosmetic dermatology: An update on current trends', *Clinics in dermatology*, 42: 216-20.

Escalé-Besa, Anna, Aïna Fuster-Casanovas, Alexander Börve, Oriol Yélamos, Xavier Fustà-Novell, Mireia Esquius Rafat, Francesc X Marin-Gomez, and Josep Vidal-Alaball. 2022. 'Using artificial intelligence as a diagnostic decision support tool in skin disease: protocol for an observational prospective cohort study', *JMIR research protocols*, 11: e37531.

Escalé-Besa, Anna, Josep Vidal-Alaball, Queralt Miró Catalina, Victor Hugo Garcia Gracia, Francesc X MarinGomez, and Aïna Fuster-Casanovas. 2024. "The use of artificial intelligence for skin disease diagnosis in primary care settings: a systematic review." In *Healthcare*, 1192. MDPI.

Escalé-Besa, Anna, Oriol Yélamos, Josep Vidal-Alaball, Aïna Fuster-Casanovas, Queralt Miró Catalina, Alexander Börve, Ricardo Ander-Egg Aguilar, Xavier Fustà-Novell, Xavier Cubiró, and Mireia Esquius Rafat. 2023. 'Exploring the potential of artificial intelligence in improving skin lesion diagnosis in primary care', *Scientific Reports*, 13: 4293.

Escalé Besa, Anna. 2025. 'Using artificial intelligence as a diagnostic decision support tool to help the diagnosis of skin disease in primary healthcare in Catalonia'.

Felmingham, Claire, Samantha MacNamara, William Cranwell, Narelle Williams, Miki Wada, Nikki R Adler, Zongyuan Ge, Alastair Sharfe, Adrian Bowling, and Martin Haskett. 2022. 'Improving Skin cancer Management with ARTificial Intelligence (SMARTI): protocol for a preintervention/postintervention trial of an artificial intelligence system used as a diagnostic aid for skin cancer management in a specialist dermatology setting', *BMJ open*, 12: e050203.

Fliorent, Rebecca, Brian Fardman, Alicia Podwojniak, Kiran Javaid, Isabella J Tan, Hira Ghani, Thu M Truong, Babar Rao, and Candrice Heath. 2024. 'Artificial intelligence in dermatology: advancements and challenges in skin of color', *International Journal of Dermatology*, 63: 455-61.

Gomolin, Arieh, Elena Netchiporouk, Robert Gniadecki, and Ivan V Litvinov. 2020. 'Artificial intelligence applications in dermatology: where do we stand?', *Frontiers in medicine*, 7: 100.

Gordon, Emily R, Megan H Trager, Despina Kontos, Chunhua Weng, Larisa J Geskin, Lydia S Dugdale, and Faramarz H Samie. 2024. 'Ethical considerations for artificial intelligence in dermatology: a scoping review', *British Journal of Dermatology*, 190: 789-97.

Goyal, Manu, Thomas Knackstedt, Shaofeng Yan, and Saeed Hassanpour. 2020. 'Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities', *Computers in biology and medicine*, 127: 104065.

Grzybowski, Andrzej, Kai Jin, and Hongkang Wu. 2024. 'Challenges of artificial intelligence in medicine and dermatology', *Clinics in dermatology*, 42: 210-15.

Han, Seung Seog, Young Jae Kim, Ik Jun Moon, Joon Min Jung, Mi Young Lee, Woo Jin Lee, Chong Hyun Won, Mi Woo Lee, Seong Hwan Kim, and Cristian Navarrete-Dechent. 2022. 'Evaluation of artificial intelligence– assisted diagnosis of skin neoplasms: a single-center, paralleled, unmasked, randomized controlled trial', *Journal of Investigative Dermatology*, 142: 2353-62. e2.

Han, Seung Seog, Ilwoo Park, Sung Eun Chang, Woohyung Lim, Myoung Shin Kim, Gyeong Hun Park, Je Byeong Chae, Chang Hun Huh, and Jung-Im Na. 2020. 'Augmented intelligence dermatology: deep neural networks empower medical professionals in diagnosing skin cancer and predicting treatment options for 134 skin disorders', *Journal of Investigative Dermatology*, 140: 1753-61.

Haykal, Diala. 2024. 'The suitability of AI in dermatology for enhanced skin care', *Journal of Aesthetic Nursing*, 13: 240-51.

