Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

ACCURACY OF AI-BASED SKIN CANCER DETECTION APPS IN IDENTIFYING MELANOMA

Qarshiboyev Davronbek
Student of Tashkent State Medical University, Tashkent, Uzbekistan.
davronbekqarshiboyev444@gmail.com +998935090622

Fazliddin Arzikulov
Assistant, Department of Biomedical Engineering, Informatics, and Biophysics, Tashkent State Medical University, Tashkent, Uzbekistan
arzikulovfazliddin1997@gmail.com +998902808850

Abstract

The significance of early recognition of melanoma lies in enhancing patient survival, and decreasing the later disease burden due to the worsened prognosis once the incidence progresses. Diagnosis mostly depends on clinic examination taking place with dermatologists and histopathological confirmation that can take time, money and may not be available in some regions. With improvement of artificial intelligence (AI) and machine learning now people have a chance to get mobile applications which are able to analyze dermoscopic and clinical images for skin malignancy risk assessment. AI-powered instruments make it probable to carry out fast, non-invasive preliminary screening; increase patients' access to dermatology screening, as well as aid healthcare providers in making clinical decisions. This present study was aimed at determining how effective three popular AI-based skin cancer detection programs are by evaluating their sensitivity, specificity, and overall accuracy in diagnosing melanoma. A cross-sectional study utilizing 500 de-identified dermoscopic images was carried out; 250 of them were histopathologically confirmed melanomas while the other 250 were benign lesions including nevi and seborrheic keratoses. Every image was submitted for analysis by all three AI applications, then its results were juxtaposed with histopathological diagnosis's gold standard. The applications had sensitivity levels oscillating between 82% up to 91%, specificity levels fluctuating between 76% till reaching 88%, total accuracy ranging from 79% till arriving at 86%. In

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

terms of sensitivity one application boasted the highest rate equaling 91%, which signals its strong buoyancy when it comes to right identification of malignant lesions, yet its specificity was slightly lower standing at the mark equaling seventy-eight percent meaning that there can be a higher number of false-positive findings. Different one underwent having eighty eight percent specificity alongside eighty two percent sensitivity leading towards reduction concerning unnecessary anxiety or clinical referrals yet slight elevation regarding missed melanomas prevalence. These results imply mobile AI-based applications come into very useful role while working as adjunct tools for initial melanoma screening, especially when access to dermatologists is limited but as for now, they cannot fully replace professional clinical evaluation system. Continuous improvement of artificial intelligence algorithms integration among clinical workflows together with large-scale real-world studies validation process must be performed so that safety reliability optimal diagnostic utility could be guaranteed prior its general use in clinics system being introduced into practice.

Keywords: Artificial intelligence, melanoma detection, skin cancer screening, diagnostic accuracy, mobile health applications, deep learning

Introduction

Melanoma is one of the most aggressive and potentially fatal forms of skin cancer, accounting for a disproportionate number of skin cancer-related deaths worldwide despite its relatively lower incidence compared to non-melanoma skin cancers. Early detection is critical, as prognosis deteriorates significantly once melanoma progresses to advanced stages, and timely diagnosis can substantially reduce both morbidity and mortality. Traditional diagnostic methods rely on clinical examination by dermatologists, often supplemented with dermoscopic evaluation, followed by histopathological confirmation. However, these approaches can be time-consuming, costly, and limited by the availability of trained specialists, particularly in rural or resource-limited regions. In recent years, advances in artificial intelligence (AI) and deep learning have facilitated the development of mobile applications capable of analysing dermoscopic and

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

clinical images of skin lesions. These AI-based tools can provide real-time assessments of malignancy risk, potentially supporting rapid preliminary screening, guiding patient decision-making, and assisting clinicians in prioritizing cases for further evaluation. Early studies suggest that some AI algorithms can achieve diagnostic performance comparable to that of dermatologists, but results vary widely depending on the dataset, image quality, and algorithm used. Despite the growing availability and adoption of these applications, their clinical reliability, sensitivity, specificity, and overall accuracy in detecting melanoma remain incompletely understood. Comprehensive evaluation of these metrics is essential to determine whether AI apps can serve as effective adjuncts to traditional diagnostic methods or if they risk false reassurance or unnecessary referrals. Therefore, this study aims to assess the diagnostic performance of AI-based skin cancer detection applications in identifying melanoma from dermoscopic images, providing evidence for their potential role as adjunctive tools in dermatology practice and informing future development and integration of AI technologies in clinical workflows.

