MEDWAVE
PEER REVIEWED MEDICAL JOURNAL

REVIEWS

Artificial intelligence for skin lesion classification and diagnosis in dermatology: A narrative review

Paula Andrea Ramos Chaparro[a], Erwin Hernando Hernández Rincón[b*], Nicolas Melo Sierra[a], Gabriela Alejandra Osorio Betancourt[a], Diana Marcela Díaz Quijano[c], Samuel David Barbosa Ardila[d], Mariam Rolón Cadena[e], Alejandro Rueda Rodríguez[a]

[a]Universidad de La Sabana, Chía, Colombia; [b]Department of Family Medicine and Public Health, Universidad de La Sabana, Chía, Colombia; [c]Department of Epidemiology, Universidad de La Sabana, Chía, Colombia; [d]Universitat Oberta de Catalunya, Barcelona, Spain; [e]Pathology Department, Fundación Santa Fe de Bogotá, Bogotá, Colombia

ABSTRACT

INTRODUCTION Artificial intelligence (AI) is increasingly present in dermatology, demonstrating accuracy levels comparable to, or even superior to, those of dermatologists in diagnosing skin lesions from clinical and dermoscopic images. This review provides an overview of AI's role in the automated classification and monitoring of skin lesions.

OBJECTIVE To explore and map the existing literature on the use and benefits of AI in dermatology.

METHODS This narrative review focused on exploring the use and benefits of AI in dermatology, utilizing MeSH/DeCS terms such as "Dermatology," "Artificial Intelligence," "Diagnosis," and "Computer Aided Diagnosis." Three databases (PubMed, Lillacs, and Scopus) from 2008 to 2024. We excluded articles that did not focus on dermatology, lacked the topic of Artificial Intelligence, or presented theoretical designs without practical application or evidence, resulting in a final selection of forty-four articles.

RESULTS The results strongly support AI's effectiveness, displaying its precision in diagnosis, comparable to or exceeding that of human dermatologists across diverse tasks. The evolution of AI in dermatology implies a substantial transformation in care, extending its applications from skin cancer to various dermatological pathologies. While emphasizing the vital collaboration between AI and healthcare professionals, a critical gap remains in the real-world clinical validation of AI. Ethical considerations, especially in automated decision-making, need careful attention.

CONCLUSIONS This narrative review highlights the crucial role of AI in dermatology, emphasizing its potential to enhance diagnostic accuracy for skin lesions.

KEYWORDS Dermatology, Artificial Intelligence, Diagnosis, Computer Aided Diagnosis

INTRODUCTION

Artificial intelligence (AI) is transforming dermatology, achieving accuracy comparable to or exceeding dermatologists in diagnosing skin lesions from clinical and dermoscopic images. Despite these advances, a gap remains in real-world clinical validation. This narrative review explores AI applications in dermatological image analysis, assessing capabilities, challenges, failure modes, and performance evaluation metrics.

AI has the potential to enhance access and quality of care, with studies showing it can match or surpass experts in diagnosing multiple skin conditions, highlighting its transformative potential [1–4].

Dermatology faces increasing complexity and a need for rapid, accurate decisions. AI has evolved to create multimodal models integrating medical records, lab reports, and dermoscopic images. While early research focused on skin cancer, especially melanoma, current efforts aim to improve diagnostic efficiency across various conditions and support therapeutic recommendations [5,6].

In parallel, the increasing burden of skin diseases in primary care has generated notable interest in the application of machine learning models as collaborative tools. Statistics reveal that dermatological consultations account for around 20% of total annual visits to primary care, with around 35% being referred to dermatology specialists [7]. Machine learning models, with their ability to analyze and optimize complex

*** Corresponding author** erwinhr@unisabana.edu.co
Citation Ramos Chaparro PA, Hernández Rincón EH, Melo Sierra N, Osorio Betancourt GA, Díaz Quijano DM, Barbosa Ardila SD, et al. Artificial intelligence for skin lesion classification and diagnosis in dermatology: A narrative review. Medwave 2025;25(10):e3120
DOI 10.5867/medwave.2025.10.3120
Submitted Jun 18, 2025, **Accepted** Aug 26, 2025,
Published Nov 20, 2025
Postal address Puente del Comun University Campus, Km 7, North Highway of Chía 250001, Colombia

MAIN MESSAGES

- Artificial intelligence (AI) is revolutionizing dermatology by offering accurate diagnoses of skin lesions, comparable to or superior to those of dermatologists.
- This review highlights the growing applicability of AI beyond skin cancer, including multiple dermatological pathologies.
- There is a significant limitation in clinical validation in real-life settings, which restricts its widespread implementation.
- The main finding shows that AI improves diagnostic accuracy and could optimize dermatological care in the future.
- Greater attention is needed to the ethical implications of automated decision-making in healthcare.

datasets [8] as shown in Table 1, are presented as an effective way to improve efficiency in the detection and classification of skin diseases in this environment. The convergence of these technological developments highlights the importance of a comprehensive vision ranging from primary to specialized care, integrating the application of artificial intelligence in dermatology with the growing need to optimize the management of skin pathologies [9].

OBJECTIVE

The objective of this review is to comprehensively explore and map the existing literature on the use and benefits of artificial intelligence (AI) in the field of dermatology. We aim to identify and synthesize the diverse research and approaches that have addressed the integration of AI in the diagnosis, treatment, and management of skin diseases. In this way, it provides an overview that serves as an essential reference for future research and development in the application of artificial intelligence in dermatology.

METHODS

This review was conducted following the guidelines from the Preferred Reporting Items for Systematic Review and Meta-Analysis adapted from Arksey for narrative reviews (PRISMA-ScR) [10] and refined by Levac et al [11], a systematic review was conducted following these steps: 1) identifying the research question; 2) searching and identifying relevant studies; 3) selecting studies; 4) data analysis; 5) summarizing and reporting results. We adhered to the reporting guidelines for systematic reviews and meta-analyses tailored for narrative reviews (PRISMA-ScR) [12]. We searched three databases (PubMed, Lilacs and Scopus). To review the available evidence on the use and benefits of artificial intelligence in dermatology.

For the above, MeSH/DeCS terms such as "Dermatology", "Artificial Intelligence", "Diagnosis", and "Computer Aided Diagnosis" were used, in their translation into English, Portuguese, and Spanish, limiting the publication date from January 1st, 2008. It was filtered by articles that were reflective texts or research studies including topics such as the application of AI in dermatology and the comparison at the time of diagnosis between AI and dermatologists.

The eligibility criteria included publications with empirical and observational data that studied or discussed the use of artificial intelligence in dermatology, in English, Portuguese, and Spanish, between 2008 and 2024. We excluded articles with theoretical data, clinical management guidelines, protocols, and publications without access to abstracts and full documents. For data extraction, four independent authors reviewed and selected titles and abstracts of candidate publications. Regular meetings were held with all authors to resolve discrepancies and reach a consensus on the articles to be included in the review.Two additional independent researchers resolved further disagreements. Duplicates were removed, and the selected documents in full text were obtained through the University of La Sabana library. Following the final selection of documents, the following data were extracted from eligible studies: authors, document type, population characteristics, objective, year of publication, authors' country, main findings, and limitations, presented in Table 2.

RESULTS

Of 329 articles that were found in databases, 102 duplicates were discarded. By title and abstract, 118 were discarded because dermatology was not the central topic, did not include the topic Artificial Intelligence, or featured theoretical designs without application or evidence. The remaining 109 were downloaded and stored in PDF format for complete review by the research group. Subsequently, 64 articles were discarded for not meeting the inclusion criteria after review, leaving a total of 44 articles, as shown in Figure 1. This figure summarizes the identification process, including the understanding of eligibility criteria based on originality, methodology, and relevance, as well as the inclusion of articles.

The selected studies included several types of articles, which allowed us to obtain a spectrum of various aspects that must be evaluated in the study of artificial intelligence in dermatology. Among these, the vast majority are experimental studies where tests on sensitivity and specificity are carried out. When making diagnoses of the machines in front of diverse types of people (general practitioners not trained in dermatology, in some cases, and other dermatologists with vast experience). The study type was primarily experimental, with 22 articles in total; followed by seven observational studies, five cross-sectional analyses, two systematic reviews & meta-analyses, two retrospective studies, two reviews and perspectives, two comparative studies, and two literature/narrative reviews.

Table 1. Types of artificial intelligence [8].