Hogarty, Daniel T, John C Su, Kevin Phan, Mohamed Attia, Mohammed Hossny, Saeid Nahavandi, Patricia Lenane, Fergal J Moloney, and Anousha Yazdabadi. 2020. 'Artificial intelligence in dermatology—where we are and the way to the future: a review', *American journal of clinical dermatology*, 21: 41-47.

Jain, Ayush, David Way, Vishakha Gupta, Yi Gao, Guilherme de Oliveira Marinho, Jay Hartford, Rory Sayres, Kimberly Kanada, Clara Eng, and Kunal Nagpal. 2021. 'Development and assessment of an artificial intelligence–based tool for skin condition diagnosis by primary care physicians and nurse practitioners in teledermatology practices', *JAMA network open*, 4: e217249.

Jairath, Neil, Vartan Pahalyants, Rohan Shah, Jason Weed, John A Carucci, and Maressa C Criscito. 2024. 'Artificial intelligence in dermatology: a systematic review of its applications in melanoma and keratinocyte carcinoma diagnosis', *Dermatologic Surgery*, 50: 791-98.

Jartarkar, Shishira R, Clay J Cockerell, Anant Patil, Martin Kassir, Mahsa Babaei, Beate Weidenthaler-Barth, Stephan Grabbe, and Mohamad Goldust. 2023. 'Artificial intelligence in Dermatopathology', *Journal of Cosmetic Dermatology*, 22: 1163-67.

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

Jartarkar, Shishira R, Anant Patil, Uwe Wollina, Michael H Gold, Henner Stege, Stephan Grabbe, and Mohamad Goldust. 2021. 'New diagnostic and imaging technologies in dermatology', *Journal of Cosmetic Dermatology*, 20: 3782-87.

Jutzi, Tanja B, Eva I Krieghoff-Henning, Tim Holland-Letz, Jochen Sven Utikal, Axel Hauschild, Dirk Schadendorf, Wiebke Sondermann, Stefan Fröhling, Achim Hekler, and Max Schmitt. 2020. 'Artificial intelligence in skin cancer diagnostics: the patients' perspective', *Frontiers in medicine*, 7: 233.

Kania, Barbara, Karen Montecinos, and David J Goldberg. 2024. 'Artificial intelligence in cosmetic dermatology', *Journal of Cosmetic Dermatology*, 23: 3305-11.

Karampinis, Emmanouil, Dimitrios Mantzaris, and Anna Mavrophorou. 2026. 'Can Artificial Intelligence Replace Dermatologists?' in, *Artificial Intelligence Applications in Dermatology: Dermatology ex Machina: Clinical applications* (Springer).

Karampinis, Emmanouil, Christina-Marina Zoumpourli, Aimilios Lallas, Zoe Apalla, John Paoli, Bengü Nisa Akay, Cristian Navarette-Dechent, Behera Biswanath, Nkechi Enechuwku, and Peter Chai. 2026. 'Perspectives on Artificial Intelligence in Dermatology: An International Cross-Sectional Study', *Medicina*, 62: 759.

Koka, Sanjay Satya-Akunuri, and Craig G Burkhart. 2023. 'Artificial intelligence in dermatology: current uses, shortfalls, and potential opportunities for further implementation in diagnostics and care', *The Open Dermatology Journal*, 17.

Kumar, Sahil, Anirudh Gupta Faisal Saeed, and Sudhair Abbas Bangash. 2024. 'Artificial intelligence in dermatology', *Rheumatology*, 41: 2345-52.

Li, Cheng-Xu, Chang-Bing Shen, Ke Xue, Xue Shen, Yan Jing, Zi-Yi Wang, Feng Xu, Ru-Song Meng, Jian-Bin Yu, and Yong Cui. 2019. "Artificial intelligence in dermatology: past, present, and future." In, 2017-20. Chinese Medical Journals Publishing House Co., Ltd. 42 Dongsi Xidajie ....

Li, Chengxu, Je-Ho Mun, Paola Pasquali, Hang Li, H Peter Soyer, and Yong Cui. 2021. 'Progress and Prospects on Skin Imaging Technology, Teledermatology and Artificial Intelligence in Dermatology', *Frontiers in medicine*, 8: 757538.

