Published in final edited form as:
Curr Oncol Rep. 2023 June ; 25(6): 635–645. doi:10.1007/s11912-023-01407-3.

Recent Advances in Melanoma Diagnosis and Prognosis Using Machine Learning Methods

Sarah Grossarth[1], **Dominique Mosley**[2], **Christopher Madden**[3,4], **Jacqueline Ike**[3,5], **Isabelle Smith**[3,6], **Yuankai Huo**[7], **Lee Wheless**[3,8,9]

[1]Quillen College of Medicine, East Tennessee State University, Johnson City, TN, USA

[2]Vanderbilt University School of Medicine, Nashville, TN, USA

[3]Department of Dermatology, Vanderbilt University Medicine Center, Nashville, TN, USA

[4]State University of New York Downstate College of Medicine, Brooklyn, NY, USA

[5]Meharry Medical College, Nashville, TN, USA

[6]Vanderbilt University, Nashville, TN, USA

[7]Department of Computer Science and Electrical Engineering, Vanderbilt University, Nashville, TN 37235, USA

[8]Department of Medicine, Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN, USA

[9]Tennessee Valley Healthcare System VA Medical Center, Nashville, TN, USA

Abstract

Purpose of Review—The purpose was to summarize the current role and state of artificial intelligence and machine learning in the diagnosis and management of melanoma.

Recent Findings—Deep learning algorithms can identify melanoma from clinical, dermoscopic, and whole slide pathology images with increasing accuracy. Efforts to provide more granular annotation to datasets and to identify new predictors are ongoing.

Summary—There have been many incremental advances in both melanoma diagnostics and prognostic tools using artificial intelligence and machine learning. Higher quality input data will further improve these models' capabilities.

✉Lee Wheless, lee.e.wheless@vumc.org.

Human and Animal Rights and Informed Consent The work included does not contain any new experiments conducted by the authors on human or animal subjects.

Conflicts of Interest The authors declare no conflicts of interest.

Disclaimer These funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The contents are solely the responsibility of the authors and do not necessarily represent official views of the Department of Veteran Affairs or the United States Government.

Keywords

Artificial intelligence; Melanoma; Machine Learning; Melanoma; Deep learning; Digital pathology; Dermoscopy

Introduction

Sparked by Estreva et al. declaring in 2017 that deep neural networks could achieve dermatologist-level classification of skin lesions, there has been a wave of new research into using deep learning and other artificial intelligence methods for melanoma diagnosis and prognosis [1]. Unlike most other cancers, skin cancers including melanoma are diagnosed primarily visually, which lend themselves to multiple imaging modalities, including "naked-eye" clinical images, dermoscopy, reflectance confocal microscopy (RCM), and optical coherence tomography (OCT) (see Table 1) [2, 3]. These are in addition to the widening use of pathology whole slide images (WSI) in clinical medicine [4]. In this review, we will discuss recent applications of deep learning (DL) and other machine learning (ML) methods to these imaging modalities, as well as future directions and barriers to deployment. Broadly, ML refers to the use of algorithms on structured data to identify patterns to guide classification or prediction without being explicitly programmed to produce the output. That is, when encountering new data, ML algorithms can predict an outcome based on prior training. Familiar examples include linear or logistic regression models. DL is a subset of ML that specifically uses neural networks to perform this task. Several excellent reviews have been written on this topic, including one describing the basics of the computational methodologies [5, 6•, 7]. The present review will focus primarily on studies that have been reported in the two years since these others. Moreover, while many earlier models were developed with the intent of making a binary classification, e.g., melanoma vs. benign nevus, most now have adopted a multiclass learner to identify numerous lesion types. As a result, the development of melanoma-specific models is now uncommon, and most of the studies included evaluate the respective model's ability to identify not only melanoma but also other malignant and benign skin lesions. We will focus the use of DL on several different imaging modalities individually and only briefly discuss models based on genetic or other non-imaging data.

Methods

An electronic literature search was performed on July 5, 2022, with the PubMed, Web of Science, and Embase databases. Articles published in English from October 31, 2019 through July 5, 2022 were included in the search. These were supplemented by hand searches through reference lists and expanding beyond the search window for key papers. A second search was conducted on October 18, 2022 to include papers published since the initial search. Search terms included *melanoma* or *skin neoplasm(s)* combined with *diagnosis* combined with *artificial intelligence* or *machine learning*. The full search strategy, including MeSH terms and keywords, can be found in "Appendix."

Clinical Image Analysis

The most accessible application of DL to diagnose melanoma uses clinical photographs. This method involves using a pretrained network such as ResNet, VGG-16, or InceptionV3, or training a convolutional neural network (CNN) on images sourced from large datasets such as ImageNet (https://www.image-net.org/) and the International Skin Imaging Collaboration (ISIC) Archive (https://www.isic-archive.com/) or individually created datasets and then providing clinical photograph as inputs for the CNN to diagnose [8]. Even the oldest published DL models have shown success in being able to distinguish melanoma from benign nevi, achieving high accuracies, sensitivities, and specificities [9–11].

Early studies trained CNN to discriminate melanoma from benign nevi [12–14]. Clinically, it can often be difficult to discern melanocytic lesions from solar lentigines or seborrheic keratoses, among others, and these simple binary models would perform poorly in practice if other lesion types were encountered. Moreover, at this time, there was no benchmark against which to test CNN performance to compare models to each other and, more importantly, to the clinical performance of trained dermatologists. This led Brinker et al. to create the first melanoma classification benchmark (MClass) and to provide the images publicly for direct comparison with other algorithms [15]. This standard, however, was a reference for only melanocytic images of light-skinned Western populations. Many of these early publications did not test algorithms in clinical settings and identified this as a limitation [8, 16]. Therefore Han et al. utilized a CNN trained with over 200,000 clinical images in a multi-ethnic cohort to classify images including melanoma and 11 other common malignant or benign skin lesions [14]. Binary and multiclass classification was used to evaluate the performance of the algorithm compared to clinicians and AI-assisted clinicians. The results showed how even subtle differences in lighting in images and lack of clinical information could negatively impact the ability to classify lesions [16]. Another group recognized this gap and trained their CNN on clinical images that were artificially darkened [17]. Their results suggest that even artificial darkening of images can significantly improve a CNN's ability to recognize melanoma in patients with darker skin tones.

