Volume **8** (2) 2026 https://doi.org/10.5204/lthj.4233

Direct-to-Consumer AI Health Services: Precision Healthcare Needs Precision Consent

Hui Chia
The University of Melbourne, Australia
Daniel Beck
RMIT University, Australia
Jeannie Marie Paterson
The University of Melbourne, Australia
Julian Savulescu
University of Oxford, United Kingdom

Abstract

Artificial Intelligence (AI) technology has the potential to replace a visit to the doctor, with AI companies offering health services directly to the public that, until recently, could only be performed by humans. AI apps that claim to detect disease, such as skin cancer, envisage a hopeful future where everyone can access expert medical care on their phone, at a fraction of the cost of traditional healthcare. However, the advent of AI in medical contexts also raises anxiety that AI may not be as reliable as claimed, and that laypersons are ill-equipped to make these decisions by themselves.

AI companies have a duty of care to provide their consumers with adequate notice of the risks and limitations of the AI system. But given the complex and technical nature of this information, how can such notice be made salient for consumers? The main argument of this article is that, if AI companies effectively offer a health service traditionally performed by doctors, then they should be guided by the values that govern the doctor-patient relationship regarding the information provided to consumers about the limitations of the AI system.

We propose Precision Consent, an interdisciplinary framework that draws connections between the legal and ethical duties of doctors towards their patients, and capabilities in the field of AI. We discuss how values that guide doctors, such as respect for patient autonomy, non-maleficence and personalised warnings, can be incorporated into how notice is provided to consumers regarding the accuracy of AI health services.

***Keywords**:* Artificial intelligence; healthcare; legal liability; consumer protection.

Introduction

The emergence of medical AI with capabilities that are said to match or even surpass human medical experts has resulted in services marketed direct-to-consumers that potentially replace a visit to the doctor.[1] This disrupts the traditional practice in which medical services are provided within the confines of a doctor-patient relationship. Consumers may monitor their heart health through an AI-powered app,[2] seek mental health support through a chatbot therapist,[3] or investigate concerns about their skin through an app using computer vision.[4]

[1] Cohen, "Direct-to-Consumer Digital Health."
[2] Heartscan, "Heart Rate Monitor."
[3] Haque, "Overview of Chatbot-Based Mobile Mental Health Apps."
[4] SkinVision, "Accurate Skin Cancer Detection Made Simple."

Direct -o-consumer medical AI disrupts the traditional practice that medical services are provided within the confines of a doctor-patient relationship. Instead, the service is provided directly to the consumer. Removing the medical practitioner from medical AI diagnosis introduces risks for consumers. Consumers are laypeople who lack expert medical knowledge and may not adequately understand the risks and potential errors of AI technology. When AI service providers take over medical tasks from doctors, such as diagnosis and advice, what obligations do AI service providers have in helping consumers understand the medical decisions that flow from that diagnosis and the limitations of the AI technology used to make the diagnosis?

The provider of an AI health service will be subject to a duty to use reasonable care arising in tort, contract and under legislation.[5] These duties go to the care with which the service is developed and offered. They also apply to the information provided to consumers about the scope of the service and the risks inherent in it, which is the focus of this article. Where AI automates the diagnostic role of a doctor, the AI service provider must use care and skill in ensuring the performance of the AI is consistent with what consumers would reasonably expect. However, no service, AI or human, will be perfect, or completely risk free. The diagnosis will have a margin of error and may be less accurate in its application to some cohorts. The AI service should reach a reasonable standard of accuracy to be made available on the market. Consumers should be warned about residual risks and provided with information about how best to use the service. Where this is done, consumers are said to have consented to use of the service and its associated limitations.

We propose the concept of 'Precision Consent' using AI technology to perform the task of obtaining patient consent in a way that not only satisfies the legal obligations of the AI provider in using reasonable care and skill in providing the service and informing consumers about possible risks, but also meets the ethical obligations that should arise in offering health care. We discuss how this kind of care in informing the consumer about the service should not be viewed as a barrier to the development of medical AI services. In fact, there is exciting potential to develop AI technology capable of performing functions similar to what doctors traditionally do for their patients – personalised warnings, respecting patient autonomy, and practicing non-maleficence – and this is Precision Consent.

The value that Precision Consent contributes to the area of legal regulation of emerging AI technologies is that we offer a framework for determining a standard of care for AI service providers in providing notice and obtaining consent about the scope of their service that is legally compliant, ethical compliant and technologically feasible. This is significant in the context of healthcare, where there are compelling reasons to encourage AI innovation that would address the shortage of health resources.

Part 1 of this article will outline the legal obligations of an AI service provider when they seek to deliver health services direct-to-consumers, including the information provided about the service, its risks and its limitations. In Part 2, we look at a case study of Google's DermAssist app, to illustrate the challenges that an AI service provider faces when attempting to offer a health service direct to consumers, in particular the difficulty with communicating limitations of the AI service. Part 3 delves into further detail about the concept of 'accuracy' in AI, and why it is a far more complicated concept than laypersons can be expected to understand on their own. Finally, in Part 4 we explain how techniques in AI can be used by AI service providers to provide notice about the limitations in the AI service that are salient for the individual consumer and therefore meet the providers' legal and ethical duties.

Part 1. Legal Obligations of an AI Service Provider

Developments in AI technology have brought a surge in AI systems marketed direct-to-consumers for medical or health related purposes. Some examples include 'mySugr,' an app which helps people with diabetes manage their blood sugar levels including calculating their dosage of insulin,[6] 'SkinVision,' a smartphone app that claims to detect skin cancer using photos from a consumer's phone,[7] and 'Ada,' a text-based symptom checker that allows consumers to input their symptoms and receive a suggested diagnosis.[8]

There have been legislative responses that aim to regulate transparency regarding the use of AI technology in healthcare. For example, the state of California has passed laws requiring that healthcare providers disclose to patients where AI has been used to generate a clinical communication to the patient,[9] and the providers of AI healthcare tools are prohibited from falsely

[5] *Competition and Consumer Act 2010* (Cth) sch 2, s 60.
[6] mySugr, "The mySugr App."
[7] SkinVision, "Accurate Skin Cancer Detection Made Simple."
[8] Ada, "More than a Symptom Checker."
[9] *AB 3030 Health Care Services: Artificial Intelligence*, 2023-2024 Leg. Sess. (Cal. 2024), s 1339.75.

indicating that output generated from an AI system has been created by a human doctor.[10] The concern here is that there must be transparency for patients about whether the healthcare they are receiving has come from a human doctor or an AI system.

That is a related but distinct concern to the scenario that is the focus of this article. Our discussion concerns direct-to-consumer AI services where it is clearly communicated that the health service is from an AI system. The question is whether the average consumer will understand what that really means, given the complexity of the risks of AI technology. The providers of these direct-to-consumer AI health services have a duty to inform consumers about the limitations of their AI service, but what would that involve practically?

1.1 Duties of Care in Providing the Service and in Providing Information about the Service

There are two main elements of the service being provided:

(i) The health service, i.e., how well the AI performs, and
(ii) Information provided about the service, i.e., claims about how well the AI performs.

This article focuses on the problem of where there is a mismatch between how well consumers might reasonably expect the AI to perform versus how well the AI system actually performs. To satisfy its duty of care, the provider of the AI service will need to provide information to the consumer about how to use the service, the scope of the service (i.e., what it can reasonably be used to diagnose), and limitations on its accuracy (generally and for specific cohorts). Reasonable care in providing information about the AI service should provide consumers with a sufficient understanding of the of the AI system to make an informed choice about whether they want to rely on it.