AI type	Characteristics	Applications	Limitations
Deep learning	Uses deep neural networks to solve complex problems	Pattern recognition in images, natural language processing, and gaming	Requires substantial amounts of data, computationally intensive
Neural networks	Computational model inspired by the human brain, composed of interconnected nodes	Speech recognition, pattern recognition, image processing	Requires extensive training, complexity in interpretation
Reinforcement learning	An agent learns through interaction with an environment and receives feedback in the form of rewards or penalties	Gaming, robotics, and autonomous decision-making	Requires extended training time, design complexity
Feedforward neural networks	Information flows in one direction, without cycles or loops	Pattern recognition in images, natural language processing	Does not handle sequential data well, such as time series
Recurrent neural networks	Include cyclical connections to handle sequential data and retain information over time	Natural language processing, automatic translation, and time series	May face issues of vanishing or exploding gradients
Convolutional neural networks	Specifically designed for image processing, with shared weight layers and pooling layers	Object recognition in images, and medical diagnosis based on images	Requires enormous amounts of training data, specialized hardware
Decision trees	Tree-shaped models that make decisions based on features at each node	Medical diagnosis, customer classification, decision-making	May be prone to overfitting, difficulty capturing complex relationships
Random forests	An ensemble method that combines multiple decision trees to improve accuracy	Classification, regression, anomaly detection	Complexity in interpretation, resource-intensive
Support vector machines	Classifies data points by finding the hyperplane that best separates classes	Text classification, speech recognition, bioinformatics	May not perform well on large datasets, sensitive to the choice of kernel
Logistic regression	Linear model used for binary classification problems	Risk prediction, disease classification, and marketing	Limited to binary classification problems, does not handle nonlinear relationships

Source: prepared by the authors based on the results of the study.

Evaluating the performance of artificial intelligence in dermatology

In evaluating the performance of AI in dermatology, a critical aspect is comparing its effectiveness to that of human professionals' diagnostic capacity. Numerous studies consistently demonstrate that AI models exhibit a classification accuracy for skin lesions comparable to, and in some cases, superior to that of dermatologists. This marks a significant milestone in dermatological diagnosis, suggesting a transformative potential for enhancing efficiency in clinical decision-making. For instance, an experimental study conducted in 2020 assessed the diagnostic performance of a convolutional neural network in differentiating between combined naevi and melanomas, comparing dermatologists' results with those of the convolutional neural network. The results revealed notable outcomes: the convolutional neural network's sensitivity was 97.1%, its specificity was 78.8%, dermatologists' sensitivity was 90.6%, and dermatologists' specificity was 71.0%. Notably, a particular benefit was observed in 'beginners' when aided by convolutional neural network verification, with a diagnostic odds ratio of 98, indicating improved diagnostic performance. This highlights the potential of AI, particularly in assisting less experienced dermatologists in accurately differentiating lesions [16].

The remarkable advancements in AI's diagnostic accuracy in dermatology are not confined to recent breakthroughs. The roots of this progress can be traced back to 2009, when Rajpara et al. conducted a study evaluating the diagnostic accuracy of dermoscopy and digital dermoscopy/artificial intelligence in melanoma diagnosis. In this study, various dermoscopic algorithms were compared with each other and with digital dermoscopy/artificial intelligence for melanoma detection. The findings revealed a pooled sensitivity for artificial intelligence that slightly surpassed that of dermoscopy performed by dermatologists (91% vs. 88%, $p = 0.076$). Although dermoscopy demonstrated significantly better pooled specificity compared to artificial intelligence (86% vs. 79%, $p < 0.001$), the pooled diagnostic odds ratio showed no significant difference (51.5 for dermoscopy and 57.8 for artificial intelligence, $p = 0.783$). This historical perspective highlights the continuous evolution and validation of AI in dermatological diagnostics over the years. In contrast to these advancements, novel investigations reveal a clear progression [55].

Implications for clinical decision making

The rapid advancement of technology has enabled the integration of multimodal data, including clinical information, laboratory reports, and visual data, into AI models in dermatology [49]. This opens the door to more complete and accurate

Table 2. Detailed description of the selected articles.

Article reference	Authors - Country of origin	Type of article	Objective	Main Findings
[13]	Han SS., et al. Korea 2020	Experimental study	Train an algorithm for: Malignancy prediction. Suggesting primary treatment options.	The algorithm showed high AUCs for malignancy detection (Edinburgh: 0.928; SNU: 0.937) and treatment suggestion (AUCs up to 0.918). Multi-class classification accuracy reached up to 56.7% (top 1) and 92.0% (top 5).
[14]	Gouabou AC., et al. France 2022	Experimental study	To develop a new framework for automated melanoma diagnosis using artificial intelligence, specifically CNNs.	Achieved an area under the receiver operating curve of 0.93 for melanoma, 0.96 for nevus, and 0.97 for benign keratosis.
[15]	Guimarães, et al. Germany 2022	Experimental study	To propose a fully automatic approach based on CNNs for the analysis of multiphoton tomography data in dermatology, specifically for atopic dermatitis (AD).	The proposed algorithm correctly diagnosed AD in 97.0 ± 0.2% of all images presenting living cells. Achieved a sensitivity of 0.966 ± 0.003, specificity of 0.977 ± 0.003, and F-score of 0.964 ± 0.002.
[16]	Fink A., et al. Germany 2020	Diagnostic test accuracy study	To assess the diagnostic performance of a CNN in comparison with dermatologists for the differentiation of melanomas from combined naevi.	CNN sensitivity: 97.1% CNN specificity: 78.8% Dermatologists' sensitivity: 90.6% Dermatologists' specificity: 71.0% Benefit observed in 'beginners' with CNN verification (DOR = 98).
[17]	Brinker, et al. Germany 2019	Experimental (CNN training)	To demonstrate the training of an image-classifier CNN that outperforms the winner of the ISBI 2016 CNNs challenge using open-source images exclusively.	Achieved an average precision of 0.709 (vs. 0.637 of the ISBI winner) and an area under the receiver operating curve of 0.85.
[18]	Yoo KH., et al. Korea 2022	Comparative study	Investigate errors in estimating BSA in psoriasis by comparing physicians' results with those of computer-assisted.	The mean proportion of correct assessments by physicians was 49.4%. Physicians tended to overestimate the BSA of psoriatic lesions by 8.76% ± 8.82% compared to the CAIA.
[19]	Hekler A., et al. Germany 2019	Comparative study	Investigate the potential benefit of combining human and artificial intelligence for skin cancer classification.	For the multiclass task, the combination of human and machine achieved an accuracy of 82.95%, 1.36% higher than the best individual classifier (CNN).
[20]	Winkler, et al. Germany 2021	Cross-sectional analysis	To investigate whether scale bars in dermoscopic images are associated with the diagnostic accuracy of a market-approved CNN.	In images without scale bars, the CNN had a sensitivity of 87.0%, specificity of 87.9%, and ROC-AUC of 0.953. Superimposed scale bars may impair the CNN's diagnostic accuracy by increasing false-positive diagnoses.
[21]	Sies K., et al. Germany 2020	Cross-sectional study	Head-to-head comparison of a market-approved CNN (Moleanalyzer-Pro™, developed in 2018) to a CIA (Moleanalyzer-3™/Dynamole™; developed in 2004).	CNN sensitivity: 77.6%, specificity: 95.3%, ROC-AUC: 0.945.CIAI sensitivity: 53.4%, specificity: 86.6%, ROC-AUC: 0.738.Pairwise comparisons favor CNN, indicating clear outperformance (all $p < 0.001$).
[22],[23]	Chanki Yu, et al. Korea 2018	Experimental	To evaluate the usefulness of a CNN in the early diagnosis of acral melanoma and benign nevi from dermoscopy images on the hands and feet, comparing with dermatologist.	The CNN achieved an accuracy of 83.51% and 80.23%, higher than the non-expert's evaluation (67.84%, 62.71%) and close to that of the expert (81.08%, 81.64%).
(31)[24]	Ba W., et al. China 2022	Experimental	To evaluate the impact of CNN assistance on dermatologists for the interpretation of clinical images of cutaneous tumors.	The CNN achieved an overall accuracy of 78.45% and kappa of 0.73 in classifying cutaneous tumors. Dermatologists with CNN assistance showed significantly higher accuracy and kappa compared to unassisted dermatologists in interpreting clinical images.