Incremental advances have been achieved by including additional clinical data such as age, part of the body where the lesion is located, symptoms present, and change in size or elevation [18, 19]. Maclellan et al. conducted a study to evaluate three different models in the diagnosis of melanoma and compared those to the performance of a teledermoscopist and local dermatologists [19]. While several DL models had comparable or lower diagnostic performance compared to teledermoscopy (84.5% sensitivity, 82.6% specificity) or face-to-face evaluation by a dermatologist (96.6% sensitivity, 32.2% specificity), the best performing one was FotoFinder Moleanalyzer Pro (88.1% sensitivity, 78.8% specificity). The authors concluded that the lesion size and location, as well as Fitzpatrick skin type, could impact classification, and that the use of DL in the diagnosis of melanoma currently should be as an ancillary tool rather than a standalone one [20].

In line with this assessment, others have investigated the use of DL as a triage system for the identification and prioritization of malignant lesions [21]. In fact, many publications have suggested ancillary use of DL augments the diagnostic ability of dermatologists and non-dermatologists [16, 19, 22•, 23, 24]. Ba et al. specifically investigated the

difference in accuracy, sensitivity, and specificity of AI-assisted dermatologists vs. unassisted dermatologists [8]. They observed that AI-assisted dermatologists achieved a significantly higher accuracy (76.60% vs. 62.78%, $P < 0.001$), sensitivity (89.56% vs. 83.21%, $P < 0.001$), and specificity (87.90% vs. 80.92%, $P < 0.001$) compared to unassisted dermatologists (Table 2).

Multiple patient-driven smartphone applications ("apps") exist for detecting melanoma, such as SkinVision® and TeleSkin, both of which are approved in Europe as certified medical products for use as diagnostics, and DermaCompare, which is approved by the US Food and Drug Administration (FDA) as a class 1 medical device that is not for use as a diagnostic standalone tool. Multiple recent publications have evaluated the ability of these and other smartphone applications at detecting melanoma [25–27, 28•]. The conclusion of each of them was that the apps performed far worse in practice than the published metrics. Moreover, no patients in one study preferred melanoma screening based solely on the results from an app, which could severely restrict the broader acceptance of this tool in practice [28•]. As with dermatologist-driven methods, the current recommended use of DL in patient-driven screening for melanoma is to assist the triage and assessment by a trained dermatologist.

Dermoscopy

Dermoscopy is a real-time, in vivo imaging technique that allows for visualization of substructures in the skin surface and superficial dermis [29]. This involves the use of a handheld magnifier that can emit both polarized and non-polarized light. Polarization decreases the light reflection at the surface, thereby revealing the superficial features beneath (Fig. 1B and E). Several large datasets include dermoscopic images that have been used extensively for building deep learning models [30–32]. DL models based on these datasets have tended to achieve excellent classification of melanoma and other skin lesions, often with sensitivity greater than 90% [15, 33–41].

In actual practice, however, these models attain far worse accuracy and are highly sensitive to variation in how the image is obtained and oriented. One group found up to 22% misclassification with simple perturbations in image composition [42]. Others showed how slight changes in the red, green, and blue values for each pixel that would be imperceptible to the naked eye can dramatically impair an algorithm's ability to discriminate melanoma from benign nevus, even after training the network on similarly distorted images [43]. This study also showed how rotation or horizontal or vertical shifting was enough to cause misclassification of melanoma as benign nevus. Still, others showed that the use of skin markers, or dermatoscopes with different sized scale bars imprinted on the lens could greatly diminish classification ability [44, 45]. Finally, a large study from ISIC found a balanced accuracy of only 58.8% for multi-class classification in a dataset developed to reflect more real-world lesions that a dermatologist would encounter [46•]. This suboptimal accuracy was encountered even with algorithms designed to flag images that scored poorly on all classes and were likely outside of the trained expertise of the models. The authors concluded that such algorithms could potentially be dangerous to deploy in their current state as standalone methods to identify melanoma. Others, however, found that a DL system approved for

market in Europe performed on par with dermatologists on classifying a small set of lesions that included melanomas and could improve a dermatologist's ability to classify skin lesions when used as a diagnostic aid [41, 47].

Beyond changes in the image characteristics, change in the lesion itself over time is one of the hallmark features of melanoma. Two DL algorithms showed markedly improved performance when comparing sequential dermoscopy imaging to a one-time image [48]. The standard of care for high-risk patients remains serial observation, and DL has the potential to identify more subtle changes over time to augment the clinical exam. This especially has the potential for easy deployment in clinics using total body photography (colloquially referred to as "mole mapping"), with the potential for direct interface with the photography software [49, 50].

More recently, investigators have acknowledged that further improvements in models can come only with improvements in the data, both in terms of volume and granularity. As with clinical image analysis, the addition of simple clinical data such as patient age and sex can have small impacts on the classification of dermoscopic images [32, 51–53]. Ongoing efforts to identify new features to encode into training images or to concatenate into DL models are likely to produce further progress toward validated, clinically useful tools for the diagnosis of melanoma dermoscopically.

Reflectance Confocal Microscopy and Optical Coherence Tomography

Reflectance confocal microscopy (RCM) and optical coherence tomography (OCT) both use near-infrared light to provide high resolution of the epidermis and papillary dermis in vivo up to a depth of 2 mm [54, 55]. Focusing the light source up or down can produce a 3-dimensional stack of images. RCM uses a shorter wavelength that gives excellent resolution at the cost of shallower penetration, while OCT can image deeper into the skin, but at lower resolution. As the microscope scans over the lesion, it creates a mosaic of image tiles that are pieced together into a larger image (Fig. 1C). A particular benefit of RCM is that the method is not only FDA approved but also has its own reimbursement codes from the US Centers for Medicare and Medicaid Services (Current Procedural Terminology codes 96,931–96,936), paving the way for expanded clinical use [56]. RCM has even been shown to be superior to dermoscopy for the diagnosis of melanoma in a meta-analysis [57].

