

# DEVELOPMENT AND ACCURACY OF AN ARTIFICIAL INTELLIGENCE ALGORITHM FOR ACNE EVALUATION

B. DRENO<sup>1</sup>, D. MOYAL<sup>2</sup>, M. BENZAQUEN<sup>3</sup>, A. KHAMMARI<sup>1</sup>, G. LE DANTEC<sup>2</sup>, S. SEITE<sup>2</sup>

<sup>1</sup>Department of Dermatology, CHU Nantes, CIC1413, CRCINA, University Nantes, Nantes, France

<sup>2</sup>La Roche-Posay Dermatological Laboratories, Levallois-Perret, France

<sup>3</sup>Department of Dermatology, Hôpital Nord, Assistance Publique - Hôpitaux de Marseille, Aix-Marseille University, Marseille, France

## INTRODUCTION

Smartphone applications have been proposed as diagnostic self-monitoring tools. Acne is a very common chronic multi-factorial dermatosis, evolving by flare-ups and affecting mainly young adults. The assessment of acne severity and lesions identification is important for the therapeutic choice (treatments depend on severity and type of acne), the follow-up of patients improvement/evolution and the evaluation of treatment efficacy. However, the access to dermatologists is sometimes difficult, the follow-up of patients irregular and the observance of treatments poor.

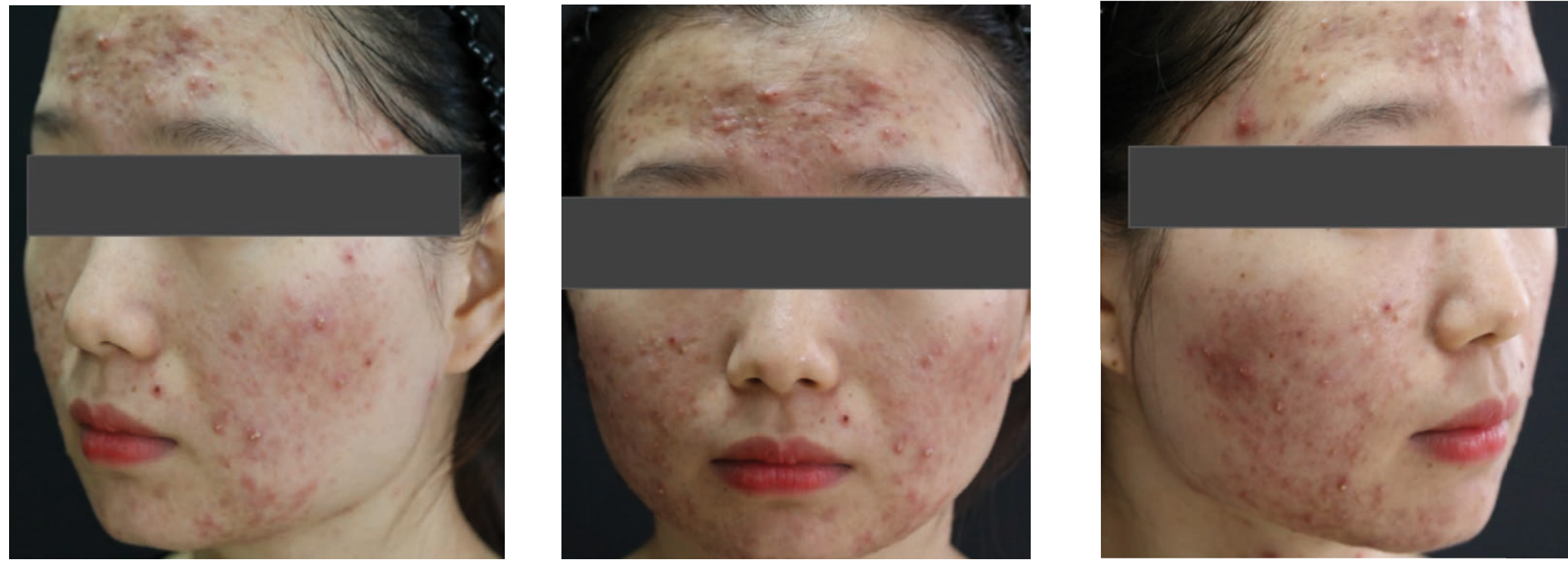
## OBJECTIVE

We have developed an artificial intelligence algorithm for acne severity assessment based on GEA scale<sup>1</sup> and for acne lesions identification which was evaluated compared to clinical diagnosis by trained dermatologists.

