Artificial Intelligence in Dermatology: A Review of Literature and Application to Pediatric Dermatology

Joshua Burshtein, MD¹, Maria Gnarra Buethe, MD, PhD²

¹ Department of Dermatology, Mount Sinai Icahn School of Medicine, New York, New York
² Department of Dermatology, University of California San Diego, San Diego, CA

ABSTRACT

Background: Artificial intelligence (AI) is increasingly investigated for use in dermatologic conditions. We review recent literature on AI, its potential application for pediatric dermatology, and its impact on the underserved community.

Objective: To evaluate the current state of AI in dermatology and its application to pediatric patients.

Methods: Literature search was performed in PubMed and Google Scholar using the following key terms in combination with "pediatric", and "dermatology": "artificial intelligence," "AI," "machine learning," "augmented intelligence," "neural network," and "deep learning".

Results: Current research is based on images from adult databases, with minimal delineation of patient age. Most literature on AI and dermatologic conditions pertains to melanoma and non-melanoma skin cancers, reporting accuracy from 67-99%. Other commonly studied diseases include psoriasis, acne vulgaris, onychomycosis, and atopic dermatitis, having varying accuracy, sensitivity, and specificity. A recently developed AI algorithm for diagnosis of infantile hemangioma found 91.7% accuracy. AI may be a means to increase access to pediatric dermatologic care, yet challenges remain for its use in underserved communities.

Conclusion: Literature on AI systems for dermatologic diseases continues to grow. Further research may tailor AI algorithms for pediatric patients and those of diverse skin color to decrease algorithm bias and increase diagnostic accuracy.

INTRODUCTION

Research into medical application of artificial intelligence (AI) has risen rapidly over the past several years. Its potential for use in dermatology has been increasingly explored, with research evaluating efficacy of various diagnostic algorithms. With a shortage of pediatric dermatologists and wait times of several months for appointments¹,², such technological advancements may be seen as a way to increase access to care.

Understanding AI technology is the first step to utilization in the pediatric dermatology setting. AI consists of several components - machine learning (ML) refers to algorithms and statistical models that can recognize patterns after learning from training data³. The form of AI that has become most prevalent in healthcare is deep learning (DL)⁴. DL utilizes multiple layers of neural networks which are series of interconnected nodes that receive input data and provide an output of information, that are built into layers with adjustable strength of connections to
A type of DL is convolutional neural networks (CNN), where information from a data set of images is transmitted through multi-layer nodes and passes through several layers that act as filters. As DL and CNN technologies improve, there is a rise in diagnostic accuracy. Recent evidence demonstrates that CNN algorithms have performed similar to healthcare professionals in diagnosing image-based diseases.

As a field of medicine with a visual component to its diagnosis, dermatology has seen a rapid rise in AI-related research in recent years. Research has largely focused on skin lesions such as melanoma and non-melanoma skin cancers, and these have shown more promising results compared to other conditions such as inflammatory dermatoses. In addition, most studies have centered around delineating between binary outcomes, such as malignant vs benign skin cancers.

AI in pediatric dermatology has not been well studied. Here, we aim to explore current pediatric dermatology literature on AI, including accuracy, sensitivity, and specificity of AI among commonly studied diseases, as well as its potential application to the pediatric dermatology clinic and its impact on the underserved community.

A review of the literature on artificial intelligence in adult and pediatric dermatology was completed. PubMed and Google Scholar were queried using a combination of terms including "pediatric", and "dermatology": "artificial intelligence," "AI," "machine learning," "augmented intelligence," "neural network," and "deep learning". The review covered literature published after 2012. Articles were individually evaluated for their relevance to artificial intelligence in adult and pediatric dermatology.

### Diagnostic Accuracy, Sensitivity, and Specificity

Most literature on AI and dermatologic conditions pertains to melanoma and non-melanoma skin cancers, though recent studies have explored diseases such as psoriasis, acne vulgaris, onychomycosis, and atopic dermatitis. Majority of analyses in the AI literature reflect the adult population, and it is important to consider its implications to the pediatric population. Diagnostic accuracy, sensitivity, and specificity are important to assess the current state of AI algorithms for dermatologic patients.

