Heliyon 10 (2024) e34602

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Predictive model for novel subtypes of patients undergoing lower extremity amputation for peripheral artery disease: An unsupervised machine learning study

Yuanliang Ma [1], Lin Zhang [1], Que Li, Xiao Qin [*]

The First Affiliated Hospital of Guangxi Medical University, PR China

A R T I C L E I N F O

Keywords:
Peripheral artery disease
Unsupervised machine learning
Amputation
Predictive model
Precision medicine

A B S T R A C T

Background: Peripheral artery disease (PAD) represents the frequently seen circulatory condition related to a risk of critical limb ischemia and amputation. Critical lower extremity ischemia may require amputation, and the outcomes vary. In this study, we developed an artificial intelligence (AI)-driven predictive model for PAD subtypes to assess risk among patients more precisely and accurately to predict disease progression.
Methods: The present retrospective study examined clinical data in PAD patents undergoing lower extremity amputation. The data were analyzed using an unsupervised machine learning algorithm (UMLA) for subgroup identification and risk stratification. The clustering result accuracy was validated by analyzing the follow-up data of clusters. Finally, we built the prediction model with binary logistic regression.
Results: In total, we enrolled 507 cases into this work. Two distinct subgroups, consisting of Clusters 1 and 2, were identified by UMLA; those from Cluster 1 showed markedly poorer conditions and prognostic outcomes compared with those from Cluster 2. With regard to the new PAD subtype, we established a nomogram with eight predictive factors, including gender, age, smoking history, diabetes and coronary heart disease history, albumin levels, endovascular intervention, and amputation level. The nomogram could accurately categorize patients into two identified clusters, and the area under receiver operating characteristic curve was 0.861 (95 % confidence interval: 0.830–0.893).
Conclusion: In this study, UMLA was used to identify new phenotypic subgroups among PAD cases who showed different risks of amputation. Our constructed AI-driven predictive model for PAD subtypes showed that it can be used for risk stratification and clinical management with high accuracy and reliability.

1. Introduction

Peripheral artery disease (PAD) represents the frequently seen circulatory disease that is related to a risk of critical limb ischemia and amputation. PAD can result in arterial stenosis and occlusion [1], eventually resulting in chronic limb-threatening ischemia (CLTI) [2]. Some studies have reported a high incidence of peripheral arterial ischemic disease, especially in individuals who are over 70 years

* Corresponding author. Department of Vascular Surgery Ward, No.6 of Shuangyong Road, Nanning, Guangxi, 530021, PR China.
E-mail address: dr_qinxiao@hotmail.com (X. Qin).
[1] First Author:Yuanliang Ma and Lin Zhang contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e34602
Received 22 April 2024; Received in revised form 12 July 2024; Accepted 12 July 2024
Available online 19 July 2024
2405-8440/

old [3]; the incidence of the disease is around 15–20 % among Western countries and approximately 15.91 % in China [4,5]. PAD represents a major factor causing lower limb amputation, which occupies 40–60 % of total amputated cases [6]. Post-amputation complications such as wound infections and poor healing may lead to longer hospital stays or even necessitate a second amputation. Some patients succumb to sepsis, multiple organ failure, or other complications after amputation [7]. Therefore, medical practitioners need to comprehensively evaluate the condition of PAD patients who have undergone lower extremity amputation and perform precise risk stratification.

As artificial intelligence (AI) develops, researchers have used machine learning in diagnosing, classifying and treating illnesses [8]. An unsupervised machine learning algorithm (UMLA) is a category of AI in which the model is responsible for learning structures and patterns from unlabeled data with no explicit guidance or labeled outcomes [9]. The UMLA is widely used in clinical research because it is capable of clustering patients according to corresponding disease features and efficiently classifying heterogeneous cohorts precisely and rationally [10].

In this study, we applied the UMLA for identifying and clustering pre-operative data from PAD cases who underwent amputation surgery and analyzed their post-operative data to categorize PAD into different subtypes. By integrating comprehensive clinical data, we established a novel PAD subtype prediction model. This innovative model can accurately classify patients and greatly contributes to the personalized and effective clinical management of PAD.

2. Methods

2.1. Patients

The Ethics Committee of The First Affiliated Hospital of Guangxi Medical University (2023-E314–01) approved our study protocols.