## METHODS

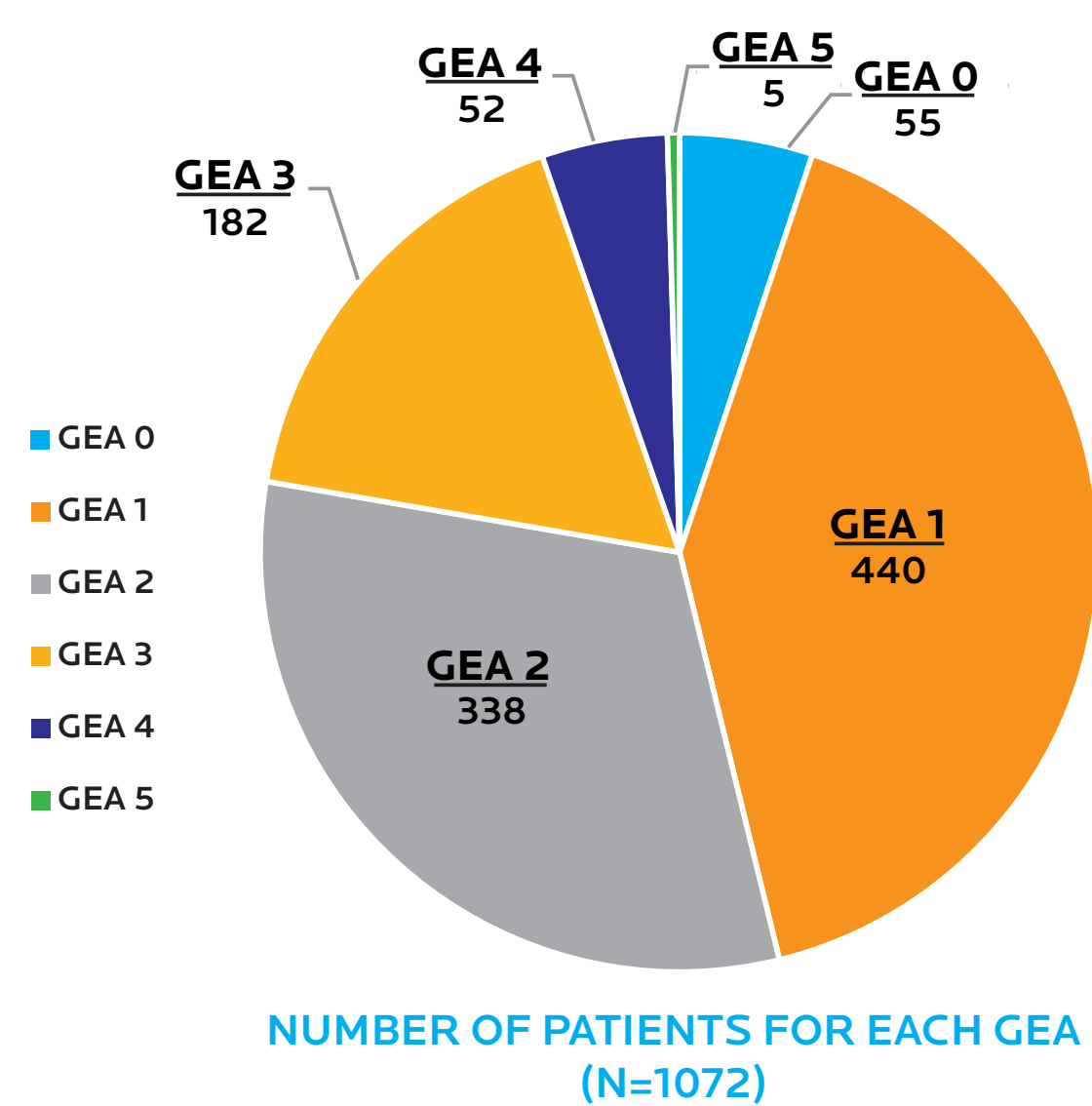
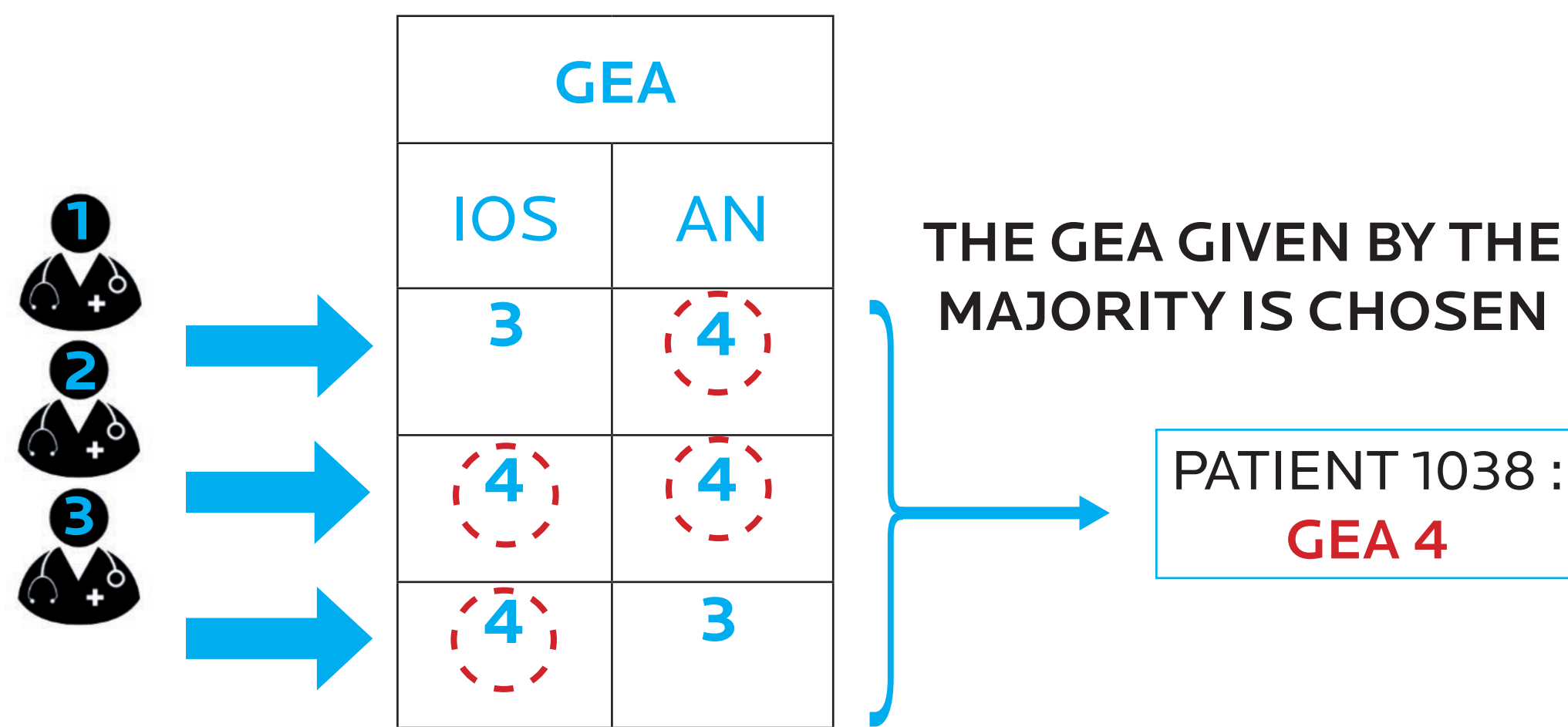
The development of the algorithm was done in five steps.