Studies on AI diagnosis for melanoma and non-melanoma skin cancer report accuracy rates from 67-99%, sensitivity rates from 77-96%, and specificity rates from 70-96%. Specifically for melanoma, a systematic review by Jones et al. found a mean accuracy of 89.5% (95% CI 88.2–90.8%) and mean sensitivity of 84.2% (95% CI 81.6–86.8%)\(^{13}\). This study also found a mean specificity of 89.1% (95% CI 87.1–91.0%), mean positive predictive value (PPV) of 81.4% (95% CI 76.9–85.9%), and mean negative predictive value (NPV) of 92.9% (95% CI 90.9–94.9%)\(^{13}\). Another systematic review found that accuracy of studies on melanoma was generally >90%\(^{14}\). For data pertaining to basal cell carcinoma (BCC), mean accuracy of AI systems was found to be 87.6% (95% CI 80.7–94.6%), while sensitivity and specificity were 83.7% (95% CI 79.2–88.3%) and 88.7% (95% CI 78.3–99.0%), respectively\(^{13}\). Notably, two studies did find 100% accuracy, sensitivity, and specificity for BCC\(^{15, 16}\), whereas two other studies found
accuracy and sensitivity were quite low at 72% and 38%, respectively\textsuperscript{17}. With regard to squamous cell carcinoma (SCC), mean accuracy and sensitivity were 85.3% (95% CI 77.3–93.3%) and 60.3 (95% CI 39.6–81.0%), respectively\textsuperscript{13}. Literature demonstrates high accuracy for melanoma and BCC detection, with slightly lower accuracy for SCC\textsuperscript{13-16}, which indicates promise for future integration into clinical settings.

There has been increasing research for AI diagnosis and lesion differentiation for psoriasis. Shrivastava et al. completed several studies demonstrating accuracy of psoriasis diagnosis around 99\%\textsuperscript{8, 18, 19}. Lu et al. completed a three-way classification of scaling lesions and reported sensitivity of 72.3-81% and specificity of 87-91\%\textsuperscript{20}. Huang et al. completed a 6-way classification differentiating psoriasis, lichen planus, pityriasis, dermatitis, seborrheic dermatitis, pityriasis rubra pilaris, and found an accuracy of 95.38\%\textsuperscript{21}. Another 6-way classification showed an even higher accuracy of 100\%\textsuperscript{22}. Additionally, studies have looked at segmenting psoriatic lesions to identify borders, and one found an accuracy of 94.8\%, sensitivity of 89.6\%, and specificity of 97.6\%\textsuperscript{23}.

Several other diseases have been recently studied for their potential use with AI\textsuperscript{14}. Min et al. found sensitivity for acne vulgaris to be 66.7-77.3\% for blackheads, whiteheads, papules, nodules, and pustules, with specificity of 21.1-77.6\%\textsuperscript{24}. Khan et al. found a higher sensitivity of 89.7\% and specificity of 93.2\% for acne vulgaris, while accuracy was 92.6\%\textsuperscript{25}. Two studies on AI’s use for onychomycosis found that a six-way classification had a sensitivity of 82.7-96\% and specificity of 69.3-96.7\%\textsuperscript{10}. Lastly, for plantar and common warts, Khozeimeh et al. found that AI had an accuracy of 80\% for selection of cryotherapy treatment and an accuracy of 83.33\% for selection of immunotherapy\textsuperscript{26}.

Accurate and efficient integration to the clinical setting continues to pose a challenge\textsuperscript{5}. A recent AI system approved for use in the European market performed comparable to dermatologists in a setting similar to store-and-forward dermatology\textsuperscript{27}. Research into AI analysis of dermoscopic images and dermatopathology has also yielded promising results\textsuperscript{28}. Several studies on dermoscopic images for melanoma and non-melanoma skin cancer found CNN algorithms performed on par with dermatologists\textsuperscript{28}. High accuracy was also demonstrated in several studies using ML for histologic images\textsuperscript{28}. A recent study was the first to explore AI in relation to pediatric dermatology, specifically studying a proof-of-concept AI algorithm for diagnosing infantile hemangiomas (IH)\textsuperscript{29}.

Majority of research on AI and dermatologic conditions is not specific for pediatric patients, and therefore future studies may seek to delineate their analyses based on age group. Published studies have also used a wide variation of methods and varying data inputs to train ML algorithms, and therefore, results should be carefully evaluated.

