ARTIFICIAL INTELLIGENCE-BASED DIAGNOSIS: DISTINGUISHING EARLY SYPHILIS FROM OTHER SEXUALLY TRANSMITTED INFECTIONS (STIS)

Authors:
Sun J[1,2], Li Y[3], Yu Z[1,2,7], Towns J[1,2], Soe N[1,2], Latt P[1,2], Zhang L[4,5], Ge Z[6,7,8], Fairley C[1,2], Ong J[1,2], Zhang L[1,2,9]

[1]Melbourne Sexual Health Centre, Alfred Health, Melbourne, Australia, [2]Central Clinical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia, [3]School of Artificial Intelligence, Xidian University, Xi'an, China, [4]Suzhou Industrial Park Monash Research Institute of Science and Technology, Suzhou, China, [5]School of Public Health and Preventative Medicine, School of Medicine, Nursing and Health Sciences, Monash University, Australia, [6]AIM for Health Lab, Monash University, Melbourne, VIC, Australia [7]Monash Medical AI, Monash University, Melbourne, VIC, Australia, [8]Faculty of IT, Monash University, Melbourne, VIC, Australia, [9]China-Australia Joint Research Center for Infectious Diseases, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China

Background:
Early diagnosis of syphilis is vital for its effective control. We aimed to develop an Artificial Intelligence (AI) diagnostic model, designed based on radiomics technology, to distinguish early syphilis from other Sexually Transmitted Infections (STIs).

Methods:
260 images with different skin lesions caused by Sexually Transmitted Infections (STIs) were collected, including 115 syphilis and 145 other skin lesions. 80% of the dataset was used for developing the Artificial Intelligence (AI) model with 5-fold cross-validation, and the remaining 20% was used for the hold-out test. The exact lesion region was manually segmented as Region of Interest (ROI) in each image with the help of STI experts. 102 radiomics features were extracted from each ROI, and fed into 11 different classifiers (including SVM, Logistic Regression and so on) after deleting the redundant features using Pearson correlation coefficient. Different image filters like Laplacian of Gaussian and Wavelet were also investigated to improve the model performance. The area under the ROC curve (AUC) was used for the evaluation of the model and the Shapley Additive exPlanations (SHAP) was used for interpreting the developed model.

Results:
Among the 11 classifiers, the Gradient Boosted Decision Trees (GBDT) classifier with applying the wavelet filter on the images demonstrated the best performance, offering the stratified 5-fold cross-validation AUC of 0.832±0.042 and accuracy of 0.735±0.043. On the hold-out test dataset, the model shows an AUC, accuracy, sensitivity, specificity, and f1-score of 0.792, 0.750, 0.609, 0.778, and 0.683, respectively. The SHAP analysis shows that the low-high grayscale run length matrix (GLRLM) run entropy and original shape 2D sphericity were the most predictive radiomics features for diagnosising early syphilis.

Conclusion:

The proposed AI-based model could be used to assist in the diagnosis of early syphilis from other STIs.

Disclosure of Interest Statement:
There is no financial interest to report.