Received: 25 May 2022 | Revised: 6 October 2022 | Accepted: 17 October 2022

DOI: 10.1111/ajd.13946

Australasian Journal of Dermatology

REVIEW ARTICLE

Informing a position statement on the use of artificial intelligence in dermatology in Australia

Liam J. Caffery PhD[1,2,3] | **Monika Janda PhD**[2] | **Robert Miller MBBS**[4] | **Lisa M. Abbott MBBS, LLM**[4,5] | **Chris Arnold BComm, MBA**[4,6,7] | **Tony Caccetta MBBS**[4] | **Pascale Guitera MD, PhD**[4] | **Stephen Shumack MBBS**[4,8] | **Pablo Fernández-Peñas MD, PhD**[4,8] | **Victoria Mar PhD, MBBS**[4,9,10] | **H. Peter Soyer MD**[3]

[1]Centre for Online Health, The University of Queensland, Brisbane, Queensland, Australia

[2]Centre for Health Services Research, The University of Queensland, Brisbane, Queensland, Australia

[3]The University of Queensland Diamantina Institute, The University of Queensland, Dermatology Research Centre, Brisbane, Queensland, Australia

[4]The Australasian College of Dermatologists, Sydney, New South Wales, Australia

[5]Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia

[6]Australasian Society of Cosmetic Dermatologists, South Yarra, Victoria, Australia

[7]BioGrid Australia, Parkville, Victoria, Australia

[8]The University of Sydney, Sydney, New South Wales, Australia

[9]Victorian Melanoma Service, Alfred Health, Victoria, Australia

[10]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia

Correspondence
Liam J. Caffery, Ground Floor, Building 33, Princess Alexandra Hospital, 199 Ipswich Road Woolloongabba QLD 4102, Australia.
Email: l.caffery@uq.edu.au

Funding information
Australasian College of Dermatologists

Abstract

Artificial Intelligence (AI) is the ability for computers to simulate human intelligence. In dermatology, there is substantial interest in using AI to identify skin lesions from images. Due to increasing research and interest in the use of AI, the Australasian College of Dermatologists has developed a position statement to inform its members of appropriate use of AI. This article presents the ACD Position Statement on the use of AI in dermatology, and provides explanatory information that was used to inform the development of this statement.

KEYWORDS

artificial intelligence, dermatology, guidelines, machine learning, position statement

INTRODUCTION

Artificial Intelligence (AI) is the ability for computers to simulate human intelligence.[1] It is a broad discipline covering fields such as natural language processing, robotics, and computer vision. Computer vision enables computers (i.e. AI models) to interpret and derive information from visual inputs (e.g. images or video).

In dermatology, there is substantial interest in using AI to identify skin lesions from images. As visual pattern recognition plays an important role in dermatology, it is an ideal clinical application of AI.[1] The use of AI to diagnose or predict the risk of a disease from an image is known as image classification. Image classification has been at the forefront of AI research in dermatology with encouraging findings.[2]

 © 2022 Australasian College of Dermatologists.

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Due to increasing research and interest in the use of AI, the Australasian College of Dermatologists (ACD) has developed a position statement to inform its members of appropriate use of AI. The purpose of this article is to present the ACD Position Statement on the use of AI in dermatology; and to provide explanatory information that was used to inform the development of this statement.

METHODS

The ACD Position Statement was developed collaboratively by the ACD Digital Health Committee and The University of Queensland. Several international dermatology professional bodies have published position papers on the use of AI in dermatology. Other medical image-centric specialties have also published position statements on the use of AI in their disciplines. These publications have been reviewed to identify consensus and help ensure completeness of the ACD's position statement. Australian and international regulatory guidelines and ethical position papers were also reviewed.

BACKGROUND

Methods of AI

AI uses different methods (Figure 1), with machine learning (ML) and its subfield deep learning (DL) being most relevant to dermatology. The most common DL architecture is a convolutional neural network (CNN) which is well suited to image analysis. ML methods use generic software known as algorithms which are mathematical functions that learn from data without explicit programming. As such, a generic ML algorithm could be trained to identify dog breeds from images of dogs, or to identify an image of a skin lesion as benign or malignant. Training occurs by inputting images and an associated label (e.g. image of a dog and a label of the breed, or an image of a skin lesion and the diagnosis). Labels are also referred to as the ground truth. Image labels or the ground truth may come from different sources (e.g. dermatologist clinical diagnosis, consensus diagnosis by multiple dermatologists, histopathologic diagnosis). Data used to train an AI algorithm are known as the training dataset. Once an algorithm has been trained on a dataset, it becomes an AI model.

Developing an AI model for image classification task

Training an AI model on a specific task using the training dataset is known as supervised learning. During supervised learning, an ML algorithm will determine the mathematical functions necessary to map a set of images to their correct label.

