Submission Id: 2647

Title

The use of innovative modeling strategies to predict the likelihood of infectious mononucleosis among college students

Priority 1 (Research Category)

Survey research or cross-sectional study

Presenters

Xinyan Cai, MSPH, Mark Ebell, MD, MS

Abstract

Context: Individual symptoms or signs of infectious mononucleosis (IM) are of limited value in determining the presence of the disease. Objective: To develop and to internally validate decision support tools using three different innovative statistical methods for assessing the likelihood of infectious mononucleosis (IM) among university students using symptoms, clinical signs, and hematologic parameters. Study Design and Analysis: From 2015 to 2019, structured data were extracted from the electronic health records of a university health center and were divided into derivation and validation cohorts. Independent predictors for the diagnosis of IM were identified by univariate analysis in the derivation cohort. Each statistical method (classification and regression tree (CART), fast and frugal tree (FFT), and artificial neural network (ANN)) was used to develop two models: one with only symptoms and signs (IM-Nolab) and one adding hematologic parameters to the model (IM-Lab). The performance of each model was then internally validated using the validation cohort. Setting: secondary analysis of a previously collected dataset. The investigator obtained the de-identified data from the university health center for all patients in which IM was clinically suspected between 2015 and 2019, based on the fact that a heterophile antibody test for IM was ordered. Results: The derivation cohort had 1498 patients (243 with IM) and the validation cohort had 844 patients (126 with IM). The IM-Nolab CART model has only two predictors; the IM-Lab CART model, IM-Nolab FFT model, and IM-Lab FFT model have three predictors; and the ANN models have five predictors. The probability of IM in the low- and high-risk groups in the validation cohort was 7.3% and 32.2% for IM-Nolab CART model (AUC=0.69); 5.9% and 61.8% for IM-Lab CART model (AUC=0.93); 8.2% and 33.5% for IM-Nolab FFT model (AUC=0.71); 5% and 68.2% for IM-Lab FFT model (AUC=0.94); 8.8% and 50.4% for IM-Nolab ANN model (AUC=0.70); and 4.4% and 69.3% for IM-Lab ANN model (AUC=0.97). Conclusions: The derived IM-Lab and IM-Nolab models based on CART, FFT and ANN methods provided useful tools to help clinicians make rapid diagnosis of IM. Each model was internally validated, and the discrimination plots showed good discriminations for the IM-Nolab models and excellent discriminations for the IM-Lab models.