1.2 The Significance of Consent

In the context of healthcare, the risk of harm of a new technology or service must be weighed against the harm of no service at all. Health resources are expensive and scarce. There are not enough qualified human medical practitioners to meet the health needs of the public.[11] If AI technology can potentially perform some of the work that human medical practitioners perform, this could greatly benefit society. Thus, the risks of AI health services should be considered in the context that the alternative might be no health service at all.

It is not inherently a problem for AI providers to offer consumers a lower level of health service than human doctors, if the consumer adequately understands *and consents* to a lower level of service, because the AI is more accessible and affordable. The problem arises when a person decides to rely on an AI service instead of seeking a human medical practitioner, based on an *incorrect* understanding of how well the AI performs.

Traditionally, the types of services that could be delivered direct-to-consumers in the health area were relatively low risk, for example exercise and diet tracking apps. Any service that potentially carried greater risks if incorrect, such as disease diagnosis, would need to be performed by a doctor. There was no technology capable of performing tasks like diagnosis direct-to-consumers. As such, medical professionals acted as gatekeepers to health interventions that carried higher levels of risks.[12] With AI technology it is possible to automate more complex services, such as diagnosis, which can cause great harm if the consumer relies on an incorrect response. This exposes consumers to a higher level of risk than would have occurred prior to AI technology.

An incorrect understanding may have catastrophic consequences. If a consumer relies on an incorrect diagnosis from an AI medical service, this could cause grave harm by delaying the discovery of disease and medical treatment. The more serious the risk of harm, the greater the precautions required in ensuring the consumer understands the relevant risks.[13]

1.2.1 Salient Information

We argue that for an AI service provider to meet their duty of reasonable care, they need to provide information such that their target consumer will have an adequate understanding of the limitations of the AI, and will be able to validly consent to that level of service. If representations made by the AI service provider give their target audience an overly high expectation of how

[10] *AB 489 Health Care Professions: Deceptive Terms or Letters: Artificial Intelligence*, 2025-2026 Reg. Sess. (Cal. 2025), s 4999.9.
[11] Mathieu, "Global Health Workforce Stock and Distribution in 2020 and 2030."
[12] Husgen, "Product Liability Suits Involving Drug or Device Manufacturers and Physicians."
[13] Discussed in the case *Wyong Shire Council v Shirt* [1980] HCA 12 [47]-[48], that the reasonable person's response calls for a consideration of the magnitude of the risk and degree of the probability of its occurrence.

well the AI system performs, then this would breach their duty of care. Moreover, there is a challenge in making the information provided salient to consumers. This issue goes to the complexity of the information in question and the lack of expertise of most consumers.

1.2.2 Complexity of Information

Another factor to be considered with direct-to-consumer AI health services is that laypersons are being expected to understand highly complex information and concepts about a new technology that the average person has little knowledge about. As we will expand upon in Part 3 of this article, the 'accuracy' of an AI system can be technically correct but misleading in giving consumers the impression that the AI performs better than it actually does. The AI system can have a high accuracy as technically defined, but actually be unreliable in its real-world performance.

We highlight the fact that historically, disadvantaged groups are particularly likely to encounter the problem that the AI system's stated accuracy rate, while technically correct, is not reliable for them as an individual, because they have been under-represented in the training data of the AI.[14] This factor needs to be considered when analysing the standard of care for AI service providers concerning how claims about the 'accuracy' of AI systems are presented to lay consumers.

1.2.3 Consumer Inexperience

A significant point about the direct-to-consumer context is that laypeople are making health decisions without a human doctor acting as an intermediary. One of the roles of doctors is to explain the risks associated with a diagnosis in a way their patients can understand. Where a doctor is not involved and the medical service is provided directly to consumers, then this kind of information may be hidden in fine print terms. Consumers frequently do not read, or make sense of, fine print disclosures because they do not have time to invest in them or lack the technical understanding to act on the information provided.[15] To be more impactful, we suggest information about the service and warning about its limitations should be personalised to the individual consumer. Instead of overwhelming consumers with a long list of every possible risk, which is recognised can be counter-productive to consumer understanding,[16] it is likely to be far more effective to advise consumers about which risks apply to them specifically.

Part 2. Case Study: Google's DermAssist

We examine a case study of an attempt to develop an AI health service direct-to-consumers, which illustrates the significant challenges that AI service providers face regarding how to properly communicate the limitations of the service. The service in question was Google's DermAssist app, an app intended for identification of skin diseases. This was initially offered to the public as a testing phase product in 2021,[17] with Google ceasing to offer the service in 2024. We selected this example as an interesting case study, firstly because the application of AI and smartphone technology for skin disease detection is an area of keen research interest;[18] and secondly, we chose to focus on the Google app because of the significance of Google as a globally recognised brand in technology; and thirdly, detailed information about the app's design and development was publicly available in an academic journal.

In this section of the article, we detail how Google originally marketed the DermAssist app to the public in 2021, and the problems with how it was presented to consumers when first launched. We observe that Google faced the following challenges in offering an AI health service direct-to-consumers:

> (i) Consumers may treat the service as a medical diagnosis, despite warnings that the service is not meant to substitute a diagnosis from a medical practitioner;
> (ii) The service was open to criticisms about inequity due to the lack of representation of darker skin tones in the training data; and

[14] California Attorney General, "Legal Advisory on Artificial Intelligence in Healthcare," 4, states that an AI system may be discriminatory if it makes less accurate predictions about demographic groups whose information is under-represented in datasets.
[15] See discussions about what should reasonably be expected from consumers in the modern context of the vast volume of standard form contracts that consumers are confronted with, in Ayres, "The No-Reading Problem in Consumer Contract Law," and Adar, "Ending the License to Exploit."
[16] Discussed in *Cotton v. Buckeye Gas Products Company,* 840 F.2d 935 (D.C. Cir 1988), that the problem of including too many warnings is that each extra item dilutes the punch of every other item; they get lost in the fine print.
[17] Bui, "Using AI to Help Find Answers to Common Skin Conditions."
[18] Sangers, "Validation of AI Mobile Health App," Ouellette, "Usefulness of Smartphones in Dermatology," and Smak Gregoor, "Artificial Intelligence Based App for Skin Cancer Detection."

(iii) The service was developed using data collected from a healthcare setting, which is different to data the app would encounter in the real world direct-to-consumer setting.

2.1 The Limitations of Warnings in the Consumer Context

Google's DermAssist app was launched in 2021 as a service claiming to assist users with identifying medical skin conditions, by analysing photos of their skin with their smartphone camera. The DermAssist website included a clearly visible warning that the app's results do not constitute a medical diagnosis, that it is instead a 'search journey.'[19]

However, there was the risk that some consumers would treat the search results as a diagnosis, despite the warning that they should not do so. Consumers may place too much reliance on AI technology, as indicated in a pilot study where participants were offered the use of an AI-based mobile phone app before consulting their GP, which found that 54% of patients with a benign skin lesion and low risk rating would be reassured and cancel their GP visit.[20]

Regarding the overall impression that a consumer was likely to take from the DermAssist website as it was first launched, whilst there were warnings that DermAssist was not intended to replace a medical diagnosis, there was the conflicting impression that it could.