Lim, BCW, and G Flaherty. 2019. 'Artificial intelligence in dermatology: are we there yet?', *British Journal of Dermatology*, 181: 190-91.

Lim, K, G Neal-Smith, C Mitchell, J Xerri, and P Chuanromanee. 2022. 'Perceptions of the use of artificial intelligence in the diagnosis of skin cancer: an outpatient survey', *Clinical and Experimental Dermatology*, 47: 542-46.

Liopyris, Konstantinos, Stamatios Gregoriou, Julia Dias, and Alexandros J Stratigos. 2022. 'Artificial intelligence in dermatology: challenges and perspectives', *Dermatology and Therapy*, 12: 2637-51.

Manole, Ionela, and George-Sorin Tiplica. 2024. 'Integrating Artificial Intelligence in dermatology: progress, challenges and perspectives', *Romanian Medical Journal*, 71.

Mar, VJ, and HP Soyer. 2018. 'Artificial intelligence for melanoma diagnosis: how can we deliver on the promise?', *Annals of Oncology*, 29: 1625-28.

Maron, Roman C, Jochen S Utikal, Achim Hekler, Axel Hauschild, Elke Sattler, Wiebke Sondermann, Sebastian Haferkamp, Bastian Schilling, Markus V Heppt, and Philipp Jansen. 2020. 'Artificial intelligence and its effect on dermatologists' accuracy in dermoscopic melanoma image classification: web-based survey study', *Journal of medical Internet research*, 22: e18091.

Frontier in Medical & Health Research
ISSN: (e) 3007-1607 (p) 3007-1593

Marri, Shiva Shankar, Arun C Inamadar, Ajit B Janagond, and Warood Albadri. 2023. 'Analyzing the Predictability of an artificial intelligence app (Tibot) in the diagnosis of dermatological conditions: A crosssectional study', *JMIR dermatology*, 6: e45529.

Nahm, William J, Nayyab Sohail, Joshua Burshtein, Mohamad Goldust, and Maria Tsoukas. 2025. 'Artificial intelligence in dermatology: a comprehensive review of approved applications, clinical implementation, and future directions', *International Journal of Dermatology*, 64: 1568-83.

Nelson, Caroline A, Swapna Pachauri, Rosie Balk, Jeffrey Miller, Rushan Theunis, Justin M Ko, and Carrie L Kovarik. 2021. 'Dermatologists' perspectives on artificial intelligence and augmented intelligence—a crosssectional survey', *JAMA dermatology*, 157: 871-74.

Nelson, Caroline A, Lourdes Maria Pérez-Chada, Andrew Creadore, Sara Jiayang Li, Kelly Lo, Priya Manjaly, Ashley Bahareh Pournamdari, Elizabeth Tkachenko, John S Barbieri, and Justin M Ko. 2020. 'Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study', *JAMA dermatology*, 156: 501-12.

Oscar, ZAAR, Alexander Larson, Sam Polesie, Karim Saleh, Mikael Tarstedt, Antonio Olives, Andrea Suarez, Martin Gillstedt, and Noora Neittaanmäki. 2020. 'Evaluation of the diagnostic accuracy of an online artificial intelligence application for skin disease diagnosis', *Acta Dermato-Venereologica*, 100: 5873.

Partridge, Brad, Monika Janda, Nicole Gillespie, Carina Vasconcelos Silva, Chris Arnold, Lisa Abbott, Tony Caccetta, and H Peter Soyer. 2025. 'Attitudes towards the use of artificial intelligence in dermatology: a survey of Australian dermatologists', *Australasian Journal of Dermatology*, 66: e279-e86.

Patel, Shaan, Jordan V Wang, Kiran Motaparthi, and Jason B Lee. 2021. 'Artificial intelligence in dermatology for the clinician', *Clinics in dermatology*, 39: 667-72.

Phillips, Michael, Helen Marsden, Wayne Jaffe, Rubeta N Matin, Gorav N Wali, Jack Greenhalgh, Emily McGrath, Rob James, Evmorfia Ladoyanni, and Anthony Bewley. 2019. 'Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions', *JAMA network open*, 2: e1913436.

Rundle, Chandler W, Parker Hollingsworth, and Robert P Dellavalle. 2021. 'Artificial intelligence in dermatology', *Clinics in dermatology*, 39: 657-66.

