Clinical trial title

Estudio de validación clínica de un sistema DAO con algoritmos de inteligencia artificial para la detección temprana no invasiva de melanoma cutáneo in vivo.

Clinical validation study of a CAD system with artificial intelligence algorithms for early noninvasive detection of in vivo cutaneous melanoma.

Protocol code: LEGIT_MC_EVCDAO_2019

Protocol version: Version 3.0, date October 28th 2021

ÍNDICE

1. GENERAL INFORMATION

1.1 Study identification

Title: Clinical validation study of a CAD system with artificial intelligence algorithms for early noninvasive detection of in vivo cutaneous melanoma.

Code: LEGIT_MC_EVCDAO_2019

Study design: Cross-sectional analytical observational study of clinical case series.

Product under Research: Legit.Health Plus

Version and date: Version 3.0, date October 28, 2021

1.2 Principal Investigators/Collaborators and Involved Centers

Principal Investigators: Dr. Jesús Gardeazabal García[a], Dra. Rosa María Izu Belloso[b]

Collaborators: Dr. Juan Antonio Ratón Nieto[a], Dra. Ana Sánchez Díez[b], Alfonso Medela[c], Mireya Fernández[c], Andy Aguilar[c], Taig Mac Carthy[c]

Involved Centers: [a] Servicio de Dermatología, Hospital Universitario Cruces, [b] Servicio Dermatología, Hospital Universitario Basurto, [c] AI LABS GROUP SL

2. INTRODUCTION AND JUSTIFICATION

The Spanish Society of Medical Oncology (SEOM) has estimated that the number of new cancer cases diagnosed in Spain in 2019 will reach 277,234, 12% more than in 2015 when 247,771 were diagnosed. Among the different types of cancer, cutaneous melanoma (CM) is the type of skin cancer that causes the most deaths, with a significant increase in incidence and mortality in recent decades. It is characterized by a rapidly increasing incidence rate among Caucasian populations and tens of thousands of people worldwide die each year from this cancer. Melanoma is one of the most aggressive malignancies and rapidly metastasizes to distant organs.

When it progresses to the metastatic stage, it establishes powerful mechanisms to resist chemotherapy and radiotherapy, which hinders the efficacy of current medical therapies[1]. However, when detected early, melanoma is treatable in almost all cases with simple surgical excision[2]. On the other hand, there are also benign types of pigmented, such as; skin lesions, moles, which are natural parts of the skin, benign and malignant pigmented lesions that share similar visual characteristics that make the difference between them a challenging problem for specialists, especially non-specialists. This problem is particularly significant during a naked eye examination, because early stage melanomas often resemble benign lesions. A person with a suspicious pigmented skin lesion will go through several steps before a definitive diagnosis of melanoma: self-assessment, evaluation by a primary care physician, evaluation by a specialist, excision, and evaluation by histopathology. Due to low public awareness of the importance of skin cancer prevention and insufficient access to dermatologists in many regions worldwide, melanoma is often diagnosed only after a tumor grows to a medium size.

In light of the above data, prevention and early diagnosis of melanoma have become an extremely important issue. In recent years, there has been increasing demand to develop **computer-aided diagnostics** (CAD) and systems that facilitate the early detection of melanoma that could be applied by non-experts and the general public [3].

Advances in image recognition and artificial intelligence have set in motion innovations in the diagnosis of skin lesions. With appropriate development and proper evaluation, technology could improve diagnostic accuracy. It has been demonstrated that through artificial intelligence (AI) algorithms it is possible to classify photographs of lesions,