The DermAssist website provides a clearly visible warning:

> DermAssist is intended for informational purposes only and does not provide a medical diagnosis.

The website further states that DermAssist is 'designed for use by people not seeking a diagnosis.' Given that the warnings were placed prominently on the main website, it would be reasonable to state that the average consumer should be expected to have read this warning.

However, there were conflicting statements on the website that may have given the impression that DermAssist could be used as a diagnosis:

> A whole new way to help identify your skin conditions.
>
> From your phone or computer, upload three photos of your skin condition and answer a few questions. Using what it has learned from millions of skin-related images, DermAssist then looks for signs of various skin conditions in your submitted photos and information.

Claims made on the website about the accuracy of DermAssist suggested that its accuracy was as good or better than a human doctor:

> DermAssist is the culmination of years of machine learning research, dermatologist-reviewed content, user testing, and product development.
>
> Trained using millions of skin images, DermAssist can identify more than 90 percent of the most commonly searched-for skin conditions, and research demonstrates that the underlying technology can help clinicians better identify skin conditions across all populations.

Another claim made on the DermAssist website that could be misunderstood by consumers is the claim that the DermAssist tool had been validated and certified by external bodies.

The DermAssist website stated:

> DermAssist is validated through testing.

[19] The original website for DermAssist by Google, "Identify Skin Conditions with DermAssist - Google Health," has been removed since Google ceased to offer the tool, but can still be viewed using the website archive WayBack Machine. To view the original DermAssist website go to https://web.archive.org/, then enter the URL https://health.google/consumers/dermassist/, and view examples of the website archived between 2022 and 2023.
[20] Smak Gregoor, "Artificial Intelligence in Mobile Health."

> DermAssist is CE-marked as a Class 1 Medical Device in the EU. We work with diverse partners and clinicians to guide our research and product development, review the data collected, validate product performance, and comply with regulatory requirements.

These statements might give the impression that the DermAssist app has been independently validated by external bodies when, in fact, the requirement for CE marking for Class 1 Medical Devices is self-certification.[21] The CE marking simply means that Google self-declares that it complies with EU regulation – there is no external validation by any independent body.

In addition to statements on the DermAssist website, there was also an announcement by Google introducing the DermAssist app that made even stronger claims about what DermAssist could do:

> Our landmark study, featured in Nature Medicine, debuted our deep learning approach to assessing skin diseases and showed that our AI system can achieve accuracy that is on par with U.S. board-certified dermatologists.[22]

The subsequent withdrawal of the DermAssist app in 2024 illustrates how it is very difficult for AI service providers to give effective warnings about health services in the direct-to-consumer context.[23] Despite Google's display of clearly visible warnings that the tool was not intended to replace diagnosis by a medical professional, we contend that it was still potentially below the standard of care in communicating the risks of the app to laypersons, given the conflicting statements, and the difficulty for laypersons to properly understand AI 'accuracy.'

The end result however has been a setback for health innovation, which is that Google has ceased to develop the DermAssist tool and it is no longer offered to the public. With the serious shortage of dermatology specialists to meet public needs,[24] this is not the ideal outcome. It is a loss for society if AI service providers find it too risky to explore AI innovations that could potentially provide medical benefits, because it is unclear how they could meet their legal obligations to warn consumers about the limitations of their service.

2.2 Under-Representation in Training Data

Another challenge for the development of AI health services is that how well an AI system performs greatly depends upon whether the AI system has been trained on data that is representative of prospective users. This issue is of particular significance in the context of skin health conditions. The visual appearance of the same disease can look very different on different skin tones. Multiple studies have identified that AI diagnosis of skin conditions have lower accuracy for dark skin tones compared to light skin tones, due to biased training data.[25]

The DermAssist app was an example of this problem, where training data was not representative of the general population. Following its launch in 2021, it was criticised for the lack of representation of darker skin tones in its training data.[26] In the published research article detailing how the DermAssist AI system was developed and trained, the researchers acknowledged a lack of diversity in the training data as a shortcoming of the AI system:[27]

> Importantly, our validation data were limited with respect to the uncommon skin types (0.2% type I, 2.7% type V and 0% type VI); further validation on these skin types will be needed to complement the race and ethnicity analysis.[28]

[21] *Regulation (EU) 2017/745 of the European Parliament and of the Council*, *of 5 April 2017 on Medical Devices,* Article 52 (7) and Annex VIII (4.1).
[22] Bui, "Using AI to Help Find Answers to Common Skin Conditions." (This claim by Google is referring to the study published in Nature detailing the AI system behind DermAssist, see Liu, "Deep Learning System for Differential Diagnosis of Skin Diseases." As far as we know there have not been any studies disproving the claim.)
[23] We note that Google did not publicly state a reason for withdrawal of the app.
[24] Duniphin, "Limited Access to Dermatology Specialty Care."
[25] Daneshjou, "Lack of Transparency and Potential Bias," and Adamson, "Machine Learning and Health Care Disparities."
[26] Khatun, "Technology Innovation to Reduce Health Inequality" and Savulescu, "Ethics of Artificial Intelligence."
[27] Liu, "Deep Learning System for Differential Diagnosis."
[28] Liu, "Deep Learning System for Differential Diagnosis," 907.

Table 1. Skin Type Representation in Training Data of DermAssist[29]

Skin type (Fitzpatrick scale)	Training data Number of Patients	Test data Number of Patients
Type I - Pale white skin	46 (0.3%)	9 (0.2%)
Type II - Fair skin	2,807 (17.4%)	383 (10.2%)
Type III - Darker white skin	6,641 (41.2%)	2,412 (64.2%)
Type IV - Light brown skin	5,040 (31.3%)	724 (19.3%)
Type V - Brown skin	510 (3.2%)	101 (2.7%)
Type VI - Dark brown or black skin	46 (0.3%)	1 (0.0%)
Undetermined	1,024 (10.2%)	126 (3.4%)

The table above summarises the proportions of different skin tones in the data that was used to train the DermAssist model. The training data contained a very low proportion of brown and black skin tones, 3.2% and 0.3 % respectively. The testing data, which is the dataset used to measure the accuracy of the AI model, is even less representative, with brown skin tones constituting 2.7% of the test dataset and black skin tones accounting for 0% of the test dataset. To see how this bias in the training data impacts the accuracy of the AI system for different skin tones, we looked at the supplementary information accompanying the article.

Table 2. Performance of DermAssist by Skin Type[30]

Skin type (Fitzpatrick scale)	Top-1 Accuracy	Top-3 Accuracy
Type I - Pale white skin	0.78 [0.56, 1.00]	0.89 [0.67, 1.00]
Type II - Fair skin	0.83 [0.79, 0.87]	0.97 [0.96, 0.99]
Type III - Darker white skin	0.82 [0.81, 0.84]	0.98 [0.97, 0.99]
Type IV - Light brown skin	0.82 [0.79, 0.85]	0.97[0.96, 0.98]
Type V - Brown skin	0.84 [0.77, 0.91]	0.98 [0.95, 1.00]
Type VI - Dark brown or black skin	1.00*	1.00*
Undetermined	0.76 [0.68, 0.84]	0.97 [0.93, 1.00]

* There was only 1 case labelled as Type VI, so confidence intervals were not meaningful.