**Step 1 - Get data:** Collection of 5972 images (face, right and left profiles) from 1072 acne patients using 1 or 2 types of smartphone for each patient - IOS (n=2933) and Android (n=3039) systems. Patients were included from France (32.7%), South Africa (46.5%), China (14.4%), India (5.1%) and Brazil (1.3%) after having signed an informed consent.



Patients were 35% men and 65% women, aged of 23.9+/-9.2 years old and from 3 main types of ethnicity (Asian, Black African and Caucasian) with all phototypes.

**Step 2 - Clean, prepare and manipulate data:** Three trained dermatologists assessed GEA for each patient on images. For each patient, the GEA given by the majority was used to train the algorithm :



Lesions (retentional and inflammatory) and PIHP (postinflammatory hyperpigmentation scars) identification was performed by a dermatologist on images using a tagging tool.

	NUMBER OF AREAS TAGGED BY DERMATOLOGIST
Non-inflammatory lesions	N= 7603
Inflammatory lesions	N= 2939
PIHP	N= 5702

**Step 3 - Train model:** 4958 images corresponding to 903 patients associated with their GEA grade and images with tagged lesions and PIHP were used to train the algorithm. Participants with an acne grade of "4" or "5" were pooled in Grade "4+". Image classification and segmentation techniques based on deep learning were applied to develop the algorithm.

**Step 4 - Test data:** The algorithm was tested internally on a set of 1014 images from 169 acne patients. For lesions and PIHP identification, results were submitted to the dermatologist for correction.