DL methods use an architecture called artificial neural networks which are interconnected layered networks where the output of one layer is the input of the next layer. Data pass through each layer of the neural network where a mathematical operation is applied to the data. During training, each layer is adjusted or tuned to improve accuracy of the data matching the label. A convolutional neural network (CNN) is a type of artificial neural network where each layer applies filters to specific areas within an image. CNNs have proven to be effective and consistently superior to other architectures for image classification.[1,3] The conclusions of evidence reviews evaluating

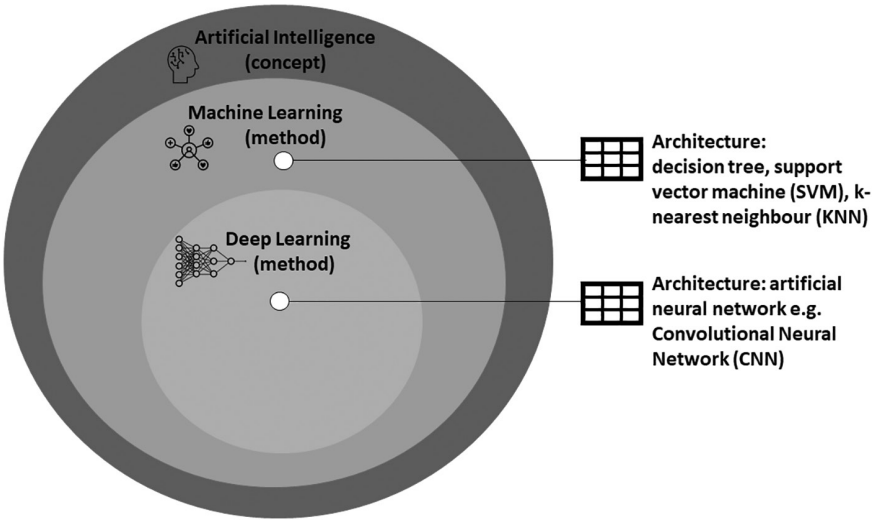

FIGURE 1 Methods of AI

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the effectiveness of AI in dermatology are summarised in Table S1–S3.

An important consideration is how many images are needed to train an AI model using ML. As a rule of thumb, the more the images, the more accurate the model will be. Some commentators have suggested that 1000–5000 images per output class are needed.[4] Others have suggested it is not a linear relationship, and the number of training images depends on the complexity of the problem, and the ratio between classifications and the full dataset.[5]

Testing and validation

AI models need to be tested and validated, which is an iterative process of determining how well the model's output matches the ground truth. Validation and testing are important steps in developing an AI model, as they help identify issues such as overfitting and underfitting. Testing also informs statistical reporting of the trained model's performance.

Underfitting and overfitting

Underfitting and overfitting are problems that occur during supervised learning that reduce the accuracy and generalisability of AI model. Poor generalisability means the AI model does not perform well on data it has not been trained with. Overfitting occurs when the AI model has high accuracy on training data but low accuracy on unseen (test) data; that is, the model has learnt the dataset "too well". Underfitting results in low accuracy on both the training and test data. Further explanation of underfitting and overfitting is available in Table S2.

Once training is optimised, the model is tested using a test dataset of images that have been held out from the training dataset. The results of this testing phase are used for statistical reporting of the accuracy of the model.

Generalisability

External validation is used to determine the generalisability of the AI model and is typically assessed using images from sources external to those used for training. AI models are highly context-specific to the data on which they were trained. To illustrate this point and the importance of external validation, an AI model developed by Han et al.[6] was shown to have considerably lower sensitivity when applied to a different population.[7] This was despite both studies using test datasets containing images from Caucasian patients. In a further study, the area under the receiver operating characteristic curve (AUROC) dropped 27%–36% compared to the model's original test results when validated with an image dataset with diverse skin tones.[8] Training data bias on gender, ethnicity and race means that models are not generalisable for people of an under-represented gender, ethnicity or race.[3,9]

Binary versus multidisease classifiers

AI models can be binary classifiers or multidisease classifiers. Most AI models in dermatology have been developed for binary classification of melanoma.[2] Binary classifiers have two outputs (e.g. benign or malignant, normal or not normal, or classification of a lesion into melanoma or naevus). Binary classifiers will convert a threshold into a binary label. For example, if a binary classifier for melanoma versus naevus uses a 50% threshold, any risk prediction score of greater than 50% will be classified as melanoma. The number of false positives increases as the threshold is lowered. Developers of AI models with high risk aversion may intentionally use a low threshold for diagnosis resulting in a high number of false positives. This may result in undue burden on the healthcare system through unnecessary doctor visits, and over-testing of people who are otherwise not at risk for disease.[10]

Multidisease or multiclass classifiers will provide risk prediction for a range of diseases. For example, the International Skin Imaging Collaborative (ISIC) 2018 challenge used seven disease classes (actinic keratosis and Bowen disease, basal cell carcinoma, benign keratinocytic tumours including solar lentigo and seborrheic keratosis, dermatofibroma, melanoma, melanocytic nevus and vascular lesion).[11] A further study trained a multiclass CNN to classify 40 different skin diseases.[12] Multiclass image classifiers will often output a risk prediction score for top predictions (e.g. top-1, top-3, or top-5 disease condition predictions).

