

Protocol

# Using Artificial Intelligence as a Diagnostic Decision Support Tool in Skin Disease: Protocol for an Observational Prospective Cohort Study

Anna Escalé-Besa<sup>1</sup>, MD; Aïna Fuster-Casanovas<sup>2,3</sup>, RPh, MSc; Alexander Börve<sup>4,5</sup>, MD; Oriol Yéllamos<sup>6</sup>, MD, PhD; Xavier Fustà-Novell<sup>7</sup>, MD; Mireia Esquiús Rafat<sup>7</sup>, MD; Francesc X Marin-Gomez<sup>8</sup>, MD, PhD; Josep Vidal-Alaball<sup>2,3,9</sup>, MPH, MD, PhD

<sup>1</sup>Centre d'Atenció Primària Navàs-Balsareny, Institut Català de la Salut, Navàs, Spain

<sup>2</sup>Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain

<sup>3</sup>Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain

<sup>4</sup>iDoc24 Inc, San Francisco, CA, United States

<sup>5</sup>Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

<sup>6</sup>Dermatology Department, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>7</sup>Fundació Althaia de Manresa, Manresa, Spain

<sup>8</sup>Servei d'Atenció Primària Osona, Gerència Territorial de Barcelona, Institut Català de la Salut, Vic, Spain

<sup>9</sup>Faculty of Medicine, University of Vic-Central University of Catalonia, Vic, Spain

**Corresponding Author:**

Aïna Fuster-Casanovas, RPh, MSc

Unitat de Suport a la Recerca de la Catalunya Central

Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina

C/Pica d'Estats 13-15

Sant Fruitós de Bages, 08272

Spain

Phone: 34 93 693 0040

Email: [afuster.cc.ics@gencat.cat](mailto:afuster.cc.ics@gencat.cat)

## Abstract

**Background:** Dermatological conditions are a relevant health problem. Each person has an average of 1.6 skin diseases per year, and consultations for skin pathology represent 20% of the total annual visits to primary care and around 35% are referred to a dermatology specialist. Machine learning (ML) models can be a good tool to help primary care professionals, as it can analyze and optimize complex sets of data. In addition, ML models are increasingly being applied to dermatology as a diagnostic decision support tool using image analysis, especially for skin cancer detection and classification.

**Objective:** This study aims to perform a prospective validation of an image analysis ML model as a diagnostic decision support tool for the diagnosis of dermatological conditions.

**Methods:** In this prospective study, 100 consecutive patients who visit a participant general practitioner (GP) with a skin problem in central Catalonia were recruited. Data collection was planned to last 7 months. Anonymized pictures of skin diseases were taken and introduced to the ML model interface (capable of screening for 44 different skin diseases), which returned the top 5 diagnoses by probability. The same image was also sent as a teledermatology consultation following the current established workflow. The GP, ML model, and dermatologist's assessments will be compared to calculate the precision, sensitivity, specificity, and accuracy of the ML model. The results will be represented globally and individually for each skin disease class using a confusion matrix and one-versus-all methodology. The time taken to make the diagnosis will also be taken into consideration.

**Results:** Patient recruitment began in June 2021 and lasted for 5 months. Currently, all patients have been recruited and the images have been shown to the GPs and dermatologists. The analysis of the results has already started.

**Conclusions:** This study will provide information about ML models' effectiveness and limitations. External testing is essential for regulating these diagnostic systems to deploy ML models in a primary care practice setting.

**KEYWORDS**

machine learning; artificial intelligence; data accuracy; computer-assisted diagnosis; neural network computer; support tool; skin disease; cohort study; dermatology

## ***Introduction***

Health care systems in Western countries are increasingly exposed to new challenges: a high volume of demand, aging populations, chronic diseases, a high degree of comorbidity, and the global pandemic situation. These factors, together with the lack of professionals, particularly general practitioners (GPs), generate the need to find new solutions to improve the quality of care and the workflow of professionals [1].

