



OPEN

## Predicting outcomes following lower extremity open revascularization using machine learning

Ben Li<sup>1,2,3,4</sup>, Raj Verma<sup>5</sup>, Derek Beaton<sup>6</sup>, Hani Tamim<sup>7,8</sup>, Mohamad A. Hussain<sup>9</sup>, Jamal J. Hoballah<sup>10</sup>, Douglas S. Lee<sup>11,12,13</sup>, Duminda N. Wijeyesundera<sup>12,13,14,15</sup>, Charles de Mestral<sup>1,2,12,13,15</sup>, Muhammad Mamdani<sup>3,4,6,12,13,15,16</sup> & Mohammed Al-Omran<sup>1,2,3,4,8,15,17</sup>✉

Lower extremity open revascularization is a treatment option for peripheral artery disease that carries significant peri-operative risks; however, outcome prediction tools remain limited. Using machine learning (ML), we developed automated algorithms that predict 30-day outcomes following lower extremity open revascularization. The National Surgical Quality Improvement Program targeted vascular database was used to identify patients who underwent lower extremity open revascularization for chronic atherosclerotic disease between 2011 and 2021. Input features included 37 pre-operative demographic/clinical variables. The primary outcome was 30-day major adverse limb event (MALE; composite of untreated loss of patency, major reintervention, or major amputation) or death. Our data were split into training (70%) and test (30%) sets. Using tenfold cross-validation, we trained 6 ML models. Overall, 24,309 patients were included. The primary outcome of 30-day MALE or death occurred in 2349 (9.3%) patients. Our best performing prediction model was XGBoost, achieving an area under the receiver operating characteristic curve (95% CI) of 0.93 (0.92–0.94). The calibration plot showed good agreement between predicted and observed event probabilities with a Brier score of 0.08. Our ML algorithm has potential for important utility in guiding risk mitigation strategies for patients being considered for lower extremity open revascularization to improve outcomes.

Peripheral artery disease (PAD) is a chronic atherosclerotic disorder that primarily causes decreased perfusion to the lower extremities, manifesting in claudication, rest pain, and tissue loss<sup>1</sup>. Affecting over 200 million people worldwide, PAD is a major contributor to decreased quality of life, rising health care costs, limb loss, and death<sup>2–5</sup>. Lower extremity open revascularization is a surgical treatment option for PAD that has been recently demonstrated in the BEST-CLI trial to achieve superior outcomes compared to endovascular therapy for chronic

<sup>1</sup>Department of Surgery, University of Toronto, Toronto, Canada. <sup>2</sup>Division of Vascular Surgery, St. Michael's Hospital, Unity Health Toronto, University of Toronto, 30 Bond Street, Toronto, ON M5B 1W8, Canada. <sup>3</sup>Institute of Medical Science, University of Toronto, Toronto, Canada. <sup>4</sup>Temerty Centre for Artificial Intelligence Research and Education in Medicine (T-CAIREM), University of Toronto, Toronto, Canada. <sup>5</sup>School of Medicine, Royal College of Surgeons in Ireland, University of Medicine and Health Sciences, Dublin, Ireland. <sup>6</sup>Data Science & Advanced Analytics, Unity Health Toronto, University of Toronto, Toronto, Canada. <sup>7</sup>Faculty of Medicine, Clinical Research Institute, American University of Beirut Medical Center, Beirut, Lebanon. <sup>8</sup>College of Medicine, Alfaisal University, Riyadh, Kingdom of Saudi Arabia. <sup>9</sup>Division of Vascular and Endovascular Surgery and the Center for Surgery and Public Health, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. <sup>10</sup>Division of Vascular and Endovascular Surgery, Department of Surgery, American University of Beirut Medical Center, Beirut, Lebanon. <sup>11</sup>Division of Cardiology, Peter Munk Cardiac Centre, University Health Network, Toronto, Canada. <sup>12</sup>Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada. <sup>13</sup>ICES, University of Toronto, Toronto, Canada. <sup>14</sup>Department of Anesthesia, St. Michael's Hospital, Unity Health Toronto, Toronto, Canada. <sup>15</sup>Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, Canada. <sup>16</sup>Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Canada. <sup>17</sup>Department of Surgery, King Faisal Specialist Hospital and Research Center, Riyadh, Kingdom of Saudi Arabia. ✉email: mohammed.al-omran@unityhealth.to

limb threatening ischemia (CLTI) in patients with an adequate great saphenous vein conduit<sup>6</sup>. Nevertheless, open revascularization carries a high risk of complications, with major adverse limb event (MALE) or death occurring in over 40% of the surgical group in the BEST-CLI trial after a median follow-up of 2.7 years<sup>6</sup>. Others have shown that over 30% of patients will suffer a major adverse event within 30 days following lower extremity bypass<sup>7</sup>. As a result, the Global Vascular Guidelines recommend careful assessment of surgical risk when considering patients for revascularization<sup>8</sup>.

There are currently no widely used clinical tools to predict adverse events following lower extremity open revascularization. In the research setting, current models are limited to trauma patients<sup>9</sup>, Japanese and Finnish cohorts<sup>10,11</sup>, and prediction of groin wound infections<sup>12</sup>. Furthermore, tools such as the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) surgical risk calculator<sup>13</sup> and Vascular Study Group of New England (VSGNE) Cardiac Risk Index (CRI)<sup>14</sup> use modelling techniques that require manual input of clinical variables, which deters routine use in busy medical settings<sup>15</sup>. Therefore, there is an important need to develop better and more practical surgical risk prediction tools that overcome existing limitations with automated modelling techniques, inclusion of more geographically diverse cohorts with atherosclerotic disease, and assessment of more clinically relevant outcomes such as MALE or death.

Machine learning (ML) is a rapidly advancing technology that allows computers to learn from data and make predictions<sup>16</sup>. This field has been driven by the explosion of electronic medical record data combined with increasing computational power<sup>17</sup>. Previously, ML has been applied to the ACS NSQIP database to develop algorithms that predict peri-operative complications in a pooled dataset of over 2900 unique procedures, including patients undergoing day surgery to those requiring intensive care unit admission<sup>18</sup>. Given that this cohort represents a heterogeneous surgical population, better predictive performance may be achieved by developing ML algorithms specific to patients undergoing lower extremity open revascularization. In this study, ML was applied to the ACS NSQIP database to predict 30-day MALE or death and other outcomes following lower extremity open revascularization using pre-operative data.

## Methods

### Ethics

All methods were carried out in accordance with the World Medical Association Declaration of Helsinki<sup>19</sup>. Institutional research ethics board review and informed patient consent were not required as the data came from a large, deidentified registry, which is an accepted practice for studies based on ACS NSQIP data<sup>20</sup>.

### Design

We conducted a multicenter retrospective cohort ML-based prognostic study and findings were reported based on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement<sup>21</sup>.

### Dataset

Created in 2004, the ACS NSQIP database contains demographic, clinical, and 30-day outcomes data on surgical patients across over 700 hospitals in approximately 15 countries worldwide<sup>22</sup>. The information is prospectively collected from electronic health records by trained and certified clinical reviewers and regularly audited by ACS for accuracy<sup>23</sup>. In 2011, targeted NSQIP registries for vascular operations were developed by vascular surgeons, which contain additional procedure-specific variables and outcomes<sup>24</sup>.

### Cohort

All patients who underwent scheduled and unscheduled lower extremity infrainguinal open revascularization for chronic atherosclerotic disease between 2011 and 2021 in the ACS NSQIP targeted vascular database were included. This information was merged with the main ACS NSQIP database for a complete set of generic and procedure-specific variables and outcomes. Patients treated for lower extremity aneurysmal disease, acute limb ischemia, trauma, dissection, or malignancy, as well as those with unreported symptom status (CLTI, claudication, or asymptomatic) or undergoing concurrent major amputation were excluded.