An important point to observe is that the accuracy between the different skin tones does not vary that much, varying between 0.78 - 0.84 for the range of light to brown skin tones. But what does vary significantly is the confidence interval, which is the range of values that the accuracy could truly be. This means that even if the accuracy for brown skin tones is calculated to be 84%, the true accuracy of the AI system could actually be lower at 77%. The margin for error is greater for the skin tones that were less represented in training data. For the darkest skin tone type 6, the table quotes an accuracy of 100%, but that is because there was only one dark skin tone sample in the test set, and the model got that one example correct. There is no meaningful way to ascertain confidence in the accuracy when there is only one example.

The complexity of information about the DermAssist AI system that we have presented above, coupled with the fact that we had to read a scientific journal article and its supplementary materials to obtain this information, demonstrates that it is very challenging for an AI service provider to communicate the limitations of an AI system in the consumer context. A proper understanding of the limitations of an AI system requires knowledge about AI technology, and an investment of time to read the specifics of that particular AI service, both of which are unrealistic for the average consumer. Even if Google had disclosed all of the accuracy rates as shown in the tables above, this could still be misleading to consumers, because they do not have the

[29] Liu, "Deep Learning System for Differential Diagnosis," 902, Table 1. The training and test data was split temporally. The researchers had data from 2010 to 2018, the dataset was split 2010-2017 for training, and 2017-2018 for testing, 80%/20% split, 901.
[30] Liu, "Deep Learning System for Differential Diagnosis," 36 – 37, Supplementary Information, Table 13. The model gave a list of most likely diagnosis in descending probability. Top-1 accuracy is whether the model was correct for its top diagnosis, Top-3 accuracy means the correct diagnosis was within the model's top-3 answers. Numbers in square braces indicate 95% confidence intervals.

necessary knowledge about AI to understand that the accuracy is *uncertain*. In Part 4 of the article, we discuss how Precision Consent could address some of these concerns by providing safeguards for consumers when the uncertainty of the AI is too high.

2.3 Accuracy is Specific to Testing Data

Another risk with data-based AI systems is that the AI system's accuracy can be brittle, in the sense that small shifts in the type of data inputs to the AI system can have a large impact on the accuracy rate.[31] The AI system's accuracy must be understood as being applicable only to the specific type of data on which the system was trained and tested. The accuracy rate cannot be assumed to be the same for a different type of data, even if the differences in the data might seem small. This is called a 'domain shift' in machine learning.[32] Importantly, both the accuracy rates and the confidence interval for that accuracy are at higher risk to be wrong when there is a domain shift in data.[33]

The DermAssist app is an example of how seemingly small differences in the real-world usage, compared to how training data was collected, can create risks for consumers. The accuracy rates calculated at testing phase of the app may not be correct in its real-world application, because of the shift in type of data. The original goal of the DermAssist AI system was that it would be used by medical practitioners as a tool to assist them in making a diagnosis. It was not designed with the goal of being a direct-to-consumer service.[34] As such, the training data in the original research project was from a primary healthcare setting,[35] where primary healthcare providers (GPs) who were not sure about a diagnosis for a suspected skin care condition, would take photographs and send the photographs to a dermatologist, and the dermatologist's diagnosis was taken to be the correct diagnosis.

This difference in intended usage of the AI model means that the training data for the AI model is different to the data that the DermAssist app would receive from consumers. For instance, photographs that laypeople have taken of themselves in their bedroom will not be of the same quality as photographs taken by a doctor or nurse of patients in a clinic setting. This shift of visual features of the data from clinical setting done by health professionals, to do-it-yourself by laypersons, means that the accuracy rates calculated based on data from the clinical setting might not be accurate for consumers using the AI system in its real-world application.[36]

This represents another challenge for the development of direct-to-consumer AI medical services. Data that is available for training AI systems is likely to have come from healthcare settings, not the consumer context. This also reveals another layer of complexity in understanding the 'accuracy' of AI systems – that the accuracy rates are brittle to the specific type of data on which the AI system was trained and tested.

Part 3. Understanding 'Accuracy'

This section of the article will delve further into the concept of AI 'accuracy' as it is defined in machine learning. The purpose of analysing AI accuracy in greater depth is to appreciate the complexity of the risks that laypersons are being expected to understand. Accuracy is often used as a simple measurement to convey how reliable an AI model is, or how well the AI can be expected to perform. 'Accuracy' as it is defined technically is the number of correct predictions divided by the total number of predictions, i.e., the percentage of correct predictions.[37] Most consumers would have a general perception that the higher the accuracy of an AI model, the more likely that the predictions of the model will be correct. However, what 'accuracy' really means is much more nuanced than the average consumer is likely to be aware of. Consider the example below, which shows the performance of an AI prediction model with an accuracy of 94.6%.

[31] Zhou, "Domain Generalization: A Survey."
[32] Guo, "Approaches to Preserve Machine Learning Performance."
[33] Ulmer, "Trust Issues."
[34] Liu, "Deep Learning System for Differential Diagnosis," 903, 904 and 906.
[35] Liu, "Deep learning System for Differential Diagnosis," 901 and 903, Table 1.
[36] Kilim, "Physical Imaging Parameter Variation."
[37] Zheng, Evaluating Machine Learning Models, 8.

Table 3. Example Error Matrix[38]

True Positive: 1	False Positive: 1
False Negative: 897	True Negative: 15737

Accuracy = True Positives + True Negatives (1 + 15737)
Total no. of predictions (1 + 1 + 897 + 15737)
= 0.946

An accuracy of 94.6% seems impressive. Most consumers would get the impression that this model is very accurate and that they could rely upon the predictions of the model. But accuracy only measures whether predictions are correct or incorrect, without taking into account that some errors are more significant than others. When we look at the model's predictions in more detail, we see that the number of True Positives is 1, and the number of False Negatives is 897. This means that for detection of Positives, i.e., of the 898 Positive input samples, the model only correctly predicted Positive for 1 out of 898.

If an AI model for disease detection had this performance matrix, this would mean that of the 898 people who had the disease, the model would only correctly predict disease for one person, and for the other 897 people who had the disease, it would wrongly predict that the person does not have the disease. This illustrates how the 'accuracy' of a model can give a misleading impression about how reliable the model really is. The model can have a high accuracy because of the high number of True Negatives, but the model has low True Positives, which is often the more consequential class that we want to correctly predict. An important concept in machine learning that can give users a better understanding of model performance is uncertainty estimation (UE).[39] Uncertainty estimation is a measurement of how confident the AI model is that its prediction is correct.[40] Consider for instance an AI model that is trained to make a diagnostic prediction based on an image of whether a lesion is cancerous or not cancerous. A model that does not have UE will simply output a prediction, i.e,. cancer or not cancer. A model with UE will give a prediction, and give its confidence in that prediction, e.g., cancer with X% confidence.[41]

Uncertainty Estimation gives the user more information about the reliability of the AI model, as the user receives not only the model's prediction, but also a measure of how confident the model is that its prediction is correct. Importantly, UE enables further techniques that address some of the concerns regarding how reliable the predictions of a model really are.

We will discuss three of these techniques to illustrate how an AI model designed for the purpose of disease detection could be improved, both in terms of actual performance and in terms of communicating the reliability of the model: Model Calibration, Cost Sensitive Classification, and Reject Option.

3.1 Model Calibration

Model Calibration is any transformation of Uncertainty Estimation (UE) to make the confidence estimate more likely to be correct, or to alert the user as to when the model is likely to be giving an incorrect confidence estimate.[42] An important application of Model Calibration is detecting Out-Of-Distribution input data,[43] also called domain shift as discussed earlier. This is where the model can alert the user that the input data has been under-represented in the model's training, and that the model has had no or very few similar examples.[44] Therefore, not only is the model's prediction more likely to be wrong, the model's confidence in its prediction is also more likely to be wrong.