Dermatological conditions are a relevant health problem, and skin disease is one of the principal reasons why patients visit their GPs. Every person has on average 1.6 skin diseases per year [2]. About 20% of all GP visits are related to a dermatological concern, and 15% of all telehealth visits are related to dermatology [3,4]. About 7.6% of the total population of Catalonia visit a primary care center (PCC) due to skin concerns every year, and around 35% are referred to a dermatology specialist [5,6]. Nowadays, in the health care area of central Catalonia, teledermatology consultations are commonly used to refer patients to a hospital-based dermatologist. It is estimated that more than 70% of all PCC patients with a skin problem can be effectively triaged with teledermatology and do not need a face-to-face visit with a dermatologist [7,8].

The use of computer-assisted diagnosis in medicine dates to the 1960s in radiology. The initial description of artificial intelligence (AI) in dermatopathology dates to 1987, when the text-based system TEGUMENT was produced [9]. TEGUMENT included a semantic tree with 986 potential diagnoses used to assist the dermatologist in the histopathologic differential diagnosis of diseases and tumors of the skin. Computer-aided melanoma diagnosis was introduced in the early 2000s in dermatology using rule-based classifiers, which use predefined features to classify images into desired categories [10].

The application of teledermatology worldwide has increased over the years. It is used in many PCC settings and has been well established and backed by extensive research that it is a viable method of triage, particularly for skin cancer lesions [11]. Studies comparing the general accuracy of face-to-face dermatology consultation versus teledermatology have different results. In general, face-to-face consultations achieve higher diagnostic accuracy than teledermatology. However, some studies did report the high accuracy of teledermatology diagnoses for skin cancer [12]. Nevertheless, it is necessary to first ensure that the clinicians have high interrater reliability; without this, it is difficult to tell whether the limited agreement in diagnoses is related to the use of the technology itself or differences in clinical opinion that ordinarily exist in practice. In this context, studies have compared the diagnostic agreement between GPs using telemedicine and dermatologists. The results of the studies showed an overall diagnostic agreement of

65.52%, showing that GPs tend to overdiagnose some diseases [13]. The concordance obtained for teledermatology was 94.7%. Even though this technique showed merits in triage quality, it presented low accuracy in inflammatory problems [13]. Teledermatology has the potential to increase access by facilitating referrals and offering convenience and decreased waiting times, as well as providing diagnostic support and improved satisfaction for both patients and providers [8,14-17]. To achieve the correct implementation of AI in primary care, it is important to know the real needs and developed an easy-to-use interface, which can help reduce resistance to change from traditional to touch-based interfaces in current clinical setups [18].

In recent years, AI has been developed, researched, and applied in many medical disciplines. Images are the most commonly used form of data for AI development, such as electrocardiograms or radiologic images [19-21]. Dermatopathology is particularly suited for deep learning algorithms, because pattern recognition in scanning magnification is fundamental for diagnosis [10,22-24]. Furthermore, machine learning (ML) is increasingly being applied to dermatology, particularly focused on skin cancer detection using image analysis with ML models that include deep convolutional neural networks (CNNs) [25,26]. Algorithms and models that include CNNs were introduced in the 1980s [23], but it was not until 2012 that the ImageNet competition demonstrated their potential for image analysis. Since then, CNN has become a popular ML approach in several disciplines including dermatology [27]. There are also ML studies that have investigated the use of a wider classification of skin diseases that could be used in primary health care [28]. The evolution in ML came around 2010 with deep learning [10], and it has revolutionized tasks such as image classification and segmentation and speech recognition.

Even though GPs see a lot of skin ailments [5,29], few studies have been conducted in primary health care settings prospectively. However, some studies have included GPs along with dermatologists as readers for the comparison group to compare the performance of ML with clinicians [11,28,30] and have concluded that AI tools could be used in primary care [28]. For all these reasons, the main objective of the study is to perform a prospective validation in real primary care practice settings of an ML model as a diagnostic decision support tool for the diagnosis of dermatological conditions in a rural area of Catalonia (Spain).