(Continued)

(Continued)

Article reference	Authors - Country of origin	Type of article	Objective	Main Findings
[25]	Hekler A., et al. Germany 2019	Experimental	To directly compare the performance of a CNN with 11 histopathologists in classifying histopathological melanoma images.	The CNN achieved a significantly higher mean sensitivity, specificity, and accuracy (76%, 60%, 68%) compared to the 11 histopathologists (51.8%, 66.5%, 59.2%) in classifying cropped histopathological melanoma images.
[26]	Zunair H., et al. Canada 2020	Experimental	To address the class imbalance problem in skin lesion datasets and improve melanoma detection by proposing a two-stage framework.	The proposed approach demonstrates superiority over standard baseline methods, achieving significant performance improvements in skin lesion classification.
[27]	Birkner, et al. Germany 2022	Experimental study	To develop a deep CNN for analyzing wound photographs to aid in the diagnosis of pyoderma gangrenosum and differentiate it from conventional leg ulcers.	A CNN was trained using 422 expert-selected pictures of PG and LU. In a comparison between man and machine, 33 PG pictures and 36 LU pictures were presented for diagnosis to 18 dermatologists and the CNN.
[28]	Yanagisawa, et al. Japan 2019	Methodological study	To develop a CNN segmentation model that can automatically extract lesions and generate segmented images of the skin area from non-standardized conventional photo images.	The CNN segmentation model achieved approximately 90% of sensitivity and specificity in differentiating atopic dermatitis from malignant diseases and complications.
[29]	Brinker, et al. Germany 2019	Observational study	To compare the performance of a CNN, exclusively trained with dermoscopic images, with dermatologists in the clinical image classification of melanoma.	Dermatologists achieved a mean sensitivity of 89.4% and a mean specificity of 64.4% with clinical images.
[30]	Felmingham C., et al. Australia 2022	Prospective clinical study.	Validate the performance of a CNN developed by MoleMap Ltd and Monash eResearch in diagnosing skin cancers in a real-world clinical setting.	The impact of the AI algorithm on diagnostic and management decisions will be evaluated by: (1) comparing the initial management decision of the registrar with their AI-assisted decision, and (2) comparing the benign to malignant ratio (for lesions biopsied) between the preintervention and postintervention periods.
[31]	Tognetti, et al. Italy 2021	Experimental	To develop a deep CNN model to support dermatologists in the classification and management of atypical melanocytic skin lesions (aMSL) by integrating dermoscopic images with clinical data.	The deep CNN_aMSL model achieved the best accuracy, with an AUC of 90.3%, sensitivity of 86.5%, and specificity of 73.6%. It was particularly effective in supporting management decisions, reducing inappropriate excisions.
[32]	Burlina, et al. USA 2019	Experimental	To develop deep learning approaches using deep convolutional neural networks for detecting acute Lyme disease from erythema migrans images.	The machine achieved an accuracy of 86.53%, ROC-AUC of 0.9510, and Kappa of 0.7143 for detecting erythema migrans.
[33]	Filipescu, et al. Romania 2021	Experimental	To classify dermoscopic images from distinct categories of lesions, distinguishing between malignant and benign cases.	The model achieved an overall accuracy of 78.11%. Out of 5031 samples in the test subset, 3958 were correctly classified.
[6]	Young, et al. USA 2020	Literature review	Review dermatological applications of deep learning in artificial intelligence.	AI accuracy matches or exceeds dermatologists in skin lesion diagnosis. Lack of real-world clinical validation. Addresses applications in teledermatology, clinical assessment, and dermatopathology.

(Continued)

(Continued)

Article reference	Authors - Country of origin	Type of article	Objective	Main Findings
[34]	Pai V., et al. India 2021	Narrative review	To assess the application of artificial intelligence in dermatology, emphasizing its impact on medical diagnosis, medical statistics, robotics, and human biology.	The application of deep learning methods, especially artificial neural networks, in the analysis of medical images, such as pigmented skin lesions, is highlighted.
[35]	Lee, et al. Korea 2020	Cross-sectional analysis	Develop a deep learning framework to determine the Severity of Alopecia Tool (SALT) score for measurement of hair loss in patients with alopecia areata.	Improved accuracy and interrater reliability in the challenge study with computer-assisted approach Increased explanatory power of SALT score for predicting hair regrowth
[36]	Koo, et al. Korea 2021	Experimental	To develop a deep learning-based autodetection model for quicker, more convenient, and consistent detection of hyphae in superficial fungal infections using microscopy images.	The model achieved high sensitivity and specificity, with values ranging from 93.2% to 100%, indicating reliable accuracy in detecting hyphae in microscopic images.
[37]	S.S. Han, et al. Korea 2018	Observational study	To test the use of a deep learning algorithm for the classification of clinical images of various skin diseases, comparing its performance with dermatologists.	The algorithm demonstrated high accuracy (area under the curve) for the diagnosis of basal cell carcinoma, squamous cell carcinoma, intraepithelial carcinoma, and melanoma across different datasets.
[38]	S.S. Han, et al. Korea 2018	Observational study	To evaluate the diagnostic accuracy of deep learning models in the diagnosis of onychomycosis using standardized nail images and compare it with dermatologists' assessments.	The AI (ensemble model) demonstrated high sensitivity, specificity, and area under the curve values for different datasets.
[3]	Fujisawa Y., et al. Japan 2018	Observational study	To determine the efficiency of a skin tumor classification system using deep-learning technology with a small dataset of clinical images.	The deep CNN achieved an overall classification accuracy of 76.5%, with high sensitivity (96.3%) and specificity (89.5%).
[39]	Thomsen, et al. Denmark 2019	Systematic review	To evaluate the utilization of machine learning in dermatology by conducting a systematic review of existing literature.	Identified eight major categories where machine learning tools were tested in dermatology. Most systems involved image recognition tools primarily aimed at binary classification of malignant melanoma (MM).
[40]	Koller S., et al. Austria 2011	Observational study	To investigate the applicability of an automated image analysis system using a machine learning algorithm for diagnostic discrimination of benign and malignant melanocytic skin tumors in RCM	The classification tree analysis correctly classified 93.60% of melanoma and 90.40% of nevi images in the learning set. When applied to the independent test set, the system classified 46.71% of images in benign lesions as 'malignant' and 55.68% in malignant lesions
[41]	Gareau, et al. USA 2017	Observational study	To develop an automated approach for generating imaging biomarkers, applying machine learning algorithms to create an overall risk	The approach achieved 98% sensitivity and 36% specificity for melanoma detection, approaching the sensitivity/specificity of expert lesion evaluation.
[42]	Bruland, et al. Germany 2015	Experimental	To develop a software tool: PIACS - Prurigo Image Analyzing and Comparing System, for automatic detection, categorization, and comparison of scratch-related skin lesions.	The software detected and categorized lesions with an overall sensitivity of 95.7% and an accuracy of 75.3%.

(Continued)

(Continued)

Article reference	Authors - Country of origin	Type of article	Objective	Main Findings
[43],[44]	Gareau, et al. USA 2010	Analytic diagnostic accuracy study	To assess the potential of in-vivo reflectance confocal microscopy (RCM) for the early detection of superficial spreading melanoma by quantifying specific markers	The pattern recognition algorithm identified PMs in all five superficial spreading melanoma and none in the nevi.
44]	Lim K., et al. UK 2021	Cross-sectional questionnaire study	To obtain patient opinions on the use of AI in a dermatology setting, specifically in aiding the diagnosis of skin cancers.	47% of respondents were not concerned if AI technology was used by a skin specialist to aid skin cancer diagnosis.81% considered it important for a dermatologist to examine and confirm a diagnosis and to be present for the discussion of a cancer diagnosis.
[45]	Samaran R., et al. France 2021	Cross-sectional study using a questionnaire.	Assess difficulties faced by GPs in diagnosing non-pigmented skin tumors (NPSTs). Determine the interest of GPs in using artificial intelligence software as a diagnostic tool for NPSTs.	147 respondents (98%) had faced difficulties diagnosing NPSTs. 86% agreed that an AI diagnostic tool could be useful in a GP's office. 83% agreed that AI could change their practice. 68% would not be willing to pay for this kind of software.
[46]	S.S. Han, et al. Korea 2020	Experimental (algorithm development and evaluation)	To evaluate whether an algorithm could automatically locate suspected areas and predict the probability of a skin lesion being malignant.	The algorithm achieved an area under the receiver operating characteristic curve of 0.910 in the validation data set. At a high-sensitivity cutoff threshold, the model demonstrated a sensitivity of 76.8% and specificity of 90.6%.
[47]	Hurault G., et al. UK 2020	Experimental study	To develop a mechanistic machine learning model that predicts the patient-specific evolution of atopic dermatitis AD severity scores on a daily basis.	The model successfully predicted the daily evolution of AD severity scores at the individual level, improving chance-level forecast by 60%.
[48]	A Jain, et al. USA 2021	Multiple-reader, multiple-case diagnostic study.	To evaluate an AI based tool that assists with the diagnosis of dermatologic conditions by primary care physicians and 20 nurse practitioners.	AI assistance was significantly associated with higher agreement with reference diagnoses. For primary care physicians, the increase in diagnostic agreement was 10% (from 48% to 58%). For nurse practitioners, the increase was 12% (from 46% to 58%).
[49]	R. Moore et al. USA 2021	Observational study	To assess whether automated digital analysis of TILs improves the accuracy of predicting disease-specific survival (DSS) in early-stage melanomas.	ADTA, when applied to a validation cohort, stratified patients into high and low-risk groups, demonstrating significant differences in Kaplan-Meier analysis ($p \leq 0.001$). Multivariable Cox proportional hazards analysis showed that ADTA contributed to DSS prediction
[50]	K. Sohn, et al. USA 2021	Retrospective study	Develop an algorithm to aid surgeons in detecting BCC on frozen sections during Mohs micrographic surgery	The final model achieved an AUC of ROC of 0.753, with a sensitivity of 70.6% at a specificity of 79.1%.
[51]	Chang W., et al. Taiwan 2013	Retrospective study	Evaluate the feasibility of using CADx for diagnosing both melanocytic and non-melanocytic skin lesions.	CADx system performance: Area Under the Receiver Operating Characteristic (ROC) curve Az = 0.949, sensitivity = 85.63%, specificity = 87.65%, maximum accuracy = 90.64%.
[52]	Blum A., et al. Germany 2020	Review and perspective	To examine how AI can allow medical work to focus on skin cancer patients and expedite treatment.	Advantages of AI in dermato-oncology include optimized medical focus on skin cancer patients, enhanced treatment speed, and user learning opportunities. Problematic aspects in AI use encompass medicolegal issues and remuneration challenges.