Evaluating AI models

AI models are evaluated using common statistical measures such as overall accuracy, sensitivity, specificity, and AUROC. Table S2 summarises AI reader studies in dermatology that use CNNs. Most studies report highly accurate performance of AI models which are often described as equivalent or superior to human readers.[13]

This good performance may hide the fact that AI models produce false-negative and false-positive results whilst still reporting high accuracy, sensitivity, specificity, and AUC. For example, Haenssle et al.[13] tested their AI model on two datasets. The resultant sensitivity, specificity, and AUROC

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

were 95%, 63.8%, and 0.86, 95%, 80%, and 0.95, respectively. However, the AI model's risk prediction for invasive melanoma ranged from 0% to 100%, meaning a number of invasive melanomas were classified as having 0% probability of melanoma, that is, were missed by the AI (false negative). Similarly, false positives were also reported.

There are a number published guidelines that have been developed to assist with evaluation and reporting of AI studies. CLEAR Derm is a consensus-based guideline that can be used for a structured evaluation of image-based AI studies in dermatology.[14] CONSORT-AI and SPIRIT-AI provide guidance to researchers on the transparent reporting of AI studies.[15]

One of the well-recognised issues with testing and validation of AI models in dermatology is they are typically done retrospectively under highly artificial conditions, for example, images have been collected because they were noteworthy for some reason. Therefore, performance of an AI model in clinical practice remains largely unknown and prospective testing in real-world settings is needed to help answer this question.

Black box

There are criticisms that neural networks function as 'black boxes' with potential unanticipated and hard-to-explain failures such as false positives and false negatives.[1] AI models produce outputs without transparency or explanation meaning that end-users (e.g. clinicians) often have no knowledge of how an AI model derives its output. The inputs (e.g. images) and output (e.g. risk prediction of disease) of an AI model are visible. However, how the input is transformed into the output may not be understandable. Explainable AI may be used to mitigate the black box nature of AI models. Explainable AI is the term used for techniques that identify and present salient features or areas of images used to derive the model output to the end user and may thereby lessen the 'black box' concerns. Saliency maps or class activation maps are two examples of Explainable AI that use colour overlays to identify the area of the image that was used by the model to generate output. If a class activation map identified that areas outside the border of a skin lesion or ink markings were important in deriving its output, then the confidence in the prediction may be lessened. Further explanation of Explainable AI is available in Table S3.

Automation bias

Automation bias is a tendency for a human to accept the decision made by a computer even if it contradicts their own intuition.[16] One such example of automation bias was demonstrated when clinicians altered their initial diagnosis of a skin lesion after seeing deliberately inaccurate AI risk predictions scores.[17] Notably, in this study all clinicians ranging from inexperienced to expert altered their initial diagnosis. Using an AI model outside of the context in which it was trained is likely to increase inaccurate disease predictions and the possibility for clinicians to accept the incorrect diagnosis. Automation bias can be caused by automation-induced complacency and insufficient monitoring of automation output.[18] Continual monitoring of the output of the AI model is one strategy to reduce automation bias.

AI use in different settings

AI models may be integrated in imaging systems for use by clinicians or to assist consumers directly in their at-home self-management. Clinician AI models are already integrated in multiple commercial imaging systems that dermatologists and other skin cancer doctors use in their practice. AI algorithms are also integrated into mobile phone apps, many of which are directed at consumers to support self-assessment of concerning skin lesions.[19]

AI POSITION STATEMENTS FROM DERMATOLOGY PROFESSIONAL GROUPS

Two published position papers from dermatology professional bodies, namely the American Academy of Dermatology (AAD) (May 2019)[20] and the British Association of Dermatologists (BAD) (February 2021)[21] were identified and reviewed. Notably, both position papers advocate the development of AI for dermatology. However, there is hesitancy in recommending the use of AI in clinical practice. To this end, the AAD recommend that AI models are externally validated by "prospective clinical trials" which at the time of writing did not exist. The BAD identified that current evidence for AI was limited and the use of AI for dermatology was currently in a "study phase". The AAD uses the terminology Augmented Intelligence (AuI) as opposed to AI. AuI focusses on artificial intelligence's assistive role in the physician/patient relationship as opposed to replacement of it.

AI REGULATION

The position statements from the AAD and BAD both call for appropriate regulation of AI models for dermatology.