## Methods

### Study Design

#### *Trial Design*

This is a prospective study that aims to evaluate an ML model's performance, comparing its diagnostic capacity with GPs and dermatologists. A secure, anonymous, and stand-alone web interface that is compatible to any mobile device was integrated with the Autoderm application programming interface (API; iDoc24 Inc).

To conduct this study, the following procedure were carried out until the required number of samples was reached: (1) a suitable patient with skin concern was asked to participate and sign the patient study agreement; (2) GPs diagnosed the skin condition; (3) GPs took 1 good-quality image of the skin condition; (4) GPs sent the photograph as a teledermatology consultation following the current workflow; (5) the image were entered into the Autoderm ML interface; and (6) dermatologists diagnosed the skin condition.

The satisfaction of the health care professionals using the ML tool were assessed using 3 questions embedded in the questionnaire. The questions relate to the potential usefulness of the tool to help the diagnosis or consider further diagnosis not contemplated initially and the potential use of the tool to avoid a dermatology referral.

#### *Study Population, Site Participation, and Recruitment*

The study was conducted in PCCs managed by the Catalan Health Institute (the main primary care services provider in Catalonia) in central Catalonia, which includes the regions of Anoia, Bages, Moianès, Berguedà, and Osona. The reference population included in the study was around 512,050 habitants. The recruitment of prospective subjects was done consecutively.

### Data Collection and Sources of Information

#### *Patients, Data Collection, Sources of Information, and Intervention*

GPs collected data from consecutive patients who met the inclusion criteria after obtaining written informed consent. The collected data were reported exclusively in a case report form.

The GP diagnosed the skin condition and filled in a questionnaire. For each patient, the GP used a smartphone camera to take a close-up good-quality image of the skin problem. The image is anonymous, as it is not possible to identify the patients. The GP then used the Autoderm ML interface to upload the anonymized image and filled in the questionnaire with the top 5 diagnoses generated by the ML model.

This evaluation study of the Autoderm API tool is intended as a validation study of a tool to support the diagnosis of skin lesions in real clinical practice conditions in primary care. Therefore, although the tool uses a closed source code, this study is intended to be a starting point to see if similar tools can be suitable for use as working tools in real clinical conditions. Autoderm is a research-backed, Conformité Européenne–marked

dermatology search engine using ML technology to help provide faster and more accurate skin diagnosis. The current ML model can screen for 44 different skin disease types, which includes inflammatory skin diseases, skin tumors, and genital skin concerns, and can be accessed via an API. For this study, a user web interface was developed for the easy upload of images from the smartphone library or those taken with the smartphone camera. From just a smartphone photo, the model generates the top 5 ranked skin diseases in order of probability. The life cycle of this ML model is estimated to be around 3 months. After this period, the model will be upgraded to a more accurate model that will possibly include more skin diseases.

At its current stage, the ML model uses a 34-layer pretrained ResNet model provided by TorchVision (PyTorch) that is used for applications such as computer vision and natural language processing. In addition, the model has been trained using transfer learning on a proprietary data set of 55,364 images for the training set and 13,841 images for testing. The average accuracy of the model used is 31.7% for the top 1 diagnosis and 68.1% for the top 5. Some skin diseases have higher accuracy and some have lower accuracy, which is a consequence of the number of images the ML was trained on and the fact that some skin diseases are more distinct and certain anatomic locations make diagnosis more difficult. Before deployment, the ML model was also manually tested with a data set collected from various websites that provided images of skin disease taken with a mobile camera. The ML model was deployed when it was deemed to be robust. The 44 different skin disease classes represent about 90% of what the general public are concerned about and consults for.

To get a second opinion, the GP incorporated the anonymized image and an accurate description of the skin lesion into the patient's medical history following the current teledermatology workflow. The dermatologist then filled in the "Assessment by teledermatology" questionnaire after receiving the information. The response time was expected to be about 2-7 days.

In case of a dermatology referral, the GP filled in the "Assessment by in-person dermatologist" questionnaire by accessing the electronic health records as they become available. The average waiting time for a dermatology referral ranges from 30-90 days.