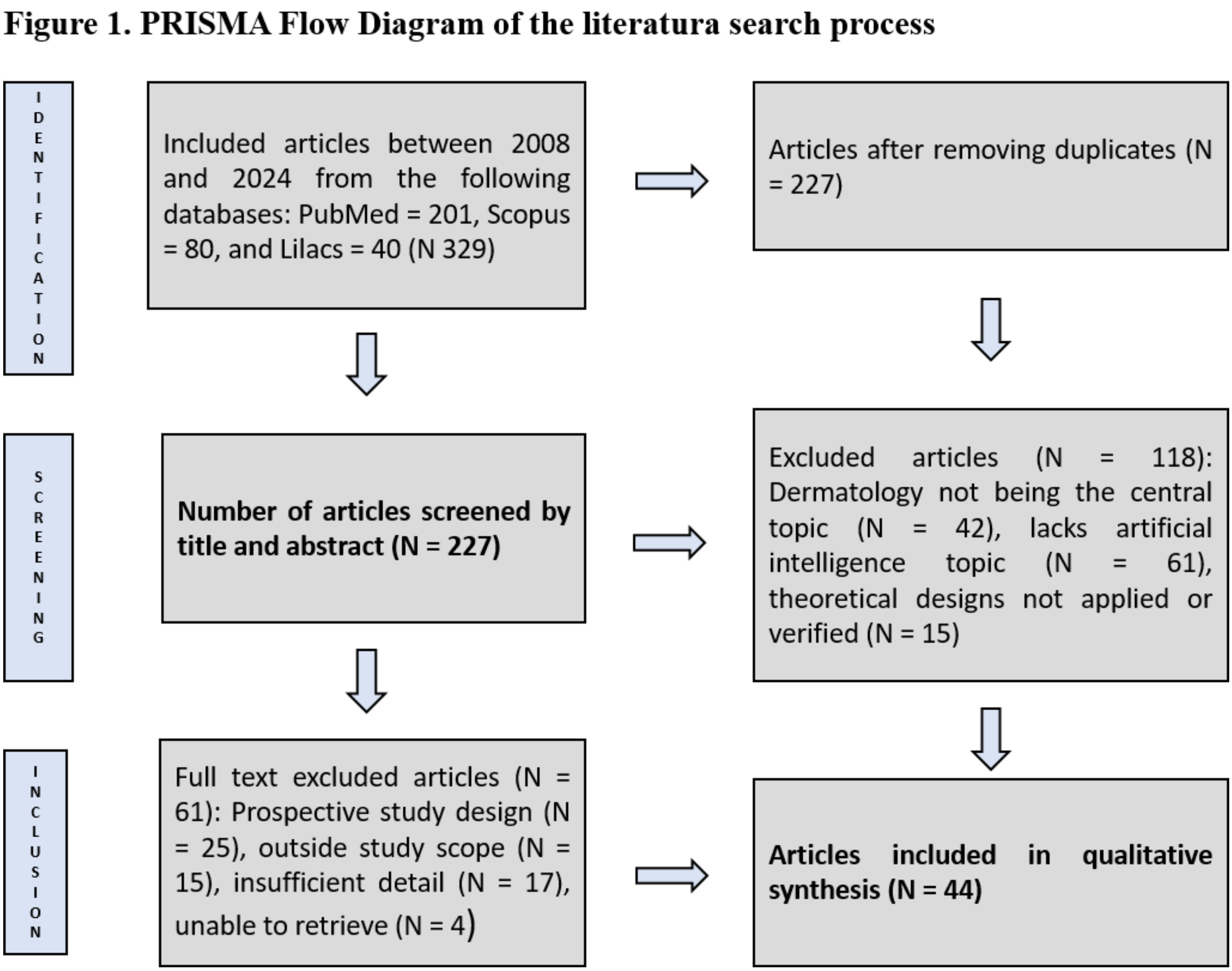
(Continued)

(Continued)

Article reference	Authors - Country of origin	Type of article	Objective	Main Findings
[53]	Goyal M., et al. USA 2020	Review	To review and discuss advancements in digital image-based AI solutions for the diagnosis of skin cancer.	While claims suggest AI systems may achieve higher accuracy than dermatologists, their clinical application for assisting in skin cancer diagnosis is still in its early stages.
[54]	S.S. Han, et al. Korea 2022	Single-center, paralleled, unmasked, randomized controlled trial	To validate whether AI could augment the accuracy of nonexpert physicians in real-world settings.	The accuracy of the AI-assisted group was significantly higher than the unaided group, particularly benefiting nondermatology trainees.
[55]	Rajpara S., et al. UK 2009	Systematic review and meta-analysis.	To compare the diagnostic accuracy of different dermoscopic algorithms with each other and with digital dermoscopy/artificial intelligence for the detection of melanoma.	Pooled sensitivity for artificial intelligence was slightly higher than for dermoscopy (91% vs. 88%; $p = 0.076$). Pooled specificity for dermoscopy was significantly better than artificial intelligence (86% vs. 79%; $P < 0.001$).

CNN: Convolutional Neural Networks. AUC: Area Under the Curve. AD: Atopic dermatitis. BSA: Body Surface Area. ROC: Receiver Operating Characteristic. PG: Pyoderma Gangrenosum. LU: Leg Ulcers. AI: Artificial intelligence. SALT: Severity of Alopecia Tool. MM: Malignant Melanoma. PIACS: Prurigo Image Analyzing and Comparing System. RCM: Reflectance Confocal Microscopy. NPST: Non-Pigmented Skin Tumors. DSS: Disease-Specific Survival. USA: United States of America. UK: United Kingdom. DOR: Diagnostic Odds Ratio. ISBI: International Society for Burn Injuries. CIA: Conventional Image Analyzer. aMSL: atypical Melanocytic Skin Lesions. GP: General Practitioner. CAIA: Computer-Assisted Image Analysis. PMs: Pigmented Melanoma. TILs: Tumor-Infiltrating Lymphocytes. ADTA: Automated Digital (TIL) Analysis. BCC: Basal Cell Carcinoma. CADx: Computer-Aided Diagnosis.
Prepared by the authors based on the results of the study.

Figure 1. PRISMA Flow Diagram of the literature search process.