Clinical data were obtained from PAD cases undergoing lower extremity amputation at First Affiliated Hospital of Guangxi Medical University between January 2012 and June 2023. Patients below were included: those developing lower extremity PAD who underwent major amputation surgery based on their symptoms, physical signs, ultrasound examination results, contrast-enhanced computed tomography, and digital subtraction angiography. The exclusion criteria were as follows: (a) patients who underwent amputation due to trauma, malignant tumors, congenital vascular diseases, acute limb ischemia, thromboangiitis obliterans, or vasculitis; and (b) patients who had insufficient case information, inadequate follow-up length, or loss-to-follow-up.

Both pre-operative and follow-up data were recorded for all patients. Pre-operative data included gender, age, BMI, smoking history, presence of comorbidities, whether the patient underwent endovascular intervention or open lower limb revascularization before amputation, Rutherford classification, amputation level, and distance to the artery occlusion. We also collected data on pre-operative blood test indicators, such as the platelet count, white blood cell count, hemoglobin level, albumin level, D-dimer and C-reactive protein (CRP) levels. The follow-up data included intraoperative bleeding, intensive care unit (ICU) stay, hospital stay, the occurrence of multiple organ dysfunction syndrome (MODS), septicemia, acute renal failure (ARF), wound infection and operative ulcers, whether the patient underwent secondary amputation, and whether the patient died during hospitalization.

2.2. Construction of PAD subtypes using UMLA

We used K-means cluster algorithm, an supervised machine learning algorithm for splitting and rearranging a dataset into distinct clusters, for clustering PAD cases. The primary objective of using this algorithm is for grouping similar data points and maximizing the between-cluster dissimilarity [11]. A scale function was used to normalize the data of the PAD patients via R 'factoextra' package [12]. In addition, we utilized 'Fpc' package to determine the best cluster number (k value) through determining silhouette coefficient (SC) [13,14]. SC has been the metric adopted for assessing clustering technique effectiveness in UMLA, which reflects how well each data point fits into the assigned cluster by considering cohesion within the cluster and separation from other clusters [15]. SC can be calculated using the following equation:

$$\mathrm{SC}\ (i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}$$

Here, a(i) stands for mean distance from a data point to additional points within an identical cluster, whereas b(i) indicates the shortest average distance from data point to points within a different cluster, minimized over clusters. The SC(i) ranges from −1 to 1, in which −1, 0 and 1 stand for incorrect, overlapping and well-separated and distinct clusters, separately. This metric reflects clustering quality, and high values indicate well-defined clusters [16].

According to pre-operative data, we categorized UMLA cases as distinct clusters. By analyzing follow-up data between distinct clusters, the precision of UMLA clustering was determined, which helped in correctly placing the PAD patients in different subtype categories.

Predictors that exhibited statistical significance were used to conduct binary logistic regression. Also, we used the coefficients from the logistic regression model for nomogram construction. Logistic regression coefficients were mapped by the nomogram to the 0–100 scale for visually representing predicted probabilities. In addition, receiver operating characteristic (ROC) and calibration curve analyses were performed to evaluate nomogram performance.