The questionnaire case number was predefined before the initiation of the data collection phase and was the same for all questionnaires, making it impossible to identify the patient.

#### *Inclusion Criteria*

Patients visiting for reasons related to a cutaneous disease at a participating PCC, who provided written informed consent and were aged  $\geq 18$  years, were included in the prospective study.

#### *Exclusion Criteria*

Patients with a cutaneous lesion that could not be photographed with a smartphone or had conditions associated with a risk of poor protocol compliance were excluded from the study. Images with poor quality were also excluded from the study.

## Statistical Analysis

### Calculation of Sample Size

To compare the performance of the ML model with those of the GPs and dermatologists, a sample size of 100 images of skin diseases from patients who meet the inclusion criteria is required. The proposed sample size is based on sample size calculation used in similar research studies [31-33].

### Planned Analysis

The validation data set will include about 100 cases, consisting of an image and 3 or 4 assessments: the face-to-face assessment by a GP, the assessment made by teledermatology, the top 5 differential diagnoses from the ML model ordered by probability, and the assessment by the face-to-face dermatologist (in cases with a referral). The ML model assessment will be limited to 44 skin diseases classes. A confusion matrix will be used to calculate the precision, sensitivity (recall), specificity, and accuracy of the ML model. For each individual skin disease, the number of true positives, true negatives, false positives, and false negatives will be calculated. To evaluate the ML multiclass classifier, data will be treated as a collection of binary problems, 1 for each skin disease class. Area under the curve and receiver operating characteristics curve for N number of skin diseases classes will be calculated using one-versus-all methodology. Macro- and micro-averaging measures will be considered to highlight the performance of infrequent skin disease classes (weighted by prevalence). Precision, recall, and *F*-measure will be calculated independently for each skin disease class, and the results will be combined to obtain the average precision and *F*-score. The accuracy of the top 3 diagnoses of the ML model will be also calculated.

### Ethics Approval

The Institut Universitari d'Investigació en Atenció Primària (University Institute for Research in Primary Health Care) Jordi Gol i Gurina ethics committee approved the trial study protocol (code 20-159P). Written informed consent was sought from all patients participating in the study.

## Results

The results will be represented globally and individually for each skin disease class using a confusion matrix and one-versus-all methodology. The time taken to make the diagnosis will also be taken into consideration. The satisfaction of the professionals with the use of this ML tool will be assessed.

Patient recruitment began in June 2021 and lasted for 5 months. Currently, all patients have been recruited and the images have been shown to the GPs and dermatologists. The analysis of the results has already started. We hope that sufficient evidence can be obtained to validate this image analysis ML model. We believe the results will be used in clinical practice on patients with skin diseases to make a GP's workflow more efficient and safer for the patient. This study is a first approach to designing larger ML model validation studies.

It has to be considered that even if the ML model does not provide a better diagnosis than the doctor's, it is expected to help the practitioner consider other differential diagnoses.

## Discussion

This study aims to perform a prospective validation of an ML model as a diagnostic decision support tool for the diagnosis of dermatological conditions. It would also assess the diagnostic accuracy and efficacy of a ML model in a PCC setting. In this context, this study may provide added value for both patients and primary care physicians, increasing the effectiveness and efficiency of the system, and will provide information about ML models' effectiveness and limitations. External testing is essential to regulate these diagnostic systems and deploy ML models in real PCC settings.

First, the most relevant limitation of this study is the number of image samples used for the evaluation of the performance of the ML model. As Autoderm assesses only 44 skin diseases and that the prevalence of a substantial number of these skin conditions represents less than 1% to 5% of the images, the sample data of each class may be unbalanced and some skin conditions may not be evaluated, causing an insufficient confidence level and therefore, less conclusive results for these specific conditions.

Second, due to the sample size and consecutive case recruitment, we will probably not obtain representative results for less common diseases. As class imbalance may be an issue in the 100 patients recruited, we will focus on the *F*-Score for the analysis, as otherwise having 90% of the most common skin lesions may overestimate the quality of the model when considering accuracy, sensitivity, and specificity. It has to be taken into consideration that this study will be done in real practice conditions, and we will not be able to select the patients.