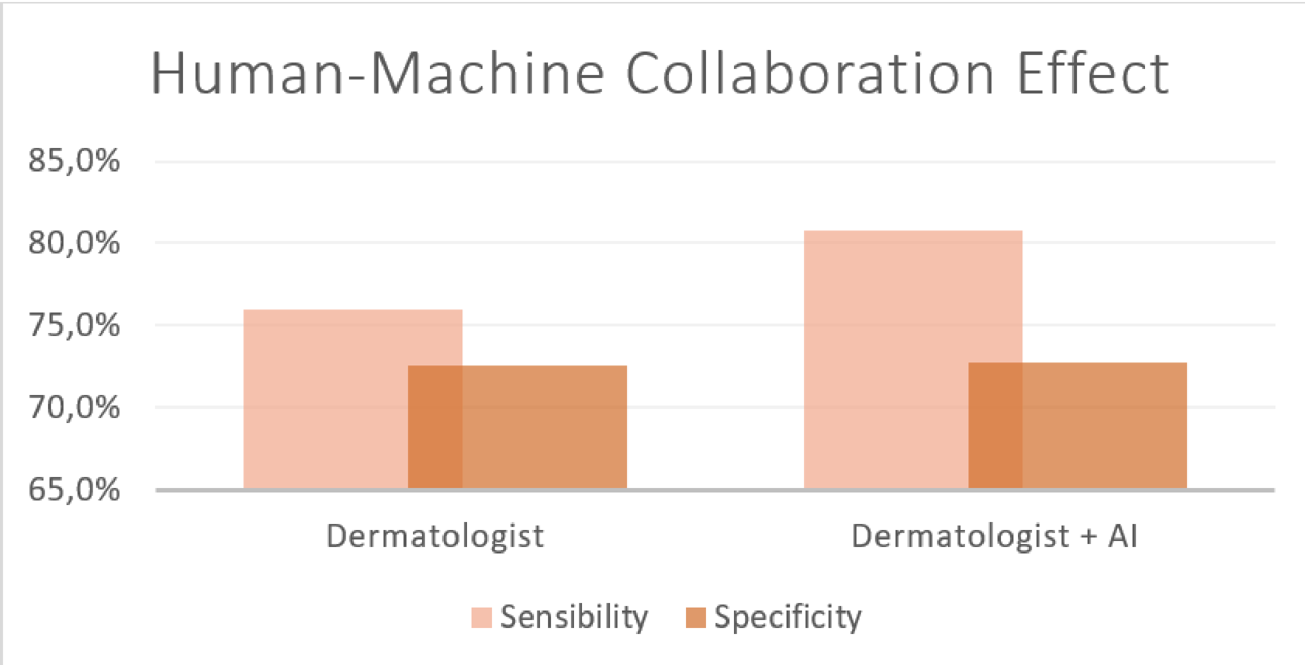
Source: Prepared by the authors of this study.

decision-making, supported by comprehensive data evaluation. However, there arises a need to carefully address the ethical implications and responsibility in automated decision-making in the medical field. The study by Hekler et al. states that during the clinical examination, the doctor has additional data to make the diagnosis that goes beyond the simple visual examination of the area of skin explored, such as the data obtained in interrogation [19]. Therefore, the study by Birkner et al. concludes that comparing artificial intelligence and visual diagnoses from human experts may be perceived as unequal, given that, in clinical practice, skin specialists tend to consider non-visual information when formulating a diagnosis [27].

Current applications and challenges in dermatopathology

As artificial intelligence advances, its application in dermatopathology has expanded beyond the initial focus on skin cancer, encompassing a range of dermatological pathologies. It is imperative to emphasize that AI aims to complement, not replace, the expertise and clinical judgment of dermatologists. Collaboration between AI and healthcare professionals is pivotal for the effective development of deep learning algorithms in dermatopathology. In 2017 study, the top-ranked computer algorithm demonstrated an area under the receiver operating characteristic curve of 0.87, surpassing dermatologists (0.74) and residents (0.66) significantly ($p < 0.001$ for all comparisons) but demonstrated that integrating computer algorithm classifications into dermatologist evaluations elevated dermatologist sensitivity and specificity from 76.0% to 80.8% and specificity from 72.6% to 72.8% as shown in Figure 2. The mounting evidence suggests that deep neural networks hold the potential to accurately classify skin images of melanoma and its benign mimickers, offering a promising avenue to enhance human performance [56].

Figure 2. Human-machine collaboration effect.



Notes: Integrating computer algorithm classifications into dermatologist evaluations increased dermatologists sensitivity from 76% to 80.8%, and specificity from 72.6% to 72.8%. [18].
Source: Prepared by the authors based on the results of the study.

Artificial intelligence´s performance in primary care

Using AI in primary care can accelerate the diagnosis process, contribute to earlier and more effective care, and reduce the care burden and improve the quality of dermatological care in primary care. In a study conducted by Jain et al. in 2021, which involved primary care physicians without dermatology experience and nurse practitioners, it was found that AI assistance was significantly compared to higher performance at diagnosis, in the case of physicians, the increase in diagnostic agreement was 10% (95% CI, 8% to 11%; p < 0.001), from 48% to 58%; for nurses, the increase was 12% (95% CI, 10% to 14%; p < 0.001), from 46% to 58%. In this way, the desire rates for biopsies and referrals to a dermatologist are expected to decrease with the use of AI [48].

Similarly, in the study conducted by Samaran et al. in 2021 in France, which surveyed 147 general practitioners, 98% said they had difficulty diagnosing non-pigmented skin tumors, and 86% agreed that an AI diagnostic tool could be useful in a general practice setting; however, 68% would not be willing to pay for this type of software [45]. Finally, a study conducted by Han et al. in 2022, which evaluated the use of AI in medical students and dermatology residents, showed that the accuracy of the AI-assisted group was significantly higher than that of the unassisted group, particularly benefiting non-dermatology students [54].

Areas for improvement

It is crucial to acknowledge the current limitations of AI in dermatology as identified within the reviewed articles. A considerable number of studies highlight a major limitation – the lack of research on real-world applications [14,17,26,31,32,46,47,53]. Additionally, there is a call for broader studies, as the evaluated algorithms show potential in small groups [3,49]. Similarly, several articles mention that the performance of studies was assessed on specific datasets, which could limit the generalizability of the findings. Therefore, there is a need to expand the characteristics of datasets and clinical scenarios [25].

Another noteworthy limitation is the variability in dermatologists' performance, which is based on their experience and familiarity with technology in assisting convolutional neural networks [24]. Furthermore, practical algorithm training requires high-quality photos for result interpretation [13]. Lastly, ethical concerns, particularly in clinical decision-making, must be taken into consideration [34]. Identifying these limitations guides future research towards areas focused on enhancing AI models.

DISCUSSION

The narrative review highlights the critical role that artificial intelligence (AI) plays in improving diagnostic capacities while illuminating the changing landscape of AI applications in dermatology. The results provide a strong body of data in favor of AI's ability to diagnose skin lesions as accurately as human dermatologists, if not surpassing them. Noteworthy meta-analyses and studies consistently highlight AI's effectiveness in various diagnostic tasks, marking a transformative milestone in dermatological care. The exploration of AI's applications extends beyond its initial focus on skin cancer, encompassing diverse dermatological pathologies. The study underscores the collaborative nature of AI integration, emphasizing its role as a complementary tool to enhance rather than replace the clinical judgment of dermatologists.

The narrative review highlights the critical role that artificial intelligence (AI) plays in enhancing diagnostic capabilities while shedding light on the evolving landscape of AI applications in dermatology. The results provide a substantial body of data in favor of AI's ability to diagnose skin lesions as accurately as human dermatologists, if not surpassing them. Noteworthy meta-analyses and studies consistently highlight the effectiveness of AI in various diagnostic tasks, marking a transformative milestone in dermatological care. The exploration of AI's applications extends beyond its initial focus on skin cancer, encompassing a diverse range of dermatological pathologies. The study highlights the collaborative nature of AI integration, emphasizing its role as a complementary tool to enhance, rather than replace, the clinical judgment of dermatologists.

The evaluation of AI's performance in dermatology demonstrates its effectiveness in lesion classification, with several studies showing accuracy comparable to, and in some cases superior to, that of human dermatologists. The narrative review delves into historical perspectives, highlighting the continuous evolution and validation of AI in dermatological diagnostics since 2009. The comparison of diagnostic accuracy between AI and dermoscopy, as performed by dermatologists, reveals a nuanced landscape, with AI showing slightly higher sensitivity but, in some cases, lower specificity. The discussion acknowledges the ethical implications of automated decision-making, emphasizing the importance of considering additional non-visual information during clinical examinations. The integration of multimodal data into AI models presents opportunities for more comprehensive decision-making, but ethical considerations and responsibilities must be carefully addressed.