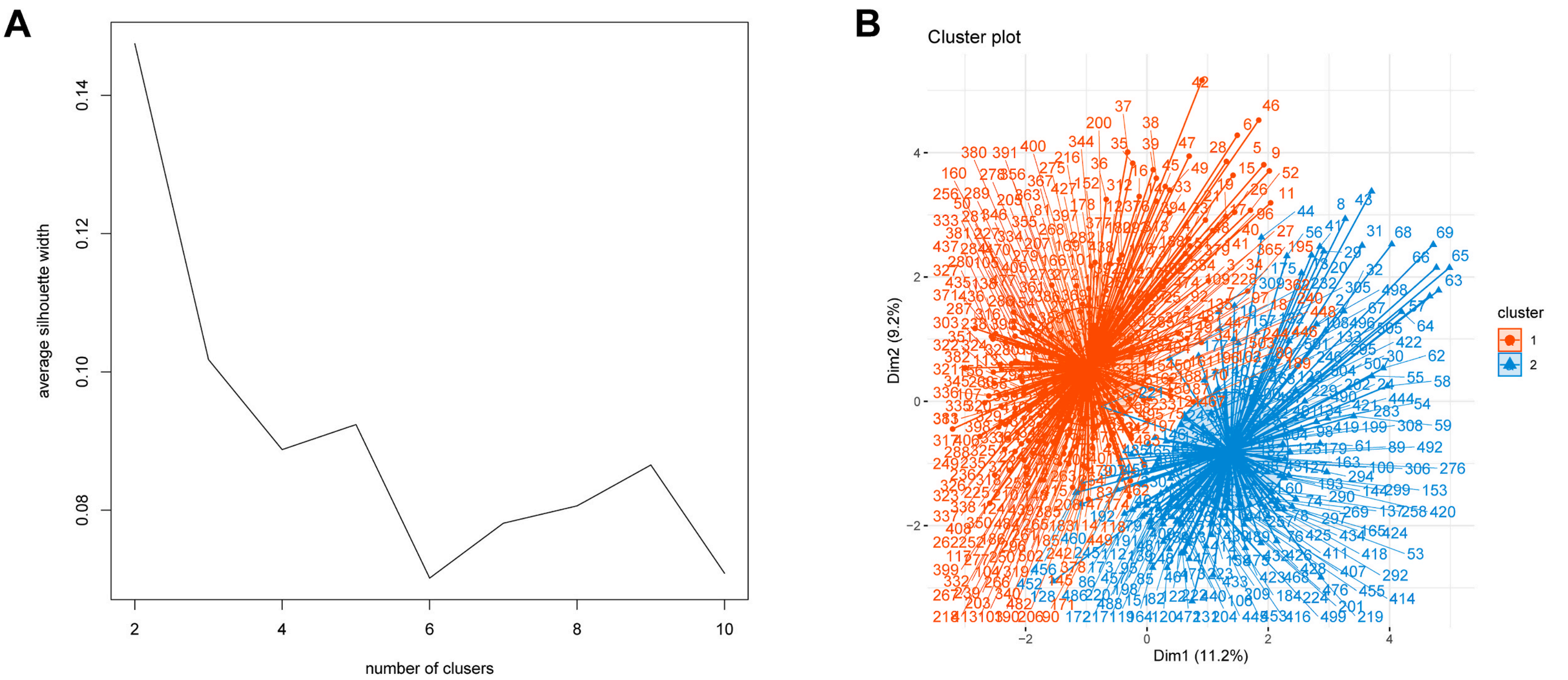

Fig. 1. Work flow and result of unsupervised machine learning.
(A) Optimal clustering number of the K-means clustering algorithm was determined by Silhouette coefficient (SC). The peak of the curve is the best value for the Silhouette coefficient (Y-axis); the best number of clusters was equal to 2 (X–axis). (B) Scatter plots of patients' clinical data. Scatter points on the graph represent each patient. The K-means algorithm divides patients into two clusters. The red scatter represents cluster 1 and the blue scatter represents cluster 2.

Table 1
Pre-operative conditions of the study patients by clusters.

