Performance Analysis of an AI-Based Medical Device for Optimizing Clinical Workflow in Patients with Pigmented Lesions

Miguel Sánchez-Viera[1], Pablo López[1], Isabel del Campo[1], Miguel Sanchez-Illescas[1], Taig Mac Carthy[2], Andy Aguilar[2], Jordi Barrachina[2] y Alfonso Medela[3]

[1]Instituto de Dermatología Integral (IDEI), Madrid, Spain
[2]Department of Clinical Endpoint Innovation, Legit.Health, Bilbao, Spain
[3]Department of Medical Data Science, Legit.Health, Bilbao, Spain

Abstract

- Artificial Intelligence (AI) has frequently demonstrated high levels of diagnostic accuracy in image-based diagnosis. Nevertheless, many **AI algorithms remain underutilized** in clinical settings **due to regulatory non-compliance**.
- We present the initial results of a project (**DERMATIA**) to evaluate the performance of an AI-Based medical device (**Legit.Health**). In this phase, we evaluated the device's performance in identifying malignant lesions from dermatoscopic images, achieving an area under the curve (AUC) of 0.95, with a **sensitivity of 100%** and a **specificity of 71%**. **Dermatologists** had an AUC of 0.96, with a **sensitivity of 100%** and a **specificity of 32%**.
- The medical device demonstrated high sensitivity and specificity in detecting malignant lesions through dermatoscopic images. Its implementation could streamline the diagnosis of malignant lesions and **improve clinical workflow**, potentially increasing the survival and quality of life of affected patients.

Introduction and objective

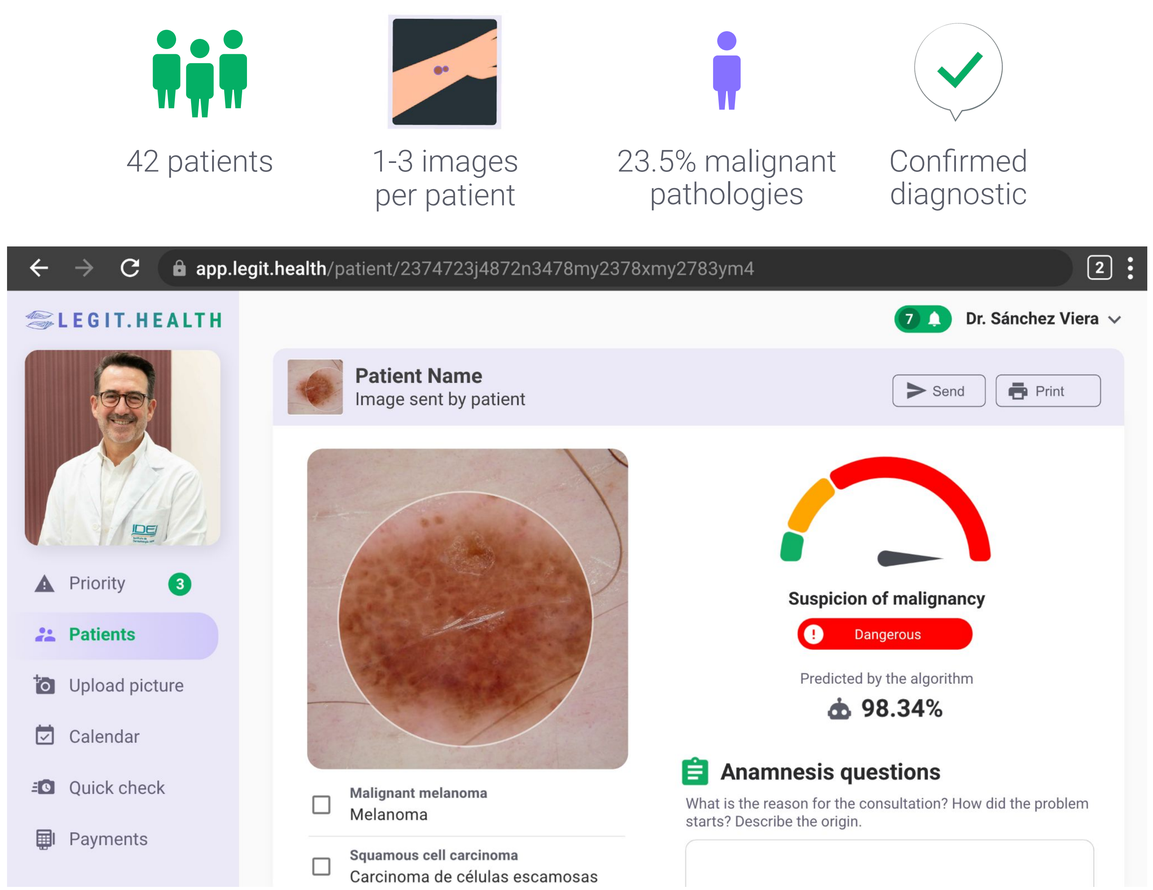
Artificial Intelligence (AI) has frequently demonstrated high level of diagnostic accuracy in various aspects of image-based diagnosis. Nevertheless, many AI algorithms remain underutilized in clinical settings due to regulatory non-compliance.