### Features

Thirty-seven pre-operative variables were used as input features for our ML models. Demographic variables included age, sex, body mass index, race, ethnicity, and origin status. Comorbidities included hypertension, diabetes, smoking status, congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), end stage renal disease (ESRD) requiring dialysis, functional status, and physiologic high-risk factor [defined as at least one of the following: (1) end stage renal disease, (2) age > 80, (3) New York Heart Association CHF class III/IV, (4) left ventricular ejection fraction < 30%, (5) unstable angina within 30 days prior to surgery, or (6) myocardial infarction (MI) within 30 days prior to surgery]. Medications included antiplatelets, statins, and beta blockers. Pre-operative laboratory investigations included serum sodium, blood urea nitrogen (BUN), serum creatinine, albumin, white blood cell count, hematocrit, platelet count, international normalized unit (INR), and partial thromboplastin time (PTT). Limb hemodynamics based on ankle brachial index (ABI), toe pressure, and palpability of pedal pulses, as well as anatomic high-risk factors (defined as a prior bypass or endovascular intervention involving the currently treated segment) were recorded. Concurrent procedures recorded included minor amputation (below the ankle) and endovascular iliac or infrainguinal revascularization. Other pre-procedural characteristics recorded were symptom status [asymptomatic, claudication, or CLTI (defined as rest pain or tissue loss)], primary procedure including inflow, outflow, and conduit, urgency of surgery (elective,

urgent, or emergent), and American Society of Anesthesiologists (ASA) class. A complete list of features and definitions can be found in Supplementary Table 1.

### Outcomes

The primary outcome was 30-day MALE (composite of untreated loss of patency, major reintervention, or major amputation) or death. Untreated loss of patency was defined as a loss of graft patency on imaging or physical exam with no subsequent open or endovascular revascularization procedure. Major reintervention was defined as a new or revision lower extremity bypass, interposition graft revision, or bypass graft thrombectomy/thrombolysis. Major amputation was defined as a transtibial or more proximal amputation on the ipsilateral leg. Death was defined as all-cause mortality. This composite outcome was chosen because it is frequently reported as a primary outcome in landmark studies, including the BEST-CLI trial<sup>6</sup>.

Thirty-day secondary outcomes included individual components of the primary outcome, major adverse cardiovascular event (MACE), individual components of MACE, wound complication, bleeding requiring transfusion or secondary procedure, other morbidity, non-home discharge, and unplanned readmission. MACE was defined as a composite of MI (ischemic electrocardiogram changes, troponin elevation, or physician/advanced provider diagnosis), stroke (motor, sensory, or cognitive dysfunction persisting for > 24 h in the setting of suspected stroke), or death. Wound complication was defined as a non-healing or open wound at the surgical incision, dehiscence, or cellulitis. Other morbidity was defined as a composite of pneumonia, unplanned reintubation, pulmonary embolism (PE), failure to wean from ventilator (cumulative time of ventilator-assisted respirations > 48 h), acute kidney injury (AKI; rise in creatinine of > 2 mg/dL from pre-operative value or requirement of dialysis in a patient who did not require dialysis pre-operatively), urinary tract infection (UTI), cardiac arrest, deep vein thrombosis (DVT) requiring therapy, *Clostridium difficile* infection, sepsis, or septic shock. Non-home discharge was defined as discharge to rehabilitation, skilled care, or other facility. These outcomes are defined by the ACS NSQIP data dictionary<sup>25</sup>.

### Model development

Six ML models were trained to predict 30-day primary and secondary outcomes: Extreme Gradient Boosting (XGBoost), random forest, Naïve Bayes classifier, radial basis function (RBF) support vector machine (SVM), multilayer perceptron (MLP) artificial neural network (ANN) with a single hidden layer, sigmoid activation function, and cross-entropy loss function, and logistic regression. These were chosen because they demonstrate the best performance for predicting surgical outcomes in the literature<sup>26–28</sup>. XGBoost is a gradient-boosted decision-tree-based ensemble model that is highly effective at regression and classification predictive modelling<sup>29</sup>. Random forest is an ensemble learning method that operates through multiple decision trees<sup>30</sup>. Naïve Bayes classifiers apply Bayes' theorem to generate highly accurate predictions in high-dimensional datasets<sup>31</sup>. SVM's can find hyperplanes in dimensional space to distinctly separate data points and achieve binary classification<sup>32</sup>. Neural networks resemble biological neurons and consist of an input, hidden, and output layer, capable of making meaningful predictions from complex information<sup>33</sup>. Logistic regression is a traditional statistical method used to model the relationship between independent and dependent variables, assuming a linear correlation between the predictors and logit of the outcome, as well as a lack of multicollinearity between explanatory variables<sup>34</sup>. The advantage of newer ML techniques over logistic regression is that they apply more advanced analytics to better model complex, multicollinear relationships between predictors and outcomes<sup>35</sup>. Nonlinear associations are common in health care data as patient trajectories are often influenced by many clinical, demographic, and systems-level factors<sup>36</sup>. Logistic regression was therefore used as the baseline comparator to assess relative model performance because it is the most common modelling technique used in traditional risk prediction tools<sup>37</sup>.

Our data were split into training (70%) and test (30%) sets<sup>38</sup>. Ten-fold cross-validation and grid search were performed on the training set to find optimal hyperparameters for each ML model<sup>39,40</sup>. Preliminary analysis of our data demonstrated that the primary outcome was uncommon, occurring in 2349/24,309 (9.7%) of patients in our cohort. To improve class balance, Random Over-Sample Examples (ROSE) was applied<sup>41</sup>. ROSE uses smoothed bootstrapping to draw new samples from the feature space neighbourhood around the minority class and is a commonly used method to support predictive modelling of rare events<sup>41</sup>. The models were then evaluated on test set data and ranked based on the primary discriminatory metric of AUROC. Our best performing model was XGBoost, which had the following optimized hyperparameters for our dataset: number of rounds = 100, maximum tree depth = 6, learning rate = 0.3, gamma = 0, column sample by tree = 1, minimum child weight = 1, subsample = 1. The process for selecting these hyperparameters through grid search and cross validation is detailed in Supplementary Table 2. Once we identified XGBoost as the best performing ML model for the primary outcome, we trained the algorithm to predict secondary outcomes.

### Statistical analysis

Baseline demographic and clinical characteristics for patients with vs. without 30-day MALE or death were summarized as means (standard deviation), medians (interquartile range), or number (proportion). Differences in characteristics between outcome groups were assessed using independent t-test for continuous variables or chi-square test for categorical variables. Statistical significance was set at two-tailed  $p < 0.05$ .

The primary metric for assessing model performance was AUROC (95% CI), a validated method to assess discriminatory ability that considers both sensitivity and specificity<sup>42</sup>. Secondary performance metrics were accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). To further assess model performance, we plotted a calibration curve and calculated the Brier score, a measurement of the agreement between predicted and observed event probabilities<sup>43</sup>. In the final model, feature importance was determined by ranking the top 10 predictors based on the variable importance score (gain), a measure of

the relative impact of individual covariates in contributing to an overall prediction<sup>44</sup>. Feature importance was determined for the overall cohort, CLTI patients, and asymptomatic/ Claudication groups. To assess model robustness on various populations, we performed subgroup analysis of predictive performance based on age (under vs. over 70 years), sex (male vs. female), race (White vs. non-White), ethnicity (Hispanic vs. Non-Hispanic), symptom status (CLTI vs. asymptomatic/ Claudication), procedure type (femoropopliteal bypass vs. femoral to tibial/pedal bypass vs. popliteal to tibial/pedal bypass vs. femoral endarterectomy/profundoplasty), and urgency (urgent/emergent vs. elective).

Based on a validated sample size calculator for clinical prediction models, to achieve a minimum AUROC of 0.7 with an outcome rate of ~ 10% and 37 input features, the minimum sample size required is 6960 patients with 696 events<sup>45,46</sup>. Our cohort of 24,309 patients with 2349 primary events meets this sample size requirement. There was less than 5% missing data for variables of interest; therefore, complete-case analysis was applied whereby only non-missing covariates for each patient were considered<sup>47</sup>. This has been demonstrated to be a valid analytical method for datasets with small amounts of missing data (< 5%) and reflects predictive modelling of real-world data, which inherently includes missing information<sup>48,49</sup>. All analyses were performed in R version 4.2.1<sup>50</sup> with the following packages: caret<sup>51</sup>, xgboost<sup>52</sup>, ranger<sup>53</sup>, naivebayes<sup>54</sup>, e1071<sup>55</sup>, nnet<sup>56</sup>, and pROC<sup>57</sup>.