Salinas, Maria Paz, Javiera Sepúlveda, Leonel Hidalgo, Dominga Peirano, Macarena Morel, Pablo Uribe, Veronica Rotemberg, Juan Briones, Domingo Mery, and Cristian Navarrete-Dechent. 2024. 'A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis', *NPJ digital medicine*, 7: 125.

Samaran, Romain, Jean-Matthieu L'Orphelin, Brigitte Dreno, Cédric Rat, and Anne Dompmartin. 2021. 'Interest in artificial intelligence for the diagnosis of non-melanoma skin cancer: a survey among French general practitioners', *European Journal of Dermatology*, 31: 457-62.

Sangers, Tobias E, Marlies Wakkee, Folkert J Moolenburgh, Tamar Nijsten, and Marjolein Lugtenberg. 2023. 'Towards successful implementation of artificial intelligence in skin cancer care: a qualitative study exploring the views of dermatologists and general practitioners', *Archives of dermatological research*, 315: 1187-95.

Schierle, Felicitas, Jon Lindström Bolmgren, Mette Deleuran, Oliver Welter, and Vanessa Gebauer. 2026. 'Precision dermatology 2050: AI-driven personalized monitoring and individualized treatment', *Journal of the European Academy of Dermatology and Venereology*, 40: 431-39.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

Shapiro, Jonathan, and Anna Lyakhovitsky. 2024. 'Revolutionizing teledermatology: Exploring the integration of artificial intelligence, including Generative Pre-trained Transformer chatbots for artificial intelligence-driven anamnesis, diagnosis, and treatment plans', *Clinics in dermatology*, 42: 492-97.

Shen, Changbing, Chengxu Li, Feng Xu, Ziyi Wang, Xue Shen, Jing Gao, Randy Ko, Yan Jing, Xiaofeng Tang, and Ruixing Yu. 2020. 'Web-based study on Chinese dermatologists' attitudes towards artificial intelligence', *Annals of translational medicine*, 8: 698.

Stoneham, Sophie, Amy Livesey, Hywel Cooper, and Charles Mitchell. 2024. 'ChatGPT versus clinician: challenging the diagnostic capabilities of artificial intelligence in dermatology', *Clinical and Experimental Dermatology*, 49: 707-10.

Talebi-Liasi, Faezeh, and Orit Markowitz. 2020. 'Is artificial intelligence going to replace dermatologists', *Cutis*, 105: 28-31.

Thunga, Sukruthi, Marius Khan, Soo Ick Cho, Jung Im Na, and Jane Yoo. 2025. 'AI in aesthetic/cosmetic dermatology: current and future', *Journal of Cosmetic Dermatology*, 24: e16640.

Vatiwutipong, Pat, Sirawich Vachmanus, Thanapon Noraset, and Suppawong Tuarob. 2023. 'Artificial intelligence in cosmetic dermatology: a systematic literature review', *IEEE Access*, 11: 71407-25.

Wada, Miki, ZongYuan Ge, Stephen J Gilmore, and Victoria J Mar. 2020. 'Use of artificial intelligence in skin cancer diagnosis and management', *Medical Journal of Australia*, 213: 256-59. e1.

Woźniacka, Anna, Sebastian Patrzyk, and Maksym Mikołajczyk. 2021. 'Artificial intelligence in medicine and dermatology', *Advances in Dermatology and Allergology/Postępy Dermatologii i Alergologii*, 38: 948-52.

Xiong, Mulin, Jacob Pfau, Albert T Young, and Maria L Wei. 2019. 'Artificial intelligence in teledermatology', *Current Dermatology Reports*, 8: 85-90.

Ye, Shengzhen, and Mingling Chen. 2023. 'The emerging role of artificial intelligence in diagnosis and clinical analysis of dermatology', *Dermatologica Sinica*, 41: 145-52.

Young, Albert T, Mulin Xiong, Jacob Pfau, Michael J Keiser, and Maria L Wei. 2020. 'Artificial intelligence in dermatology: a primer', *Journal of Investigative Dermatology*, 140: 1504-12.

Zbrzezny, Agnieszka M, and Tomasz Krzywicki. 2025. 'Artificial intelligence in dermatology: A review of methods, clinical applic tions, and perspectives', *Applied Sciences*, 15: 7856.