Faced with the notable advances in the implementation of artificial intelligence in the medical field, various tools have been developed to ensure transparency and reproducibility in research that evaluates its application. A notable example is the APPRAISE-AI tool, which is responsible for evaluating methodological quality to support clinical decisions. This tool, composed of 24 items, analyzes in detail the methods, results, and transparency of the studies, providing researchers with the ability to generate higher-quality evidence to support their research in the field of artificial intelligence in medicine [57]. The use of this tool is recommended, given its usefulness in strengthening the methodological foundation and enhancing transparency in studies related to artificial intelligence in dermatology.

Furthermore, the narrative review addresses the current applications and challenges in dermatopathology, emphasizing the role of AI in complementing dermatologists' expertise. The collaboration between AI and healthcare professionals is highlighted as pivotal for the effective development of algorithms. Notable advancements in diagnostic accuracy, as evidenced by a 2017 study, underscore the potential of AI to classify skin images and enhance human performance accurately. The main topic encompasses the implications of AI in primary care, highlighting its potential to accelerate diagnosis, improve care quality, and alleviate the burden on healthcare providers. Insights from studies involving primary care physicians and medical students underscore the positive impact of AI assistance in diagnosis. Overall, the narrative review provides a comprehensive overview of AI's current standing in dermatology, acknowledging its transformative potential while recognizing the need for ethical considerations and ongoing collaboration between AI and healthcare professionals.

The World Health Organization's recent guidance on AI in healthcare, developed collaboratively by the Health Ethics and Governance unit and the Department of Digital Health and Innovation, provides key ethical principles for the use of AI in health. These principles include protecting human autonomy, promoting human well-being and safety, ensuring transparency, explainability, and intelligibility, fostering responsibility and accountability, ensuring inclusiveness and equity, and promoting AI that is responsive and sustainable. The guidance underscores the importance of avoiding harm, maintaining transparency in AI technologies, and holding stakeholders accountable for the responsible development and deployment of AI in healthcare. The document emphasizes the ethical use of AI to benefit public health and addresses challenges related to privacy, autonomy, and the need for valid informed consent. Additionally, it highlights the importance of inclusiveness and equity in AI applications, emphasizing the need to avoid biases that could disproportionately affect marginalized groups. As the use of AI in healthcare continues to evolve, these principles aim to guide governments, technology developers, companies, and other stakeholders in adopting ethical approaches for the responsible integration of AI in health systems worldwide [58].

The comprehensive responsibility for making a diagnosis falls on the health professional, establishing a fundamental bond that covers professional, ethical, and legal aspects with the patient. This approach differs significantly from the contribution of artificial intelligence (AI), which offers a valuable diagnostic approach but does not replace the significant role of the medical professional. In the dermatological context, the gold standard continues to be the biopsy, since, for example, although the lesions may be highly suggestive of carcinoma or melanoma, it is the biopsy that definitively determines the malignant or benign nature of the lesion.

The comprehensive responsibility for making a diagnosis rests with the health professional, establishing a fundamental bond that encompasses professional, ethical, and legal aspects with the patient. This approach differs significantly from the contribution of artificial intelligence (AI), which offers a valuable diagnostic approach but does not replace the significant role of the medical professional. In the dermatological context, the gold standard remains the biopsy, as it is the definitive determination of the malignant or benign nature of the lesion, even when lesions are highly suggestive of carcinoma or melanoma.

In scenarios where dermatological care is not readily available, AI emerges as a promising tool for providing early guidance to care teams. This AI capability is crucial for patient prioritization, thereby mitigating costs associated with complications and delayed care. Although artificial intelligence can facilitate an initial approach, we emphasize that confirmation and exhaustive evaluation through biopsy remain the undisputed standard in dermatology.

Limitations

The limitation of this narrative review lies in the variability and lack of standardization among the included studies. Despite the growing interest in the application of artificial intelligence in dermatology over the past few years, a notable dearth of standardized methodologies and consensus remains in the field. The studies encompassed a range of AI models; each was developed independently with distinct characteristics and methods. This lack of uniformity poses challenges in assessing the reproducibility and generalizability of findings across different studies. The absence of standardized rules or a consensus framework introduces potential bias, as the quality and performance of AI models can vary significantly. This limitation highlights the need for future research efforts to establish standardized guidelines and a consensus on the application of artificial intelligence in dermatology, thereby ensuring a more robust and reliable foundation for the field.

CONCLUSIONS

In conclusion, the narrative review underscores the pivotal role of artificial intelligence (AI) in dermatology, emphasizing its ability to enhance diagnostic capabilities for skin lesions. The findings robustly support the effectiveness of AI, demonstrating its precision in diagnosis, comparable to or surpassing that of human dermatologists across various tasks. As AI evolves, a notable transformation in dermatological care is evident, with applications extending beyond skin cancer to diverse dermatological pathologies. While emphasizing the essential collaboration between AI and healthcare professionals, there is a recognition of the need to carefully address ethical implications, particularly in automated decision-making.

In conclusion, this narrative review highlights the fundamental role of artificial intelligence (AI) in dermatology, emphasizing its potential to enhance the diagnosis of skin lesions. The results strongly support the effectiveness of AI, demonstrating its diagnostic accuracy, which is comparable to or superior to that of human dermatologists in various tasks. As AI evolves, a remarkable transformation in dermatological care is becoming apparent, with applications extending beyond skin cancer to encompass various pathologies. While the essential collaboration between AI and healthcare professionals is highlighted, the need to carefully address the ethical implications of automated decision-making is recognized.

Overall, this review provides an overview of the current state of AI in dermatology, highlighting its transformative potential and underscoring the importance of ethical and collaborative integration in the future development of these technologies in the medical field.

Contributor roles All authors meet the ICMJE authorship criteria. They contributed substantially to the study conception and design, data collection and interpretation, critically revised the manuscript, approved the final version, and agreed to be accountable for all aspects of the work.

Funding Research derived from the MED-341-2023 project, developed by the Family Medicine and Population Health Research Group at Universidad de La Sabana, Colombia.

Competing interests The authors declare that there are no conflicts of interest regarding the publication of this article.

Language of submission English.

Peer review and provenance Not commissioned. Externally peer reviewed by two reviewers, in a double-anonymous modality.

REFERENCES

1. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med. 2019;25: 24–29. https://www.nature.com/articles/s41591-018-0316-z https://doi.org/10.1038/s41591-018-0316-z
2. Naylor CD. On the Prospects for a (Deep) Learning Health Care System. JAMA. 2018;320: 1099. https://jamanetwork.com/journals/jama/fullarticle/2701667 https://doi.org/10.1001/jama.2018.11103
3. Fujisawa Y, Otomo Y, Ogata Y, Nakamura Y, Fujita R, Ishitsuka Y, et al. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. Br J Dermatol. 2019;180: 373–381. https://onlinelibrary.wiley.com/toc/13652133/180/2 https://doi.org/10.1111/bjd.16924
4. Martorell A, Martin-Gorgojo A, Ríos-Viñuela E, Rueda-Carnero JM, Alfageme F, Taberner R. Inteligencia artificial en dermatología: ¿amenaza u oportunidad? Actas Dermo-Sifiliográficas. 2022;113: 30–46. https://doi.org/10.1016/j.ad.2021.07.003 https://doi.org/10.1016/j.ad.2021.07.003
5. Wells A, Patel S, Lee JB, Motaparthi K. Artificial intelligence in dermatopathology: Diagnosis, education, and research. J Cutan Pathol. 2021;48: 1061–1068. https://onlinelibrary.wiley.com/doi/full/10.1111/cup.13954 https://doi.org/10.1111/cup.13954
6. Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial Intelligence in Dermatology: A Primer. J Invest Dermatol.