## Results

### Patients and events

From an initial cohort of 25,318 patients who underwent lower extremity open revascularization in the NSQIP targeted database between 2011 and 2021, we excluded 1,009 patients for the following reasons: treatment for lower extremity aneurysmal disease (n = 669), acute limb ischemia (n = 24), trauma (n = 4), or dissection (n = 2), unreported symptom status (n = 306), and concurrent major amputation (n = 4). Overall, we included 24,309 patients. The primary outcome of 30-day MALE or death occurred in 2349 (9.3%) patients. The 30-day secondary outcomes occurred in the following distribution: untreated loss of patency (n = 457 [1.9%]), major reintervention (n = 1,111 [4.9%]), major amputation (n = 689 [2.8%]), death (n = 547 [2.3%]), MACE (n = 1,346 [5.5%]), MI (n = 771 [3.2%]), stroke (n = 225 [0.9%]), wound complication (n = 3241 [13.3%]), bleeding requiring transfusion or secondary procedure (n = 4041 [16.6%]), other morbidity (n = 1799 [7.4%]); composite of pneumonia (n = 343), unplanned reintubation (n = 380), PE (n = 62), failure to wean from ventilator (n = 223), AKI (n = 123), UTI (n = 328), cardiac arrest (n = 230), DVT (n = 185), sepsis (n = 475), septic shock (n = 190), *Clostridium difficile* infection (n = 85), non-home discharge (n = 6954 [28.6%]), and unplanned readmission (n = 3621 [14.9%]).

### Pre-operative demographic and clinical characteristics

Compared to patients without a primary outcome, those who developed 30-day MALE or death were older and more likely to be female, Black, Hispanic, and transferred from another hospital, with a greater proportion residing in nursing homes. They were also more likely to have insulin dependent diabetes, CHF, ESRD requiring dialysis, and at least 1 physiologic high-risk factor. Functionally, patients with 30-day MALE or death were more likely to be partially or totally dependent. Despite being at higher cardiovascular risk, patients with 30-day MALE or death were less likely to receive antiplatelets. Notable differences in laboratory investigations included patients with 30-day MALE or death having higher levels of creatinine and BUN. Patients with a primary outcome were more likely to have an ABI ≤ 0.39 and a previous bypass or endovascular intervention involving the currently treated segment, with a greater proportion undergoing a concurrent minor amputation or endovascular infrainguinal revascularization. Patients with 30-day MALE or death were more likely to have CLTI, undergo a bypass to a tibial/pedal target, receive urgent/emergent surgery, and be ASA class 4 or higher (Table 1).

### Model performance

Of the 6 ML models evaluated on test set data for predicting 30-day MALE or death following lower extremity open revascularization, XGBoost had the best performance with an AUROC (95% CI) of 0.93 (0.92–0.94) compared to random forest [0.92 (0.91–0.93)], Naïve Bayes [0.87 (0.86–0.88)], RBF SVM [0.85 (0.84–0.86)], MLP ANN [0.80 (0.78–0.82)], and logistic regression [0.63 (0.61–0.65)]. The other performance metrics of XGBoost were the following: accuracy 0.86 (95% CI 0.85–0.87), sensitivity 0.84, specificity 0.89, PPV 0.90, and NPV 0.83 (Table 2).

For 30-day secondary outcomes, XGBoost achieved the following AUROC's (95% CI): untreated loss of patency [0.90 (0.89–0.91)], major reintervention [0.91 (0.89–0.93)], major amputation [0.95 (0.94–0.96)], death [0.96 (0.95–0.96)], MACE [0.93 (0.92–0.94)], MI [0.88 (0.87–0.89)], stroke [0.91 (0.90–0.92)], wound complication [0.90 (0.88–0.92)], bleeding requiring transfusion or secondary procedure [0.92 (0.91–0.93)], other morbidity [0.91 (0.89–0.92)], non-home discharge [0.95 (0.95–0.96)], and unplanned readmission [0.87 (0.86–0.89)] (Table 3).

The ROC curve for prediction of 30-day MALE or death using XGBoost is demonstrated in Fig. 1. Our model achieved good calibration with a Brier score of 0.08, indicating excellent agreement between predicted and observed event probabilities (Fig. 2). The top 10 predictors of 30-day MALE or death in our XGBoost model were the following: (1) symptom status: CLTI, (2) pre-operative dialysis, (3) functional status, (4) pre-operative CHF, (5) pre-operative creatinine, (6) urgency of surgery, (7) procedure type: conduit/target/inflow, (8) physiologic high-risk factor, (9) pre-operative antiplatelet, and (10) anatomic high-risk factor (Fig. 3). On subgroup analysis based on symptom status, 9/10 of the most important predictive features were the same for patients with CLTI and those who were asymptomatic or had claudication, with the two most important predictors being functional status and pre-operative dialysis for both groups (Supplementary Fig. 1).

	Absence of MALE or death at 30 days (n = 21,960)	Presence of MALE or death at 30 days (n = 2349)	P
Demographics			
Age, years, mean (SD)	67.7 (9.9)	68.4 (10.7)	<0.001
Female	7240 (33.0)	871 (37.1)	<0.001
BMI, kg/m <sup>2</sup> , mean (SD)	27.8 (5.8)	27.7 (6.3)	0.29
Race			
White	14,165 (64.5)	1,405 (59.8)	<0.001
Black or African American	3595 (16.4)	454 (19.3)	
American Indian or Alaskan Native	72 (0.3)	6 (0.3)	
Native Hawaiian or Other Pacific Islander	18 (0.08)	0	
Asian	178 (0.8)	19 (0.8)	
Other	15 (0.07)	3 (0.1)	
Unknown or not reported	3917 (17.8)	462 (19.7)	
Hispanic ethnicity	1072 (4.9)	130 (5.5)	0.004
Origin status			
Transferred from another hospital	1534 (7.0)	277 (11.8)	<0.001
Home	19,768 (90.0)	1,956 (83.3)	
Nursing home	523 (2.4)	90 (3.8)	
Other facility	128 (0.6)	25 (1.1)	
Unknown	7 (0.03)	1 (0.04)	
Comorbidities			
Hypertension	17,924 (81.6)	1908 (81.2)	0.66
Diabetes			
Non-insulin dependent	4228 (19.3)	417 (17.8)	0.15
Insulin dependent	5993 (27.3)	690 (29.4)	0.03
Current smoker	9404 (42.8)	990 (42.1)	0.69
Congestive heart failure	762 (3.5)	138 (5.9)	<0.001
Chronic obstructive pulmonary disease	2754 (12.5)	315 (13.4)	0.24
Dialysis	1205 (5.5)	252 (10.7)	<0.001
Functional status			
Independent	20,463 (93.2)	2057 (87.6)	<0.001
Partially dependent	1310 (6.0)	247 (10.5)	
Totally dependent	106 (0.5)	27 (1.2)	
Unknown	81 (0.4)	18 (0.8)	
Physiologic high-risk factor <sup>a</sup>	4578 (20.8)	735 (31.3)	<0.001
Medications			
Antiplatelet	17,947 (81.7)	1814 (77.2)	<0.001
Statin	15,900 (72.4)	1666 (70.9)	0.08
Beta blocker	12,394 (56.4)	1368 (58.2)	0.11
Laboratory investigations			
Sodium, mmol/L, mean (SD)	138.0 (3.4)	137.0 (3.6)	<0.001
BUN, mmol/L, mean (SD)	56.2 (31.6)	60.1 (39.1)	<0.001
Creatinine, umol/L, median (IQR)	88.4 (70.7–115.0)	91.1 (70.7–124.0)	<0.001
Albumin, g/L, mean (SD)	26.3 (10.2)	26.3 (9.6)	0.92
White blood cell count, cells/mm <sup>3</sup> , mean (SD)	8.4 (4.4)	9.3 (4.6)	<0.001
Hematocrit, L/L (%), mean (SD)	37.4 (6.2)	35.9 (6.5)	<0.001
Platelet count, 10 <sup>9</sup> /L, mean (SD)	251.0 (92.3)	265.0 (111.0)	<0.001
INR, mean (SD)	1.1 (0.3)	1.2 (0.4)	<0.001
PTT, sec, mean (SD)	34.0 (13.9)	37.0 (17.2)	<0.001
Anatomy/hemodynamics			
Limb hemodynamics			
Continued			