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 TGA classification of devices used for the diagnosis of a disease

Risk	Provides information to an individual (non-healthcare professional)	Provides information to a healthcare professional
Death/severe deterioration	Class III high	Class IIb medium-high
Serious disease or condition	Class IIb medium-high	Class IIa low-medium
Other	Class IIa low-medium	Class I low

There appears to be limited international regulatory approval of AI for dermatology. SkinVision™, a melanoma detection smartphone app for consumers that is integrated with a teledermatology service, has regulatory approval in Australia as a Class I (low-risk) medical device,[22] and in Europe as a low-risk device.[19] Moleanalyzer Pro™ and Teleskin™ have also been approved for the European market.[19,23] At the time of writing, the US FDA has not approved any dermatology AI models.[24] However, by 2020 the FDA had approved 64 AI image classifiers for ophthalmology, radiology, and cardiology.[25]

In Australia, AI models are regulated as Software as a Medical Device (SaMD) by the Department of Health's Therapeutic Goods Administration (TGA). Software is categorised as a medical device if it is intended to be used for the diagnosis, prevention, monitoring, prediction, prognosis, or treatment of a disease.[26] AI models (including those implemented as smartphone apps) would meet the definition of a medical device and would therefore be regulated as an SaMD. Notably, not all devices are regulated. Unregulated SaMD may include software that are simply a source of information, consumer health products that do not provide specific treatment suggestions, communication software (such as those that enable a telehealth consultation), or systems that store or transmit patient images. Direct-to-consumer AI apps may not require TGA approval if they do not meet the definition of a medical device.

Before an SaMD AI model can be legally supplied in Australia it must have TGA pre-market approval. Products approved by the TGA are listed on the Australian Register of Therapeutic Goods. AI models are subject to both pre- and post-market regulatory oversight. Pre-market approval uses a risk-based approach that includes a risk assessment and clinical assessment commensurate with risk level. The TGA is currently undertaking a reform of the regulation of SaMD. The aim of this reform is to be consistent with the International Medical Device Regulators Forum (IMDRF) regulation of SaMD.

One notable change that the adoption of the IMDRF processes has brought is a change to the way risk is assessed. Under legacy processes the risk to patients was assessed by the physical harm that could result from the device; as an AI model was unlikely to result in physical harm they were classified as low risk. This approach has been criticised because the harm posed by AI models is not physical harm but instead related to the information provided and its subsequent use in clinical decision making.[22] The new rules consider potential harm caused by providing incorrect information to users of the medical devices. Further considerations in the risk assessment include the purpose of the device (diagnosing or screening for a disease, monitoring the progression of a disease, specifying or recommending a treatment, or providing therapy), and whether the intended user is a healthcare professional or an individual (i.e. a non-health professional). Under new risk assessment processes an AI model for the diagnosis of melanoma would have the highest risk categorisation (Table 1) and therefore the most stringent clinical evaluation. If software is intended for use by a relevant healthcare professional, it would have a lower classification than if it was intended to be used by an individual.[26]

As of February 2021, all new SaMD applications for regulatory approval will be assessed using the new risk assessment processes. Devices that currently have regulatory approval can continue to be supplied during a transition period (until 2024) but will need to be reassessed.

Post-market regulatory oversight involves the vendor maintaining conformity, monitoring ongoing performance and safety, and informing the TGA if there are any series issues, notably adverse incidents, overseas regulatory actions, and investigation by the manufacturer (e.g. further clinical studies, review of adverse events).

ETHICAL USE OF AI

The use of AI raises ethical concerns such as undermining human autonomy, accountability, and fairness. Currently there is an overabundance of guidelines on the ethical use of AI. A recent review identified 84 published ethical guidelines.[27] The same review found not one single ethical principle was common to all guidelines.

The inconsistent definition and interpretation was a limitation identified when ethical guidelines were compared.[27] A further difficulty is the lack of pragmatic guidance on appropriate redress or solutions to potential ethical issues. For example, many ethical guidelines recommend the explainabilty. Current DL models have well over 100 million parameters and are virtually incomprehensible.[28]

14409960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hence, it is unreasonable to expect a dermatologist would be able to fully comprehend and understand such an AI model. A more pragmatic way to address ethical principle of explainabilty could be the use Explainable AI techniques (e.g. saliency maps, class activation maps) that help humans understand factors that contributed to a model's decision.

The principles of transparency, justice and fairness, non-maleficence, responsibility, and privacy were identified as being common to over half of published ethical guidelines.[27] These principles form the basis of ACD guidelines (Table 2) along with other highly influential guidance namely the joint statement on *Ethics of AI in Radiology*,[28] from the American College of Radiology, European Society of Radiology, Radiological Society of North America, Society for Imaging Informatics in Medicine, European Society of Medical Imaging Informatics, Canadian Association of Radiologists, and American Association of Physicists in Medicine.

Use and sharing of data

As previously identified, large, aggregated datasets generally improve the performance of AI models. Dermatologists can improve AI models by contributing or developing datasets for ML or DL purposes. Dermatologists are encouraged to contribute data to improve AI models as data sharing can be seen as a form of altruism.[32] The joint statement on *Ethics of AI in Radiology*,[28] states:

> As physicians, radiologists have a moral duty to use the data they collect to serve patients and improve the common good, extract more information about patients and their diseases, and improve the practice of radiology.