This study presents the outcomes of the DERMATIA* project, focusing on evaluating the efficacy of a CE certified AI-Based medical device for diagnosing and triaging patients with pigmented lesions. Our main objective was to assess the device's capacity to identify malignant lesions using smartphone-captured clinical images and compare its performance to that of expert dermatologists.

Materials and methods

We collected imaging data from 42 patients presenting with suspected malignant pigmented lesions at the IDEI private dermatology clinic. During the initial evaluation, the dermatologist provided clinical assessments, annotated diagnoses, recommended further actions (e.g., follow-up, referral, or excision), and assigned a suspicion score for malignancy on a scale from 1 to 10. The dermatologist then captured 1 to 3 images of each lesion using a smartphone camera and the medical device. Additionally, the dermatologist recorded the malignancy probability detected by the device on a scale from 0 to 100.

All patients also underwent biopsy procedures to obtain histopathological diagnoses, which served as the gold standard in this study. We compared the dermatologist's clinical suspicion of malignancy with the device-generated assessments. Both clinical and device diagnoses were then juxtaposed against biopsy results to determine diagnostic accuracy metrics.



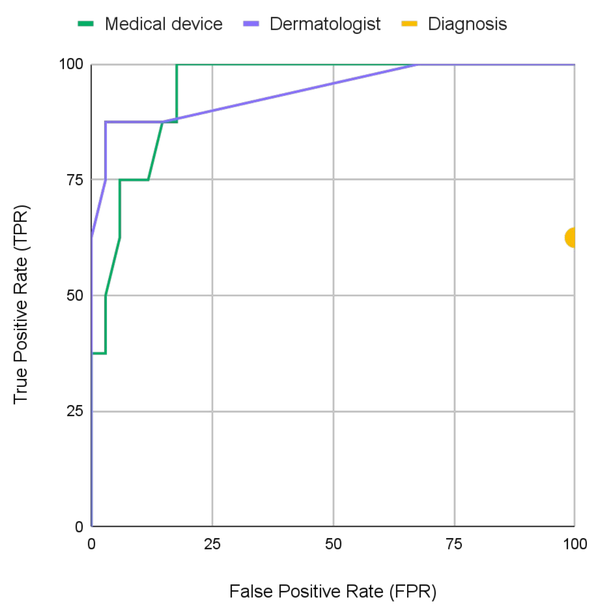
*DERMATIA (Implementation and multicenter validation of a platform with artificial intelligence for clinical decision support and patient management in Dermatology); Project co-financed by the European Union (NextGenerationEU) through the Public Business Entity Red.es (Secretary of State for Digitalization and Artificial Intelligence, Ministry of Economic Affairs and Digital Transformation) within the framework of the 2021 Call for grants aimed at research and development projects in artificial intelligence and other digital technologies and their integration into value chains C005/21-ED, with file number 2021/C005/00154001.

Results

The findings demonstrated that both the medical device and the dermatologist exhibited excellent discrimination capacity for malignancy, with areas under the curve (AUC) of 0.95 and 0.96, respectively. Notably, the device achieved a sensitivity of 100% with a specificity of 71%, while the dermatologist achieved a sensitivity of 100% with a specificity of 32%. Conversely, at 100% specificity, the device exhibited a sensitivity of 37.5%, whereas the dermatologist's sensitivity was 62.5%.

In terms of correlation with the gold standard, the device exhibited a coefficient of 0.70, indicating a strong association, while the dermatologist achieved a slightly higher coefficient of 0.81. Furthermore, the correlation between the device and the dermatologist assessments was also robust, with a coefficient of 0.71.

Figura 1. AUC-ROC graph



Note: This figure shows the performance analysis of Legit.Health in 76 patients from the Instituto de Dermatología Integral (IDEI) using the ROC Curve. The y-axis represents the true positive rate (TPR), while the x-axis shows the false positive rate (FPR).

Tabla 1. Sensitivity and Specificity

	Dermatologist		Legit.Health	
Estimator	**Diagnosis**	**Threshold = 1**	**Threshold = 15%**	**Threshold = 85%**
Sensitivity	62.50%	100.00%	100.00%	37.50%
Specificity	100.00%	32.00%	79.00%	100.00%

Note: Sensitivity and specificity of dermatologists using both their clinical diagnosis and malignancy evaluation on a scale from 0 to 10, compared with the medical device.

Conclusion

The medical device demonstrated high performance, comparable to that of an expert dermatologist, despite the dataset bias, which includes only lesions with sufficient suspicion to justify a biopsy. Our analysis concludes that it exhibited high sensitivity and specificity in detecting malignant lesions through clinical images. Its implementation could streamline the diagnosis of malignant lesions, improve clinical workflow, and potentially increase the survival and quality of life of affected patients.

Funded by the European Union
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
red.es
Plan de Recuperación, Transformación y Resiliencia

LEGIT.HEALTH

IDEI