	Absence of MALE or death at 30 days (n = 21,960)	Presence of MALE or death at 30 days (n = 2349)	P
ABI ≥ 1.3 and toe pressure ≥ 30 mmHg	164 (0.7)	20 (0.9)	< 0.001
ABI ≥ 1.3 and toe pressure < 30 mmHg	158 (0.7)	24 (1.0)	
ABI ≥ 1.3 and no toe pressure recorded	341 (1.6)	48 (2.0)	
ABI 0.90 – 1.29	599 (2.7)	60 (2.6)	
ABI 0.40 – 0.89	6861 (31.2)	502 (21.4)	
ABI ≤ 0.39	4060 (18.5)	502 (21.4)	
ABI not recorded and palpable pedal pulse	941 (4.3)	78 (3.3)	
ABI not recorded and pedal pulse not palpable	3947 (18.0)	567 (24.1)	
Not documented	4889 (22.3)	548 (23.3)	
Anatomic high-risk factor			
Prior bypass involving currently treated segment	4655 (21.2)	670 (28.5)	< 0.001
Prior endovascular intervention involving currently treated segment	4013 (18.3)	419 (17.8)	
None/not documented	13,292 (60.5)	1,260 (53.6)	
Concurrent procedures			
Minor amputation	870 (4.0)	119 (5.1)	0.01
Endovascular iliac revascularization	789 (3.6)	92 (3.9)	0.46
Endovascular infrainguinal revascularization	499 (2.3)	78 (3.3)	0.002
Other pre-procedural characteristics			
Symptom status			
Asymptomatic	431 (2.0)	21 (0.9)	< 0.001
Claudication	5808 (26.4)	296 (12.6)	
Chronic limb threatening ischemia: rest pain	6408 (29.2)	808 (34.4)	
Chronic limb threatening ischemia: tissue loss	9313 (42.4)	1,224 (52.1)	
Primary procedure			
Femoropopliteal bypass with single segment saphenous vein	6928 (31.5)	550 (23.4)	< 0.001
Femoropopliteal bypass with prosthetic, spliced vein, or composite	5649 (25.7)	450 (19.2)	
Femoral to tibial/pedal bypass with single segment saphenous vein	4337 (19.7)	575 (24.5)	
Femoral to tibial/pedal bypass with prosthetic, spliced vein, or composite	2348 (10.7)	433 (18.4)	
Popliteal to tibial/pedal bypass with single segment saphenous vein	1549 (7.1)	195 (8.3)	
Popliteal to tibial/pedal bypass with prosthetic, spliced vein, composite, or non-saphenous conduit	460 (2.1)	67 (2.9)	
Femoral endarterectomy	257 (1.2)	35 (1.5)	
Profundoplasty	69 (0.3)	5 (0.2)	
Not documented/other	363 (1.7)	39 (1.7)	
Urgency			
Elective	14,632 (66.6)	1,164 (49.6)	< 0.001
Urgent	6050 (27.6)	887 (37.8)	
Emergent	1267 (5.8)	295 (12.6)	
ASA class			
1	37 (0.2)	5 (0.2)	< 0.001
2	813 (3.7)	70 (3.0)	
3	15,594 (71.0)	1449 (61.7)	
4	5476 (24.9)	815 (34.7)	
5	15 (0.07)	8 (0.3)	
Not reported	25 (0.1)	2 (0.09)	

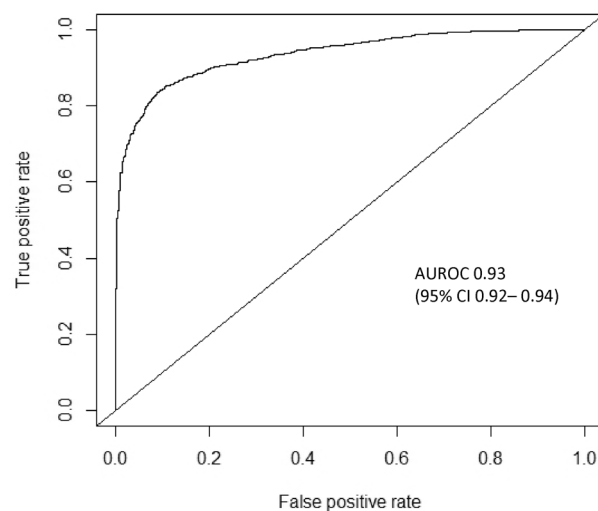
**Table 1.** Pre-operative demographic and clinical characteristics of patients undergoing lower extremity open revascularization with and without major adverse limb event or death at 30 days. Values are reported as No. (%) unless otherwise indicated. MALE (major adverse limb event), BMI (body mass index), BUN (blood urea nitrogen), INR (international normalized ratio), PTT (partial thromboplastin time), ABI (ankle brachial index), ASA (American Society of Anesthesiologists), SD (standard deviation). <sup>a</sup>At least 1 of the following: (1) end stage renal disease, (2) age > 80, (3) New York Heart Association congestive heart failure class III/IV, (4) left ventricular ejection fraction < 30%, (5) unstable angina within 30 days prior to surgery, or (6) myocardial infarction within 30 days prior to surgery.

	AUROC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity	PPV	NPV
XGBoost	0.93 (0.92–0.94)	0.86 (0.85–0.87)	0.84	0.89	0.90	0.83
Random forest	0.92 (0.91–0.93)	0.85 (0.84–0.86)	0.84	0.86	0.86	0.83
Naïve bayes	0.87 (0.86–0.88)	0.85 (0.84–0.86)	0.83	0.85	0.86	0.82
RBF SVM	0.85 (0.84–0.86)	0.77 (0.75–0.79)	0.75	0.80	0.83	0.71
MLP ANN	0.80 (0.78–0.82)	0.73 (0.70–0.75)	0.71	0.75	0.79	0.69
Logistic regression	0.63 (0.61–0.65)	0.58 (0.56–0.60)	0.55	0.71	0.60	0.56

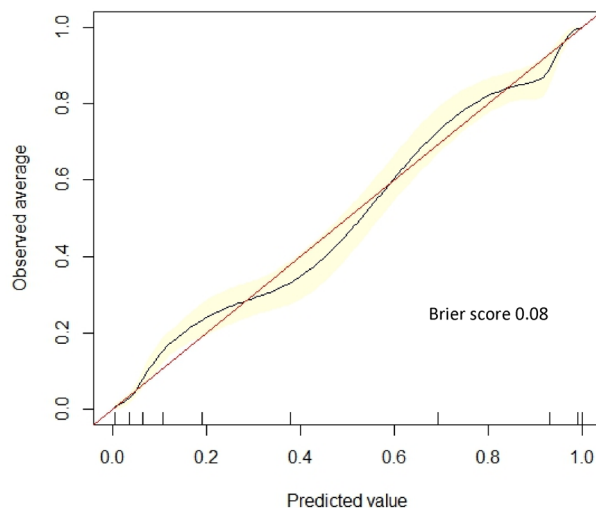
**Table 2.** Model performance on test set data for predicting 30-day major adverse limb event or death following lower extremity open revascularization using pre-operative features. Abbreviations: XGBoost (Extreme Gradient Boosting), AUROC (area under the receiver operating characteristic curve), CI (confidence interval), PPV (positive predictive value), NPV (negative predictive value), RBF SVM (radial basis function support vector machine), MLP ANN (multilayer perceptron artificial neural network).