Adoption of this method has been limited by the need for equipment that costs 2 orders of magnitude greater than a handheld dermatoscope, as well as the need for highly specialized expertise in both the production of images and the reading of them. Perhaps the greatest barrier to wider use of RCM is that many table-mounted microscope-based techniques require up to 15 min to image a single lesion, whereas dermoscopy can be done in seconds at the bedside. RCM is therefore often scheduled as a special imaging encounter where no dermatologist is seen, so if imaging indicates a need for biopsy, it necessitates creation of an additional appointment for this procedure. There has been very little investigation to date into melanoma diagnosis and deep learning with RCM images, with current research on optimizing the image acquisition and analysis protocol [58]. Proposed use cases for RCM in melanoma diagnosis and management is to map pre-surgical margins for lentigo maligna in patients undergoing Mohs, monitoring excision sites, and to assess clearance

after non-invasive treatments [59, 60]. As the technology shifts to faster, less expensive hand-held devices, there will likely be broader use of RCM and OCT in the clinical setting [61].

Pathology Whole Slide Image Analysis

Whole slide imaging (WSI), or virtual microscopy, is an imaging modality that has become popular for its ability to emulate conventional light microscopy in a virtual manner. As the name implies, the entire biopsy slide is imaged at high-resolution yielding a gigabyte-size image file. Current computational capabilities are unable to analyze these files in one piece in a DL framework, and so they are divided into individual tiles of 256 × 256 pixels, or the size required by the individual pre-trained network that is being used. There are two main approaches to image analysis: classification and segmentation. In classification, entire slides are given a label and can be classified as melanoma or not melanoma, for example. Segmentation is a much more intensive task and instead is a pixel-level or near pixel-level classification, such that the model could learn to identify the specific melanoma cells on the slide and not label uninvolved tissue. Most DL has focused on classification, although progress into segmentation of melanoma has already begun [62–64]. DL models based on WSI datasets have achieved classification of melanoma and other skin lesions, with the majority having both a sensitivity and an area under the ROC curve greater than 90% [65–72].

Several studies found that DL models based on WSI datasets had comparable sensitivity with dermatopathologists for histologic detection of melanoma [65, 66, 73]. Furthermore, one group found DL may be a useful assistive diagnostic tool for dermatologists in certain cases of melanoma. This study found that dermatologists performed better in diagnosis of nevoid melanoma cases with the assistance of DL trained with WSI, but it did not show any benefit for dermatologists when diagnosing general cases of melanomas [74]. Other studies found that DL models had significant prognostic capabilities of predicting progression-free survival in melanoma, which may provide patients with more insight into their prognosis [64, 72, 75].

There are various commonly found genetic mutations in melanomas, and mutational profiling of melanomas can be useful in determining precise treatments. One WSI study found a use for DL in predicting specific mutations on histopathology images of melanoma [63]. This study showed their DL model could detect distinct morphological changes in mutated BRAF melanomas with sensitivities greater than 0.75, which may prove beneficial in providing mutation-specific treatments. Others have tried to automate tasks such as counting mitotic figures or characterizing tumor infiltrating lymphocytes (TILs) to include both not only for staging purposes but also for inclusion in DL models [67, 76]. In terms of training methods, one study found that in most cases, the model was sufficiently accurate in predicting melanomas without requiring any additional patient data integration to aid in diagnosis [71]. Interestingly, they found, however, that incorporating patient data for lesions classified by the CNN with “low confidence” led to improved diagnostic accuracy.

Studies on DL models trained with WSI have shown impressive capabilities in classifying melanomas, but they still have significant limitations. As with many of the other early

models using clinical or dermoscopic images, several of these studies used a binary classification (i.e., melanoma vs. not melanoma) and not the many differential diagnoses a dermatopathologist would need to consider [68, 69, 77]. One of these studies found a misclassification rate of 19% for WSI, even with a binary classification model [77]. Another limitation is that training of DL models varies from one study to the next, and the models' abilities are confined to that which they are taught [68]. In this group's study, they only included pT3 and pT4 melanomas in training, which may have limited the model's ability to discriminate melanomas from nevi or even thin melanomas since it was not trained in recognizing these.

Response to Immune Checkpoint Inhibition

The use of immune checkpoint inhibitors (ICI) has had the single most beneficial impact on melanoma-specific survival in decades [78]. Not all patients will respond to this type of therapy and with high rates of immune-related adverse events, identifying those who are unlikely to respond to this type of treatment can reduce harms while maximizing benefits [79, 80]. We have shown that even simple segmentation of melanoma WSI can predict response to ICI with AUC of 0.80 [64]. A recurring theme of DL models is that better data leads to better prediction. Information on TILs, genetic signatures, and other clinical factors can help predict ICI response in melanoma and other tumors [76, 81–83]. As biomarkers for response become better characterized, incorporation of these into predictive models will produce further refinements toward precision oncology.

Discussion

The use of DL has advanced from simple binary classification and limited multi-class classification to a wide range of diagnoses in dermatology. Multiple studies from clinical and dermoscopic images as well as WSI support that the best classification from DL models comes from the use of these tools to augment human clinical judgement [16, 22•, 23, 36, 38]. To date, the evidence indicates that DL systems are not refined to the point of making highly reliable diagnoses without human input [46•, 84]. Several commercial smart phone applications exist for patient-driven point of care testing that have been approved at varying levels in Europe and North America. These apps have shifted from places to store photos and provide education to tools that interface with DL algorithms in attempts to provide diagnoses [26, 28•, 85]. These applications unfortunately have performed poorly in identifying melanoma, and none have attained FDA approval as a diagnostic.

Appropriately, it is becoming the standard for studies of new DL tools to report an uncertainty metric, akin to a confidence interval around a point estimate, in addition to the tool's class probability, and other metrics on the dataset and model [46•, 86••, 87]. This inclusion underscores the need for a trained dermatologist to review the final diagnoses before treatment decisions are made. These tools have great potential to be integrated into clinical practice, some more easily than others. For example, dermoscopic evaluation of a lesion might take 3 to 10 s, while RCM can take up to 15 min for a single lesion. Computational time for DL algorithms is also a critical feature, although a longer analytic time becomes acceptable when the patient is not sitting in the exam room waiting on an

immediate result. Image detection and classification models that can be run rapidly on a local device would be ideal for diagnostic DL to be adopted into bedside practice.

Recently, the artificial intelligence working group of ISIC established guidelines for the presentation and use of imaging in AI studies towards making these tools more reliable [86••]. The resulting checklist included reporting features of the data, the technique by which the images were generated, a technical assessment of any models used, and lastly how the authors envision the application of the AI, including the need for fairness in applicability. Fairness extends to having diverse representation in terms of skin color in the samples informing the AI as the dermatologic literature has long suffered from a relative lack of non-White images [88, 89]. The development of new datasets that specifically include non-White populations will further improve the ability to diagnose both melanoma and other skin disorders [14, 17, 90, 91].

