VALIDATION TEST OF ARTIFICIAL INTELLIGENCE ASSISTED IN PIGMENTED SKIN LESIONS DIAGNOSIS

[1]THORFUN TREEWATANAKUL, [2]CHARTURONG TANTIBUNDHIT, [3]PANLOP CHAKKAVITTUMRONG, [4]SAROJ SUVANASUTHI

[1]Division of Dermatology, Chulabhorn International College of Medicine Thammasat University, Pathum Thani, Thailand
[2]Center of Excellence in Intelligence Informatics, Speech and Language Technology, and Service Innovation(CILS), Faculty of Engineering, Thammasart university, Thailand
[3]Department of Dermatology, Faculty of Medicine, Thammasart University, Thailand
[4]Dermatology Clinic, Samitivej Sukhumvit Hospital, Thailand
E-mail:[1] Omamp_aha@hotmail.com, [2]tchartur@engr.tu.ac.th, [3]panlop078@yahoo.com, [4]doctorhe@gmail.com

Abstract - Artificial intelligence(AI) has been reported in high potential in classification of pigmented skin lesions(PSLs) under experimental conditions recently. Thai researchers team has trained the first AI which enable to diagnose PSLs in Thailand. AI showed high sensitivity, specificity, and accuracy in diagnosis up to 80-90%. Although the computer diagnosis of PSLs is accurate under experimental conditions, the practical value of automated diagnostic tool compared to standard is currently unknown. This study aimed to validate the diagnosis performances of AI assisted in PSLs diagnosis compared to Board-certified dermatologist using dermoscopic images.

Index Terms - Artificial intelligence, dermoscopy, pigmented skin lesions

I. INTRODUCTION

Pigmented skin lesions (PSLs) are included both benign and malignancy. They are very similar in morphologies, colors, and textures. To distinguish between malignant melanomas and benign moles is challenging task for dermatologists. Dermoscopy, a non-invasive tool for microscopic examination of pigmented skin lesions, significantly improved the diagnostic accuracy in comparison with inspection by unaided eyes but only in specialized well-trained physicians. [1]-[3]. Recently, computer aided diagnosis system has been introduced. [4]-[5] Deep learning, a neural network paradigm of artificial, has been used in several fields particularly in visual task such as face recognition, object classification, playing strategic board game like Go, and medical screening which has been shown to exceed human performances. [6]-[8].

Thai researchers team has trained AI in classification of melanoma, nevus, and seborrheic keratosis by using dermoscopic images from International Skin Imaging Collaboration (ISIC) 2017 dataset. The result showed high performance in sensitivity, specificity, and accuracy up to 80-90%. Although the computer diagnosis of PSLs is accurate under experimental conditions, the practical value of automated diagnostic tool under clinical conditions is currently unknown.

In this study, we aimed to validate the diagnostic performances of the first artificial intelligence assisted pigmented skin lesions diagnosis in Thailand and compare the diagnostic ability to Board-certified dermatologist using dermoscopic images.

II. MATERIALS AND METHODS

This study was approved by Institutional Review Board of Bangkok Hospital Medical Center (BMC-IRB). Our researchers trained AI using DenseNet-121 algorithm with dermoscopic images of pigmented skin lesions(PSLs) from International Skin Imaging Collaboration (ISIC) Challenge 2017 dataset (see Fig. 1) We conducted cross-sectional study using 150 randomly selected dermoscopic images of PSLs as test dataset; 78 malignant melanomas, 42 seborrheic keratosis, and 30 nevi from ISIC Challenge 2017 source of data. All images were saved in JPG format. Inadequate image qualities such as poor focus or too much artifacts and images which are included multiple lesions were excluded. No overlap exists between test dataset and trained dataset for AI. Each dermoscopic image was provided one correct answer out of three choices; melanoma, seborrheic keratosis, and nevus. The test dataset was read by our AI system vs Board-certified dermatologist. No additional clinical information was provided to the dermatologist.

III. STATISTICAL ANALYSIS

Sensitivity, specificity, and accuracy were calculated according to standard formulae. Computers submitted predictions of each dermoscopic image with one out of three choices including 0, 1, 2.
Each score was checked with correct (1.0) and incorrect
(0.0) answer.
Statistical analyses used SPSS (version 10.0).

IV. RESULTS

A. Diagnostic performances of Board-certified dermatologists

Table 1 Diagnostic performances of Board-certified dermatologist in diagnosis of PSLs

	MM	SK	NV
Sensitivity (%)	90.0	69.0	48.7
Specificity (%)	58.8	92.9	98.2
Accuracy (%)	65.2	83.9	69.6

Abbreviations used:
MM: Melanoma
SK: Seborrheic keratosis
NV: Nevus

From Table 1, sensitivities of Board-certified dermatologist in melanoma, SK, and NV diagnosis were 90%, 69%, and 48.7%. Specificities were 58.8%, 92.9%, and 83.9% respectively. Accuracies were 65.2%, 83.9%, and 69.6%.