Characteristic	cluster			p-value
	Overall, N = 507	1, N = 298	2, N = 209	
Age				<0.001
Mean ± SD	64 ± 17	69 ± 14	57 ± 17	
Median (IQR)	66 (52, 77)	71 (61, 80)	56 (44, 70)	
BMI				0.637
Mean ± SD	22.6 ± 3.8	22.6 ± 3.6	22.7 ± 4.0	
Median (IQR)	22.4 (20.1, 24.5)	22.4 (20.0, 24.3)	22.5 (20.3, 24.8)	
Rutherford category				0.235
Mean ± SD	5.01 ± 0.74	4.99 ± 0.62	5.03 ± 0.89	
Median (IQR)	5.00 (5.00, 6.00)	5.00 (5.00, 5.00)	5.00 (4.00, 6.00)	
Hemoglobin(g/L)				0.216
Mean ± SD	114 ± 39	118 ± 45	109 ± 27	
Median (IQR)	111 (90, 127)	111 (90, 130)	111 (90, 124)	
White blood cell(10ˆ9/L)				0.396
Mean ± SD	10.9 ± 5.8	11.0 ± 6.2	10.8 ± 5.1	
Median (IQR)	9.5 (6.5, 13.6)	9.3 (6.4, 13.6)	9.6 (6.6, 13.7)	
CRP(mg/L)				0.123
Mean ± SD	83 ± 61	85 ± 61	79 ± 61	
Median (IQR)	66 (40, 120)	66 (46, 101)	61 (35, 129)	
Platelet(10ˆ9/L)				0.720
Mean ± SD	335 ± 162	337 ± 163	331 ± 160	
Median (IQR)	322 (214, 432)	325 (217, 436)	321 (213, 428)	
D-dimer(mg/L)				0.571
Mean ± SD	532 ± 882	470 ± 349	620 ± 1306	
Median (IQR)	430 (235, 621)	424 (231, 588)	445 (249, 667)	
Albumin(g/L)				<0.001
Mean ± SD	31 ± 7	30 ± 7	33 ± 8	
Median (IQR)	31 (26, 36)	30 (26, 34)	32 (28, 39)	
Amputation distance (cm)				0.053
Mean ± SD	31 ± 18	32 ± 18	29 ± 18	
Median (IQR)	25 (20, 40)	25 (20, 40)	25 (20, 40)	
Gender				<0.001
Female	174 (34.3 %)	70 (23.5 %)	104 (49.8 %)	
Male	333 (65.7 %)	228 (76.5 %)	105 (50.2 %)	
Smoking				<0.001
No	237 (46.8 %)	103 (34.6 %)	134 (64.1 %)	
Yes	270 (53.3 %)	195 (65.4 %)	75 (35.9 %)	
Diabetes				<0.001
No	316 (62.3 %)	163 (54.7 %)	153 (73.2 %)	
1	191 (37.7 %)	135 (45.3 %)	56 (26.8 %)	
Hypertension				0.165
No	354 (69.8 %)	201 (67.5 %)	153 (73.2 %)	
Yes	153 (30.2 %)	97 (32.6 %)	56 (26.8 %)	
CD				0.390
No	441 (87.0 %)	256 (85.9 %)	185 (88.5 %)	
Yes	66 (13.0 %)	42 (14.1 %)	24 (11.5 %)	
CKD				0.944
No	436 (86.0 %)	256 (85.9 %)	180 (86.1 %)	
Yes	71 (14.0 %)	42 (14.1 %)	29 (13.9 %)	
CAD				0.001
No	371 (73.2 %)	202 (67.8 %)	169 (80.9 %)	
Yes	136 (26.8 %)	96 (32.2 %)	40 (19.1 %)	
Endovascular therapy				<0.001
No	298 (58.8 %)	200 (67.1 %)	98 (46.9 %)	
Yes	209 (41.2 %)	98 (32.9 %)	111 (53.1 %)	
Amputation level				<0.001
Above knee	182 (35.9 %)	86 (28.9 %)	96 (45.9 %)	
Below knee	210 (41.4 %)	135 (45.3 %)	75 (35.9 %)	
Ankle	115 (22.7 %)	77 (25.8 %)	38 (18.2 %)	
Revascularization				0.003
No	454 (89.5 %)	277 (93.0 %)	177 (84.7 %)	
Yes	53 (10.5 %)	21 (7.0 %)	32 (15.3 %)	

Amputation distance: distance between amputation level and artery occlusion; BMI: Body mass index; CAD: Coronary artery disease; CD:Cerebrovascular disease; CKD: Chronic kidney disease; CRP: C-reactive protein; Endovascular therapy: It comprises minimally invasive procedures performed within blood vessels to restore blood flow and treat vascular diseases; Revascularization: open lower limb revascularization; Rutherford category: It classifies the severity of PAD into stages, ranging from asymptomatic (category 0) to major tissue loss (category 6).

2.3. Statistical analysis

R4.2.1 software and IBM SPSS 26.0 were utilized in the statistical analysis. Clinical results were indicated by mean (SD) and median (P25, P75). Student's t-test, chi-square test and Mann-Whitney *U* test were adopted according to different data types. $P < 0.05$ stood for significant difference.