There is a major difference in the clinical questions that each of these image analysis methods addresses. Lesions seen clinically do not necessarily need a proper diagnosis. Rather, the question that must be answered in that moment is “Does this lesion need to be biopsied?” In instances of technically non-malignant lesions such a severely dysplastic nevus, the answer can still be yes, and models should not be penalized for “misclassification” as a melanoma. Similarly, for a basal cell carcinoma, the answer could be no particularly if it is asymptomatic, in an elderly patient, and unlikely to invade underlying structures if left untreated [92]. Understanding that the goal of clinical image analysis is to remove the lesions that are troublesome and leave the ones that are not is one area where AI can improve in dermatology overall, not just in melanoma diagnosis.

For biopsied lesions with pathology slides available, there are several clinical questions: “What is the diagnosis? What is the prognosis of this? Does this need additional treatment? Which treatment is best?” At the time of writing, several whole slide imaging systems have been FDA approved for diagnostic use (Leica, Phillips IntelliSite Pathology Solutions), paving the way for augmented or even automated diagnoses. The issues of treatment and prognosis are intertwined. The lesion with narrowly clear margins and a low-risk profile could perhaps be safely monitored, whereas high-risk ones could require re-excision or sentinel lymph node biopsy. For melanoma specifically, there are poor data to guide the decision to pursue SLNB in T1a and T1b lesions, and a model that could predict the likelihood of nodal or distant metastases would greatly improve the ability to offer this procedure to those most likely to benefit from it while minimizing harms [93].

Future Goals and Barriers

1. Datasets with more benign lesions and more skin types: one of the most common complaints from patients is a changing lesion and concern for melanoma. The vast majority of these are not melanocytic but instead are either seborrheic keratoses or solar lentigines. Datasets that include predominantly melanocytic or malignant lesions will therefore be unable to address this important clinical question, and one that could lend itself well to automated triage of patient-supplied images. More and more image banks are including a vast array of diagnoses, with a drive to include more images of non-White skin.

2. Datasets including multiple imaging modalities on lesions: most ISIC nevi are not biopsied. A lack of histologic data is itself a confounder for deep learning algorithms, such that more benign-appearing lesions histologically could pose some difficulty to deep learning algorithms if they are essentially unseen before.

3. Datasets including melanoma subtypes and sun-exposed skin: melanomas are a heterogeneous group of skin cancers ranging from the aggressive nodular melanoma, to the more indolent lentigo maligna. Even within the four recognized histologic subtypes (these two plus superficial spreading and acral lentiginous), there can be additional subtypes such as nevoid, pigmented epithelioid, desmoplastic, and spindle cell. Datasets in the future should include labels with this level of granularity to aid in model development. For example, while lentigo maligna is a subtype of melanoma in situ, its development on extensively sun-damaged skin has a different histologic pattern, behavior, and often treatment.

The formation of large, high-quality datasets for DL models will likely require multi-institutional or multi-national collaborations that include both public and access-restricted data. There are multiple public access databases for skin images, though these tend to be only clinical and dermoscopic images [94]. Accordingly, the ISIC Archive currently allows for users to upload their own clinical and dermoscopic images, although there is the potential to add digital pathology in the future. Pooled resources such as this database can help overcome the limitations of sample size and data quality to advance the use of artificial intelligence and machine learning in melanoma.

Conclusion

There have been many incremental advances in both melanoma diagnostics and prognostic tools using artificial intelligence and machine learning. Higher quality input data will further improve these models' capabilities. While these tools hold potential to augment diagnosis by a trained dermatologist, they are not currently ready for fully automated deployment.

Acknowledgements

We would like to extend thanks to Ms. Rachel Lane Walden, MLIS at Vanderbilt's Annette and Irwin Eskind Family Biomedical Library for her help in identifying search terms and strategies. Photos are courtesy of Dr. Eric Tkaczyk, MD, PhD and the Vanderbilt Dermatology Translation Research Clinic (VDTRC.org).

Funding

Dr. Wheless is supported by a grant from the Skin Cancer Foundation, a Physician Scientist Career Development Award from the Dermatology Foundation, and VA CSR&D grant IK2 CX002452.

Appendix

PubMed

((“melanoma”[mesh] OR melanoma* [tiab] OR “skin neoplasms”[mesh] OR skin neoplasm*[tiab]) AND (“artificial intelligence”[mesh] OR artificial intelligence[tiab] OR machine learning[tiab])) AND (“Diagnosis”[Mesh] OR Diagnos*[tiab]).

Web of Science

TS = (melanoma OR “skin neoplasm*”) AND TS = (“artificial intelligence” OR “machine learning”) AND TS = (diagnos*).

Embase

(exp melanoma/ or melanoma.ab,ti. or exp skin tumor/ or skin neoplasm*.ab,ti.) and (exp artificial intelligence/ or artificial intelligence.ab,ti. or machine learning.ab,ti.) and (exp diagnosis/ or diagnos*.ab,ti.)

References

Papers of particular interest, published recently, have been highlighted as:

• Of importance

•• Of major importance

1. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8. [PubMed: 28117445]
2. Meng X, Chen J, Zhang Z, Li K, Li J, Yu Z, et al. Non-invasive optical methods for melanoma diagnosis. Photodiagnosis Photodyn Ther. 2021;34:102266. [PubMed: 33785441]
3. Shahriari N, Rabinovitz H, Oliviero M, Grant-Kels JM. Reflectance confocal microscopy: melanocytic and nonmelanocytic. Clin Dermatol. 2021;39(4):643–56. [PubMed: 34809769]
4. Serag A, Ion-Margineanu A, Qureshi H, McMillan R, Saint Martin M-J, Diamond J, et al. Translational AI and deep learning in diagnostic pathology. Front Med. 2019;1(6):185.
5. Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial intelligence in dermatology: a primer. J Invest Dermatol. 2020;140(8):1504–12. [PubMed: 32229141]
6•. Murphree DH, Puri P, Shamim H, Bezalel SA, Drage LA, Wang M, Pittelkow MR, Carter RE, Davis MDP, Bridges AG, Mangold AR, Yiannias JA, Tollefson MM, Lehman JS, Meves A, Otley CC, Sokumbi O, Hall MR, Comfere N. Deep learning for dermatologists: Part I. Fundamental concepts. J Am Acad Dermatol. 2022;87(6):1343–1351. 10.1016/j.jaad.2020.05.056. [PubMed: 32434009] Good overview of the more technical aspects of deep learning written for a clinical audience
7. Puri P, Comfere N, Drage LA, Shamim H, Bezalel SA, Pittelkow MR, et al. Deep learning for dermatologists: part II. Current applications. J Am Acad Dermatol [Internet]. 2020 May 16; Available from: http://www.sciencedirect.com/science/article/pii/S019096222030918X
8. 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;169:156–65. [PubMed: 35569282]
9. Soenksen LR, Kassis T, Conover ST, Marti-Fuster B, Birkenfeld JS, Tucker-Schwartz J, Naseem A, Stavert RR, Kim CC, Senna MM, Avilés-Izquierdo J, Collins JJ, Barzilay R, Gray ML. Using deep learning for dermatologist-level detection of suspicious pigmented skin lesions from wide-field images. Sci Transl Med. 2021;13(581):eabb3652. 10.1126/scitranslmed.abb3652 [PubMed: 33597262]
10. Huang H-W, Hsu BW-Y, Lee C-H, Tseng VS. Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. J Dermatol. 2021;48(3):310–6. [PubMed: 33211346]
11. 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 boardcertified dermatologists in skin tumour diagnosis. Br J Dermatol. 2019;180(2):373–81. [PubMed: 29953582]

12. Marchetti MA, Codella NCF, Dusza SW, Gutman DA, Helba B, Kalloo A, et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. J Am Acad Dermatol. 2018;78(2):270–277.e1. [PubMed: 28969863]

13. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol. 2018;29(8):1836–42. [PubMed: 29846502]

14. 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. J Invest Dermatol. 2018;138(7):1529–38. [PubMed: 29428356]

15. Brinker TJ, Hekler A, Hauschild A, Berking C, Schilling B, Enk AH, et al. Comparing artificial intelligence algorithms to 157 German dermatologists: the melanoma classification benchmark. Eur J Cancer. 2019;111:30–7. [PubMed: 30802784]

16. 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. J Invest Dermatol. 2020;140(9):1753–61. [PubMed: 32243882]

17. Aggarwal P, Papay FA. Artificial intelligence image recognition of melanoma and basal cell carcinoma in racially diverse populations. J Dermatolog Treat. 2022;33(4):2257–62. [PubMed: 34154489]

18. Pacheco AGC, Krohling RA. The impact of patient clinical information on automated skin cancer detection. Comput Biol Med. 2020;116:103545. [PubMed: 31760271]

19. MacLellan AN, Price EL, Publicover-Brouwer P, Matheson K, Ly TY, Pasternak S, et al. The use of noninvasive imaging techniques in the diagnosis of melanoma: a prospective diagnostic accuracy study. J Am Acad Dermatol. 2021;85(2):353–9. [PubMed: 32289389]

20. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol. 1988;124(6):869–71. [PubMed: 3377516]

21. Majidian M, Tejani I, Jarmain T, Kellett L, Moy R. Artificial Intelligence in the evaluation of telemedicine dermatology patients. J Drugs Dermatol. 2022;21(2):191–4. [PubMed: 35133107]

22•. 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(9):2353–2362.e2. [PubMed: 35183551] (This was perhaps the first prospective clinical trial on the impact of augmented intelligence in dermatology and skin cancer detection)

23. Kim YJ, Na J-I, Han SS, Won CH, Lee MW, Shin J-W, et al. Augmenting the accuracy of trainee doctors in diagnosing skin lesions suspected of skin neoplasms in a real-world setting: a prospective controlled before-and-after study. PLoS One. 2022;17(1):e0260895. [PubMed: 35061692]

24. Jain A, Way D, Gupta V, Gao Y, de Oliveira MG, 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(4):e217249. [PubMed: 33909055]

25. Udrea A, Mitra GD, Costea D, Noels EC, Wakkee M, Siegel DM, et al. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. J Eur Acad Dermatol Venereol. 2020;34(3):648–55. [PubMed: 31494983]

26. Sangers T, Reeder S, van der Vet S, Jhingoer S, Mooyaart A, Siegel DM, et al. Validation of a market-approved artificial intelligence mobile health app for skin cancer screening: a prospective multicenter diagnostic accuracy study. Dermatology. 2022;238(4):649–56. [PubMed: 35124665]

27. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. Br J Dermatol. 2022;186(4):744–6. [PubMed: 34811727]

28•. Jahn AS, Navarini AA, Cerminara SE, Kostner L, Huber SM, Kunz M, Maul JT, Dummer R, Sommer S, Neuner AD, Levesque MP, Cheng PF, Maul LV. Over-Detection of Melanoma-Suspect Lesions by a CECertified Smartphone App: Performance in Comparison to Dermatologists, 2D and 3D Convolutional Neural Networks in a Prospective Data Set of 1204 Pigmented Skin Lesions Involving Patients' Perception. Cancers (Basel). 2022;14(15):3829. 10.3390/cancers14153829 [PubMed: 35954491] This study showed how even approved smart phone applications greatly underperformed compared to their published metrics when tested against an external dataset.

29. Argenziano G, Soyer HP. Dermoscopy of pigmented skin lesions–a valuable tool for early diagnosis of melanoma. Lancet Oncol. 2001;2(7):443–9. [PubMed: 11905739]

30. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data. 2018;5(1):1–9. [PubMed: 30482902]

31. Combalia M, Codella NCF, Rotemberg V, Helba B, Vilaplana V, Reiter O, Carrera C, Barreiro A, Halpern AC, Puig S, Malvehy J BCN20000: dermoscopic lesions in the wild. arXiv e-prints. 2019. arXiv:1908. 02288.