[38] Esposito, "GHOST: Adjusting the Decision Threshold," Figure 1. Note that this AI prediction model was not a health application, but the figure is useful to illustrate how the math works, i.e., that an AI model can have high accuracy but perform poorly for detecting True Positives. True Positive and True Negative means the AI correctly predicted whether the sample was in the Positive class or Negative class. False Negative means the AI predicted Negative, when the sample was actually Positive. False Positive means the AI predicted Positive when the sample was actually Negative.
[39] Bhatt, "Uncertainty as a Form of Transparency."
[40] Abdar, "Review of Uncertainty Quantification."
[41] Dolezal, "Uncertainty-Informed Deep Learning Models."
[42] Levi, "Evaluating and Calibrating Uncertainty Prediction."
[43] Yang, "Generalized Out-of-Distribution Detection."
[44] Mehrtash, "Confidence Calibration."

This scenario poses the greatest risk of harm to consumers, where the AI model makes an incorrect prediction, and the model is also over-confident in its prediction. For example, where a model has been trained to detect skin disease mostly on fair skin tones, model calibration can be used to alert a user with dark skin tones that their image is under-represented in the model's training data. This means that the model is more likely to be wrong in its prediction, and also more likely to be overconfident.

3.2 Cost Sensitive Classification

Cost Sensitive Classification is a concept where the real world costs of incorrect predictions are taken into account when the AI model makes a prediction.[45] As discussed in the example at the start of this section, where a model could have 94.6% accuracy but would actually have been very poor at identifying positive cases, disease detection is a context where it is beneficial for the design of the model to take into account the degree of harm that would be caused by false negatives versus false positives. Hence Cost Sensitive Classification could be beneficial in medical applications of AI.[46]

Consider again the example of an AI model trained to identify whether a lesion is cancerous (positive) or non-cancerous (negative). The model will output a probability of the lesion being cancerous (positive) between 0 to 1, i.e., probability is between 0% - 100%. By default, the Decision Threshold is set at 0.5.[47] This means if the model predicts there is a more than 50% likelihood the lesion is cancerous, it will output the class label of cancerous. Less than 50% and the model will output the label non-cancerous.

The default setting of a 50% Decision Threshold may not be optimal, if the cost of a false negative (predicting non-cancerous when the lesion is actually cancerous) is greater than the cost of a false positive (predicting the lesion is cancerous when it is not). What might be considered the best Decision Threshold is a trade-off between the cost of false negatives versus false positives, i.e., missing a diagnosis of cancer versus a false scare of cancer. This is a personal choice of giving weight to different risks.

The key point regarding Cost Sensitive Classification for the aims of this article is that deciding what is the relative weight to be given to the cost of missing a cancer diagnosis (false negatives), versus the costs of unnecessary anxiety and further testing (false positives), is a value judgement. It should be conveyed to consumers so that they can make this choice themselves, rather than the AI developers making this value judgement for them by default.

3.3 Reject Option

The Reject Option is a technique where the model can be designed to abstain from giving a prediction, if the model's confidence is below a certain threshold.[48] The model will still have a class label that is most probable, but if that probability is below a certain threshold of certainty, the model will not output a prediction. Instead, the model will alert the user that there is high uncertainty and the user should seek alternative advice. For example, where the user's data is under-represented in the training dataset, the model should report high predictive uncertainty.[49]

In the context of AI models that are being marketed direct-to-consumers, the Reject Option could be highly useful as it does not require the consumer to read or understand any complex information. The model simply abstains from giving a prediction and lets the user know that they should seek the advice of a human professional. From a harm minimisation perspective, this minimises the harm that could be caused when the AI model is not confident about its prediction.

To summarise this section on understanding AI 'accuracy,' these discussions about Uncertainty Estimation, Model Calibration, Cost Sensitive Classification and Reject Option are only a small glimpse into the layers of complexity regarding how the performance of an AI system can be understood. These are factors that a user would need to know to make a fully informed choice about whether they should rely on the predictions of the AI model. But these concepts are highly complex and require a significant amount of time to understand.

This level of complexity would be equivalent to expecting consumers to read the reports of clinical trials for drugs, and to decide for themselves whether a particular drug is appropriate for them. We do not expect that from consumers. We expect a

[45] Ho, "Real-World-Weight Cross-Entropy Loss Function."
[46] Araf, "Cost-Sensitive Learning," Mienye, "Performance Analysis of Cost-Sensitive Learning Methods," and Gan, "Integrating TANBN for imbalanced Data."
[47] Esposito, "Ghost: Adjusting the Decision Threshold," 2623.
[48] Hendrickx, "Machine Learning with Reject Option."
[49] Kompa, "Second Opinion Needed."

medical practitioner to read the reports, and to highlight the aspects of risk that are most relevant for an individual patient and explain those risks to them. Where an AI system is delivering a health service directly to consumers, it is very challenging for consumers to properly understand the risks of using the AI system, as there is no one to explain how these complex risks apply to their individual circumstances.

Part 4. Precision Healthcare Needs Precision Consent

In this final section, we describe how the techniques that are available in machine learning discussed in Part 3 could be utilised to provide consumers with a much better understanding of how well the AI performs and how reliable it is, i.e., fulfilling the service provider's duty of care in informing the consumer about the limitations of the service. We propose, moreover, how techniques in AI can be utilised to deliver personalised information about the scope and limitations of the service which are likely to be more effective in informing the consumers' understanding than general or boilerplate information, thus bolstering consent. This approach brings the information that AI service providers need to disclose to consumers, closer to the approach that would be taken by medical practitioners in disclosing risks to patients. We use the key principles of medical ethics to guide our proposed approach.

To summarise our arguments about the legal obligations of AI service providers to this point:

> (i) AI service providers can offer a medical service that is at a lower level than a human medical practitioner, if they properly communicate this to consumers and the consumer consents to the lower level of service; and
> (ii) The concept of Precision Consent is our proposed method for AI service providers to achieve this standard of care, by using techniques in AI to personalise information about risk to consumers.

We recognise that the Precision Consent framework is at this stage a theoretical proposal. The aim of this article is to examine the obligations of AI health providers pertaining to disclosure of risks, and to explore how AI technology could be utilised to meet those obligations. Further research is required to establish how this can be implemented. Also, there are challenges to regulatory compliance with the framework that are beyond the scope of this article; more research is needed to explore practicalities from a regulatory perspective.

The next sections describe the features of Precision Consent. In this discussion, we find it useful to draw an analogy with the ethical duties of a medical practitioner as a way of grounding the approach.

4.1 Non-Maleficence

In some cases, providing information about the limitations of a particular AI service will not be sufficient to satisfy the providers' duty of care because the service is unsuitable in any circumstances. This limitation expresses a core tenet of medicine, that the first obligation of a doctor to their patient is that the health intervention does not cause more harm than the disease itself.[50]

The Reject Option is a method of implementing this value. The AI service provider, practising the value that the AI system should not cause more harm to the consumer than if the consumer did not use the AI, sets a limit to the amount of risk that a user will be exposed to by using the AI system. The Reject Option enables the AI service provider to decide that above a particular threshold of uncertainty, such as one where a reasonable person would consider the output of the AI too risky to rely upon, the AI system will abstain from giving an answer and recommend the user seek alternative advice. This is analogous to a human doctor telling a patient that the medical task is outside the scope of their ability or specialty, and that the most they can do is recommend that the patient seek advice from another medical practitioner.