Lesions identification by the algorithm



Correction of the algorithm's results by the dermatologist

In addition, a clinical study was also achieved on 53 acneic patients in order to compare GEA grading and lesions identification, performed by 3 dermatologists in face to face, on photos and by the algorithm.

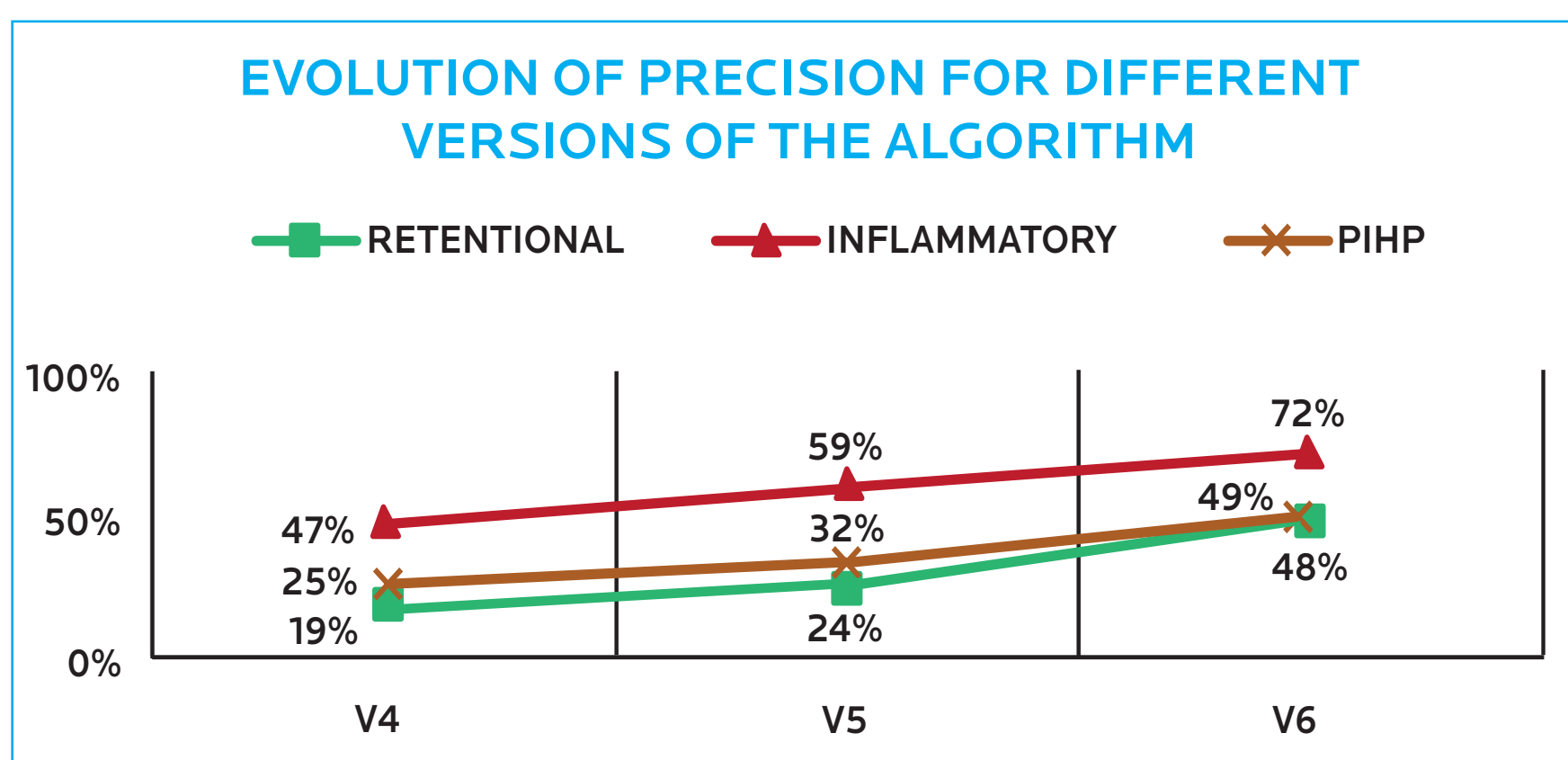
**Step 5 - Improvement of the algorithm:** based on the results obtained in step 4, new versions of algorithms were proposed and tested.