	AUROC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity	PPV	NPV
MALE or death	0.93 (0.92–0.94)	0.86 (0.85–0.87)	0.84	0.89	0.90	0.83
Untreated loss of patency	0.90 (0.89–0.91)	0.83 (0.81–0.84)	0.81	0.85	0.86	0.79
Major reintervention	0.91 (0.89–0.93)	0.84 (0.82–0.86)	0.83	0.85	0.87	0.80
Major amputation	0.95 (0.94–0.96)	0.88 (0.87–0.89)	0.86	0.90	0.91	0.85
Death	0.96 (0.95–0.96)	0.89 (0.87–0.90)	0.87	0.90	0.90	0.87
MACE	0.93 (0.92–0.94)	0.85 (0.84–0.86)	0.84	0.86	0.88	0.85
Myocardial infarction	0.88 (0.87–0.89)	0.80 (0.79–0.82)	0.79	0.81	0.82	0.78
Stroke	0.91 (0.90–0.92)	0.83 (0.81–0.84)	0.82	0.84	0.85	0.80
Wound complication	0.90 (0.88–0.92)	0.82 (0.83–0.87)	0.82	0.83	0.86	0.79
Bleeding requiring transfusion or secondary procedure	0.92 (0.91–0.93)	0.84 (0.82–0.85)	0.82	0.86	0.87	0.81
Other morbidity	0.91 (0.89–0.92)	0.82 (0.81–0.84)	0.81	0.84	0.85	0.79
Non-home discharge	0.95 (0.95–0.96)	0.89 (0.88–0.90)	0.86	0.92	0.92	0.85
Unplanned readmission	0.87 (0.86–0.89)	0.80 (0.78–0.82)	0.80	0.77	0.77	0.81

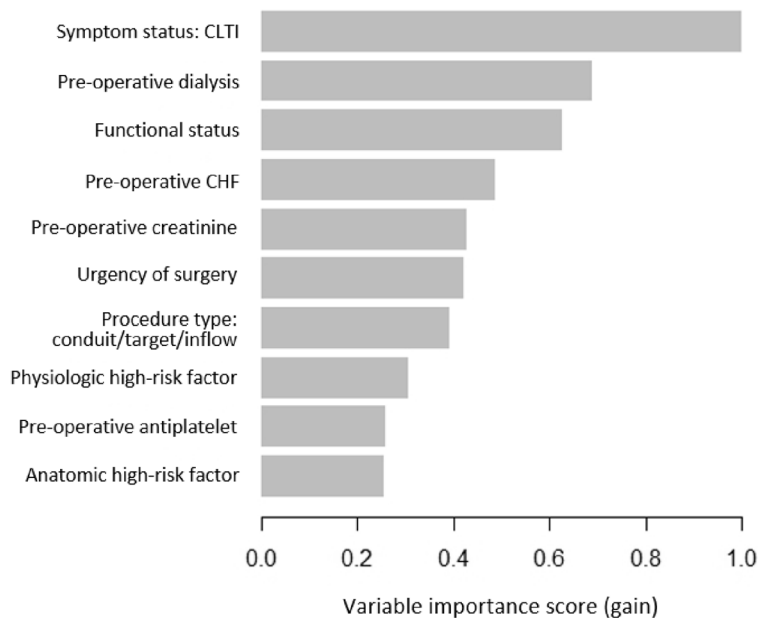
**Table 3.** XGBoost performance on test set data for predicting 30-day primary and secondary outcomes following lower extremity open revascularization using pre-operative features. XGBoost (Extreme Gradient Boosting), AUROC (area under the receiver operating characteristic curve), CI (confidence interval), PPV (positive predictive value), NPV (negative predictive value), MALE (major adverse limb event; composite of untreated loss of patency, major reintervention, or major amputation), MACE (major adverse cardiovascular event; composite of myocardial infarction, stroke, or death).



**Figure 1.** Receiver operating characteristic curve for predicting 30-day major adverse limb event or death following lower extremity open revascularization using Extreme Gradient Boosting (XGBoost) model. AUROC (area under the receiver operating characteristic curve), CI (confidence interval).



**Figure 2.** Calibration plot with Brier score for predicting 30-day major adverse limb event or death following lower extremity open revascularization using Extreme Gradient Boosting (XGBoost) model.



**Figure 3.** Variable importance scores (gain) for the top 10 predictors of 30-day major adverse limb event or death following lower extremity open revascularization in the Extreme Gradient Boosting (XGBoost) model. Abbreviations: CLTI (chronic limb threatening ischemia), CHF (congestive heart failure).

### Subgroup analysis

Our XGBoost model performance for predicting 30-day MALE or death remained excellent on subgroup analysis of specific demographic and clinical populations with the following AUROC's (95% CI): age < 70 [0.93 (0.92–0.94)] and age > 70 [0.94 (0.93–0.95)] (Supplementary Fig. 2), males [0.94 (0.93–0.95)] and females [0.93 (0.91–0.94)] (Supplementary Fig. 3), White patients [0.93 (0.92–0.94)] and non-White patients [0.93 (0.92–0.94)] (Supplementary Fig. 4), Hispanic patients [0.93 (0.92–0.94)] and non-Hispanic patients [0.93 (0.92–0.94)] (Supplementary Fig. 5), CLTI patients [0.93 (0.92–0.94)] and asymptomatic/claudication groups [0.94 (0.93–0.95)] (Supplementary Fig. 6), femoropopliteal bypass [0.94 (0.93–0.95)], femoral to tibial/pedal bypass [0.93 (0.92–0.94)], popliteal to tibial/pedal bypass [0.93 (0.91–0.95)], and femoral endarterectomy/profundoplasty [0.93 (0.89–0.96)] (Supplementary Fig. 7), and urgent/emergent surgery [0.94 (0.93–0.95)] and elective surgery [0.93 (0.92–0.94)] (Supplementary Fig. 8).



## Discussion

### Summary of findings

Using data from the ACS NSQIP targeted vascular files between 2011 and 2021 consisting of 24,309 patients who underwent lower extremity open revascularization for atherosclerotic disease, we developed ML models that accurately predict 30-day MALE or death with an AUROC of 0.93 using pre-operative variables. Furthermore, our algorithms predicted 30-day untreated loss of patency, major reintervention, major amputation, death, MACE, MI, stroke, wound complication, bleeding, other morbidity, non-home discharge, and readmission with AUROC's between 0.87 and 0.96. There were several other key findings. First, patients who develop 30-day MALE or death represent a high-risk population with several predictive features at the pre-operative stage. Specifically, they are older with more comorbidities, have poorer functional status, and are more likely to have high-risk physiologic and anatomic factors. In addition, a greater proportion of patients with 30-day MALE or death had CLTI, underwent tibial/pedal bypasses, and required concurrent minor amputation or endovascular revascularization. Despite these differences, they were less likely to receive optimal medical therapy including antiplatelets. This represents an important opportunity to improve medical management of PAD patients. Second, we trained 6 ML models to predict 30-day MALE or death using pre-operative features and showed that XGBoost achieved the best performance. Our model was well-calibrated, achieving a Brier score of 0.08, and remained robust on subgroup analysis based on age, sex, race, ethnicity, symptom status, procedure type, and urgency of surgery. Finally, we identified the top 10 predictors of 30-day MALE or death in our ML models. These features can be used by clinicians to identify factors that contribute to risk predictions, thereby guiding patient selection and pre-operative optimization. For example, patients with modifiable high-risk factors could be further evaluated and optimized through pre-operative consultations with anesthesiologists or cardiologists to mitigate adverse events<sup>58,59</sup>. Overall, we have developed a robust ML-based surgical risk prediction tool that can help guide clinical decision-making to improve outcomes and reduce costs from complications, reinterventions, and readmissions associated with lower extremity open revascularization.

### Comparison to existing literature

Bertges et al. developed the VSGNE CRI to predict in-hospital major adverse cardiac events in patients undergoing major vascular procedures including lower extremity bypass, carotid endarterectomy, and aortic aneurysm repair<sup>14</sup>. Using logistic regression, their model achieved an AUROC of 0.71<sup>14</sup>. Applying ML techniques to a more up-to-date cohort specifically consisting of patients undergoing lower extremity open revascularization, we achieved better performance with an AUROC of 0.93.