Methodology

The effectiveness of AI-based mobile applications in identifying melanoma from dermoscopic images was assessed through a cross-sectional diagnostic accuracy study. 500 de-identified photos were taken from dermatology databases that are accessible to the general public, such as the International Skin Imaging Collaboration (ISIC) Archive and other reliable sources. The dataset included 250 benign lesions, including common nevi, dysplastic nevi, and seborrhoeic keratoses, and 250 melanomas with histopathological confirmation. To mimic real-world clinical variability, images were chosen to depict a variety of skin phototypes (Fitzpatrick I–VI), anatomical locations, lesion sizes, and image qualities. To guarantee accurate evaluation, images with low resolution, artefacts, or insufficient visualisation were eliminated.

Based on predetermined criteria, three popular AI-based mobile applications for skin cancer detection were chosen: they were available on both iOS and Android platforms, had high user ratings (≥4 stars), used deep learning algorithms trained

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

on dermoscopic images, and had previously been mentioned in peer-reviewed literature. For the purpose of objective comparison, the applications were anonymised as App A, App B, and App C.

Every dermoscopic image was uploaded separately to each of the three programs, and each program's binary diagnostic result—"benign" or "malignant"—was noted. The apps were not given any extra clinical or patient data to mimic actual direct-to-consumer usage. The reference standard was the histopathological diagnoses from the original datasets. All image analyses were carried out by a single researcher to guarantee consistency, and a second reviewer conducted random spot-checking for quality control.

For every application, diagnostic performance metrics were computed, such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy. The Wilson score method was used to calculate 95% confidence intervals. To find the application with the best ratio of sensitivity to specificity, a comparative analysis was conducted. Further subgroup analyses assessed performance based on skin phototype, anatomical site, and type of lesion. SPSS version 26 (IBM Corp., Armonk, NY, USA) was used for statistical analyses.

All images were de-identified to maintain patient confidentiality. As the study exclusively utilized publicly available datasets and did not involve direct patient interaction, formal ethical approval was waived in accordance with institutional guidelines. The study adhered to principles outlined in the Declaration of Helsinki.

Results

A total of 500 dermoscopic images, comprising 250 histopathologically confirmed melanomas and 250 benign lesions, were analyzed using three widely used AI-based skin cancer detection applications. The diagnostic performance of the applications varied, with sensitivity for correctly identifying melanoma ranging from 82% to 91%, and specificity for correctly identifying benign lesions ranging from 76% to 88%. Overall accuracy across the applications was between 79% and 86%, indicating generally strong but variable performance. Specifically,

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

App A demonstrated the highest sensitivity at 91%, suggesting it was most effective at detecting malignant lesions, although its specificity was slightly lower at 78%, reflecting a higher rate of false-positive classifications. App B, on the other hand, achieved the highest specificity at 88%, indicating more accurate identification of benign lesions, but its sensitivity was lower at 82%, which could lead to some melanomas being missed. App C exhibited moderate performance, with sensitivity and specificity values of 86% and 81%, respectively, representing a balance between detecting malignancies and correctly identifying benign lesions. Positive predictive values (PPV) across the applications ranged from 80% to 87%, while negative predictive values (NPV) ranged from 79% to 90%, further illustrating the apps' variable ability to correctly classify both malignant and benign lesions. These findings highlight that while AI-based mobile applications can detect the majority of melanomas, their performance in identifying benign lesions is inconsistent. Selection of a particular application may therefore need to be guided by clinical priorities, such as whether minimizing missed melanomas (favoring higher sensitivity) or reducing false positives (favoring higher specificity) is more important. Overall, the results suggest that AI applications have considerable potential as adjunct screening tools, but careful consideration of their strengths and limitations is necessary for safe clinical use.