Third, a diagnosis made with only 1 image with the most optimal composition may present inherent limitations compared to diagnoses made in a clinical setting. Our ML algorithm output was based on a single photograph, which differs from other ML algorithms that consider more than 1 photograph and even those with the same algorithm available for the general public that considers 2 images.

Fourth, another limitation is that our data will not include additional testing and only a subset of suspected malignancies will have a biopsy confirmation. Instead, our golden standard for each case is based on aggregating the differential diagnoses of a panel of dermatologists. Ambiguities in diagnosis do exist in clinical practice, which makes it challenging to evaluate the accuracy of clinicians and deep learning systems, especially for conditions such as rashes, which are not typically biopsied.

Fifth, our ML algorithm did not include additional clinical metadata (past medical history, symptoms, appearance, and the texture), which is a probable grievance when comparing ML versus physicians' diagnostic accuracy.

Lastly, the clinicians were requested to provide just the top 3 diagnosis, even if they had other potential options.

## Data Availability

Our manuscript is based on confidential and sensitive health data. However, to support scientific transparency, we will publish deidentified data for reviewers or for replication purposes. The data will be deposited and made available in our publicly accessible Mendeley repository.

## Authors' Contributions

All authors contributed to the design and content of the study protocol. AEB is responsible for the coordination of the study. AEB, JVA, AFC, and FXMG are responsible for the design and writing of the initial draft of the manuscript. AEB, OY, MER, and XFN are responsible for data collection, and AEB and JVA are responsible for data processing and exploitation. All authors have read and approved the final version of the manuscript.

## Conflicts of Interest

AB is the chief executive officer and majority shareholder of iDoc24 Inc and iDoc24 AB. He provided the technology but did not take part in the data collection or any clinical validation.