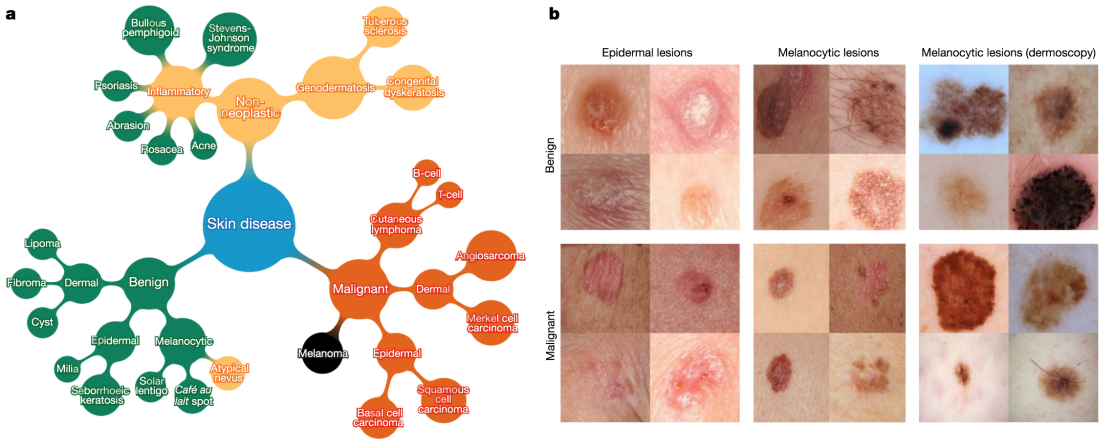
---

[1] «M. Campioni, D. Santini, G. Tonini et al., "Role of Apaf-1, a key regulator of apoptosis, in melanoma progression and chemoresistance,"» Experimental Dermatology, vol.14, no.11, pp. 811–818, 2005

[2] «M. Campioni, D. Santini, G. Tonini et al., "Role of Apaf-1, a key regulator of apoptosis, in melanoma progression and chemoresistance,"» Experimental Dermatology, vol.14, no.11, pp. 811–818, 2005

[3] «R. Tadeusiewicz, How intelligent should be system for image analysis?» in Innovations in Intelligent Image Analysis, H. Kwásnicka and L. C. Jain, Eds., vol. 339 of Studies in Computational Intelligence, pp. V–X, Springer Verlag, Berlin, Heidelberg, 2011.

including melanoma, with a level of competence comparable to that of dermatologists. [4] [5]



Schematic illustration of the diagnostic taxonomy of image sets.[6].

The advent of machine vision has revolutionized our understanding of this pathology and presents an enormous opportunity for diagnosis.

Among the trends that have caused an improvement in patient survival, preventive activities and early diagnosis campaigns stand out, among other factors. Therefore, based on our previous research, we propose a **system developed through Artificial Intelligence to evaluate the malignancy of a skin lesion,** as well as to differentiate between micro melanomas and different moles and lesions developed on the skin among them nevus and lentigines.

This study aims to clinically validate the diagnosis of cutaneous melanoma through machine vision and machine learning applications.

[4] «Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks», Nature. 2017; 542(7639): 115-118. https://dx.doi.org/10.1038/nature21056

[5] «Haenssle HA, Fink C, Schneiderbauer R, et al; Reader study level-I and level-II Groups. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists.», Ann Oncol. 2018;29(8):1836-1842. https://dx.doi.org/10.1093/annonc/mdy166

[6] Dermatologist-level classification of skin cancer with deep neural networks Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

3. OBJECTIVE AND PURPOSE OF THE PROJECT

3.1 Hypothesis

The DAO system with machine vision allows early and non-invasive diagnosis of cutaneous melanoma in-vivo.

3.2 Main objective

A. To validate that the artificial intelligence algorithm developed by AI LABS GROUP SL for the identification of cutaneous melanoma in images of lesions taken with a dermatoscopic camera achieves the following values:
    a. an AUC greater than 0.8
    b. a **sensitivity** of 80% or higher
    c. a **specificity** of 70% or higher

3.3 Secondary objectives

A. To compare the performance of the artificial intelligence algorithm developed by AI LABS GROUP SL with the performance of healthcare professionals of different specializations.:
    a. Dermatologists
    b. Primary care physicians

    In this first phase, only validation by dermatologists will be carried out..
B. Validate the usefulness and feasibility of the artificial intelligence algorithm developed by AI LABS GROUP SL in adverse environments with severe technical limitations, such as lack of instrumentation or lack of internet connection..

4. METHODOLOGY

4.1 Type Design

This is an analytical observational case series study for the performance of a diagnostic test study. Measurements are performed in a single case, so it is a cross-sectional study.

4.2 Study Period

This study is estimated to have a recruitment period of 10 months for the inclusion of the first 40 patients. The recruitment period is extended by 12 months for the inclusion of patients up to a total of 200, with a minimum of 40 melanomas.

The total duration of the study is estimated at 36 months, including the time required after recruitment of the last subject for closing and editing the database, data analysis and preparation of the final study report.

4.3 Study Population

Patients with skin lesions with suspected malignancy seen at the Dermatology Department of the Hospital Universitario Cruces and Hospital Universitario Basurto.

4.4 Selection Criteria

4.4.1 Inclusion Criteria

- Patients with skin lesions with suspected malignancy
- Age over 18 years old
- Patients who consent to participate in the study by signing the Informed Consent form.