3. Results

3.1. UMLA results

From January 2012 to June 2023, our center treated a total of 3881 patients with CLTI, 618 of whom underwent major amputations (15.9 %). In line with our eligibility criteria, we enrolled altogether 507 cases into this work. The UMLA was used to cluster the pre-operative data of PAD amputees. Typically, the best cluster number for K-means algorithm is shown in Fig. 1A. The peak of the curve in the figure indicates the optimal value of SC, suggesting that the best cluster number was two. K-means algorithm efficiently clustered pre-operative variables as two groups (Fig. 1B). The clustering results of the pre-operative data are presented in Table 1. There was a significant difference in gender distribution in both clusters ($p < 0.001$), and males occupied an increased percentage in Cluster 1 (76.5 %) compared with in Cluster 2 (50.2 %). A significantly greater number of cases from Cluster 1 compared with Cluster 2 were older, smokers, and diabetic ($p < 0.001$). Pre-operative albumin levels and the proportion of patients who underwent pre-operative endovascular intervention and open lower limb revascularization markedly decreased among patients in Cluster 1 compared with Cluster 2 ($p < 0.001$). The incidence of coronary heart disease markedly decreased in Cluster 1 relative to Cluster 2 ($p = 0.001$). Additionally, patients who had undergone above-knee amputation showed a markedly increased proportion in Cluster 2, representing 45.9 % of the total. Conversely, Cluster 1 had a greater percentage of patients who underwent below-knee amputation or ankle disarticulation, accounting for 45.3 % and 25.8 %, respectively ($p < 0.001$).

3.2. Comparison of follow-up data between the clusters

The differences in follow-up data between Cluster 1 and Cluster 2 are presented in Table 2. Patients in Cluster 1 had significantly

Table 2
Post-operative followed-up conditions of patients in two clusters.

Characteristic	cluster			p-value
	Overall, N = 507	1, N = 298	2, N = 209	
Bleeding volum(ml)				<0.001
Mean ± SD	238 ± 286	302 ± 314	147 ± 210	
Median (IQR)	100 (30, 300)	250 (50, 500)	50 (25, 250)	
ICU stays(day)				0.016
Mean ± SD	4.0 ± 11.7	5.7 ± 14.8	1.6 ± 3.3	
Median (IQR)	0.0 (0.0, 4.0)	0.0 (0.0, 5.8)	0.0 (0.0, 1.0)	
Hospital stays(day)				<0.001
Mean ± SD	21 ± 29	27 ± 36	12 ± 9	
Median (IQR)	12 (8, 18)	14 (12, 30)	7 (7, 16)	
MODS				0.127
No	493 (97.2 %)	287 (96.3 %)	206 (98.6 %)	
Yes	14 (2.8 %)	11 (3.7 %)	3 (1.4 %)	
ARF				0.045
No	490 (96.7 %)	284 (95.3 %)	206 (98.6 %)	
Yes	17 (3.4 %)	14 (4.7 %)	3 (1.4 %)	
Infection				<0.001
No	408 (80.5 %)	221 (74.2 %)	187 (89.5 %)	
Yes	99 (19.5 %)	77 (25.8 %)	22 (10.5 %)	
Operative ulcer				<0.001
No	434 (85.6 %)	234 (78.5 %)	200 (95.7 %)	
Yes	73 (14.4 %)	64 (21.5 %)	9 (4.3 %)	
Secondary amputation				<0.001
No	442 (87.2 %)	241 (80.9 %)	201 (96.2 %)	
Yes	65 (12.8 %)	57 (19.1 %)	8 (3.8 %)	
Death in hospital				0.015
No	487 (96.1 %)	281 (94.3 %)	206 (98.6 %)	
Yes	20 (3.9 %)	17 (5.7 %)	3 (1.4 %)	
Septicemia				0.178
No	494 (97.4 %)	288 (96.6 %)	206 (98.6 %)	
Yes	13 (2.6 %)	10 (3.2 %)	3 (1.4 %)	

ARF: Acute renal failure, it refers to a sudden decline in kidney function, resulting in an inability to maintain fluid, electrolyte, and acid-base balance; ICU: Intensive care unit; MODS: multiple organ dysfunction syndrome, it is a serious, life-threatening condition characterized by progressive failure of two or more organ systems. Septicemia: A serious bloodstream infection characterized by the presence of bacteria or their toxins in the blood.

greater blood loss, a longer duration of hospitalization, greater rates of surgical site infection and ulcer occurrence, and greater rates of secondary amputation relative to those of Cluster 2 patients ($p < 0.001$). Cluster 1 patients also had a longer intensive care unit stay and had greater renal failure and in-hospital mortality rates than those of Cluster 2 ($p < 0.05$). Consequently, Cluster 1 patients showed poorer conditions and prognostic outcomes compared with Cluster 2. Differences in the follow-up data between both clusters showed that UMLA clustering was accurate. We found that the UMLA accurately classified PAD patients who underwent amputation into two groups. The prognosis of the disease differed between the groups according to patient preoperative clinical data.