## References

1. Sánchez-Sagrado T. Are there too many or too few physicians in Spain? migration: the eternal resource. *Rev Clin Esp (English Ed)* 2013 Oct;213(7):347-353. [doi: [10.1016/j.rceng.2013.07.004](https://doi.org/10.1016/j.rceng.2013.07.004)]
2. Lim HW, Collins SAB, Resneck JS, Bologna JL, Hodge JA, Rohrer TA, et al. The burden of skin disease in the United States. *J Am Acad Dermatol* 2017 May;76(5):958-972.e2. [doi: [10.1016/j.jaad.2016.12.043](https://doi.org/10.1016/j.jaad.2016.12.043)] [Medline: [28259441](https://pubmed.ncbi.nlm.nih.gov/28259441/)]
3. Schofield JK, Fleming D, Grindlay D, Williams H. Skin conditions are the commonest new reason people present to general practitioners in England and Wales. *Br J Dermatol* 2011 Nov;165(5):1044-1050. [doi: [10.1111/j.1365-2133.2011.10464.x](https://doi.org/10.1111/j.1365-2133.2011.10464.x)] [Medline: [21692764](https://pubmed.ncbi.nlm.nih.gov/21692764/)]
4. Tensen E, van der Heijden JP, Jaspers MWM, Witkamp L. Two decades of teledermatology: current status and integration in national healthcare systems. *Curr Dermatol Rep* 2016 Mar 28;5:96-104 [FREE Full text] [doi: [10.1007/s13671-016-0136-7](https://doi.org/10.1007/s13671-016-0136-7)] [Medline: [27182461](https://pubmed.ncbi.nlm.nih.gov/27182461/)]
5. Servei Català de la Salut. Activitat assistencial de la xarxa sanitària de Catalunya, any 2012: registre del conjunt mínim bàsic de dades (CMBD). Barcelona: Departament de Salut. 2013 Apr. URL: <http://hdl.handle.net/11351/1025> [accessed 2022-08-11]
6. Lowell BA, Froelich CW, Federman DG, Kirsner RS. Dermatology in primary care: prevalence and patient disposition. *J Am Acad Dermatol* 2001 Aug;45(2):250-255. [doi: [10.1067/mjd.2001.114598](https://doi.org/10.1067/mjd.2001.114598)] [Medline: [11464187](https://pubmed.ncbi.nlm.nih.gov/11464187/)]
7. Porta N, San Juan J, Grasa M, Simal E, Ara M, Querol I. Diagnostic agreement between primary care physicians and dermatologists in the health area of a referral hospital. *Actas Dermosifiliogr (English Ed)* 2008 Apr;99(3):207-212 [FREE Full text] [doi: [10.1016/s1578-2190\(08\)70233-6](https://doi.org/10.1016/s1578-2190(08)70233-6)] [Medline: [18358196](https://pubmed.ncbi.nlm.nih.gov/18358196/)]
8. López Seguí F, Franch Parella J, Gironès García X, Mendioroz Peña J, García Cuyàs F, Adroher Mas C, et al. A cost-minimization analysis of a medical record-based, store and forward and provider-to-provider telemedicine compared to usual care in Catalonia: more agile and efficient, especially for users. *Int J Environ Res Public Health* 2020 Mar 18;17(6):2008 [FREE Full text] [doi: [10.3390/ijerph17062008](https://doi.org/10.3390/ijerph17062008)] [Medline: [32197434](https://pubmed.ncbi.nlm.nih.gov/32197434/)]
9. Potter B, Ronan SG. Computerized dermatopathologic diagnosis. *J Am Acad Dermatol* 1987 Jul;17(1):119-131. [doi: [10.1016/s0190-9622\(87\)70183-2](https://doi.org/10.1016/s0190-9622(87)70183-2)] [Medline: [3301922](https://pubmed.ncbi.nlm.nih.gov/3301922/)]
10. Talebi-Liasi F, Markowitz O. Is artificial intelligence going to replace dermatologists? *Cutis* 2020 Jan;105(1):28-31. [Medline: [32074153](https://pubmed.ncbi.nlm.nih.gov/32074153/)]
11. Börve A, Dahlén Gyllencreutz J, Terstappen K, Johansson Backman E, Aldenbratt A, Danielsson M, et al. Smartphone teledermoscopy referrals: a novel process for improved triage of skin cancer patients. *Acta Derm Venereol* 2015 Feb;95(2):186-190 [FREE Full text] [doi: [10.2340/00015555-1906](https://doi.org/10.2340/00015555-1906)] [Medline: [24923283](https://pubmed.ncbi.nlm.nih.gov/24923283/)]
12. Finnane A, Dallest K, Janda M, Soyer HP. Teledermatology for the diagnosis and management of skin cancer: a systematic review. *JAMA Dermatol* 2017 Mar 01;153(3):319-327. [doi: [10.1001/jamadermatol.2016.4361](https://doi.org/10.1001/jamadermatol.2016.4361)] [Medline: [27926766](https://pubmed.ncbi.nlm.nih.gov/27926766/)]
13. Ferrer RT, Bezares AP, Mañes AL, Mas AV, Gutiérrez IT, Lladó CN, et al. Diagnostic reliability of an asynchronous teledermatology consultation. Article in Spanish. *Aten Primaria* 2009 Oct;41(10):552-557 [FREE Full text] [doi: [10.1016/j.aprim.2008.11.012](https://doi.org/10.1016/j.aprim.2008.11.012)] [Medline: [19500880](https://pubmed.ncbi.nlm.nih.gov/19500880/)]
14. Mounessa JS, Chapman S, Braunberger T, Qin R, Lipoff JB, Dellavalle RP, et al. A systematic review of satisfaction with teledermatology. *J Telemed Telecare* 2018 May;24(4):263-270. [doi: [10.1177/1357633X17696587](https://doi.org/10.1177/1357633X17696587)] [Medline: [28350281](https://pubmed.ncbi.nlm.nih.gov/28350281/)]
15. Vidal-Alaball J, Álamo-Junquera D, López-Aguilá S, García-Altés A. Evaluation of the impact of teledermatology in decreasing the waiting list in the Bages region (2009-2012). Article in Spanish. *Aten Primaria* 2015 May;47(5):320-321 [FREE Full text] [doi: [10.1016/j.aprim.2014.01.009](https://doi.org/10.1016/j.aprim.2014.01.009)] [Medline: [25443766](https://pubmed.ncbi.nlm.nih.gov/25443766/)]
16. Vidal-Alaball J, López Seguí F, Garcia Domingo JL, Flores Mateo G, Sauch Valmaña G, Ruiz-Comellas A, et al. Primary care professionals' acceptance of medical record-based, store and forward provider-to-provider telemedicine in Catalonia:

- results of a web-based survey. *Int J Environ Res Public Health* 2020 Jun 08;17(11):4092 [FREE Full text] [doi: [10.3390/ijerph17114092](https://doi.org/10.3390/ijerph17114092)] [Medline: [32521740](https://pubmed.ncbi.nlm.nih.gov/32521740/)]
17. Lee MK, Rich K. Who is included in human perceptions of AI?: trust and perceived fairness around healthcare AI and cultural mistrust. 2021 May Presented at: CHI '21: CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-14. [doi: [10.1145/3411764.3445570](https://doi.org/10.1145/3411764.3445570)]
  18. Calisto FM, Ferreira A, Nascimento JC, Gonçalves D. Towards touch-based medical image diagnosis annotation. 2017 Oct 17 Presented at: ISS '17: Interactive Surfaces and Spaces; October 17-20, 2017; Brighton, United Kingdom p. 390-395. [doi: [10.1145/3132272.3134111](https://doi.org/10.1145/3132272.3134111)]
  19. Calisto FM, Santiago C, Nunes N, Nascimento JC. Introduction of human-centric AI assistant to aid radiologists for multimodal breast image classification. *Int J Hum Comput Stud* 2021 Jun;150:102607. [doi: [10.1016/j.ijhcs.2021.102607](https://doi.org/10.1016/j.ijhcs.2021.102607)]
  20. Calisto FM, Nunes N, Nascimento JC. BreastScreening: on the use of multi-modality in medical imaging diagnosis. 2020 Sep Presented at: AVI '20: International Conference on Advanced Visual Interfaces; September 28 to October 2, 2020; Salerno, Italy p. 1-5. [doi: [10.1145/3399715.3399744](https://doi.org/10.1145/3399715.3399744)]
  21. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. *Eur Heart J* 2021 Dec 07;42(46):4717-4730 [FREE Full text] [doi: [10.1093/eurheartj/ehab649](https://doi.org/10.1093/eurheartj/ehab649)] [Medline: [34534279](https://pubmed.ncbi.nlm.nih.gov/34534279/)]
  22. Wells A, Patel S, Lee JB, Motaparthy K. Artificial intelligence in dermatopathology: diagnosis, education, and research. *J Cutan Pathol* 2021 Aug 26;48(8):1061-1068. [doi: [10.1111/cup.13954](https://doi.org/10.1111/cup.13954)] [Medline: [33421167](https://pubmed.ncbi.nlm.nih.gov/33421167/)]
  23. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. *J Dermatolog Treat* 2020 Aug;31(5):496-510. [doi: [10.1080/09546634.2019.1682500](https://doi.org/10.1080/09546634.2019.1682500)] [Medline: [31625775](https://pubmed.ncbi.nlm.nih.gov/31625775/)]
  24. Du AX, Emam S, Gniadecki R. Review of machine learning in predicting dermatological outcomes. *Front Med (Lausanne)* 2020 Jun 12;7:266 [FREE Full text] [doi: [10.3389/fmed.2020.00266](https://doi.org/10.3389/fmed.2020.00266)] [Medline: [32596246](https://pubmed.ncbi.nlm.nih.gov/32596246/)]
  25. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov IV. Artificial intelligence applications in dermatology: where do we stand? *Front Med (Lausanne)* 2020 Mar 31;7:100 [FREE Full text] [doi: [10.3389/fmed.2020.00100](https://doi.org/10.3389/fmed.2020.00100)] [Medline: [32296706](https://pubmed.ncbi.nlm.nih.gov/32296706/)]
  26. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017 Feb 02;542(7639):115-118 [FREE Full text] [doi: [10.1038/nature21056](https://doi.org/10.1038/nature21056)] [Medline: [28117445](https://pubmed.ncbi.nlm.nih.gov/28117445/)]