2020;140: 1504–1512. http://www.jidonline.org/article/S0022202X2031201X/fulltext https://doi.org/10.1016/j.jid.2020.02.026
7. Escalé-Besa A, Fuster-Casanovas A, Börve A, Yélamos O, Fustà-Novell X, Esquius Rafat M, et al. Using Artificial Intelligence as a Diagnostic Decision Support Tool in Skin Disease: Protocol for an Observational Prospective Cohort Study. JMIR Res Protoc. 2022;11. https://doi.org/10.2196/37531
8. In: Deep Learning [Internet]. 8 Jan 2024. https://www.deeplearningbook.org/
9. Luo N, Zhong X, Su L, Cheng Z, Ma W, Hao P. Artificial intelligence-assisted dermatology diagnosis: From unimodal to multimodal. Comput Biol Med. 2023;Avaliable from. https://doi.org/10.1016/j.compbiomed.2023.107413https://doi.org/10.1016/j.compbiomed.2023.107413
10. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8: 19–32. https://www.tandfonline.com/doi/abs/10.1080/1364557032000119616 https://doi.org/10.1080/1364557032000119616
11. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010;5: 69. https://pubmed.ncbi.nlm.nih.gov/20854677/ https://doi.org/10.1186/1748-5908-5-69
12. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169: 467–473. https://doi.org/10.7326/m18-0850 https://doi.org/10.7326/M18-0850
13. Han SS, Park I, Eun Chang S, Lim W, Kim MS, Park GH, et al. 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. 2020;140: 1753–1761. http://www.jidonline.org/article/S0022202X20301366/fulltext https://doi.org/10.1016/j.jid.2020.01.019
14. Foahom Gouabou AC, Collenne J, Monnier J, Iguernaissi R, Damoiseaux J-L, Moudafi A, et al. Computer Aided Diagnosis of Melanoma Using Deep Neural Networks and Game Theory: Application on Dermoscopic Images of Skin Lesions. Int J Mol Sci. 2022;23. https://www.mdpi.com/1422-0067/23/22/13838/htm https://doi.org/10.3390/ijms232213838
15. Guimarães P, Batista A, Zieger M, Kaatz M, Koenig K. Artificial Intelligence in Multiphoton Tomography: Atopic Dermatitis Diagnosis. Sci Rep. 2020;10. https://www.nature.com/articles/s41598-020-64937-x https://doi.org/10.1038/s41598-020-64937-x
16. Fink C, Blum A, Buhl T, Mitteldorf C, Hofmann-Wellenhof R, Deinlein T, et al. Diagnostic performance of a deep learning convolutional neural network in the differentiation of combined naevi and melanomas. Acad Dermatol Venereol. 2020;34: 1355–1361. https://onlinelibrary.wiley.com/toc/14683083/34/6 https://doi.org/10.1111/jdv.16165
17. Brinker TJ, Hekler A, Enk AH, von Kalle C. Enhanced classifier training to improve precision of a convolutional neural network to identify images of skin lesions. PLoS ONE. 14. https://pubmed.ncbi.nlm.nih.gov/31233565/ https://doi.org/10.1371/journal.pone.0218713
18. Yoo KH, Jeong GJ, Park JH, Park SH, Li K. Estimation error of the body surface area in psoriasis: a comparative study of physician and computer-assisted image analysis (ImageJ). Clin Experimental Derm. 2022;47: 1298–1306. https://onlinelibrary.wiley.com/toc/13652230/47/7 https://doi.org/10.1111/ced.15148
19. Hekler A, Utikal JS, Enk AH, Hauschild A, Weichenthal M, Maron RC, et al. Superior skin cancer classification by the combination of human and artificial intelligence. Eur J Cancer. 2019;120: 114–121. https://doi.org/10.1016/j.ejca.2019.07.019
20. Winkler JK, Sies K, Fink C, Toberer F, Enk A, Abassi MS, et al. Association between different scale bars in dermoscopic images and diagnostic performance of a market-approved deep learning convolutional neural network for melanoma recognition. Eur J Cancer. 2021;145: 146–154. https://pubmed.ncbi.nlm.nih.gov/33465706/ https://doi.org/10.1016/j.ejca.2020.12.010
21. Sies K, Winkler JK, Fink C, Bardehle F, Toberer F, Buhl T, et al. Past and present of computer-assisted dermoscopic diagnosis: performance of a conventional image analyser versus a convolutional neural network in a prospective data set of 1,981 skin lesions. Eur J Cancer. 2020;135: 39–46. http://www.ejcancer.com/article/S0959804920302483/fulltext https://doi.org/10.1016/j.ejca.2020.04.043
22. Yu C, Yang S, Kim W, Jung J, Chung K-Y, Lee SW, et al. Acral melanoma detection using a convolutional neural network for dermoscopy images. PLoS ONE. 2018;13. https://doi.org/10.1371/journal.pone.0193321
23. Yu C, Yang S, Kim W, Jung J, Chung KY, Lee SW, et al. Acral melanoma detection using a convolutional neural network for dermoscopy images. PLoS ONE. https://pubmed.ncbi.nlm.nih.gov/29513718/
24. Ba W, Wu H, Chen WW, Wang SH, Zhang ZY, Wei XJ, et al. Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images. Eur J Cancer. 2022;169: 156–165. https://pubmed.ncbi.nlm.nih.gov/35569282/ https://doi.org/10.1016/j.ejca.2022.04.015
25. Hekler A, Utikal JS, Enk AH, Solass W, Schmitt M, Klode J, et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. Eur J Cancer. 2019;118: 91–96. https://pubmed.ncbi.nlm.nih.gov/31325876/ https://doi.org/10.1016/j.ejca.2019.06.012
26. Zunair H, Ben Hamza A. Melanoma detection using adversarial training and deep transfer learning. Phys Med Biol. 65: 135005. https://pubmed.ncbi.nlm.nih.gov/32252036/ https://doi.org/10.1088/1361-6560/ab86d3
27. Birkner M, Schalk J, von den Driesch P, Schultz ES. Computer-Assisted Differential Diagnosis of Pyoderma Gangrenosum

and Venous Ulcers with Deep Neural Networks. J Clin Med. 2022;11. https://www.mdpi.com/2077-0383/11/23/7103/htm https://doi.org/10.3390/jcm11237103

28. Yanagisawa Y, Shido K, Kojima K, Yamasaki K. Convolutional neural network-based skin image segmentation model to improve classification of skin diseases in conventional and non-standardized picture images. J Dermatol Sci. 2023;109: 30–36. https://pubmed.ncbi.nlm.nih.gov/36658056/ https://doi.org/10.1016/j.jdermsci.2023.01.005
29. Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, et al. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer. 2019;111: 148–154. https://pubmed.ncbi.nlm.nih.gov/30852421/ https://doi.org/10.1016/j.ejca.2019.02.005
30. Felmingham C, MacNamara S, Cranwell W, Williams N, Wada M, Adler NR, et al. 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. 2022;12. https://doi.org/10.1136/bmjopen-2021-050203
31. Tognetti L, Bonechi S, Andreini P, Bianchini M, Scarselli F, Cevenini G, et al. A new deep learning approach integrated with clinical data for the dermoscopic differentiation of early melanomas from atypical nevi. J Dermatol Sci. 2021;101: 115–122. https://pubmed.ncbi.nlm.nih.gov/33358096/ https://doi.org/10.1016/j.jdermsci.2020.11.009
32. Burlina PM, Joshi NJ, Ng E, Billings SD, Rebman AW, Aucott JN. Automated detection of erythema migrans and other confounding skin lesions via deep learning. Comput Biol Med. 2019;105: 151–156. https://pubmed.ncbi.nlm.nih.gov/30654165/ https://doi.org/10.1016/j.compbiomed.2018.12.007
33. Filipescu SG, Butacu AI, Tiplica GS, Nastac DI. Deep-learning approach in the study of skin lesions. Skin Res Technol. 2021;27: 931–939. https://onlinelibrary.wiley.com/toc/16000846/27/5 https://doi.org/10.1111/srt.13045
34. Pai VV, Pai RB. Artificial intelligence in dermatology and healthcare: An overview. IJDVL. 87: 457–467. https://www.ijdvl.com/artificial-intelligence-in-dermatology-and-healthcare-an-overview/ https://doi.org/10.25259/IJDVL_518_19
35. Lee S, Lee JW, Choe SJ, Yang S, Koh SB, Ahn YS, et al. Clinically Applicable Deep Learning Framework for Measurement of the Extent of Hair Loss in Patients With Alopecia Areata. JAMA Dermatol. 2020;156: 1018. https://doi.org/10.1001/jamadermatol.2020.2188
36. Koo T, Kim MH, Jue MS. Automated detection of superficial fungal infections from microscopic images through a regional convolutional neural network. PLoS ONE. 2021;16: e0256290. https://pubmed.ncbi.nlm.nih.gov/34403443/ https://doi.org/10.1371/journal.pone.0256290
37. Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. Journal of Investigative Dermatology. 2018;138: 1529–1538. https://pubmed.ncbi.nlm.nih.gov/29428356/ https://doi.org/10.1016/j.jid.2018.01.028
38. Han SS, Park GH, Lim W, Kim MS, Na JI, Park I, et al. Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. PLoS ONE. 13: e0191493. https://doi.org/10.1371/journal.pone.0191493
39. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. Journal of Dermatological Treatment. 2020;31: 496–510. https://www.tandfonline.com/doi/abs/10.1080/09546634.2019.1682500 https://doi.org/10.1080/09546634.2019.1682500
40. Koller S, Wiltgen M, Ahlgrimm-Siess V, Weger W, Hofmann-Wellenhof R, Richtig E, et al. *In vivo* reflectance confocal microscopy: automated diagnostic image analysis of melanocytic skin tumours . Acad Dermatol Venereol. 2011;25: 554–558. https://onlinelibrary.wiley.com/toc/14683083/25/5 https://doi.org/10.1111/j.1468-3083.2010.03834.x
41. Gareau DS, Correa da Rosa J, Yagerman S, Carucci JA, Gulati N, Hueto F, et al. Digital imaging biomarkers feed machine learning for melanoma screening. Exp Dermatol. 2017;26: 615–618. https://onlinelibrary.wiley.com/toc/16000625/26/7 https://doi.org/10.1111/exd.13250
42. Bruland P, Hänse W, Schedel F, Ständer S, Fritz F. PIACS: A System for the Automatic Detection, Categorization and Comparison of Scratch-Related Skin Lesions in Dermatology. Stud Health Technol Inform. 2015;216. https://doi.org/10.3233/978-1-61499-564-7-1042
43. Gareau DS, Hennessey RJ, Wan E, Pellacani G, Jacques SL. Automated detection of malignant features in confocal microscopy on superficial spreading melanoma versus nevi. J Biomed Opt. 15: 1. https://pubmed.ncbi.nlm.nih.gov/21198161/ https://doi.org/10.1117/1.3524301
44. Lim K, Neal-Smith G, Mitchell C, Xerri J, Chuanromanee P. Perceptions of the use of artificial intelligence in the diagnosis of skin cancer: an outpatient survey. Clin Experimental Derm. 2022;47: 542–546. https://onlinelibrary.wiley.com/toc/13652230/47/3 https://doi.org/10.1111/ced.14969
45. Samaran R, L'Orphelin JM, Dreno B, Rat C, Dompmartin A. Interest in artificial intelligence for the diagnosis of non-melanoma skin cancer: a survey among French general practitioners. Eur J Dermatol. 2021;31: 457–462. https://www.jle.com/fr/revues/ejd/e-docs/interest_in_artificial_intelligence_for_the_diagnosis_of_non_melanoma_skin_cancer_a_survey_among_french_general_practitioners_321088/article.phtml?tab=texte https://doi.org/10.1684/ejd.2021.4090