Data include images, labels, and metadata. The contributor should fully describe the data, including the patient population and how the ground truth label was determined so

TABLE 2 Ethical principles adopted in the ACD position statements on the use of AI in dermatology

Ethical principle	Description
Beneficence	Beneficence is the obligation of a physician to act for the benefit of the patient.[29] To this end the ACD recommends that AI be used to improve patient outcomes
Non-maleficence	A fundamental of AI ethics is it should not cause harm. To this end ACD has recommended the use of AI that has been demonstrated to be equivalent or superior in performance to clinicians, or to significantly improve clinician performance. Appropriate contextual use and continual monitoring of AI algorithms are further strategies that support non-maleficence. Ensuring patient privacy is another construct of non-maleficence. Data used to generate AI models must be handled in accordance with privacy laws. In Australia, the Privacy Act 1988 (Commonwealth) regulates Commonwealth public sector entities and private sector entities. Most Australian states have privacy legislation which apply to public sector bodies.[30] The Privacy Act is based on principles, and Australian Privacy Principle 11 states that entities must take reasonable steps to protect personal information. This means taking steps to minimise the risk to privacy (e.g. data breach)
Justice	One of the most immediate ethical concerns in the dermatology context is the lack of representation of underserved populations (e.g. skin of colour) in data used to train AI models and resultant reduction in generalisability of AI models for these underrepresented populations.[23] Two recent reviews highlight the lack of patient diversity and restricted population representation in publicly available datasets of skin images that are frequently used for AI model development.[9,31] The ACD advocates for the development and deployment of AI to improve outcomes for Aboriginal and Torres Strait Islander peoples. To this end, ACD encourages members to take images of skin of colour and to share these images (and associated labels) for the purpose of training, validating, and testing AI models
Responsibility	According to the Organisation for Economic Co-operation and Development (OECD) principles, organisations or individuals who develop, implement or adopt AI are accountable for the proper functioning of the AI system, which ultimately means the dermatologists are responsible for final diagnosis if they use AI. However, AI technology vendors also assume responsibility. In case of malpractice arising where AI was involved, it is currently unclear to what extent each involved party bears responsibility.[32]
Transparency	Transparency has broad definitions in the AAD position statement for the use of AI.[20] Firstly, transparency is used in the context of Explainable AI, specifically the AAD states they support efforts to gain transparency into the process of how augmented intelligence algorithms reach their conclusions. Secondly, it promotes transparency with patients that AI has been used in clinical practice by stating there should be transparency on how augmented intelligence technologies are utilised in their care process. Other commentators use the latter definition of transparency regarding informing the patient that their diagnosis was obtained with the help of AI.[33] The ACD has adopted the definition of transparency as informing patients of AI use in clinical practice. To this end, the ACD recommends that adopters disclose to their patients the use of AI in clinical practice

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that models developed using the data can subsequently be applied in appropriate clinical scenarios.

The ISIC host a publicly available on-line image database for the purpose of AI development and education.[34] Anyone, who agrees to the terms of use, can upload or download images to/from this public resource.

Patient perspectives

There is limited research that investigates patient perspectives on the use of their data in artificial intelligence research and development. Available research investigates perceptions on the use of deidentified data and not visually identifiable data that may occur in some types of dermatology imaging. There is significant variation in patient perception and knowledge of health data and AI, but generally patients were strongly in favour of developing AI that offer improvements in diagnosis and treatment.[35] Patients did have significant concerns about sharing their data when it may ultimately be sold to external organisations to use for profit.[36]

Consent, deidentification, and data governance

Dermatologists should obtain informed consent from their patients to use their data for ML purposes. The consent process must inform the patient of how their data will be used and of any risks to their privacy. For example, details should be specified such as whether their demographic information remains part of the dataset, and whether identifying features such as fingerprints, face, distinguishable lesions, jewellery, or tattoos are included in the dataset. In some circumstances, clinical data from non-consenting patients can be used for ML purposes, for example, Human Research Ethics Committees (HRECs) in Australia and Austria had waived written informed consent for retrospectively collected and deidentified dermoscopy images used in an AI reader study.[11] This occurs when the HREC has determined the clinical data are being used for the public good and retrospective consent is impractical.[30]

Dermatologists should assess the risk to patient privacy that may arise from data sharing. Before being used for ML purposes, data should have patient identifiers removed. This includes deidentification of visual features but only if this does not negatively affect the usefulness of the data.[37] If data are high risk (e.g. the patient remains visually identifiable) the contributor should be satisfied by the data custodian that there are stringent privacy safeguards implemented. As such, visually identifiable data should have Australian data residency and data sovereignty.