B. Diagnostic performances of artificial intelligence (AI)

Table 2 Diagnostic performances of artificial intelligence in diagnosis of PSLs

	MM	SK	NV
Sensitivity (%)	83.3	85.7	73.1
Specificity (%)	80.9	96.5	89.7
Accuracy (%)	81.4	92.9	80.8

Our AI's diagnostic performances showed sensitivity of 83.3% in melanoma diagnosis, 85.7% in SK diagnosis, and 73.1% in NV diagnosis. Specificities in diagnosis of melanoma, SK, and NV were 80.9%, 96.5%, and 89.7% respectively. Accuracies were 81.4%, 92.9%, and 80.8%. (see Table 2)

C. Comparison of diagnostic performances of Board-certified dermatologist vs artificial intelligence system

Sensitivity

In melanoma diagnosis, AI system showed lower sensitivity compared to Board-certified dermatologist (83.3% vs 90.0%) In contrast, AI has higher sensitivities in diagnosis of SK and NV. (see Fig. 1)

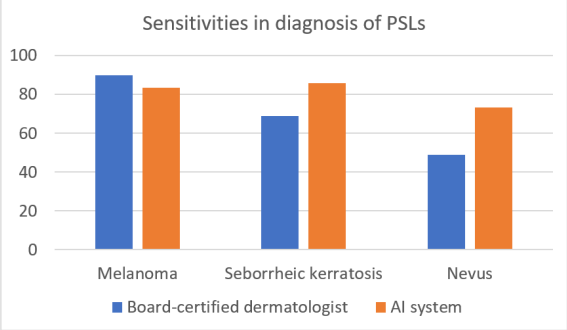


Fig.1 Sensitivities in diagnosis of PSLs between two groups

Specificity

Specificities in diagnosis of melanoma, AI showed better performance. Same as SK diagnosis. However, AI's specificity in diagnosis of NV was lower compared to dermatologist. (see Fig. 2)

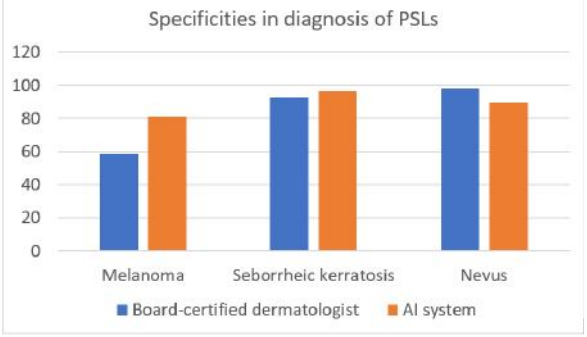


Fig.2 Specificities in diagnosis of PSLs between two groups

Accuracy

AI system showed higher accuracies in diagnosis all types of PSLs compared to dermatologist. (see Fig. 3)

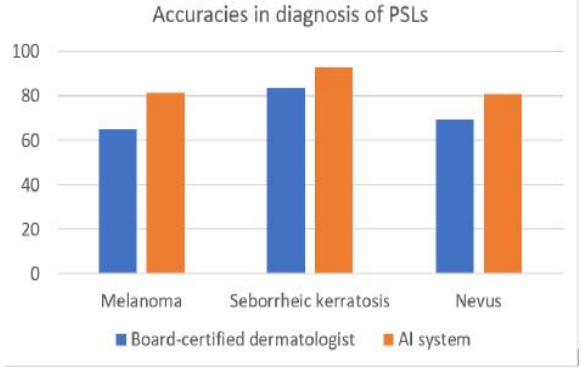


Fig.3 Accuracies in diagnosis of PSLs between two groups

DISCUSSION

This is the first research to validate AI assisted in PSLs diagnosis compared to Board-certified dermatologist in Thailand. Our validation test shows that AI system has higher diagnostic performances compared to Board-certified dermatologist in almost all parameters, except for sensitivity in melanoma diagnosis and specificity in nevus diagnosis.The more images you've trained AI, the more accuracy you will get in the future.

CONCLUSION

Our artificial intelligence system achieves performance in diagnosis of pigmented skin lesions with a level competence comparable to Board-certified dermatologists. I suggest that other types of PSLs should be considered to be included in the validation test. Also, larger sample size and clinical dermoscopic images may be necessary for the further studies.

REFERENCES

[1] Skvara H, Teban L, Fiebiger M, Binder M, Kittler H. Limitations of dermoscopy in the recognition of melanoma. 2005:155.

[2] Carli P, De Giorgi V, Crocetti E, Mannone F, Massi D, Chiarugi A, et al. Improvement of malignant/benign ratio in excised melanocytic lesions in the 'dermoscopy era': a retrospective study 1997-2001. The British journal of dermatology. 2004;150(4):687-92.

[3] Piccolo D, Ferrari A, Peris K, Daidone R, Ruggeri B, Chimenti S. Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study. British Journal of Dermatology. 2002;147(3):481-6.

[4] Masood A, Al-Jumaily AA. Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. Int J Biomed Imaging. 2013;2013:323268.

[5] Mellersen K, Kirchesch H, Zortea M, Schopf TR, Hindberg K, Godtliebsen F. Computer-aided decision support for melanoma detection applied on melanocytic and nonmelanocytic skin lesions: a comparison of two systems based on automatic analysis of dermoscopic images. BioMed Research International. 2015.

[6] Marchetti MA, Codella NCF, Dusza SW, Gutman DA, Helba B, Kalloo A, et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. Journal of the American Academy of Dermatology. 2017.

[7] Andrew J, Anant M. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of Pathology Informatics, Vol 7, Iss 1, Pp 29-29 (2016). 2016(1):29.

[8] 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;546(7639).

★★★