4.2 Personalised Risks Assessment

Generic information about the possible uses of an AI service will provide very little guidance about the application of that risk to an individual consumer. By comparison, a medical practitioner providing information about the limitations of a diagnosis will not satisfy their duty of care by giving the patient a generic warning applicable to the population generally. The medical

[50] Gillon, "'Primum Non Nocere' and the Principle of Non-Maleficence." See also Varkey, "Principles of Clinical Ethics," 18, discussing how the practical application of nonmaleficence is for the physician to weigh the benefits against burdens of all interventions and treatments.

practitioner must take into account the individual characteristics of the patient, so the practitioner can give the patient a personalised warning about risks.[51]

Our description in Part 3 about Uncertainty Estimation could be used to provide a far more precise warning of the level of risk that applies to the individual user. For example, if we consider again the case study of DermAssist, the model could calculate an accuracy rate that is specific to the skin tone of the user, based on the number of images in its training database of similar skin tones. Skin tone is one factor; other factors like age, gender, and existing medical conditions will also be relevant where the quantity of similar data in the training database will enable a more precise estimation of accuracy for the individual user. The user could be presented with a personalised accuracy rate, i.e., a single percentage number, that has been calculated to account for any bias in the training dataset regarding the proportion of samples that share relevant characteristics with the user. This would be much easier for consumers to understand, compared to a generic accuracy rate and fine print about whether the training database had sufficient samples of people from particular backgrounds.

4.3 Respecting Patient Autonomy

Most health decisions involve a trade-off between benefits and harms, and every action and inaction in healthcare carries risks and potential benefits. In making a decision about these trade-offs, patient autonomy is paramount. Thus, a key ethical duty of a medical practitioner is to empower their patient to make a choice that is consistent with the patient's own values. The doctor does this by providing the patient with enough information about the medical issue, such that the patient can understand the pros and cons of a decision.[52] But once the patient has enough knowledge that they meaningfully comprehend what the risks are, it is for the patient to decide how much weight to assign to those risks.[53]

Medical duty-to-warn cases point out that the probability of a risk occurring does not determine whether a medical practitioner is obliged to disclose the risk; even a risk that has very low probability should be disclosed if the individual patient would attach particular significance to it.[54] Courts have recognised that warnings about risk cannot be reduced to a discussion of probabilities, but depend upon the significance which the individual patient attaches to that particular risk of harm.[55] Where there is no easy answer and the patient is choosing between alternatives that all carry different risks, the weight that is attached to the different risks is a value judgement. Courts have upheld the principle that it is not for doctors to make a value judgement about the relative weight of different risks; it is for the person assuming the risk of being harmed, the patient, who gets to decide how much weight to assign to a particular risk.[56]

Cost Sensitive Classification has the potential for AI systems to embody this value if the Cost Sensitive Classification is implemented in a dynamic system, such that users can select how much weight they choose to assign to particular risks, e.g., for disease detection the user can assign the relative weight of false positives versus false negatives, and the model can adjust its predictions accordingly. Precision Consent thus fits the legal trend towards prioritising patient autonomy.

We note that the discussion of how machine learning techniques could be utilised to implement non-maleficence, personalised risk assessments, and respecting patient autonomy, is not intended to be an exhaustive description of how AI design can be used to fulfil legal or ethical obligations. It is intended as an illustration of how these particular obligations could be met with AI technology. Further obligations would require other methods.

[51] Discussed by the Supreme Court of the United Kingdom in *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, para [73], that the doctor's duty of care takes its precise content from the needs, concerns and circumstances of the individual patient. See also Gounder, "Informed Consent and the Duty to Warn," 326 – 327, arguing that the duty to warn of medical practitioner's has a subjective limb that requires a patient-specific approach to risk disclosure, including disclosure of risks that the particular patient might unreasonably be concerned about.

[52] Varkey, "Principles of Clinical Ethics," 19. Varkey argues that respecting the principle of autonomy obliges the physician to disclose medical information and treatment options that are necessary for the patient to exercise self-determination, and supports informed consent, truth telling, and confidentiality.

[53] Tickner, "Rogers v Whitaker," 114, discussing how the doctor must provide relevant information for the patient to have a meaningful choice.

[54] *Rogers v Whitaker* [1992] HCA 58, para [16], established that risk of blindness from the surgery was 1 in 14,000, but the court held this risk was within the duty to warn because the patient had expressed they were particularly afraid of that risk.

[55] *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, para [89].

[56] *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, para [87] – [91].

4.4 How Precision Consent Could Work

Finally, we outline how our concept of Precision Consent could be utilised by AI providers of a direct-to-consumer health service, to improve the standard of care in how the AI provider communicates the risks and limitations of their service. Firstly, this could occur by implementing a Reject Option, where the model will abstain from giving a result if the uncertainty is too high. For instance, surpassing a threshold of uncertainty where a reasonable person would consider the output of the AI too risky to rely upon would mean that the AI would abstain from giving a result. This is a safeguard that recognises consumers, as laypersons, do not have the expert knowledge necessary, and that AI service providers have a duty to limit the amount of risk they are exposing the consumer to.

Second, the quoted 'accuracy' of the AI system should be personalised to the individual user. Assuming that the model uncertainty is not too high, and the AI model proceeds to give the user a prediction, then the user should also be shown what is the likely 'accuracy' of that prediction, but personalised to them. This means an accuracy rate that is calculated based on training data of people who have similar relevant characteristics to the individual user, for example, skin tone, age, gender, and other characteristics.

Where the user has characteristics that were under-represented in the training data of the model, this would show as a lower personalised accuracy. This is a more effective way of communicating risk to consumers than expecting consumers to read a scientific article or a list of exclusions and warnings. A personalised accuracy rate is a single number, but it succinctly communicates whether the AI system is less accurate for this individual person, because of the training data used.
Third, for a disease detection service, a significant factor is what relative weight to assign to false positives versus false negatives. Both are harmful, but the relative degree of harm and the trade-off between the two are value judgments that a consumer should be entitled to make according to their own values. This could be achieved by asking users to assign how much weight they place on the potential harm of a false positive versus a false negative, and implementing this in the Cost Sensitive Classification of the AI's prediction.

Whilst machine learning, accuracy, confidence intervals, and uncertainty estimation are complicated subjects that most consumers will not understand, ordinary people can understand how much risk they personally want to take regarding a disease diagnosis. That is a value judgement that does not require technical knowledge about AI. If users can be asked in plain language what their values are regarding those risks, then these values can be respected with user-determined Cost Sensitive Classification.

The following table summarises our interdisciplinary framework, Precision Consent, with examples of how it brings together three key components: identifying harm to consumers, comparisons to medical practitioners, and selecting machine learning techniques that can address the harm.

Table 4. Precision Consent, Applied to Direct-to-Consumer AI Health Services

Elements of Precision Consent		**Examples**		
1	**Identify potential harm of AI service to consumer**	AI does more harm than good	AI performs worse for this individual, due to biased training	AI does not align with values of the consumer
2	**Consider comparable duty of a medical practitioner**	Non-Maleficence	Personalised warnings	Respect for patient autonomy
3	**Select machine learning technique that can best fulfill that duty**	Reject Option	Uncertainty Estimation, Model Calibration	Cost Sensitive Classification

Conclusion

The law should not discourage innovation in healthcare. There is a global shortage of specialist medical care. If people could access medical expertise through their smartphone, at a fraction of the cost, that would be greatly beneficial to society. It is a loss for society if the potential benefits of AI technology cannot be harnessed for public health because of uncertainty surrounding the legal obligations of AI service providers.