46. Han SS, Moon IJ, Lim W, Suh IS, Lee SY, Na J-I, et al. Keratinocytic Skin Cancer Detection on the Face Using Region-Based Convolutional Neural Network. JAMA Dermatol. 2020;156: 29. https://pubmed.ncbi.nlm.nih.gov/31799995/ https://doi.org/10.1001/jamadermatol.2019.3807
47. Hurault G, Domínguez-Hüttinger E, Langan SM, Williams HC, Tanaka RJ. Personalized prediction of daily eczema severity scores using a mechanistic machine learning model. Clin Exp Allergy. 2020;50: 1258–1266. https://pubmed.ncbi.nlm.nih.gov/32750186/ https://doi.org/10.1111/cea.13717
48. Jain A, Way D, Gupta V, Gao Y, de Oliveira Marinho G, Hartford J, et al. Development and Assessment of an Artificial Intelligence-Based Tool for Skin Condition Diagnosis by Primary Care Physicians and Nurse Practitioners in Teledermatology Practices. JAMA Netw Open. 2021;4. https://doi.org/10.1001/jamanetworkopen.2021.7249 https://doi.org/10.1001/jamanetworkopen.2021.7249
49. Moore MR, Friesner ID, Rizk EM, Fullerton BT, Mondal M, Trager MH, et al. Automated digital TIL analysis (ADTA) adds prognostic value to standard assessment of depth and ulceration in primary melanoma. Sci Rep. 2021;11. https://pubmed.ncbi.nlm.nih.gov/33531581/ https://doi.org/10.1038/s41598-021-82305-1
50. Sohn GK, Sohn JH, Yeh J, Chen Y, Brian Jiang SI. A deep learning algorithm to detect the presence of basal cell carcinoma on Mohs micrographic surgery frozen sections. J Am Acad Dermatol. 2021;84: 1437–1438. https://pubmed.ncbi.nlm.nih.gov/32590033/ https://doi.org/10.1016/j.jaad.2020.06.080
51. Chang W-Y, Huang A, Yang C-Y, Lee C-H, Chen Y-C, Wu T-Y, et al. Computer-Aided Diagnosis of Skin Lesions Using Conventional Digital Photography: A Reliability and Feasibility Study. PLoS ONE. 8: e76212. https://doi.org/10.1371/journal.pone.0076212
52. Blum A, Bosch S, Haenssle HA, Fink C, Hofmann-Wellenhof R, Zalaudek I, et al. Künstliche Intelligenz und Smartphone-Programm-Applikationen (Apps). Hautarzt. 2020;71: 691–698. https://doi.org/10.1007/s00105-020-04658-4
53. Goyal M, Knackstedt T, Yan S, Hassanpour S. Comput Biol Med. http://creativecommons.org/licenses/by/4.0/
54. Han SS, Kim YJ, Moon IJ, Jung JM, Lee MY, Lee WJ, et al. Evaluation of Artificial Intelligence-Assisted Diagnosis of Skin Neoplasms: A Single-Center, Paralleled, Unmasked, Randomized Controlled Trial. J Invest Dermatol. 2022;142: 2353–2362. https://pubmed.ncbi.nlm.nih.gov/35183551/ https://doi.org/10.1016/j.jid.2022.02.003
55. Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. British Journal of Dermatology. 2009;161: 591–604. http://blackwell-synergy.com/doi/abs/10.1111/bjd.2009.161.issue-3 https://doi.org/10.1111/j.1365-2133.2009.09093.x
56. Marchetti MA, Liopyris K, Dusza SW, Codella NCF, Gutman DA, Helba B, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: Results of the International Skin Imaging Collaboration 2017. J Am Acad Dermatol. 2020;82: 622–627. http://www.jaad.org/article/S0190962219323734/fulltext https://doi.org/10.1016/j.jaad.2019.07.016
57. Kwong JCC, Khondker A, Lajkosz K, McDermott MBA, Frigola XB, McCradden MD, et al. APPRAISE-AI Tool for Quantitative Evaluation of AI Studies for Clinical Decision Support. JAMA Netw Open. 2023;6. https://doi.org/10.1001/jamanetworkopen.2023.35377 https://doi.org/10.1001/jamanetworkopen.2023.35377
58. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: WHO guidance. World Health Organization. 2021;Available from: 1–148. https://www.who.int/publications/i/item/9789240029200

Inteligencia artificial para la clasificación y el diagnóstico de lesiones cutáneas en dermatología: una revisión narrativa

RESUMEN

INTRODUCCIÓN La inteligencia artificial (IA) tiene una presencia cada vez mayor en dermatología, demostrando niveles de precisión comparables o incluso superiores a los de los dermatólogos en el diagnóstico de lesiones cutáneas a partir de imágenes clínicas y dermatoscópicas. Esta revisión ofrece una visión general del papel de la IA en la clasificación automatizada y el seguimiento de las lesiones cutáneas.
OBJETIVO Explorar y cartografiar la literatura existente sobre el uso y los beneficios de la IA en dermatología.
MÉTODOS Esta revisión narrativa se centró en el uso y las ventajas de la IA en dermatología, utilizando los términos MeSH/DeCS: "Dermatology", "Artificial Intelligence", "Diagnosis" y "Computer Aided Diagnosis". Se consultaron tres bases de datos (PubMed, LILACS y Scopus) para el periodo comprendido entre 2008 y 2024. Se excluyeron los estudios que no se enfocaban en dermatología, carecían de relación con inteligencia artificial o presentaban diseños teóricos sin aplicación práctica ni evidencia, lo que permitió obtener una selección final de cuarenta y cuatro artículos.
RESULTADOS Los hallazgos respaldan sólidamente la efectividad de la IA y evidencian una precisión diagnóstica comparable o superior a la del dermatólogo humano en diversas tareas. La evolución de la IA en dermatología indica una transformación sustancial en la atención, ampliando sus aplicaciones desde el cáncer de piel hasta diversas patologías dermatológicas. Si bien se destaca la importancia de la colaboración entre la IA y los profesionales de la salud, persiste una brecha crítica en la validación clínica de la IA en contextos reales. Además, las consideraciones éticas, especialmente en la toma de decisiones automatizadas, requieren una atención rigurosa.
CONCLUSIONES Esta revisión narrativa resalta el papel fundamental de la IA en dermatología, y subraya su capacidad para mejorar la precisión diagnóstica de las lesiones cutáneas.