INTRODUCTION

Blood transfusion during or after coronary artery bypass graft (CABG) surgery is associated with an increased risk for morbidity and mortality. There is a need to develop patient-specific risk prediction tools for blood transfusions in order to improve perioperative patient optimization, patient education, patient selection, patient outcomes, and clinical guidelines. Therefore, the purpose of this study is to use artificial intelligence to 1) identify patient factors that influence the risk for requiring a perioperative blood transfusion and 2) develop and assess the performance of blood transfusion risk prediction models.

METHODS

This retrospective, multicenter analysis included all patients who underwent CABG surgery (identified through CPT codes) in the 2015-2022 datasets from the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP). The variable of interest for prediction included any blood transfusion from the surgical start time up to 72 hours postoperatively. Eight typical classification models were selected to train and test the data, including Decision Tree, SVM, Gaussian Naive Bayes, Logistic Regression, Gradient Boosting, XGBoost, KNN, and Random Forest. Many iterations of different data processing techniques, feature selection, and engineering methods were performed. Several evaluation metrics (accuracy score, root mean square error, F1 score, and ROC-AUC score) were utilized to evaluate model performances.

RESULTS

Across all data science and machine learning methods in our experiments, Gradient Boosting, Random Forest, and XGBoost consistently lead the performance metrics. Among these three models across all iterations, model accuracies ranged from 67.07 - 90.80, root-mean-square error (RMSE) 0.31 – 0.57, F1-score 0.67 – 0.80, and ROC-AUC scores from 0.71 - 0.93. The synthetic data generation method enabled significant improvements and increased the ROC-AUC score to the highest 0.93 (Gradient Boosting). By ranking, the top five most important features for the Gradient Boosting model using the synthetic data generation method were disseminated cancer, chronic steroid use, diabetes, age, and preoperative albumin.

CONCLUSION

To the best of our knowledge, this is one of the first studies to demonstrate that AI-based models can perform well in predicting which patients will need a blood transfusion within the intraoperative or acute postoperative period following CABG surgery. Additionally, this study showed that AI can be used to identify patient risk factors for a perioperative blood transfusion. Further studies are needed to continually improve model performances in order to increase the likelihood that they improve patient outcomes and are cost-effective. These models additionally need to be externally validated prior to clinical translation.