Bonde et al. trained ML algorithms on a cohort of NSQIP patients undergoing >2900 unique procedures to predict peri-operative complications, achieving AUROC's of 0.85–0.88<sup>18</sup>. Given that patients undergoing lower extremity open revascularization for atherosclerotic disease represent a unique population with a high number of vascular comorbidities, the applicability of general surgical risk prediction tools may be limited. By developing ML algorithms specific to patients undergoing lower extremity open revascularization, we achieved AUROC's >0.90. Additionally, we included limb- and graft-related outcomes such as major amputation, major reintervention, and untreated loss of patency, which are of clinical importance to vascular surgeons. Therefore, there is value in building procedure-specific ML models, which can increase accuracy and clinical applicability.

Prediction models specific to patients undergoing lower extremity revascularization remain limited. Miyata et al. (2021) applied logistic regression to predict 30-day major amputation or death in a cohort of 2906 patients identified through the Japan Critical Limb Ischemia Database, achieving an AUROC of 0.82<sup>10</sup>. Using a cohort of 24,309 patients in the multi-national NSQIP database, we achieved an AUROC >0.90 for predicting MALE or death with ML techniques. This demonstrates the benefits of applying advanced analytical techniques to larger and more diverse datasets.

### Explanation of findings

There are several explanations for our findings. First, patients who develop MALE or death following lower extremity revascularization represent a high-risk group with multiple vascular risk factors, as corroborated by previous literature<sup>60</sup>. The use of antiplatelet therapy is a Grade IA recommendation by multiple societal guidelines for all patients regardless of symptom status (asymptomatic, claudication, or CLTI)<sup>8,61–63</sup>, yet patients who developed MALE or death in our cohort were less likely to receive antiplatelets. The suboptimal rates of best medical therapy for PAD patients are further demonstrated in the recently published BEST-CLI trial<sup>6</sup>. Therefore, there are important opportunities to improve care for patients by understanding their perioperative risk and medically optimizing them prior to revascularization. Second, our ML models performed better than existing tools for several reasons. Compared to traditional logistic regression, advanced ML techniques can better model non-linear, complex relationships between inputs and outputs<sup>64,65</sup>. This is especially important in health care data, as patient outcomes can be influenced by many demographic and clinical factors<sup>66</sup>. Our best performing algorithm was XGBoost, which has unique advantages including avoiding overfitting and faster computing while maintaining precision<sup>67–69</sup>. Furthermore, XGBoost works well with structured data, which may explain why it performed better than more complex algorithms such as neural networks on our dataset<sup>70</sup>. Third, our model performance remained robust on subgroup analysis of specific demographic and clinical populations. This is an important finding given that algorithm bias against underrepresented populations is a significant issue in ML studies<sup>71</sup>. We were likely able to avoid these biases due to the excellent capture of sociodemographic data by ACS NSQIP, a multi-national database that includes diverse patient populations<sup>72,73</sup>. Fourth, a small proportion of patients in our cohort underwent lower extremity open revascularization for asymptomatic disease (<2%). The reasons for these interventions are unclear from our dataset but may be related to revisions for hemodynamically

significant stenoses of previous revascularization procedures, patient preference, poor adherence to guideline-directed revascularization, or coding errors<sup>74</sup>.

### Implications

Our ML models can be used to guide clinical decision-making in several ways. Pre-operatively, a patient predicted to be at high risk of an adverse outcome should be further assessed in terms of modifiable and non-modifiable factors<sup>75</sup>. Patients with significant non-modifiable risk of adverse outcomes following open surgical revascularization may benefit from careful considerations of alternative options including medical management alone or less invasive endovascular therapy<sup>76,77</sup>. Those with modifiable risks should be referred to anesthesiologists, cardiologists, and/or internal medicine specialists for further evaluation<sup>58,59</sup>. Intra-operatively, risk predictions may inform decisions regarding anesthetic techniques such as neuraxial vs. general anesthesia<sup>78</sup>. At the post-operative stage, patients at high risk of adverse events may benefit from close monitoring in the intensive care unit<sup>79</sup>. Additionally, patients at high risk of non-home discharge or readmission should receive early support from allied health professionals to optimize safe discharge planning<sup>80</sup>. These peri-operative decisions guided by our ML models have the potential to improve outcomes and reduce costs by mitigating adverse events.

The programming code used to develop our ML models is publicly available through GitHub, a web-based platform that offers a free and integrated environment for hosting source code, documentation, and project-related web content for open-source projects<sup>81</sup>. These tools can be used by any clinician involved in the peri-operative management of patients being considered for lower extremity open revascularization. On a systems-level, our models can be readily implemented by the > 700 centres that currently participate in ACS NSQIP worldwide. They also have potential for use at non-NSQIP sites, as the input features are commonly captured variables for the routine care of vascular surgery patients<sup>82</sup>. Given the challenges of deploying prediction models into clinical practice, consideration of principles of implementation science is critical<sup>83</sup>. Our ML models have the advantage of providing automated risk predictions using many input variables, thereby improving practicality in busy clinical settings compared to traditional risk predictors that generally require manual input of variables<sup>13</sup>. Specifically, our algorithms were built to autonomously extract a patient's prospectively collected NSQIP data to make risk predictions. Ongoing efforts to link NSQIP data to electronic health records has the potential to increase the clinical utility of our model and further support fully automated risk predictions<sup>84,85</sup>. We advocate for dedicated health care data analytics teams at the institution level, as their significant benefits have been previously demonstrated and model implementation can be facilitated by these experts<sup>86,87</sup>. Through this study, we have also provided a framework for the development of robust ML models that predict lower extremity open revascularization outcomes, which can be applied by individual centers for their specific patient populations.

### Limitations

Our study has several limitations. First, our models were developed using ACS NSQIP data. Future studies should assess whether performance can be generalized to institutions that do not participate in ACS NSQIP. Second, the ACS NSQIP database captures 30-day outcomes. Evaluation of ML models on data sources with longer follow-up would augment our understanding of long-term surgical risk. Third, our dataset did not capture low-dose rivaroxaban use. Given that the VOYAGER<sup>60</sup> and COMPASS<sup>88</sup> trials demonstrated the cardiovascular and limb benefits of low-dose rivaroxaban, future prediction models on datasets that capture this variable may improve performance. Fourth, our models are limited to patients undergoing open revascularization. Future prediction tools for outcomes following endovascular therapy would be helpful to further guide clinical decision-making.

### Conclusions

In this study, we used the ACS NSQIP database to develop robust ML models that pre-operatively predict 30-day MALE or death following lower extremity open revascularization for atherosclerotic disease with excellent performance (AUROC 0.93). Our models also predicted untreated loss of patency, major reintervention, major amputation, death, MACE, MI, stroke, wound complication, bleeding, other morbidity, non-home discharge, and readmission with AUROC's between 0.87 and 0.96. Given that our ML algorithms perform better than existing tools and logistic regression, they have potential for important utility in the peri-operative management of patients being considered for lower extremity open revascularization to mitigate adverse outcomes and reduce health care costs.

### Data availability

The data used for this study comes from ACS NSQIP. Access to and use of the data requires approval through an application process available at <https://www.facs.org/quality-programs/data-and-registries/acs-nsqip/participant-use-data-file/>.

### Code availability

The complete code used for model development and evaluation in this project is publicly available on GitHub: <https://github.com/benli12345/LEO-ML-NSQIP>.

Received: 11 April 2023; Accepted: 25 January 2024

Published online: 05 February 2024

### References

1. Zemaitis, M. R., Boll, J. M. & Dreyer, M. A. Peripheral Arterial Disease. in *StatPearls* (StatPearls Publishing, 2021).