The role of data governance is a gatekeeper role to ensure what entities can ethically use the data.[38] Data governance is underpinned by data use agreements which bind data users regarding how data can and cannot be used. Data use agreements are ethically and legally fundamental for data sharing for ML purposes.[39] Dermatologists should contribute to datasets that have appropriate data governance frameworks or develop an appropriate data governance framework for newly implemented data repositories to ensure that data are used appropriately, and that patient privacy is safeguarded.

Commoditisation

In the AI era, dermatology data do have commercial value. The ethics of selling data remains contentious and it is difficult for the ACD to provide specific guidance on the commoditisation of data. This is a similar position that the Canadian Association of Radiologists have adopted.[32] Some commentators have proposed that clinical data are used to provide care and once the primary purpose for acquiring the data is fulfilled, it should be treated as a form of public good and may be used for the benefit of future patients.[40] Other commentators have suggested a health provider can recover the cost of the extraction, assembly, and curation of the dataset,[40] whilst others suggest patients should be compensated if their data are used to achieve commercial gain,[41] although how this could be administered is unclear. When data are collected specifically for research (as opposed to clinical) purposes it is generally appropriate to reimburse the costs to participants for taking part in the research.[42]

Exclusive use licences are considered unacceptable as they prevent others from using the data which are a public good asset. When receiving income for a commoditised dataset, the financial gain overrides the public good. For this reason, patients who contribute images to the dataset must have given explicit consent for the commoditisation of data.[40]

Recommendations for ACD's position paper on the use of artificial intelligence

Consistent with other professional bodies the ACD proactively encourages and supports ongoing efforts to develop AI for dermatology whilst recognising that current AI technology is not ready for clinical use or direct-to-consumer use. The ACD acknowledges that much of the current research of AI in dermatology is for skin cancer. The ACD encourages research into how AI can be used in other areas of dermatology such as inflammatory conditions. The ACD

14409960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Position statement recommendations for The College

Recommendations for The College
1. Augmenting care: The ACD supports the development of AI to enhance the practice of dermatology
2. Ethical use: The ACD recommends that dermatologists use AI ethically – this involves beneficence, non-maleficence (including ensuring patient privacy and confidentiality is maintained), transparency (informing patients that the diagnosis was augmented by AI); and utilising AI models that have attempted to reduce bias, where possible
3. Equity: The ACD supports the development of AI that can enhance outcomes for Aboriginal and Torres Strait Islander peoples
4: Collaboration: The ACD will collaborate with regulators, policymakers, industry stakeholders, and clinicians and consumer groups to help ensure development of appropriate AI policy, regulation and education to support safe and effective use of AI in clinical dermatology to improve patient care
5. Real-world evaluations: The ACD recommends that to support clinical adoption of AI, prospective, real-world evaluations are required to demonstrate that AI significantly enhances the performance of clinicians or is equivalent or superior in performance
6. Patient safety: • The ACD recommends that only AI models that have regulatory approval by the TGA using the reformed (post-25 February 2021) risk assessment (see *AI regulation* section above) model are used in clinical practice • The ACD does not endorse the use of direct-to-consumer AI models that do not have regulatory approval as a medical device • The ACD recommends that any approved AI devices should be utilised to aid the dermatologist in reaching a diagnosis; that is, AI should be used to augment, but not replace clinical judgement
7. Transparency and traceability: The ACD recommends that all artefacts of AI workflows (e.g. model output, saliency maps) be traceable and auditable through incorporation into the patient's medical record and stored for the retention period prescribed in relevant legislation

TABLE 4 Position statement recommendations for Adopters of AI

Recommendations for Adopters of AI
Dermatologists should develop knowledge and skills in the use of AI for dermatology including: • Understanding that output from AI models can produce false-positive and false-negative results • Understanding the diagnostic threshold for binary classifiers and interpretation of probabilistic outputs • Understanding of appropriate use of an AI model (i.e. the model has been trained in a similar context as the patient population on which the AI model will be used) • Understanding that their decision making may be biased by using AI
Dermatologists should select AI models that: • Fully describe how the model has been trained; and use this information to inform appropriate use • Have been internally and externally validated, and registered with the TGA • Use models based on the most appropriate architecture • Incorporate methods of Explainable AI (e.g. class activation maps)
Dermatologists should continually monitor the output of AI models in order to: • Ensure the classification or prediction is credible • Ensure performance does not degrade over time • Reduce automation bias
Dermatologists have an obligation to report to appropriate statutory bodies (including the TGA) any adverse events.

promotes the use of ethical AI and the use of AI to improve patient care. There are limitations within current regulatory, medico-legal, policy, and privacy concepts with respect to their application in AI. The ACD has a leadership role in working with stakeholders to develop, refine, and disseminate these non-technological aspects that will help ensure the safe and effective clinical use of AI in dermatology.