**Artificial Intelligence for Common Pediatric Dermatology Diseases**

There is significant potential for the application of AI to pediatric dermatology. A recent study by Zhang et al. trained a CNN to diagnose IH based on clinical images\textsuperscript{29}. IH typically has rapid growth between 1 and 3 months of age, and early diagnosis is essential to prevent complications\textsuperscript{30}. This study’s algorithm had a 91.7\% diagnostic accuracy rate for IH and reported even greater accuracy when limiting to the facial area\textsuperscript{29}. Having built the model from images of
patients with wide variety of perspectives, rather than a standardized image, this study was the first to demonstrate utility of AI in the pediatric dermatology population.\(^{29}\)

Image capture is a vital component of the AI algorithm. Currently, there is no standardized approach to taking a lesion image. Photographic details such as distance from lesion, size of boarder around the lesion, angle of camera to the lesion, lighting, and type of camera lens all have an impact on the ability of ML performance. In their study, Zhang et al. reported that AI algorithms can be utilized for non-standardized images, demonstrating its applicability to the real-world clinical setting.\(^{29}\) For IH diagnosis, future work requires algorithms that can delineate between multiple diseases rather than a binary classifier, as well as the ability to classify risk of IH.

Despite limited literature on AI’s use for pediatric dermatology conditions, studies have evaluated adult conditions that are commonly seen in pediatric patients. One such condition is psoriasis. Research demonstrates CNNs have been able to identify psoriatic vs non-psoriatic conditions and have lower misdiagnosis rates compared to dermatologists.\(^{6, 31}\) In addition to diagnosis, successful lesion segmentation has allowed ML to accurately evaluate disease severity, including extent of erythema and scaliness.\(^{31}\) This information enables risk stratification of psoriatic lesions.\(^{8, 18, 19}\)

Atopic dermatitis is a recurrent condition that typically has its onset during childhood. In a recent study, Guimarães et al. developed a CNN that analyzed multiphoton tomography data for atopic dermatitis.\(^{33}\) This algorithm had a diagnostic accuracy rate of 97%.\(^{33}\) On the other hand, another study developed an ML algorithm for an allergen-IgE screening assay, with a study population specifically including children. This model did not show significant discriminatory ability for atopic dermatitis in children.\(^{34}\) De Guzman et al. developed a multi-model, multi-level system for diagnosis of atopic dermatitis, resulting in a higher average confidence level compared to a single-model system (68.37% vs. 63.01%, respectively).\(^{35}\)

Current research is based on images from adult databases, where there is minimal delineation between patient age. This can lead to biases in algorithms that don’t explicitly indicate ages they are used for. Additional research should be completed to protect vulnerable populations such as children. Future algorithms may utilize databases that incorporate pediatric images and/or categorize images based on age.

With promising research results, AI integration into clinical dermatology practice must be anticipated. Use of ML in real-world situations is more complicated than current studies have depicted, with a challenge of how to effectively incorporate AI. As it stands, ML application to future practice remains unclear. Potential use by primary care providers or as a supplement to dermatologist diagnosis are very likely options. Implementation of diagnostic AI software may be especially advantageous for pediatricians without dermatology training, helping decrease logistical challenges and increase early diagnosis. The benefit of AI as a diagnostic aid may be paramount, with a potential to decrease wait times and increase accessibility to dermatologic care.

**Artificial Intelligence and Impact on Underserved Communities**

Technological advancements may permit wider access to care for pediatric populations that are disproportionately affected by healthcare disparities. Pediatric dermatology

(c) 2024 THE AUTHORS. Published by the National Society for Cutaneous Medicine.
patients face particular barriers to care, including financial constraints, parental ability to take time from work, childcare support, cultural beliefs, and education\textsuperscript{36}. Telehealth services have been a step toward increasing access to care for this community, as demonstrated by decreased no-show rates for pediatric dermatology appointments\textsuperscript{37, 38}. AI and ML systems may have benefits of screening and diagnostics, which can be translated into expanding dermatologic care to more patients.

