Comparison Of Machine Learning Model With Multivariate Logistic Regression In Prediction Of 30-day Major Amputation In Patients With Peripheral Arterial Disease Following Lower Extremity Bypass

Sasank Kalipatnapu, MBBS, MS, Douglas W. Jones, MD, MS

Introduction

- Prior studies have shown that machine learning models in vascular surgery have better predictive ability than traditional logistic regression techniques.
- We studied the performance of an interpretable machine learning model and logistic regression to predict 30-day major amputation in patients with peripheral arterial disease (PAD) undergoing lower extremity bypass.

Methods

American College of Surgeons National Surgical Quality Improvement Program database for 23,536 patients who underwent lower extremity open procedures for PAD from 2011 to 2020

Data from 2011 to 2018 for training(18908 patients)
Data from 2019-2020 for testing (4628 patients)

Models - random forest machine learning model using Maximum Relevance – Minimum Redundancy (MRMR) feature selection and multivariate logistic regression models (using both forward stepwise and backward stepwise selection techniques)

Outcome - Occurrence of 30-day major amputation

		AU-ROC Scores (testing set)	
Algorithm	Number of predictors	Only Pre-operative Variables	All Variables
Random Forest	Top 10	0.69	0.80
	All variables	0.72	0.81
Logistic Regression – Forward Stepwise selection	Top 10	0.73	0.83
	All variables	0.73	0.83
Logistic Regression – Backward stepwise selection	All variables	0.73	0.83

Variables included in Random Forest / MRMR model	Variables included in Forward Stepwise LR model
ASA Class = 3	Age
Discharge Destination = Home	Discharge destination
Elective Surgery = Yes	Elective Surgery
Major re-intervention on the bypass = No	Major Re-intervention on bypass
Pre-procedural Beta-Blocker = Yes	Pre-procedural hemodynamics
Procedure = Femoropoliteal bypass with single segment saphenous vein	Procedure performed
Symptomatology = Claudication	Symptomatology at presentation
Untreated loss of patency = No	Untreated Loss of patency
Wound Infection / Complication = No	Pre-operative Platelet count
Functional Status Prior to surgery = Unknown	Sepsis within 48 hours prior to the surgery

Results

- All models have comparable predictive ability.
- In their choice of variables, both models used mostly similar predictors – elective vs emergency surgery, nature of the procedure performed, pre-operative disease severity, ASA class.
- Inclusion of post-operative variables improved predictive ability of all models (major reintervention on the bypass, untreated loss of patency, discharge destination)

Conclusions

- In a dataset of patients who underwent lower extremity bypass surgery, a logistic regression model performed as well as a machine learning model in predicting 30-day major amputation.
- While machine learning techniques are promising, they are less interpretable and may not offer predictive advantages over logistic regression models
- Care should be exercised in variable selection for predictive models. Including postoperative variables may improve accuracy but will lead to limitations in clinical applicability.