The ACD position statement has included guidance for adopters on technological aspects of the AI model. Currently, there is strong evidence to suggest that CNNs are the best architecture for image classifiers. However, this may not always be the case. Adopters should understand the context of how an AI model was trained so it can be used in an appropriate clinical context. The ACD recommends that adopters understand the principles of AI model development, including model architecture. To mitigate the black box effect of AI models, adopters can use models which include Explainable AI and continually monitor the AI model's output. As a means of auditing, there is a need to store all artefacts of AI workflows (e.g. AI model output, saliency maps). Adopters should be aware of the limitations of AI including: the paucity of evidence to support the clinical use of AI due to the lack of prospective trials, the fact that AI models can produce false-negative and false-positive outputs, and AI is likely to influence their decision making.

The ACD recognises limitations of risk assessment that may have previously occurred with the regulation of AI. To mitigate this the ACD recommends its members only use AI that has been assessed using the post-reform risk assessment that accords with the IMDRF risk assessment method. Further, the ACD cautions against the use of direct-to-consumer apps that are unregulated and may cause unnecessary healthcare utilisation and costs downstream. A full list of recommendations for The College, Adopters of AI, and Contributors are included in Tables 3–5, respectively.

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 Position statement recommendations for Contributors

Recommendations for Contributors
Data used to generate AI models must be handled in accordance with relevant privacy laws
Dermatologists should obtain informed consent from their patients to use their data for training, validation, and testing of AI models. In situations where explicit consent is impractical and there is minimal risk to the privacy of the patient, the data from non-consenting patients can be contributed to image repositories. This can occur when a Human Research Ethics Committee has determined the data can be treated as a form of public good. The receipt of payment for images voids public good and therefore explicit patient consent is required. Sensitive data or data where there is a risk to patient privacy should only be contributed to repositories with Australian data residency and data sovereignty
Dermatologists are encouraged contribute images, image labels, and metadata to existing data repositories with patient consent, and/or develop data repositories for training, validation, and testing of AI models
Dermatologists should assess the risk to patient privacy from data contribution
Dermatologists should contribute to datasets or repositories where the data custodian uses appropriate data governance
Dermatologists should fully describe the data so that models developed using the data can subsequently be used in appropriate clinical scenarios
Dermatologists should not seek payment for data, other than to cover the costs of extraction, assembly, and curation of the data
Dermatologists should avoid entering into exclusive-use licences for contributed data

ACKNOWLEDGEMENTS

We acknowledge the assistance of the Australasian College of Dermatologists for their assistance in co-ordinating the study and reviewing the manuscript.

FUNDING INFORMATION

This study was funded by the Australasian College of Dermatologists. HPS holds an NHMRC MRFF Next Generation Clinical Researchers Program Practitioner Fellowship (APP1137127).

CONFLICT OF INTEREST

HPS is a shareholder of MoleMap NZ Limited and e-derm consult GmbH, and undertakes regular teledermatological reporting for both companies. HPS is a Medical Consultant for Canfield Scientific Inc, MoleMap Australia Pty Ltd, Blaze Bioscience Inc, and a Medical Advisor for First Derm. VM's institution has received clinical trial funding from MoleMap. HPS, SS, and PF-P are current AJD Editorial Board Members.

ORCID

Liam J. Caffery https://orcid.org/0000-0003-1899-7534
Monika Janda https://orcid.org/0000-0002-1728-8085
Lisa M. Abbott https://orcid.org/0000-0002-4005-8421
Pablo Fernández-Peñas https://orcid.org/0000-0003-4882-1564
Victoria Mar https://orcid.org/0000-0001-9423-3435
H. Peter Soyer https://orcid.org/0000-0002-4770-561X

REFERENCES

1. Du-Harpur X, Watt FM, Luscombe NM, Lynch MD. What is AI? Applications of artificial intelligence to dermatology. Br J Dermatol. 2020;183(3):423–30.
2. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. J Dermatolog Treat. 2020;31(5):496–510.
3. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. Diagnostics. 2021;11(8):1390.
4. Murphree DH, Puri P, Shamim H, Bezalel SA, Drage LA, Wang M, et al. Deep learning for dermatologists: Part I. Fundamental concepts. J Am Acad Dermatol. 2020 May 17:S0190-9622(20)30921-X.
5. Weikert T, Francone M, Abbara S, Baessler B, Choi BW, Gutberlet M, et al. Machine learning in cardiovascular radiology: ESCR position statement on design requirements, quality assessment, current applications, opportunities, and challenges. Eur Radiol. 2021;31(6):3909–22.
6. 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.
7. Navarrete-Dechent C, Dusza SW, Liopyris K, Marghoob AA, Halpern AC, Marchetti MA. Automated dermatological diagnosis: hype or reality? J Invest Dermatol. 2018;138(10):2277–9.
8. 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(31):eabq6147.
9. Wen D, Khan SM, Ji XA, Ibrahim H, Smith L, Caballero J, et al. Characteristics of publicly available skin cancer image datasets: a systematic review. Lancet Digit Health. 2022;4(1):e64–74.
10. Babic B, Gerke S, Evgeniou T, Cohen IG. Direct-to-consumer medical machine learning and artificial intelligence applications. Nat Mach Intell. 2021;3(4):283–7.
11. 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.
12. Pangti R, Mathur J, Chouhan V, Kumar S, Rajput L, Shah S, et al. A machine learning-based, decision support, mobile phone application for diagnosis of common dermatological diseases. J Eur Acad Dermatol Venereol. 2021;35(2):536–45.
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