32. Rotemberg V, Kurtansky N, Betz-Stablein B, Caffery L, Chousakos E, Codella N, et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data. 2021;8(1):1–8. [PubMed: 33414438]

33. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol. 2019;155(1):58–65. [PubMed: 30484822]

34. Tschandl P, Codella N, Akay BN, Argenziano G, Braun RP, Cabo H, et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. Lancet Oncol. 2019;20(7):938–47. [PubMed: 31201137]

35. Phillips M, Marsden H, Jaffe W, Matin RN, Wali GN, Greenhalgh J, et al. Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. JAMA Netw Open. 2019;2(10):e1913436. [PubMed: 31617929]

36. Lee S, Chu YS, Yoo SK, Choi S, Choe SJ, Koh SB, et al. Augmented decision-making for acral lentiginous melanoma detection using deep convolutional neural networks. J Eur Acad Dermatol Venereol. 2020;34(8):1842–50. [PubMed: 31919901]

37. Maron RC, Weichenthal M, Utikal JS, Hekler A, Berking C, Hauschild A, et al. Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. Eur J Cancer. 2019;119:57–65. [PubMed: 31419752]

38. Maron RC, Utikal JS, Hekler A, Hauschild A, Sattler E, Sondermann W, et al. Artificial intelligence and its effect on dermatologists' accuracy in dermoscopic melanoma image classification: web-based survey study. J Med Internet Res. 2020;22(9):e18091. [PubMed: 32915161]

39. Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur J Cancer. 2019;113:47–54. [PubMed: 30981091]

40. Pham T-C, Luong C-M, Hoang V-D, Doucet A. AI outperformed every dermatologist in dermoscopic melanoma diagnosis, using an optimized deep-CNN architecture with custom mini-batch logic and loss function. Sci Rep. 2021;11(1):1–13. [PubMed: 33414495]

41. Haenssle HA, Winkler JK, Fink C, Toberer F, Enk A, Stolz W, et al. Skin lesions of face and scalp - classification by a market-approved convolutional neural network in comparison with 64 dermatologists. Eur J Cancer. 2021;144:192–9. [PubMed: 33370644]

42. Young AT, Fernandez K, Pfau J, Reddy R, Cao NA, von Franque MY, et al. Stress testing reveals gaps in clinic readiness of image-based diagnostic artificial intelligence models. npj Digit Med. 2021;4(1):1–8. [PubMed: 33398041]

43. Du-Harpur X, Arthurs C, Ganier C, Woolf R, Laftah Z, Lakhan M, et al. Clinically relevant vulnerabilities of deep machine learning systems for skin cancer diagnosis. J Invest Dermatol. 2021;141(4):916–20. [PubMed: 32931808]

44. Winkler JK, Fink C, Toberer F, Enk A, Deinlein T, Hofmann-Wellenhof R, et al. Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. JAMA Dermatol. 2019;155(10):1135–41. [PubMed: 31411641]

45. 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;1(145):146–54.

46•. Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. Lancet Digit Health. 2022;4(5):e330–9. [PubMed: 35461690] (This study showed that when networks are exposed to categories not included in the training data there is a high degree of misclassificaiton, even those that are supposed to 'fail safely' by flagging images with low scores for all categories.)

47. Haenssle HA, Fink C, Toberer F, Winkler J, Stolz W, Deinlein T, et al. Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. Ann Oncol. 2020;31(1):137–43. [PubMed: 31912788]

48. Winkler JK, Tschandl P, Toberer F, Sies K, Fink C, Enk A, et al. Monitoring patients at risk for melanoma: may convolutional neural networks replace the strategy of sequential digital dermoscopy? Eur J Cancer. 2022;160:180–8. [PubMed: 34840028]

49. Betz-Stablein B, D'Alessandro B, Koh U, Plasmeijer E, Janda M, Menzies SW, et al. Reproducible naevus counts using 3D total body photography and convolutional neural networks. Dermatology. 2022;238(1):4–11. [PubMed: 34237739]

50. Rayner JE, Laino AM, Nufer KL, Adams L, Raphael AP, Menzies SW, et al. Clinical perspective of 3D total body photography for early detection and screening of melanoma. Front Med. 2018;23(5):152.

51. Sies K, Winkler JK, Fink C, Bardehle F, Toberer F, Buhl T, et al. Does sex matter? Analysis of sex-related differences in the diagnostic performance of a market-approved convolutional neural network for skin cancer detection. Eur J Cancer. 2022;164:88–94. [PubMed: 35182926]

52. Sun Q, Huang C, Chen M, Xu H, Yang Y. Skin lesion classification using additional patient information. Biomed Res Int. 2021;10(2021):6673852.

53. 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(2):115–22. [PubMed: 33358096]

54. Waddell A, Star P, Guitera P. Advances in the use of reflectance confocal microscopy in melanoma. Melanoma Manag. 2018;5(1):MMT04. [PubMed: 30190930]

55. Yang L, Chen Y, Ling S, Wang J, Wang G, Zhang B, et al. Research progress on the application of optical coherence tomography in the field of oncology. Front Oncol. 2022;25(12):953934.

56. Levine A, Markowitz O. Introduction to reflectance confocal microscopy and its use in clinical practice. JAAD Case Rep. 2018;4(10):1014–23. [PubMed: 30456275]

57. Pezzini C, Kaleci S, Chester J, Farnetani F, Longo C, Pellacani G. Reflectance confocal microscopy diagnostic accuracy for malignant melanoma in different clinical settings: systematic review and meta-analysis. J Eur Acad Dermatol Venereol. 2020;34(10):2268–79. [PubMed: 31997465]

58. Sikorska M, Skalski A, Wodzinski M, Witkowski A, Pellacani G, Ludzik J. Learning-based local quality assessment of reflectance confocal microscopy images for dermatology applications. Biocybern Biomed Eng. 2021;41(3):880–90.

59. Aleisa A, Cordova A, Alfonso A, Kose K, Rossi AM, Rajadhyaksha M, et al. Automated analysis of lentigo maligna margin mapping using hand-held reflectance confocal microscopy. Lasers Surg Med. 2022;54(Suppl. 34):S62.

