DEVELOPMENT AND EXTERNAL VALIDATION OF THE MACHINE LEARNING MODELS TO PREDICT IN-HOSPITAL CARDIAC ARREST IN THE EMERGENCY DEPARTMENT: A CROSS-COUNTRY APPROACH

Amir Hossein Mostafavi Sterabadi B.S., Eric H. Chou, MD

Anne Burnett Marion School of Medicine at TCU, Baylor Scott and White All Saints Medical Center, Fort Worth TX

RESEARCH QUESTION

In Emergency Department (ED) presenting patients within the United States, will our 6 previously internally validated machine-learning (ML) models be able to utilize patient's triage data, vitals, chief complain, and demographics to successfully identify those who have had an emergency department-based cardiac arrest (EDCA) event?

BACKGROUND

In our initial study, we were able to identify both the predictive value as well as utility of machine learning (ML) models for patients at risk of emergency department-based cardiac arrest (EDCA) who had presented in an emergency department (ED) in Taiwan. Our current cross-country study aims to build off and externally validate our initial ML models within an ED population in the United States. Our hope of the study is to provide predictive strength, reliability, and utility to our initial ML models as a clinical tool to help appropriately predict and respond to patients who present in the ED who are at risk of cardiac arrest events.

MATERIALS AND METHODS

Our studies training cohort models were developed utilizing the database of adult patients who had presented at a tertiary training hospital in Taiwan between Jan. 1, 2009, to December 31, 2005. Our training data was collected retrospectively from the ED of a tertiary teaching hospital in the United States between August 31, 2009, to December 31, 2020, to be utilized for external validation as the testing cohort of our ML models. In addition, we training 6 different ML models within the training cohort using patient features such as vitals, clinical symptoms, and other triage information. We then utilized K-fold cross validation to evaluate the performance of our models based on the area under the receiver operating characteristic curve (AUC) in the external validation cohort.

QR code for poster

QR code for more information

Machine Learning Models may have the capabilities to be a generalized and applicable tool to predict, prevent, and respond to potential Emergency Department based cardiac arrest events (EDCA) based on their discriminatory abilities described in the study.

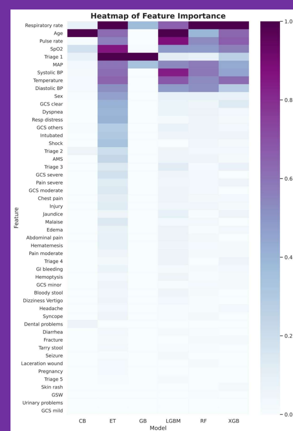

Figure 1. Representation of feature importance for the 6 ML models prior to the application of SMOTE method.

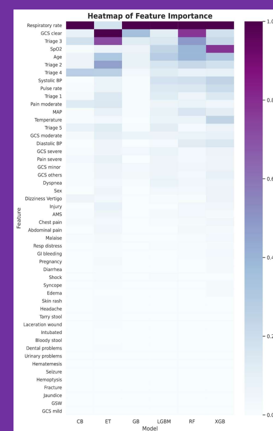

Figure 2. Representation of feature importance for the 6 ML models after the application of SMOTE method.

From the results of our data review and analysis, the top selected features we utilized to train and test our ML models on included:

- Respiratory Rate
- Consciousness
- Triage Level
- Oxygen Saturation
- Age
- Pulse Rate
- Blood Pressure
- Temperature
- Dyspnea
- Pain Severity

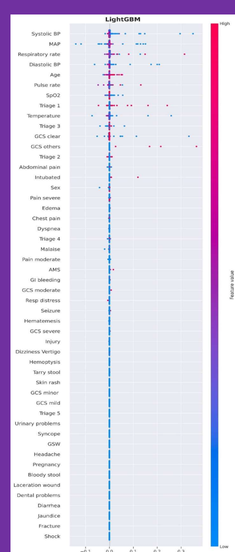

Figure 3 LGBM without SMOTE

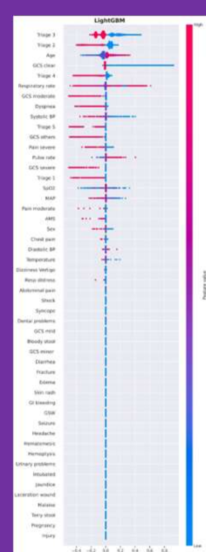

Figure 3B LGBM with SMOTE

CONCLUSION

Through our external validation study, we hope to provide ample evidence of the predictive ability and clinical utility of supervised ML models to improve patient outcomes, resource utilization, triage mechanisms, and alert systems for patients at risk of EDCA events and healthcare workers within the emergency departments. In addition, our studies findings suggest that our ML models may be a potential tool to be utilized for prediction, prevention and response to potential EDCA events based on the discriminatory abilities described within our study.

RESULTS

We included a total of 237,349 patients for our training cohort and 49,792 patients from our testing cohort. Of this population, 477 (0.2%) and 166 (0.3%) were identified to have had an EDCA event within the training and testing cohort respectively. All 6 of the ML models performed with excellent discrimination based on the AUC as noted within figure 5 below. Of the constructed ML models, light gradient-boosting machine (LGBM) achieved the best performance of AUC (0.897, 95%, 95% CI: 0.876-0.916) through utility of 7-fold cross validation. There were no significant differences noted between the constructed models.

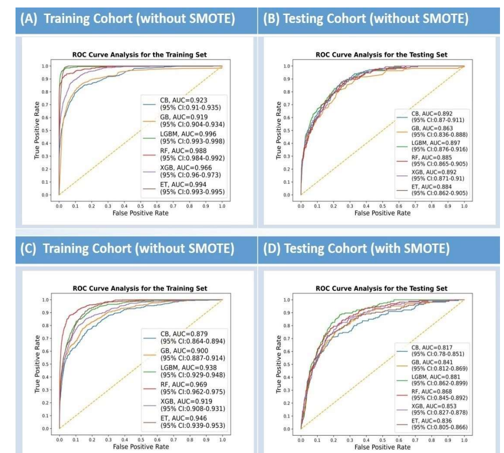

Figure 5 Comparative predictive powers of each of the six ML models in both the training (A), testing (B), training with SMOTE (C), and testing cohorts with SMOTE (D) based on their respective predictive abilities represented by the area under the receiver operating characteristics (AUPRC)

FUTURE DIRECTIONS

Through the evidence of our ML models and their predictive abilities in the setting of EDCA, future directions and avenues to build from our research may be on other disease processes that may be explored and reviewed for benefit with utility of supervised learning algorithms. To build off our ML models for prediction of EDCA, future avenues of supervised ML models may be on their prediction of disease processes with current warning systems in place, such as ICU admission risk stratification, sepsis risk stratification, or cerebrovascular event stratification. Based on patient's initial features and the utility of large data sets within the field of medicine, ML models may provide room for enhancing the field of medicine in identifying patients at risk who require higher level of care and resources.