4.4.2 Exclusion Criteria

- Patients under 18 years of age

5. SAMPLE DETERMINATION AND SAMPLING

To achieve a precision of 5.0% in the estimation of a proportion using a bilateral 95.0% Normal asymptotic confidence interval, assuming that the prevalence of melanoma in Spain is 2.40% (Las cifras del cáncer en España, 2019; SEOM), it will be necessary to include 36 patients in the study. Taking into account that the expected dropout rate is 10.0% it would be necessary to recruit 40 patients in the study.

Once the data and results of the first 40 patients included in the study have been analyzed, it is considered necessary to expand the sample to a total of 200 patients, of which at least 40 will be biopsied melanomas and the rest pigmented lesions with suspected malignancy.

6. SCIENTIFIC DESCRIPTION OF THE VARIABLES

6.1 Principal Variables

The main variable is intended to estimate the diagnostic efficacy of the DAO system with artificial vision, taking as Gold Standard the diagnosis of an expert dermatologist with more than 10 years of experience in dermatology, as is the case of the two Principal Investigators of the study.

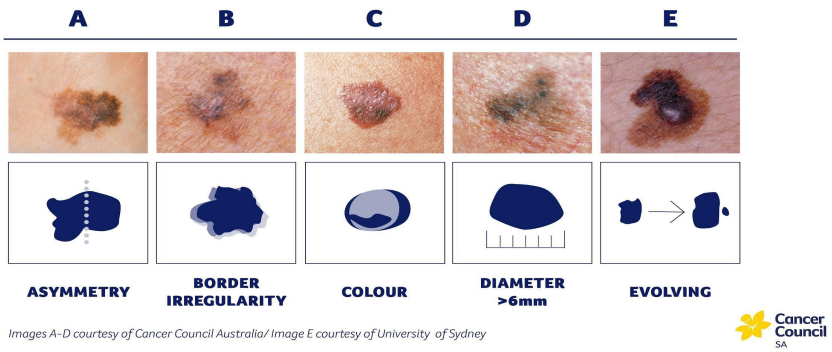
6.2 Secondary Variables

The diagnoses made by physicians on images of skin lesions in patients will be collected.

These diagnoses are based on the ABCDE[7] system, commonly used in dermatology to make a diagnosis of cutaneous melanoma.

- Asymmetry
- Irregularity of edges
- Color variation
- Diameter
- Evolution

[7] Do, Thanh-Toan & Pomponiu, Victor & Zhou, Yiren & Hoang, Tuan & Chen, Z. & Cheung, Ngai-Man & Koh, Dawn & Tan, Aaron & Tan, S. (2017). Accessible Melanoma Detection using Smartphones and Mobile Image Analysis. IEEE Transactions on Multimedia.



In addition to the ABCDE system, we analyzed one additional variable:

- Location of the lesion on the body.

A validated usability survey (System Usability Scale) and an "ad hoc" satisfaction questionnaire will be collected from all end users of the study.

7. DESCRIPTION OF CLINICAL MANAGEMENT AND STUDY PLAN

For the validation of the DAO system for the early non-invasive detection of cutaneous melanoma in vivo, we need a certain number of photographs of skin lesions from a primary source. These photographs will be taken at the dermatology offices of the researchers participating in this study, both at the Hospital Universitario Basurto and the Hospital Universitario Cruces. The research team also performs the task of labeling the photographs according to their medical diagnosis.

The research team will assist the clinical investigators in the process of taking the photographs and will supervise the task of labeling the photographs, to ensure that both the photographs and the labeling nomenclature meet the technical criteria necessary for their subsequent analysis.

The set of photographs with their respective labeled information will be referred to as "the dataset". Schematically, the process for obtaining the dataset follows the following order:

1. The research team selects the patients who meet the inclusion criteria established in point 4.4.1.

2. The research team informs the patient of the existence of the study and requests his/her voluntary participation.

3. The patient signs the Informed Consent document provided by the research team, detailing the procedures and purpose of the study, as well as his or her rights as a participant.

4. The research team assigns a unique numerical identifier (ID) to each patient.

5. The research team asks the participants to answer the following questions:

    - Demographic information
        - Age
        - Sex
        - Municipality of residence
        - Ethnicity

6. The investigating team proceeds to take a photograph of the skin lesions, following the following technical indications:

    - Uncompressed image format, such as PNG, HEIC or TIFF.
    - Taken with the DermLite Foto X dermatoscope of the 3Gen Inc. brand.
    - Taken from a smartphone with the following characteristics
        - With a camera with a minimum resolution of not less than 13 megapixels.
        - Taken with one of the following models:
            - Google Pixel 3 and Google Pixel 3 XL.