However, the prospect of direct-to-consumer AI health services does raise questions about whether consumers would be exposed to significant risks that they are not equipped to understand. The 'accuracy' of AI systems is not a simple measurement of how reliable an AI system is, but rather a highly complex and nuanced subject, and would differ for consumers based on their individual characteristics. The imposition of legal obligations on AI service providers is necessary and vital to protect consumers who often lack the expertise to have informed consent to the risks associated with AI services.

The argument of this article is that there is a path forward to balance both of these important objectives: advancing new health technologies that benefit society, and protecting consumers. There are techniques in AI that can be developed to better inform consumers of the limitations of an AI service. Generic warnings and long lists of risks should become a thing of the past. It is realistically achievable for AI systems to provide warnings that are personalised to individual consumers, guided by the principles which medical practitioners follow when advising their patients about risks.

This is not to say that automating the consent process is without its own challenges. There are risks of over-reliance on AI design to fulfil the complex and nuanced ethical role of doctors when providing information to patients, which AI technology might never be able to fully replicate. Additionally, there are practical and regulatory obstacles to implementing AI consent processes that this article leaves for future research. The goal of this article is to highlight the parallels between existing AI techniques and certain aspects of a doctor's role, in order to provide individuals seeking AI-driven healthcare with a clearer understanding of its limitations.

Bibliography

Primary Legal Material

AB 489 Health Care Professions: Deceptive Terms or Letters: Artificial Intelligence, 2025-2026 Reg. Sess. (Cal. 2025), s 4999.9. https://legiscan.com/CA/text/AB489/id/3272936.
AB 3030 Health Care Services: Artificial Intelligence, 2023-2024 Leg. Sess. (Cal. 2024), s 1339.75. https://legiscan.com/CA/bill/AB3030/2023.
Competition and Consumer Act 2010 (Cth)
Consumer Rights Act 2015 (UK)
Cotton v. Buckeye Gas Products Company, 840 F.2d 935 (D.C. Cir 1988)
Let's Go Adventures Pty Ltd v Barrett [2017] NSWCA 243
Montgomery v Lanarkshire Health Board [2015] UKSC 11
Regulation (EU) 2017/745 of the European Parliament and of the Council, *of 5 April 2017 on Medical Devices*, Article 52 (7) and Annex VIII (4.1).
Rogers v Whitaker [1992] HCA 58
Wyong Shire Council v Shirt [1980] HCA 12

Secondary Sources

Abdar, Moloud, Farhad Pourpanah, Sadiq Hussain, et al. "A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges." *Information Fusion* 76 (2021): 243-97. https://doi.org/10.1016/j.inffus.2021.05.008.
Ada, "More than a Symptom Checker." https://ada.com/app/ (accessed 4 September 2025).
Adamson, Adewole and Avery Smith. "Machine Learning and Health Care Disparities in Dermatology." *JAMA Dermatology* 154, no 11 (2018): 1247-48. https://doi.org/10.1001/jamadermatol.2018.2348.
Adar, Yehuda and Shmuel Becher. "Ending the License to Exploit: Administrative Oversight of Consumer Contracts." *Boston College Law Review,* 62 (2021): 2405-2461.
Araf, Imane, Ali Idri and Ikram Chairi. "Cost-Sensitive Learning for Imbalanced Medical Data: A Review." *Artificial Intelligence Review* 57, no 4 (2024): 80. https://doi.org/10.1007/s10462-023-10652-8.
Ayres, Ian and Alan Schwartz. "The No-Reading Problem in Consumer Contract Law." *Stanford Law Review* 66 (2014): 545-610.
Bhatt, Umang, Javier Antorán, Yunfeng Zhang, et al. "Uncertainty as a Form of Transparency: Measuring, Communicating, and Using Uncertainty." *AAI/ACM Conference on Artificial Intelligence, Ethics and Society* (2021): arXiv:2011.07586 [cs.CY], https://doi.org/10.48550/arXiv.2011.07586.
Bui, Peggy and Yuan Liu. "Using AI to Help Find Answers to Common Skin Conditions." *Google The Keyword*, May 18, 2021, https://blog.google/technology/health/ai-dermatology-preview-io-2021/ (accessed 4 September 2025).
California Department of Justice, Office of the Attorney General, *California Attorney General's Legal Advisory on the Application of Existing California Law to Artificial Intelligence in Healthcare* (January 13, 2025). https://oag.ca.gov/system/files/attachments/press-docs/Final%20Legal%20Advisory%20-%20Application%20of%20Existing%20CA%20Laws%20to%20Artificial%20Intelligence%20in%20Healthcare.pdf (accessed 8 February 2026).
Cohen, Adam, Simon Mathews, Ray Dorsey, David Bates and Kyan Safavi. "Direct-to-Consumer Digital Health." *The Lancet Digital Health* 2, no 4 (2020): e163-e65. https://doi.org/10.1016/S2589-7500(20)30057-1.
Daneshjou, Roxana, Mary Smith, Mary Sun, Veronica Rotemberg and James Zou. "Lack of Transparency and Potential Bias in Artificial Intelligence Data Sets and Algorithms: A Scoping Review." *JAMA Dermatology* 157, no 11 (2021): 1362-69. https://doi.org/10.1001/jamadermatol.2021.3129.
Dolezal, James, Andrew Srisuwananukorn, Dmitry Karpeyev, et al. "Uncertainty-Informed Deep Learning Models Enable High-Confidence Predictions for Digital Histopathology." *Nature Communications* 13, no 1 (2022). https://doi.org/10.1038/s41467-022-34025-x.
Duniphin, Darlla. "Limited Access to Dermatology Specialty Care: Barriers and Teledermatology." *Dermatol Pract Concept* 13, no 1 (January 2023). https://doi.org/10.5826/dpc.1301a31.
Esposito, Carmen, Gregory Landrum, Nadine Schneider, Nikolaus Stiefl and Sereina Riniker. "Ghost: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning." *Journal of Chemical Information and Modeling* 61, no 6 (2021): 2623-40. https://doi.org/10.1021/acs.jcim.1c00160.
Gan, Dan, Jiang Shen, Bang An, Man Xu and Na Liu. "Integrating TANBN with Cost Sensitive Classification Algorithm for Imbalanced Data in Medical Diagnosis." *Computers & Industrial Engineering* 140 (2020): 106266. https://doi.org/10.1016/j.cie.2019.106266.

Gillon, Raanan. "'Primum Non Nocere' and the Principle of Non-Maleficence." *British Medical Journal (Clinical Research ed.)* 291, no 6488 (1985): 130-1. https://doi.org/10.1136/bmj.291.6488.130.

Gounder, Rajesh. "Informed Consent and the Duty to Warn: More than the Mere Provision of Information." *Journal of Law and Medicine* 31, 2 (2024): 324-342.

Guo, Lin Lawrence, Stephen Pfohl, Jason Fries, et al. "Systematic Review of Approaches to Preserve Machine Learning Performance in the Presence of Temporal Dataset Shift in Clinical Medicine." *Applied Clinical Informatics* 12, no 04 (2021): 808-15. https://doi.org/10.1055/s-0041-1735184.

