147. Comparison of Machine Learning and Conventional Criteria in Detecting Left Ventricular Hypertrophy and Prognosis With Electrocardiography

Jui-Tzu Huang, Shih-Hsien Sung, Albert Chih-Chieh Yang, Taipei General Veterans Hospital, Taipei, Taiwan

Body

There were various criteria of electrocardiography (ECG) for left ventricular hypertrophy (LVH). Nevertheless, the accuracy and clinical utility for diagnosing echocardiographic LVH in Asian populations remain controversial. We therefore investigated the diagnostic accuracy of ECG criteria for LVH and prognostic value in Taiwanese population. We enrolled a total of 42459 patients (55.3% male; mean age: 64.3±16.5 years) who received ECG and echocardiography at cardiology out-patient department visits were prospectively enrolled during 2005-2016. Electrocardiogram criteria with Sokolow-Lyon, Cornell product, Cornell/strain index, and Framingham criterion were applied for LVH evaluation. Cardiac ECHO evidence of LVH was defined according to the Teichholz formula with M-mode data and adjusted by body height. An AI algorithm based on artificial neural network (ANN) and XGBOOST (training set 80% and test set 20%) with 5-fold cross-validation were applied for the machine learning algorithm.

The causes and date of death for the participants were confirmed via review of hospital charts and death certificates, and linkage of our database with the National Death Registry through unique personal identification numbers given to every Taiwan citizen. Pearson correlation coefficient revealed significance ($P<0.0001$) between all of the four criteria (Sokolow-Lyon $r = 0.078$, Cornell product $r = 0.036$, Cornell/strain index $r = 0.089$, and Framingham criterion $r = 0.211$) and LVH. Logistic regression of the four criteria to LVH showed Sokolow-Lyon ($P<0.0001$, $t = 5.94$), Cornell product ($P = 0.0094$, $t = 2.6$), Cornell/strain index ($P<0.0001$, $t = 9.69$), and Framingham criterion ($P<0.0001$, $t = 38.87$). The accuracy and AUC of the AI algorithms were ANN model (68.8%/0.68), XGBOOST (70.9%/0.76) , CATBOOST (75.2%/0.81). The AI algorithms based on ANN, XGBOOST and CATBOOST performed better than conventional methods on detecting LVH. Feature selection was also conducted for the CATBOOST model, which revealed gender. QRS area, QT interval and r amplitude of lead I as significant factors.

Model Performance

Variable	Sensitivity	Specificity	Accuracy	AUC (95% CI)
Sokolow-Lyon	2.3%	99.5%	62.2%	0.51 (0.50-0.51)
Cornell product	0.4%	99.9%	62.8%	0.50 (0.50-0.51)
Cornell/strain index	7%	96.9%	62.4%	0.52 (0.51-0.52)
Framingham criterion	28.4%	88.3%	65.3%	0.58 (0.58-0.59)
Logistic model	62.1%	14.0%	32.4%	0.39 (0.39-0.40)
ANN model	69.2%	71.4%	68.8%	0.68 (0.67-0.68)
XGBOOST	69.9%	70.7%	70.9%	0.76 (0.75-0.76)
CATBOOST	76.9%	71.4%	75.2%	0.81 (0.80-0.81)

Clinical Implications: My study will help enable cardiovascular clinicians to find out important factors that were not included in traditional LVH criteria, but showed significance in feature selection.