Yet, there remain challenges for application of AI technologies to different races and skin colors. Kim et al. found that only 17.3\% of research studies on DL for dermatology provide data on race and Fitzpatrick skin type\textsuperscript{39}. More specifically, only 2.1\% of all images were Fitzpatrick skin type V and VI\textsuperscript{39}. The low representation of skin of color in data sets diminishes the ability of AI technologies to be used for these populations. The difference in diagnostic accuracy was reflected in a study that showed increased error rates for identifying black individuals by commercial systems that were trained on white individuals\textsuperscript{40}, and a similar finding may potentially occur for AI in dermatologic diagnosis.

With the introduction of new technology, there is a possibility that AI exacerbates already existing healthcare disparities. For instance, literature indicates differing use of technology based on race and socioeconomic status, with evidence that minority and lower income groups are more limited in their smartphone use\textsuperscript{41}. In addition, those that have prior positive experience with AI are more likely to embrace its use in less familiar contexts, such as healthcare\textsuperscript{42}. As such, underserved communities may be at risk of having lower uptake of AI in the clinical setting. Initiatives may be implemented, similar to telehealth services, that seek to diminish barriers to AI integration. These may include patient education about AI systems and applications, community outreach initiatives, and reducing complicated registration requirements for future AI services\textsuperscript{37, 38}.

With significant potential for clinical use of AI in dermatology, it is vital for ML models to be trained on images of all skin types. This can ensure adequate diagnostic accuracy for all populations who may need dermatologic care. Its future implementation may pose substantial benefits to underserved populations, yet may come with risks of isolating these communities further if not applied properly. With an increasing need for adult and pediatric dermatologic care, there may be more willingness for uptake of AI and other technological advancements into the community.

**Challenges and Limitations of Artificial Intelligence Uptake**

There are several limitations with ML and challenges for application to the clinical setting. Generalizability is a main concern, since many ML algorithms have been trained with similar datasets of images\textsuperscript{3}. When one CNN\textsuperscript{43} was tested on a new dataset, the performance dropped significantly\textsuperscript{44}. With ML dependent upon the quality and breadth of data it receives, it is vital for future research to account for input quality. Another limitation is the standardization of images captured for dermatologic conditions\textsuperscript{31}. Images may be taken by family members, friends, physicians, nurses, or others in the community, and variability in image quality is a significant contributing factor to accuracy of ML algorithms. Images angles, zoom, sharpness, lighting exposure, color balance, and other quality characteristics may pose challenges to differentiate true lesion texture with artificial components of an image\textsuperscript{31, 45}.
Pertaining to image quality, another limitation is the analysis of a three-dimensional structure in a two-dimensional image. Lesions exist in all areas of the skin, and areas of curvature prevent even light exposure and can pose challenges for CNN algorithms.

As described previously, AI algorithms have shown to have high accuracy and sensitivity. However, the tradeoff for high sensitivity is the greater possibility of classifying benign lesions as malignant (ie, false positives). Though this may be seen as safer, it may pose strain on the medical system as more patients would seek in-person appointments, biopsies, and follow-up. An additional challenge for ML is the ability to consider clinical context and patient history. Han et al. found that when dermatologists diagnosed patients in the clinical setting, they significantly outperformed CNN. There is significant potential for AI application for dermatologic care and integration into the real-world; yet limitations and challenges need to be addressed for its uptake.

The field of dermatology has seen a rapid rise in AI-related research, however minimal literature exists for pediatric dermatology. Current research has largely focused on cancerous lesions, although other common conditions have been analyzed as well. The majority of AI databases utilize adult images, and algorithms are geared toward diagnosis of adult conditions. Most studies demonstrate positive outcomes for accuracy, sensitivity, and specificity of AI programs; however, algorithms may be biased if used for incorrect age groups. Minimal data exists for pediatric patients, and one primary study on IH shows promise. Further research may tailor AI algorithms for vulnerable populations such as children. AI demonstrates a possibility to increase access to the underserved communities if effectively employed. Implementation of AI into the clinical setting remains a challenge to be overcome.

Conflict of Interest Disclosures: None

Funding: None

Corresponding Author: Joshua Burshtein, MD
3 Tanbark Trail, Saddle River, NJ 07458
845.596.9482
E-mail: jburshtein13@gmail.com

References:

CONCLUSION