        - Samsung Galaxy Note 10, Samsung Galaxy S10, Samsung Galaxy S10E
        - iPhone X and below
    - Disabling all image post-processing, such as HDR, portrait mode, color filters or digital zoom.

7. The research team names the images with a nomenclature that follows the following format and stores them in a directory on their computer:

    ID_AAAAMMDD_hhmmss
    Example: 021_20191123_093114

8. On a monthly basis, the research team collects the images and verifies their correctness. If any image is not of sufficient quality, the investigator repeats the photograph.

9. On a monthly basis, the research team collects diagnostic data from the expert dermatologists.

10. The research team encodes the information received from the patients in ONE-HOT format for subsequent automatic processing.

11. The images with the data obtained will be saved with a randomly generated indicator in order to make the samples anonymous.

12. The pathological anatomy results of the lesions studied will be collected when appropriate according to clinical criteria and following standard clinical practice. This analysis of biopsy samples obtained in dermatology consultations is performed according to the usual pattern and no extraordinary or unusual biopsy or pathological anatomy analysis will be performed in this study.

13. At the end of the study, the research team asks two additional dermatologists to assign their own diagnoses, independently and without prior knowledge of the case. The aim is to see the concordance between different observers without diagnostic means.

8. DESCRIPTION OF SAMPLE HANDLING

Due to the absence of the need to collect biological samples for the study, this section does not require attention. It has already been specified above that no samples are collected for this study. Pathological anatomy analysis is standard practice.

9. Safety

Subjects recruited in this study will not be exposed to any procedure that could jeopardize their safety.

10. DATA COLLECTION AND MANAGEMENT

10.1 Source Data Identification

Source documents are understood to be all the observations or notes recorded on the clinical interventions, as well as all the reports and notes necessary for the reconstruction and evaluation of the study's data collection notebook.

Basically, although not exclusively, the source documents are composed of the documents and annotations that form part of the patient's Clinical History and the different surveys to be collected in the center.

Whenever possible, the original document should be kept as the source document; however, it is acceptable to provide a photocopy as long as it is clear, legible and an exact duplicate of the original document.

10.2 Data Quality Assurance

The Principal Investigator is responsible for reviewing and approving the protocol, signing the Principal Investigator commitment, guaranteeing that the persons involved in the center will respect the confidentiality of patient information and protect personal data, and reviewing and approving the final study report together with the sponsor. All the clinical members of the research team will evaluate the eligibility of the patients in the study, inform and request written informed consent, collect the source data of the study in the clinical record and transfer them to the Data Collection Notebook (CRD).

10.3 Data Management

The management of the collection and processing of the study data will be carried out through the design of a Data Collection Notebook (CRD) in paper format, in which the researchers assigned to this task will enter the source data of each patient participating in the study.

Current legislation will be complied with in terms of data confidentiality protection (European Regulation 2016/679, of 27 April, on the protection of natural persons with regard to the processing of personal data and the free movement of such data and Organic Law 3/2018, of 5 December, on the Protection of Personal Data and guarantee of digital rights). For this purpose, each patient will receive an alphanumeric identification code in the study that will not include any data allowing personal identification (coded CRD). The Principal Investigator will have an independent list that will allow the connection of the identification codes of the patients participating in the study with their clinical and personal data. This document will be filed in a secure area with restricted access, under the custody of the Principal Investigator and will never

leave the center.

Once the paper CRDs are completed and closed by the Principal Investigator, the data will be transferred to a database.

As in the CRDs, the Database will comply with current legislation in terms of data confidentiality protection (European Regulation 2016/679, of 27 April, on the protection of natural persons with regard to the processing of personal data and the free movement of such data and Organic Law 3/2018, of 5 December, on the Protection of Personal Data and guarantee of digital rights) in which no data allowing personal identification of patients will be included.

The transfer of data from the paper CRD to the electronic Database will be carried out using the double data entry technique. This will be done by the researchers collaborating in the project.

The data will be managed and tabulated with consistency rules and logical ranges to control inconsistencies during data tabulation. A validation process of the clinical data will be carried out by running computer filters based on validation rules, which will automatically identify missing values or inconsistencies of clinical data according to the protocol. Additionally, manual editing and validation will be performed using descriptive and exploratory statistical techniques to complement the detection of logical errors and inconsistent values.

The Database will be considered closed upon completion of all Data Management processes and satisfactory resolution of data discrepancies and errors. Any changes to the databases after closure can only be made after written agreement between the promoter and the technical coordinators of the project.