## RESULTS

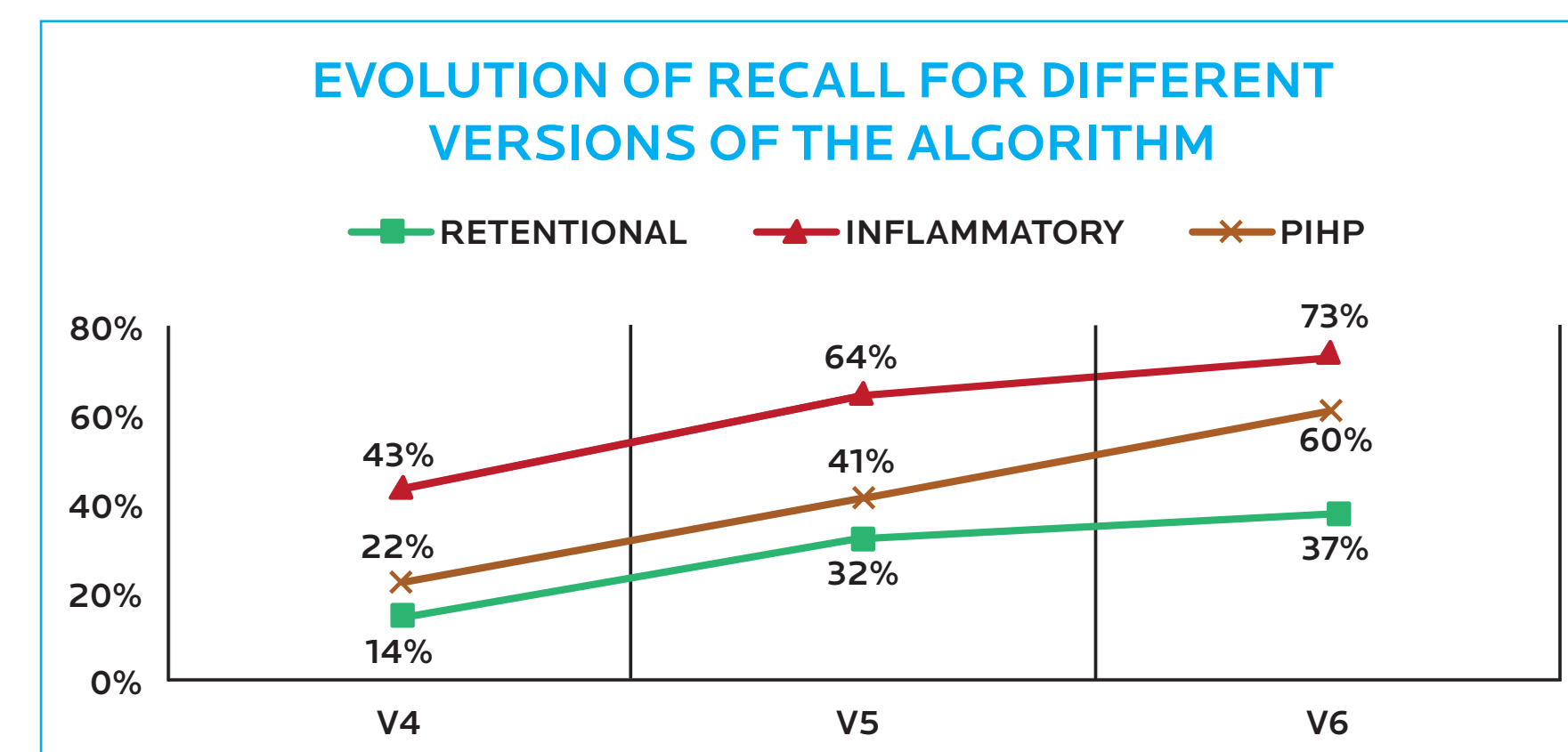
After improvement of the algorithm, the GEA grading provided by the algorithm reached 68% and was similar to the one provided by the majority (2 or 3) of dermatologists.

Precision, recall and F1 score (the weighted average of precision and recall) were evaluated for the 2 types of lesions and PIHP.

To assess the improvement of the algorithm regarding different versions, we tested it on the set of patients' photos from the clinical study.



**PRECISION:**  
how many lesions detected and identified by the algorithm are correct regarding dermatologist identification?



**RECALL:**  
how many lesions identified by the dermatologist are also correctly detected and identified by the algorithm?

The final F1 scores obtained were:

**F1 Score (0 to 1) is the weighted average of precision and recall**  
**For inflammatory lesions: 84%**  
**For PIHP: 72%**  
**For retentional lesions: 61%**

## CONCLUSION

The artificial intelligence algorithm is a promising tool for pre-evaluation of acne severity and lesion and PIHP identification.

## REFERENCE

1. B Dréno, et al. Development and evaluation of a Global Acne Severity Scale (GEA Scale) suitable for France and Europe. *Journal of the European Academy of Dermatology and Venerology* 2011, 25: 43-8