  27. Morid MA, Borjali A, Del Fiol G. A scoping review of transfer learning research on medical image analysis using ImageNet. *Comput Biol Med* 2021 Jan;128:104115. [doi: [10.1016/j.combiomed.2020.104115](https://doi.org/10.1016/j.combiomed.2020.104115)] [Medline: [33227578](https://pubmed.ncbi.nlm.nih.gov/33227578/)]
  28. Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, et al. A deep learning system for differential diagnosis of skin diseases. *Nat Med* 2020 Jun 18;26(6):900-908. [doi: [10.1038/s41591-020-0842-3](https://doi.org/10.1038/s41591-020-0842-3)] [Medline: [32424212](https://pubmed.ncbi.nlm.nih.gov/32424212/)]
  29. Servei Català de la Salut. Activitat assistencial de la xarxa sanitària de Catalunya, any 2012. Departament de Salut. URL: <http://hdl.handle.net/11351/1025> [accessed 2022-08-11]
  30. 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 Jul;20(7):938-947 [FREE Full text] [doi: [10.1016/S1470-2045\(19\)30333-X](https://doi.org/10.1016/S1470-2045(19)30333-X)] [Medline: [31201137](https://pubmed.ncbi.nlm.nih.gov/31201137/)]
  31. Kamulegeya LH, Okello M, Bwanika JM, Musunguzi D, Lubega W, Rusoke D, et al. Using artificial intelligence on dermatology conditions in Uganda: a case for diversity in training data sets for machine learning. *bioRxiv*. Preprint posted online October 31, 2019 [FREE Full text] [doi: [10.1101/826057](https://doi.org/10.1101/826057)]
  32. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, et al. Deep neural networks are superior to dermatologists in melanoma image classification. *Eur J Cancer* 2019 Sep;119:11-17 [FREE Full text] [doi: [10.1016/j.ejca.2019.05.023](https://doi.org/10.1016/j.ejca.2019.05.023)] [Medline: [31401469](https://pubmed.ncbi.nlm.nih.gov/31401469/)]
  33. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Reader study level-I and level-II Groups, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol* 2018 Aug 01;29(8):1836-1842 [FREE Full text] [doi: [10.1093/annonc/mdy166](https://doi.org/10.1093/annonc/mdy166)] [Medline: [29846502](https://pubmed.ncbi.nlm.nih.gov/29846502/)]

## Abbreviations

- AI:** artificial intelligence  
**API:** application programming interface  
**CNN:** convolutional neural network  
**GP:** general practitioner  
**PCC:** primary care center

*Edited by T Leung; submitted 24.02.22; peer-reviewed by R Kaczmarczyk, FM Calisto; comments to author 04.05.22; revised version received 11.05.22; accepted 12.05.22; published 31.08.22*

*Please cite as:*

*Escalé-Besa A, Fuster-Casanovas A, Börve A, Yélamos O, Fustà-Novell X, Esquiús Rafat M, Marin-Gomez FX, Vidal-Alaball J*  
*Using Artificial Intelligence as a Diagnostic Decision Support Tool in Skin Disease: Protocol for an Observational Prospective Cohort Study*

*JMIR Res Protoc 2022;11(8):e37531*

URL: <https://www.researchprotocols.org/2022/8/e37531>

doi: [10.2196/37531](https://doi.org/10.2196/37531)

PMID:

©Anna Escalé-Besa, Aina Fuster-Casanovas, Alexander Börve, Oriol Yélamos, Xavier Fustà-Novell, Mireia Esquiús Rafat, Francesc X Marin-Gomez, Josep Vidal-Alaball. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 31.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.