60. Shahriari N, Grant-Kels JM, Rabinovitz H, Oliviero M, Scope A. Reflectance confocal microscopy: principles, basic terminology, clinical indications, limitations, and practical considerations. J Am Acad Dermatol. 2021;84(1):1–14. [PubMed: 32553679]

61. Maguluri G, Grimble J, Mujat M, Park J, Caron A, Iftimia N. Fiber-based hand-held RCM-OCT probe for noninvasive assessment of skin lesions and therapy guidance. Translational Biophotonics. 2022;4:e202200002. 10.1002/tbio.202200002 [PubMed: 36176918]
62. Li T, Xie P, Liu J, Chen M, Zhao S, Kang W, et al. Automated diagnosis and localization of melanoma from skin histopathology slides using deep learning: a multicenter study. J Healthc Eng. 2021;26(2021):5972962.
63. Kim RH, Nomikou S, Coudray N, Jour G, Dawood Z, Hong R, et al. Deep learning and pathomics analyses reveal cell nuclei as important features for mutation prediction of BRAF-mutated melanomas. J Invest Dermatol. 2022;142(6):1650–1658.e6. [PubMed: 34757067]
64. Johannet P, Coudray N, Donnelly DM, Jour G, Illa-Bochaca I, Xia Y, et al. Using machine learning algorithms to predict immunotherapy response in patients with advanced melanoma. Clin Cancer Res. 2021;27(1):131–40. [PubMed: 33208341]
65. Ba W, Wang R, Yin G, Song Z, Zou J, Zhong C, et al. Diagnostic assessment of deep learning for melanocytic lesions using whole-slide pathological images. Transl Oncol. 2021;14(9):101161. [PubMed: 34192650]
66. Brinker TJ, Schmitt M, Krieghoff-Henning EI, Barnhill R, Beltraminelli H, Braun SA, et al. Diagnostic performance of artificial intelligence for histologic melanoma recognition compared to 18 international expert pathologists. J Am Acad Dermatol. 2022;86(3):640–2. [PubMed: 33581189]
67. Nofallah S, Mehta S, Mercan E, Knezevich S, May CJ, Weaver D, et al. Machine learning techniques for mitoses classification. Comput Med Imaging Graph. 2021;87:101832. [PubMed: 33302246]
68. De Logu F, Ugolini F, Maio V, Simi S, Cossu A, Massi D, et al. Recognition of cutaneous melanoma on digitized histopathological slides via artificial intelligence algorithm. Front Oncol. 2020;20(10):1559.
69. Wang L, Ding L, Liu Z, Sun L, Chen L, Jia R, et al. Automated identification of malignancy in whole-slide pathological images: identification of eyelid malignant melanoma in gigapixel pathological slides using deep learning. Br J Ophthalmol. 2020;104(3):318–23. [PubMed: 31302629]
70. Jackson CR, Sriharan A, Vaickus LJ. A machine learning algorithm for simulating immunohistochemistry: development of SOX10 virtual IHC and evaluation on primarily melanocytic neoplasms. Mod Pathol. 2020;33(9):1638–48. [PubMed: 32238879]
71. Höhn J, Krieghoff-Henning E, Jutzi TB, von Kalle C, Utikal JS, Meier F, et al. Combining CNN-based histologic whole slide image analysis and patient data to improve skin cancer classification. Eur J Cancer. 2021;149:94–101. [PubMed: 33838393]
72. Kulkarni PM, Robinson EJ, Sarin Pradhan J, Gartrell-Corrado RD, Rohr BR, Trager MH, et al. Deep learning based on standard H&E images of primary melanoma tumors identifies patients at risk for visceral recurrence and death. Clin Cancer Res. 2020;26(5):1126–34. [PubMed: 31636101]
73. Laverde-Saad A, Jfri A, García R, Salgüero I, Martínez C, Cembrero H, et al. Discriminative deep learning based benignity/malignancy diagnosis of dermatologic ultrasound skin lesions with pretrained artificial intelligence architecture. Skin Res Technol. 2022;28(1):35–9. [PubMed: 34420233]
74. Sturm B, Creytens D, Smits J, Ooms AHAG, Eijken E, Kurpershoek E, et al. Computer-aided assessment of melanocytic lesions by means of a mitosis algorithm. Diagnostics (Basel) [Internet]. 2022 Feb 8;12(2). Available from:10.3390/diagnostics12020436
75. 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(1):2809. [PubMed: 33531581]
76. Wang X, Barrera C, Bera K, Viswanathan VS, Azarianpour-Esfahani S, Koyuncu C, et al. Spatial interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for immune checkpoint inhibitors. Sci Adv. 2022;8(22):eabn3966. [PubMed: 35648850]

77. Hekler A, Utikal JS, Enk AH, Berking C, Klode J, Schadendorf D, et al. Pathologist-level classification of histopathological melanoma images with deep neural networks. Eur J Cancer. 2019;115:79–83. [PubMed: 31129383]

78. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. N Engl J Med. 2019;381(16):1535–46. [PubMed: 31562797]

79. Nebhan CA, Johnson DB. Predictive biomarkers of response to immune checkpoint inhibitors in melanoma. Expert Rev Anticancer Ther. 2020;20(2):137–45. [PubMed: 31997676]

80. Schneider BJ, Naidoo J, Santomasso BD, Lacchetti C, Adkins S, Anadkat M, et al. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: ASCO guideline update. J Clin Orthod. 2021;39(36):4073–126.

81. Yuan B, Miao L, Mei D, Li L, Zhou Q, Dong D, et al. Value of a signature of immune-related genes in predicting the prognosis of melanoma and its responses to immune checkpoint blocker therapies. Comput Math Methods Med. 2022;20(2022):9633416.

82. Kong J, Ha D, Lee J, Kim I, Park M, Im S-H, et al. Network-based machine learning approach to predict immunotherapy response in cancer patients. Nat Commun. 2022;13(1):1–15. [PubMed: 34983933]

83. Benzekry S, Grangeon M, Karlsen M, Alexa M, Bicalho-Frazeto I, Chaleat S, et al. Machine learning for prediction of immunotherapy efficacy in non-small cell lung cancer from simple clinical and biological data. Cancers [Internet]. 2021 Dec 9;13(24). Available from:10.3390/cancers13246210

84. Beltrami EJ, Brown AC, Salmon PJM, Leffell DJ, Ko JM, Grant-Kels JM. Artificial intelligence in the detection of skin cancer. J Am Acad Dermatol [Internet]. 2022 Aug 23; Available from:10.1016/j.jaad.2022.08.028

85. Kassianos AP, Emery JD, Murchie P, Walter FM. Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol. 2015;172(6):1507–18. [PubMed: 25600815]

86••. Daneshjou R, Barata C, Betz-Stablein B, Celebi ME, Codella N, Combalia M, et al. Checklist for evaluation of image-based artificial intelligence reports in dermatology: CLEAR derm consensus guidelines from the international skin imaging collaboration artificial intelligence working group. JAMA Dermatol. 2022;158(1):90–6. [PubMed: 34851366] (This consensus statement aims to formalize the requirements for study design, methodology, reporting, and use of AI in dermatology.)