11. STATISTICAL ANALYSIS

Each variable will be characterized using frequency distributions for qualitative variables and central tendency statistics such as mean and median and variability statistics such as standard deviation (S.D.) or interquartile range for quantitative variables according to their distributional characteristics.

Between-group and within-group comparisons will be made using parametric tests whenever the distributional characteristics of the data allow it. For intergroup comparisons, one- and two-factor Analysis of Variance techniques will be used with post-hoc comparisons if significant overall differences are detected. To evaluate intra-group changes, Student's t-test for related samples or Analysis of Variance/ANOVA with repeated measures will be used if the theoretical assumptions of the model are supported by the data. Otherwise, more flexible models (GEE) that allow incorporating different autocorrelation structures of the data will be fitted.

Comparisons between groups with respect to qualitative variables will be carried out by means of contingency tables and Fisher's exact or Chi-square tests. The probability of type I error will not be adjusted for multiple comparisons. The level of statistical significance in the contrasts (alpha) will be 5 percent with bilateral contrasts.

Comparisons between two continuous variables will be made using Pearson's or Spearman's correlation, depending on the distributional characteristics.

All the analyses described so far will be performed based on the needs of **descriptive results** of the sample. Interobserver concordance analyses will also be performed by estimating the kappa coefficient.

A **study of diagnostic tests** will be performed being the main measures in the results: sensitivity, specificity, positive and negative predictive values (PPV and NPV) and likelihood ratios (LR+ and LR-). The diagnosis of the lesion will be made according to the clinical judgment of expert dermatologists, such as the Principal Investigators, (considered the Gold Standard).

Analyses will be performed using appropriate statistical software, SPSS version 23.0 and STATA 13.0. Values of $p<0.05$ will be considered significant.

12. ETHICAL AND LEGAL CONSIDERATIONS

The conduct of the study will conform to international Good Clinical Practice guidelines, to the Declaration of Helsinki in its latest active amendment, and to international and national rules and regulations and will not be initiated until approval has been obtained from the Basque Country Committee on Drug Research Ethics (CEIm de Euskadi). Any modification of this protocol will be reviewed and approved by the Principal Investigator and must be evaluated by the CEIm of Euskadi for approval before including subjects in a modified protocol.

The study will be conducted according to European Regulation 2016/679, of 27 April, on the protection of natural persons with regard to the processing of personal data and the free movement of such data and Organic Law 3/2018, of 5 December, on the Protection of Personal Data and guarantee of digital rights with regard to data processing in which no data allowing personal identification of subjects will be included, the information being managed in an encrypted manner.

Patients will be informed orally and in writing about all the information related to the study and adapted to their level of understanding. A copy of the consent form and information sheet will be provided to the patient. The investigator should allow the patient the necessary time to ask questions about the details of the study.

The preparation of the informed consent form is the responsibility of the Principal Investigator. This form must include all the elements required by the International Conference of Harmonization (ICH), current regulatory guidelines, and comply with the GCP Guidelines and the ethical principles that originate from the Declaration of Helsinki.

The investigator or the Principal Investigator's designee will keep the original signed informed consent form in a secure restricted access area in the custody of the Principal Investigator and will never leave the facility and will provide a copy of the original signed consent form to the patient.

13. STUDY LIMITATIONS

The main limitation of machine learning lies in the quantity and quality of the images collected. Variability in illumination, color, shape, size and focus are determinants, in addition to the number of images per lesion. This means that a large variability within the same lesion and an insufficient number of images to reflect that variability can result in lower than expected accuracy.

Detection of skin lesions beyond that of cutaneous melanoma implies the need for a larger number of samples to train the artificial intelligence algorithms.

Due to the need for a balanced dataset, i.e., same number of melanoma and non-melanoma images, we consider it necessary to collect cases of nevus and/or other types of skin lesions if necessary.

For this reason, we believe that the proposed n = 40 is low considering all the types of lesions we want to analyze, for which we would need a much larger n, of approximately 200 people, of which at least 40 suffer from cutaneous melanoma.

For this reason, a preliminary study will be carried out with 40 patients and after the analysis, the need to expand the study sample will be decided. This study is intended to be a pilot study to be carried out in one year.

After analyzing the data from the pilot study with the first 40 patients included in the study, the results obtained have indicated that despite the good diagnostic capacity of the DAO system, the results of discrimination between melanoma and other pathologies is not representative of daily clinical practice, since it only and exclusively includes cases with a high suspicion of malignancy.

Therefore, it is proposed to continue with the study and extend the sample to a total of 200 patients. In this way, it will be possible to validate the system in real cases that the dermatologist receives on a daily basis and to determine its ability to discriminate beyond the extreme cases.