2. Fowkes, F. G. R. *et al.* Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: A systematic review and analysis. *Lancet Lond. Engl.* **382**, 1329–1340 (2013).
3. Agnelli, G., Belch, J. J. F., Baumgartner, I., Giovvas, P. & Hoffmann, U. Morbidity and mortality associated with atherosclerotic peripheral artery disease: A systematic review. *Atherosclerosis* **293**, 94–100 (2020).
4. Kim, M., Kim, Y., Ryu, G. W. & Choi, M. Functional status and health-related quality of life in patients with peripheral artery disease: A Cross-sectional study. *Int. J. Environ. Res. Public Health* **18**, 10941 (2021).
5. Kohn, C. G., Alberts, M. J., Peacock, W. F., Bunz, T. J. & Coleman, C. I. Cost and inpatient burden of peripheral artery disease: Findings from the national inpatient sample. *Atherosclerosis* **286**, 142–146 (2019).
6. Farber, A. *et al.* Surgery or endovascular therapy for chronic limb-threatening ischemia. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2207899> (2022).
7. Liang, P. *et al.* In-hospital versus postdischarge major adverse events within 30 days following lower extremity revascularization. *J. Vasc. Surg.* **69**, 482–489 (2019).
8. Conte, M. S. *et al.* Global vascular guidelines on the management of chronic limb-threatening ischemia. *J. Vasc. Surg.* **69**, 3S–125S.e40 (2019).
9. Perkins, Z. B. *et al.* Predicting the outcome of limb revascularization in patients with lower-extremity arterial trauma: Development and external validation of a supervised machine-learning algorithm to support surgical decisions. *Ann. Surg.* **272**, 564–572 (2020).
10. Miyata, T. *et al.* Risk prediction model for early outcomes of revascularization for chronic limb-threatening ischaemia. *Br. J. Surg.* **108**, 941–950 (2021).
11. Biancari, F. *et al.* Risk-scoring method for prediction of 30-day postoperative outcome after infrainguinal surgical revascularization for critical lower-limb ischemia: A Finnvasc registry study. *World J. Surg.* **31**, 217–225 (2007).
12. Bennett, K. M., Levinson, H., Scarborough, J. E. & Shortell, C. K. Validated prediction model for severe groin wound infection after lower extremity revascularization procedures. *J. Vasc. Surg.* **63**, 414–419 (2016).
13. Bilimoria, K. Y. *et al.* Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. *J. Am. Coll. Surg.* **217**, 833–842 (2013).
14. Bertges, D. *et al.* The vascular study group of new england cardiac risk index (VSG-CRI) predicts cardiac complications more accurately than the Revised Cardiac Risk Index in vascular surgery patients. *J. Vasc. Surg.* **52**, (2010).
15. Sharma, V. *et al.* Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. *BMJ Health Care Inform.* **28**, e100253 (2021).
16. Baştanlar, Y. & Özyüsal, M. Introduction to machine learning. *Methods Mol. Biol.* **1107**, 105–128 (2014).
17. Shah, P. *et al.* Artificial intelligence and machine learning in clinical development: A translational perspective. *NPJ Digit. Med.* **2**, 69 (2019).
18. Bonde, A. *et al.* Assessing the utility of deep neural networks in predicting postoperative surgical complications: A retrospective study. *Lancet Digit. Health* **3**, e471–e485 (2021).
19. World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA* **310**, 2191–2194 (2013).
20. Tfaily, M. A., Ghanem, P., Farran, S. H., Dabdoub, F. & Kanafani, Z. A. The role of preoperative albumin and white blood cell count in surgical site infections following whipple surgery. *Sci. Rep.* **12**, 19184 (2022).
21. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent Reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann. Intern. Med.* **162**, 55–63 (2015).
22. ACS NSQIP. ACS <https://www.facs.org/quality-programs/data-and-registries/acs-nsqip/>.
23. Shiloach, M. *et al.* Toward robust information: Data quality and inter-rater reliability in the American college of surgeons national surgical quality improvement program. *J. Am. Coll. Surg.* **210**, 6–16 (2010).
24. Cohen, M. E. *et al.* Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: Patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus. *J. Am. Coll. Surg.* **217**, 336–346.e1 (2013).
25. ACS NSQIP Participant Use Data File. ACS <https://www.facs.org/quality-programs/data-and-registries/acs-nsqip/participant-use-data-file/>.
26. Elfanagely, O. *et al.* Machine learning and surgical outcomes prediction: A systematic review. *J. Surg. Res.* **264**, 346–361 (2021).
27. Bektaş, M., Tuynman, J. B., Costa Pereira, J., Burchell, G. L. & van der Peet, D. L. Machine learning algorithms for predicting surgical outcomes after colorectal surgery: A systematic review. *World J. Surg.* <https://doi.org/10.1007/s00268-022-06728-1> (2022).
28. Senders, J. T. *et al.* Machine learning and neurosurgical outcome prediction: A systematic review. *World Neurosurg.* **109**, 476–486.e1 (2018).
29. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (2016). doi:<https://doi.org/10.1145/2939672.2939785>.
30. Rigatti, S. J. Random forest. *J. Insur. Med. N. Y. N* **47**, 31–39 (2017).
31. Zhang, Z. Naïve bayes classification in R. *Ann. Transl. Med.* **4**, 241 (2016).
32. Noble, W. S. What is a support vector machine?. *Nat. Biotechnol.* **24**, 1565–1567 (2006).
33. Schmidhuber, J. Deep learning in neural networks: an overview. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **61**, 85–117 (2015).
34. Sperandei, S. Understanding logistic regression analysis. *Biochem. Medica* **24**, 12–18 (2014).
35. Ngiam, K. Y. & Khor, I. W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* **20**, e262–e273 (2019).
36. Liew, B. X. W., Kovacs, F. M., Rügamer, D. & Royuela, A. Machine learning versus logistic regression for prognostic modelling in individuals with non-specific neck pain. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spine Res. Soc.* **31**, 2082–2091 (2022).
37. Shipe, M. E., Deppen, S. A., Farjah, F. & Grogan, E. L. Developing prediction models for clinical use using logistic regression: An overview. *J. Thorac. Dis.* **11**, S574–S584 (2019).
38. Dobbin, K. K. & Simon, R. M. Optimally splitting cases for training and testing high dimensional classifiers. *BMC Med. Genomics* **4**, 31 (2011).
39. Jung, Y. & Hu, J. A K-fold averaging cross-validation procedure. *J. Nonparametric Stat.* **27**, 167–179 (2015).
40. Adnan, M., Alarood, A. A. S., Uddin, M. I. & Ur Rehman, I. Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models. *PeerJ Comput. Sci.* **8**, e803 (2022).
41. Wibowo, P. & Faticah, C. Pruning-based oversampling technique with smoothed bootstrap resampling for imbalanced clinical dataset of Covid-19. *J. King Saud Univ.–Comput. Inf. Sci.* **34**, 7830–7839 (2022).
42. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* **4**, 627–635 (2013).
43. Redelmeier, D. A., Bloch, D. A. & Hickam, D. H. Assessing predictive accuracy: How to compare brier scores. *J. Clin. Epidemiol.* **44**, 1141–1146 (1991).
44. Loh, W.-Y. & Zhou, P. Variable importance scores. *J. Data Sci.* **19**, 569–592 (2021).
45. Riley, R. D. *et al.* Calculating the sample size required for developing a clinical prediction model. *BMJ* m441 (2020) <https://doi.org/10.1136/bmj.m441>.
46. Ensor, J., Martin, E. C. & Riley, R. D. Pmsampsize: Calculates the Minimum Sample Size Required for Developing a Multivariable Prediction Model. (2022).
47. Schafer, J. L. Multiple imputation: A primer. *Stat. Methods Med. Res.* **8**, 3–15 (1999).