87. Daneshjou R, Smith MP, Sun MD, Rotemberg V, Zou J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. JAMA Dermatol. 2021;157(11):1362–9. [PubMed: 34550305]

88. Perng P, Greer KE. An ode to kodachrome: the color film that transformed dermatologic medical education. JAMA Dermatol. 2022;158(8):928–32. [PubMed: 35704331]

89. Slaught C, Madu P, Chang AY, Williams VL, Kebaetse MB, Nkomazana O, et al. Novel education modules addressing the underrepresentation of skin of color in dermatology training. J Cutan Med Surg. 2022;26(1):17–24. [PubMed: 34340596]

90. Huang K, Jiang Z, Li Y, Wu Z, Wu X, Zhu W, et al. The classification of six common skin diseases based on Xiangya-Derm: development of a chinese database for artificial intelligence. J Med Internet Res. 2021;23(9):e26025. [PubMed: 34546174]

91. Daneshjou R, Vodrahalli K, Novoa RA, Jenkins M, Liang W, Rotemberg V, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv. 2022;8(32):eabq6147. [PubMed: 35960806]

92. Sreekantaswamy S, Endo J, Chen A, Butler D, Morrison L, Linos E. Aging and the treatment of basal cell carcinoma. Clin Dermatol. 2019;37(4):373–8. [PubMed: 31345326]

93. Coit DG, Thompson JA, Albertini MR, Barker C, Carson WE, Contreras C, et al. Cutaneous melanoma, version 2.2019, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw. 2019;17(4):367–402. [PubMed: 30959471]

94. Wen D, Khan SM, Xu AJ, Ibrahim H, Smith L, Caballero J, et al. Characteristics of publicly available skin cancer image datasets: a systematic review. Lancet Digit Health. 2022;4:e64–74. [PubMed: 34772649]

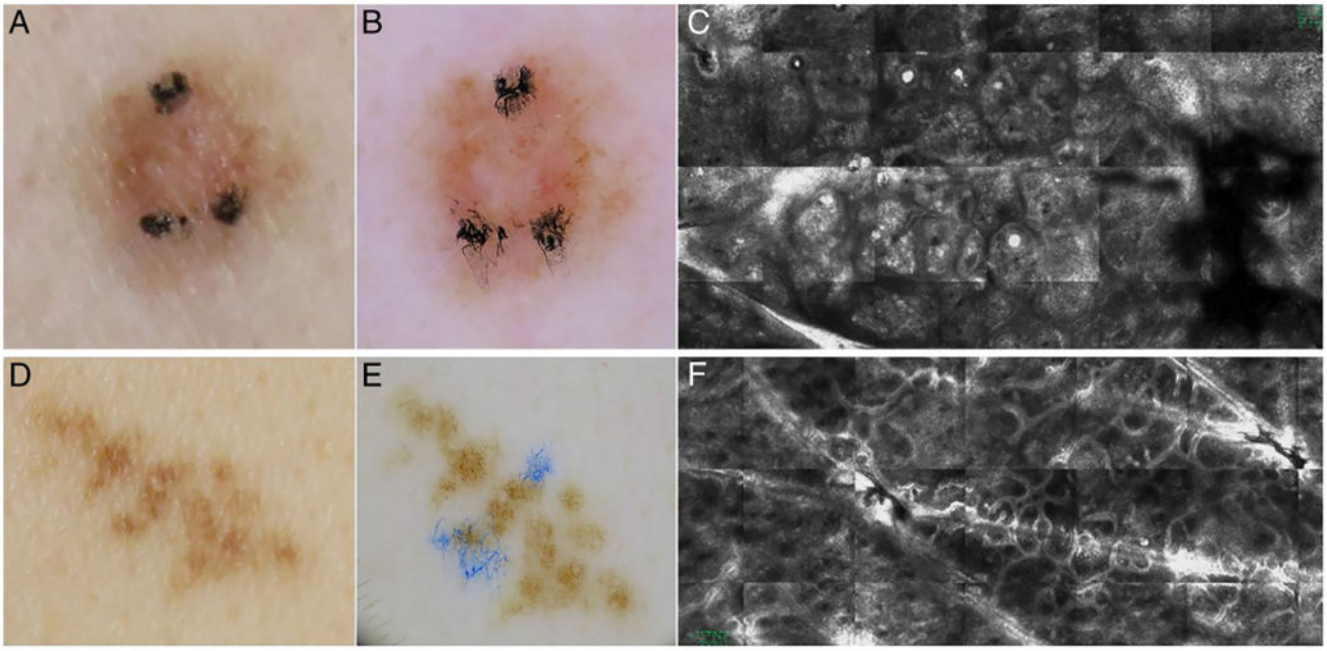

Fig. 1.

Examples of (**A**), (**D**) clinical, (**B**), (**E**) dermoscopic, and (**C**), (**F**) reflectance confocal microscopy images on two individual nevi

Table 1

In vivo skin imaging modalities used for melanoma diagnosis

Modality	Description of technology	Features
Dermoscopy	Hand-held magnifier with polarized light source	Pros: 1. Imaging: decreases light reflectance at the skin surface to allow for clearer view of epidermal structures 2. Timing: 1 to 10 s for a single lesion Cons: 1. Imaging: highly sensitive to variation in how the image is obtained and oriented for store-and-forward process
Reflectance confocal microscopy	Microscope with near-ultraviolet laser light to illuminate up to 200 microns into the skin	Pros: 1. Imaging: can provide 3-dimensional *en face* view with subcellular resolution of skin in grayscale 2. Imaging: RCM uses a shorter wavelength that gives excellent resolution at the cost of shallower penetration 3. Cost: FDA approved and has its own reimbursement codes from the US Centers for Medicare and Medicaid Services Cons: 1. Timing: up to 15 min for a single lesion 2. Cost: equipment costs 2 orders of magnitude greater than a handheld dermatoscope 3. Imaging: need for highly specialized expertise in both the production of images and the reading of them
Optical coherence tomography	Microscope with near-ultraviolet laser light to illuminate up to 2000 microns into the skin	Pros: 1. Imaging: penetrates more deeply than dermoscopy, RCM Cons: 1. Imaging: lower resolution than RCM

Table 2

Clinical imaging: sensitivity and specificity of skin lesion diagnosis among AI-assisted dermatologists vs. unassisted dermatologists using a deep learning model trained on clinical images. [8] Adapted from Ba et al.

	Sensitivity	Specificity
AI-assisted evaluation	89.56	87.90
Unassisted evaluation	83.21	80.92