Haque, Romael and Sabirat Rubya. "An Overview of Chatbot-Based Mobile Mental Health Apps: Insights from App Description and User Reviews." *JMIR Mhealth Uhealth* 11 (2023): e44838. https://doi.org/10.2196/44838.

Heartscan, "Heart Rate Monitor," https://play.google.com/store/apps/dev?id=6015445904816809789&hl=en (accessed 4 September 2025).

Hendrickx, Kilian, Lorenzo Perini, Dries Van der Plas, Wannes Meert and Jesse Davis. "Machine Learning with a Reject Option: A Survey." *Machine Learning* 113, no 5 (2024): 3073-110. https://doi.org/10.1007/s10994-024-06534-x.

Ho, Yaoshiang and Samuel Wookey. "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling." *IEEE Access* 8 (2020): 4806-13. https://doi.org/10.1109/ACCESS.2019.2962617

Husgen, Jason. "Product Liability Suits Involving Drug or Device Manufacturers and Physicians: The Learned Intermediary Doctrine and the Physician's Duty to Warn." *Missouri Medicine* 111, no 6 (2014): 478-81.

Khatun, Nazma, Gabriella Spinelli and Federico Colecchia. "Technology Innovation to Reduce Health Inequality in Skin Diagnosis and to Improve Patient Outcomes for People of Color: A Thematic Literature Review and Future Research Agenda." *Frontiers in Artificial Intelligence* 7 (2024): 1394386. https://doi.org/10.3389/frai.2024.1394386.

Kilim, Oz, Alex Olar, Tamás Joó, Tamás Palicz, Péter Pollner and István Csabai. "Physical Imaging Parameter Variation Drives Domain Shift." *Scientific Reports* 12, no 1 (2022). https://doi.org/10.1038/s41598-022-23990-4.

Kompa, Benjamin, Jasper Snoek and Andrew Beam. "Second Opinion Needed: Communicating Uncertainty in Medical Machine Learning." *npj Digital Medicine* 4, no 1 (2021): 4. https://doi.org/10.1038/s41746-020-00367-3.

Levi, Dan, Liran Gispan, Niv Giladi and Ethan Fetaya. "Evaluating and Calibrating Uncertainty Prediction in Regression Tasks." *Sensors* 22, no 15 (2022): 5540. https://doi.org/10.3390/s22155540.

Li, Yikuan, Shishir Rao, Abdelaali Hassaine, et al. "Deep Bayesian Gaussian Processes for Uncertainty Estimation in Electronic Health Records." *Scientific Reports* 11, no 1 (2021): 20685. https://doi.org/10.1038/s41598-021-00144-6.

Liu, Yuan, Ayush Jain, Clara Eng, et al. "A Deep Learning System for Differential Diagnosis of Skin Diseases." *Nature Medicine* 26 (2020): 900-908. https://doi.org/10.1038/s41591-020-0842-3.

Mathieu, Boniol, Teena Kunjumen, Tapas Sadasivan Nair, et al. "The Global Health Workforce Stock and Distribution in 2020 and 2030: A Threat to Equity and Universal Health Coverage?" *BMJ Global Health* 7, no 6 (2022): e009316. https://doi.org/10.1136/bmjgh-2022-009316.

Mehrtash, Alireza, William M. Wells, Clare M. Tempany, Purang Abolmaesumi and Tina Kapur. "Confidence Calibration and Predictive Uncertainty Estimation for Deep Medical Image Segmentation." *IEEE Transactions on Medical Imaging* 39, no 12 (2020): 3868-78. https://doi.org/10.1109/TMI.2020.3006437.

Mienye, Ibomoiye and Yanxia Sun. "Performance Analysis of Cost-Sensitive Learning Methods with Application to Imbalanced Medical Data." *Informatics in Medicine Unlocked* 25 (2021): 100690. https://doi.org/10.1016/j.imu.2021.100690.

mySugr, "The mySugr App", https://www.mysugr.com/en/diabetes-app (accessed 4 September 2025).

Ouellette, Samantha and Babar Rao. "Usefulness of Smartphones in Dermatology: A Us-Based Review." *International Journal of Environmental Research and Public Health* 19, no 6 (2022): 3553. https://doi.org/10.3390/ijerph19063553.

Sanchez, Pedro, Jeremy Voisey, Tian Xia, Hannah Watson, Alison O'Neil and Sotirios Tsaftaris. "Causal Machine Learning for Healthcare and Precision Medicine." *Royal Society Open Science* 9, no 8 (2022). https://doi.org/10.1098/rsos.220638.

Sangers, Tobias, Suzan Reeder, Sophie van der Vet, et al. "Validation of a Market-Approved Artificial Intelligence Mobile Health App for Skin Cancer Screening: A Prospective Multicenter Diagnostic Accuracy Study." *Dermatology* 238, no 4 (2022): 649-56. https://doi.org/10.1159/000520474.

Savulescu, Julian, Alberto Giubilini, Robert Vandersluis and Abhishek Mishra. "Ethics of Artificial Intelligence in Medicine." *Singapore Medical Journal* 65, no 3 (2024): 150-58. https://doi.org/10.4103/singaporemedj.SMJ-2023-279.

SkinVision, "Accurate skin cancer detection made simple." https://www.skinvision.com/ (accessed 4 September 2025).

Smak Gregoor, Anna, Tobias Sangers, Lytske Bakker, et al. "An Artificial Intelligence Based App for Skin Cancer Detection Evaluated in a Population Based Setting." *npj Digital Medicine* 6, no 1 (2023): 90. https://doi.org/10.1038/s41746-023-00831-w.

Smak Gregoor, Anna, Tobias Sangers, Just Eekhof, et al. "Artificial Intelligence in Mobile Health for Skin Cancer Diagnostics at Home (Aim High): A Pilot Feasibility Study." *eClinicalMedicine* 60 (2023). https://doi.org/10.1016/j.eclinm.2023.102019.

Tickner, Karen. "Rogers v Whitaker: Giving Patients a Meaningful Choice." *Oxford Journal of Legal Studies* 15, no 1 (Spring 1995): 109-118.

Ulmer, Dennis, Lotta Meijerink and Giovanni Cinà. "Trust Issues: Uncertainty Estimation Does Not Enable Reliable Odd Detection on Medical Tabular Data." (2021): arXiv:2011.03274 [cs.LG]. https://doi.org/10.48550/arXiv.2011.03274.

Varkey, Basil. "Principles of Clinical Ethics and Their Application to Practice." *Medical Principles and Practice: International Journal of the Kuwait University, Health Science Centre* 30, no 1 (2021): 17–28. https://doi.org/10.1159/000509119.

Yang, Jingkang, Kaiyang Zhou, Yixuan Li, and Ziwei Liu. "Generalized out-of-Distribution Detection: A Survey." *International Journal of Computer Vision* 132, no 12 (2024): 5635-62. https://doi.org/10.1007/s11263-024-02117-4.

Zheng, Alice. *Evaluting Machine Learning Models.* O'Reilly, 2015.

Zhou, Kaiyang, Ziwei Liu, Yu Qiao, Tao Xiang and Chen Change Loy. "Domain Generalization: A Survey." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, no 4 (2023): 4396-415. https://doi.org/10.1109/TPAMI.2022.3195549.