48. Ross, R. K., Breskin, A. & Westreich, D. When is a complete-case approach to missing data valid? The importance of effect-measure modification. *Am. J. Epidemiol.* **189**, 1583–1589 (2020).
49. Hughes, R. A., Heron, J., Sterne, J. A. C. & Tilling, K. Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *Int. J. Epidemiol.* **48**, 1294–1304 (2019).
50. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
51. Kuhn, M. *et al.* Caret: Classification and Regression Training. (2022).
52. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.-KDD* **16**, 785–794 (2016).
53. Wright, M. N., Wager, S. & Probst, P. Ranger: A Fast Implementation of Random Forests. (2022).
54. NaiveBayes: High Performance Implementation of the Naive Bayes Algorithm version 0.9.7 from CRAN. <https://rdr.io/cran/naivebayes/>.
55. svm function - RDocumentation. <https://www.rdocumentation.org/packages/e1071/versions/1.7-11/topics/svm>.
56. Ripley, B. & Venables, W. Nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models. (2022).
57. Robin, X. *et al.* pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
58. O'Connor, D. B. *et al.* An anaesthetic pre-operative assessment clinic reduces pre-operative inpatient stay in patients requiring major vascular surgery. *Ir. J. Med. Sci.* **180**, 649–653 (2011).
59. Davis, F. M. *et al.* The clinical impact of cardiology consultation prior to major vascular surgery. *Ann. Surg.* **267**, 189–195 (2018).
60. Bonaca, M. P. *et al.* Rivaroxaban in peripheral artery disease after revascularization. *N. Engl. J. Med.* **382**, 1994–2004 (2020).
61. Conte, M. S. *et al.* Society for vascular surgery practice guidelines for atherosclerotic occlusive disease of the lower extremities: Management of asymptomatic disease and claudication. *J. Vasc. Surg.* **61**, 2S–41S (2015).
62. Gerhard-Herman, M. D. *et al.* 2016 AHA/ACC guideline on the management of patients with lower extremity peripheral artery disease: Executive summary: A report of the American college of cardiology/American heart association task force on clinical practice guidelines. *Circulation* **135**, e686–e725 (2017).
63. Aboyans, V. *et al.* Editor's choice—2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European society for vascular surgery (ESVS). *Eur. J. Vasc. Endovasc. Surg. Off. J. Eur. Soc. Vasc. Surg.* **55**, 305–368 (2018).
64. Stoltzfus, J. C. Logistic regression: A brief primer. *Acad. Emerg. Med. Off. J. Soc. Acad. Emerg. Med.* **18**, 1099–1104 (2011).
65. Kia, B. *et al.* Nonlinear dynamics based machine learning: Utilizing dynamics-based flexibility of nonlinear circuits to implement different functions. *PLoS One* **15**, e0228534 (2020).
66. Chatterjee, P. *et al.* Nonlinear systems in healthcare towards intelligent disease prediction. *Nonlinear Syst.—Theor. Asp. Recent Appl.* <https://doi.org/10.5772/intechopen.88163> (2019).
67. Ravaut, M. *et al.* Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data. *Npj Digit. Med.* **4**, 1–12 (2021).
68. Wang, R., Zhang, J., Shan, B., He, M. & Xu, J. XGBoost machine learning algorithm for prediction of outcome in aneurysmal subarachnoid hemorrhage. *Neuropsychiatr. Dis. Treat.* **18**, 659–667 (2022).
69. Fang, Z.-G., Yang, S.-Q., Lv, C.-X., An, S.-Y. & Wu, W. Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: A time-series study. *BMJ Open* **12**, e056685 (2022).
70. Viljanen, M., Meijerink, L., Zwakhs, L. & van de Kasstele, J. A machine learning approach to small area estimation: Predicting the health, housing and well-being of the population of Netherlands. *Int. J. Health Geogr.* **21**, 4 (2022).
71. Gianfrancesco, M. A., Tamang, S., Yazdany, J. & Schmajuk, G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern. Med.* **178**, 1544–1547 (2018).
72. Mazmudar, A., Vitello, D., Chapman, M., Tomlinson, J. S. & Bentrem, D. J. Gender as a risk factor for adverse intraoperative and postoperative outcomes of elective pancreatectomy. *J. Surg. Oncol.* **115**, 131–136 (2017).
73. Halsey, J. N., Asti, L. & Kirschner, R. E. The impact of race and ethnicity on surgical risk and outcomes following palatoplasty: An analysis of the nsqip pediatric database. *Cleft Palate-Craniofac. J. Off. Publ. Am. Cleft Palate* **52**, 255. <https://doi.org/10.1177/10556656221078154> (2022).
74. Rumenapf, G., Morbach, S., Schmidt, A. & Sigl, M. Intermittent claudication and asymptomatic peripheral arterial disease. *Dtsch. Arztebl. Int.* **117**, 188–193 (2020).
75. Shaydakov, M. E. & Tuma, F. Operative Risk. in *StatPearls* (StatPearls Publishing, 2022).
76. Bevan, G. H. & WhiteSolaru, K. T. Evidence-based medical management of peripheral artery disease. *Arterioscler. Thromb. Vasc. Biol.* **40**, 541–553 (2020).
77. Biscetti, F. *et al.* Outcomes of lower extremity endovascular revascularization: Potential predictors and prevention strategies. *Int. J. Mol. Sci.* **22**, 2002 (2021).
78. Roberts, D. J. *et al.* Association between neuraxial anaesthesia or general anaesthesia for lower limb revascularisation surgery in adults and clinical outcomes: Population based comparative effectiveness study. *BMJ* **371**, m4104 (2020).
79. Gillies, M. A. *et al.* Intensive care utilization and outcomes after high-risk surgery in Scotland: A population-based cohort study. *Br. J. Anaesth.* **118**, 123–131 (2017).
80. Patel, P. R. & Bechmann, S. Discharge Planning. in *StatPearls* (StatPearls Publishing, 2022).
81. Perez-Riverol, Y. *et al.* Ten simple rules for taking advantage of git and GitHub. *PLoS Comput. Biol.* **12**, e1004947 (2016).
82. Nguyen, L. L. & Barshes, N. R. Analysis of large databases in vascular surgery. *J. Vasc. Surg.* **52**, 768–774 (2010).
83. Northridge, M. E. & Metcalf, S. S. Enhancing implementation science by applying best principles of systems science. *Health Res. Policy Syst.* **14**, 74 (2016).
84. Bronsert, M. *et al.* Identification of postoperative complications using electronic health record data and machine learning. *Am. J. Surg.* **220**, 114–119 (2020).
85. Colquhoun, D. A. *et al.* Considerations for integration of perioperative electronic health records across institutions for research and quality improvement: The approach taken by the multicenter perioperative outcomes group. *Anesth. Analg.* **130**, 1133–1146 (2020).
86. Batko, K. & Słezak, A. The use of big data analytics in healthcare. *J. Big Data* **9**, 3 (2022).
87. Leung, S. N. *et al.* Harnessing the full potential of hospital-based data to support surgical quality improvement. *BMJ Open Qual.* **10**, e001178 (2021).
88. Eikelboom, J. W. *et al.* Rivaroxaban with or without aspirin in stable cardiovascular disease. *N. Engl. J. Med.* **377**, 1319–1330 (2017).

## Acknowledgements

The ACS NSQIP and the hospitals participating in the ACS NSQIP are the source of the data used herein; they have not verified, and are not responsible for, the statistical validity of the data analysis or the conclusions derived by the authors.

## Author contributions

All authors meet all four criteria: (1) Substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of the data, (2) Drafting the work or revising it critically for important intellectual content, (3) Final approval of the completed version, and (4) Accountability for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. B.L. and M.A.O. had full access to and verified all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis. B.L., C.deM., M.M., and M.A.O. conceived of and designed the study. B.L. drafted the article, developed the models (with the support of data scientist D.B.), and performed statistical analysis (with the support of biostatistician HT). All authors (B.L., R.V., D.B., H.T., M.A.H., J.J.H., D.S.L., D.N.W., C.deM., M.M., M.A.O.) acquired, analyzed, or interpreted data, and critically revised the article for important intellectual content. C.deM., M.M., and M.A.O. provided supervision. M.A.O. had the final responsibility for the decision to submit for publication. All authors (B.L., R.V., D.B., H.T., M.A.H., J.J.H., D.S.L., D.N.W., C.deM., M.M., M.A.O.) had full access to the full data in the study, accept responsibility to submit for publication, and approve of the final manuscript.

## Funding

This research was funded partially by the Canadian Institutes of Health Research, Ontario Ministry of Health, and PSI Foundation (Dr. Li). Dr. Hussain is funded by a Brigham and Women's Hospital Heart and Vascular Center Faculty Award. The funding sources did not play a role in the design or conduct of the research.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-52944-1>.

**Correspondence** and requests for materials should be addressed to M.A.-O.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024