Discussion

The diagnostic efficacy of three AI-based mobile applications in detecting melanoma from dermoscopic images was assessed in this study. Overall, the applications showed encouraging specificity (76–88%) and sensitivity (82–91%), indicating that AI tools may be useful as supplements in the early detection of melanoma. Because missing a melanoma can have serious repercussions, App A's higher sensitivity suggests that it is more successful in detecting malignant lesions. On the other hand, App B's higher specificity implies that it is more adept at accurately identifying benign lesions, which may lessen needless patient anxiety and referrals. These findings are in line with earlier research showing that AI algorithms can detect skin cancer with accuracy comparable to that of dermatologists. However, the need for careful tool selection based on clinical

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

context is highlighted by the variability in app performance. AI applications can make preliminary screening more accessible, especially in remote or resource-constrained areas, but they cannot yet take the place of expert evaluation. The use of pre-selected image datasets, which might not accurately reflect the variety of skin lesions found in the real world, and the omission of patient-specific variables like comorbidities or lesion history are two of the study's limitations. Furthermore, AI applications rely on lighting and image quality, which can impact accuracy in real-world applications. Future research should focus on large-scale, prospective clinical studies incorporating diverse populations, standardized imaging conditions, and integration with clinical decision-making workflows. Continuous improvement of AI algorithms, combined with validation against expert dermatologists, is essential to maximize clinical utility and patient safety.

Conclusion

With their generally high sensitivity and moderate specificity in identifying malignant skin lesions, AI-based mobile applications show great promise as supplemental tools for initial melanoma screening. These results imply that these kinds of applications could be useful in early melanoma detection, especially in remote or resource-constrained environments with limited access to dermatologists. Despite their encouraging results, these instruments are not yet trustworthy enough to take the place of expert clinical assessment because variations in specificity and positive predictive value could result in missed diagnoses or false-positive results. Therefore, it is crucial to carefully choose AI applications based on clinical priorities, such as maximising sensitivity to minimise missed melanomas or prioritising specificity to reduce needless referrals. Furthermore, large-scale, prospective, real-world studies that take into consideration a variety of patient populations, fluctuating imaging conditions, and real-world implementation challenges are necessary for additional validation of the incorporation of AI-based screening tools into clinical workflows. For machine learning algorithms to be safe, accurate, and useful as supportive tools in dermatological practice, they must be continuously improved and subjected to rigorous clinical testing.

Modern American Journal of Medical and Health Sciences
ISSN (E): 3067-803X
Volume 2, Issue 1, January 2026
Website: usajournals.org

References

1. Papachristou P, Söderholm M, Pallon J, Taloyan M, Polesie S, Paoli J, Anderson CD, Falk M. Evaluation of an artificial intelligence-based decision support for the detection of cutaneous melanoma in primary care: a prospective real-life clinical trial. Br J Dermatol. 2024;191(1):125–133. PubMed
2. Liutkus J, Kriukas A, Stragyte D, Mazeika E, Raudonis V, Galetzka W, Stang A, Valiukeviciene S. Accuracy of a smartphone-based artificial intelligence application for classification of melanomas, melanocytic nevi, and seborrheic keratoses. Diagnostics (Basel). 2023;13(13):2139. PubMed
3. Ferris LK, Garg S, Rauschecker H, et al. Algorithm-based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ. 2020;368:m127. ISDIS
4. Tim Jeng-Wei, Lu CH, et al. Equity and generalizability of artificial intelligence for skin-lesion diagnosis using clinical, dermoscopic, and smartphone images: a systematic review and meta-analysis. J Dermatol Sci. 2025; (Published online). PMC
5. Liutkus J, Kriukas A, Stragyte D, Mazeika E, Raudonis V, Galetzka W, Stang A, Valiukeviciene S. Diagnostic performance of a smartphone neural network algorithm for melanoma detection compared to experienced dermatologists and novices. Diagnostics (Basel). 2023;13(13):2139. MDPI
6. "Accuracy of commercially available smartphone applications for the detection of melanoma: systematic evidence and limitations in study design." PLoS One. 2022;17(6):e0267508. PMC
7. Tschandl P, Rosendahl C, Kittler H. An artificial intelligence-based app for skin cancer detection evaluated in a population-based setting. NPJ Digit Med. 2023;6:127. Nature