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

melanoma recognition in comparison to 58 dermatologists. Ann Oncol. 2018;29(8):1836–42.
14. 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.
15. Liu XX, Rivera SC, Faes L, Keane PA, Moher D, Calvert M, et al. CONSORT-AI and SPIRIT-AI: new reporting guidelines for clinical trials and trial protocols for artificial intelligence interventions. Invest Ophth Vis Sci. 2020;61(7):1617.
16. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28(3):231–7.
17. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. Nat Med. 2020;26(8):1229.
18. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc. 2012;19(1):121–7.
19. Freeman K, Dinnes J, Chuchu N, Takwoingi Y, Bayliss SE, Matin RN, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ. 2020;368:m127.
20. Position Statement on Augmented Intelligence (AuI). 2019. Available at: https://server.aad.org/Forms/Policies/Uploads/PS/PS-Augmented%20Intelligence.pdf?
21. Position Statement on Artifical Intelligence (AI) Interventions. 2021. Available at: https://www.bad.org.uk/library-media/documents/AI%20Position%20Statement%20Final%2001.07.21.pdf
22. Moshi MR, Parsons J, Tooher R, Merlin T. Evaluation of mobile health applications: is regulatory policy up to the challenge? Int J Technol Assess. 2019;35(4):351–60.
23. Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial intelligence in dermatology: a primer. J Invest Dermatol. 2020;140(8):1504–12.
24. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med. 2021;27(4):582–4.
25. Benjamens S, Dhunnoo P, Mesko B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ Digit Med. 2020;3:3118.
26. How the TGA regulates sofwared based medical devices. 2021. Available at: https://www.tga.gov.au/sites/default/files/how-tga-regulates-software-based-medical-devices.pdf
27. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1(9):389–99.
28. Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of artificial intelligence in radiology: summary of the joint european and north american multisociety statement. Radiology. 2019;293(2):436–40.
29. Varkey B. Principles of clinical ethics and their application to practice. Med Princ Pract. 2021;30(1):17–28.
30. Paterson M, Witzleb N. The privacy-related challenges facing medical research in an era of big data analytics: a critical analysis of Australian legal and regulatory frameworks. J Law Med. 2018;26(1):188–203.
31. 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.
32. Jaremko JL, Azar M, Bromwich R, Lum A, Alicia Cheong LH, Gibert M, et al. Canadian Association of Radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. Can Assoc Radiol J. 2019;70(2):107–18.
33. Neri E, Coppola F, Miele V, Bibbolino C, Grassi R. Artificial intelligence: who is responsible for the diagnosis? Radiol Med. 2020;125(6):517–21.
34. International Skin Imaging Collaboration Archive. Available at: https://www.isic-archive.com/
35. Aggarwal R, Farag S, Martin G, Ashrafian H, Darzi A. Patient perceptions on data sharing and applying artificial intelligence to health care data: cross-sectional survey. J Med Internet Res. 2021;23(8):e26162.
36. Balthazar P, Harri P, Prater A, Safdar NM. Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. J Am Coll Radiol. 2018;15(3 Pt B):580–6.
37. De-identification decision-making framework 2017. Available at: https://publications.csiro.au/rpr/download?pid=csiro:EP173122&dsid=DS3
38. Janssen M, Brous P, Estevez E, Barbosa LS, Janowski T. Data governance: organizing data for trustworthy artificial intelligence. Government Information Quarterly. 2020;37(3):101493.
39. Banja J, Rousselle R, Duszak R Jr, Safdar N, Alessio AM. Sharing and selling images: ethical and regulatory considerations for radiologists. J Am Coll Radiol. 2021;18(2):298–304.
40. Larson DB, Magnus DC, Lungren MP, Shah NH, Langlotz CP. Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. Radiology. 2020;295(3):675–82.
41. Mitchell V. What if the companies that profit from your data had to pay you? 2018. Accessed 15 November 2022. Available at: https://theconversation.com/what-if-the-companies-that-profit-from-your-data-had-to-pay-you-100380
42. National statement on ethical conduct in human research 2007 (Updated 2018) 2018. Available at: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Caffery LJ, Janda M, Miller R, Abbott LM, Arnold C, Caccetta T, et al. Informing a position statement on the use of artificial intelligence in dermatology in Australia. Australas J Dermatol. 2023;64:e11–e20. https://doi.org/10.1111/ajd.13946

14400960, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajd.13946 by NHMRC National Cochrane Australia, Wiley Online Library on [17/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License