3.3. Prediction model establishment for new PAD subtypes

Our results indicated that UMLA accurately divided PAD patients into clusters with different prognoses. Next, the pre-operative variables that were significantly different between the two clusters were subjected to logistic regression. As a result, gender, age, smoking history, diabetes and coronary heart disease history, pre-operative albumin levels, amputation level, and presence/absence of endovascular intervention were independent clustering predictors for PAD patients in Cluster 1, i.e., the group with a poorer prognosis ($p < 0.05$) (Table 3). Based on logistic regression results, we constructed a forest plot (Fig. 2A) and constructed a nomogram in which the eight risk factors were incorporated (Fig. 2B). The area under ROC curve (AUC) was 0.861 (95 % confidence interval: 0.830–0.893) (Fig. 2C). The calibration curve verified the probability of the actual and predicted nomograms, yielding a brier score of 0.148 (Fig. 2D). Supplementary Table 1 provides additional details on various performance indicators of the predictive model. These results indicated that the novel prediction model for PAD subtype accurately predicted adverse outcome probabilities in PAD patients who underwent amputation.

4. Discussion

4.1. Clinical value of AI-based predictive model for new PAD subtypes

As post-amputation prognosis of PAD patients is complex, risk assessment is required to provide personalized treatment and make better surgical decisions for these patients. In our study, UMLA was used to classify pre-operative data obtained from PAD patients into two clusters. Subsequent analysis of the follow-up data revealed that Cluster 1 cases showed poorer clinical conditions compared with Cluster 2 counterparts. According to the logistic regression results, we developed the predictive model for identifying patients at greater risk of disease progression and adverse prognosis. The predictive performance of the model was excellent, which indicated that artificial intelligence could be used to make prognostic predictions in individuals who have undergone amputation due to PAD.

One of the primary novelties of our study is the utilization of unsupervised machine-learning techniques to identify distinct patient subtypes within the PAD population undergoing amputation. Traditionally, risk stratification in this patient cohort has been based on clinical judgment and established risk factors. However, our approach offers a data-driven methodology for identifying high-risk individuals probably benefiting from targeted interventions and closer monitoring.

In healthcare, identifying patient heterogeneity is crucial. The UMLA is utilized for identifying patient subgroups and can help in the administration of personalized and targeted treatment [17]. For example, COVID-19 symptom were investigated for their

Table 3
Results of Univariate and Multivariate Logistic regression.

Characteristic	Univariable			Multivariable		
	OR	95 % CI	p-value	OR	95 % CI	p-value
Gender						
Female	–	–		–	–	
Male	3.23	2.20, 4.72	<0.001	2.43	1.41, 4.18	0.001
Age	1.05	1.04, 1.06	<0.001	1.07	1.05, 1.09	<0.001
Smoking						
No	–	–		–	–	
Yes	3.38	2.34, 4.90	<0.001	4.85	2.74, 8.59	<0.001
Diabetes						
No	–	–		–	–	
Yes	2.26	1.54, 3.31	<0.001	2.31	1.41, 3.79	<0.001
CAD						
No	–	–		–	–	
Yes	2.01	1.32, 3.06	0.001	2.12	1.23, 3.64	0.007
Albumin	0.94	0.92, 0.97	<0.001	0.94	0.91, 0.97	<0.001
Endovascular therapy						
No	–	–		–	–	
Yes	0.43	0.30, 0.62	<0.001	0.46	0.29, 0.74	0.001
Amputation level						
Above knee	–	–		–	–	
Below knee	2.01	1.34, 3.01	<0.001	2.29	1.36, 3.84	0.002
Ankle	2.26	1.39, 3.68	<0.001	4.24	2.24, 8.02	<0.001

CAD: Coronary artery disease.

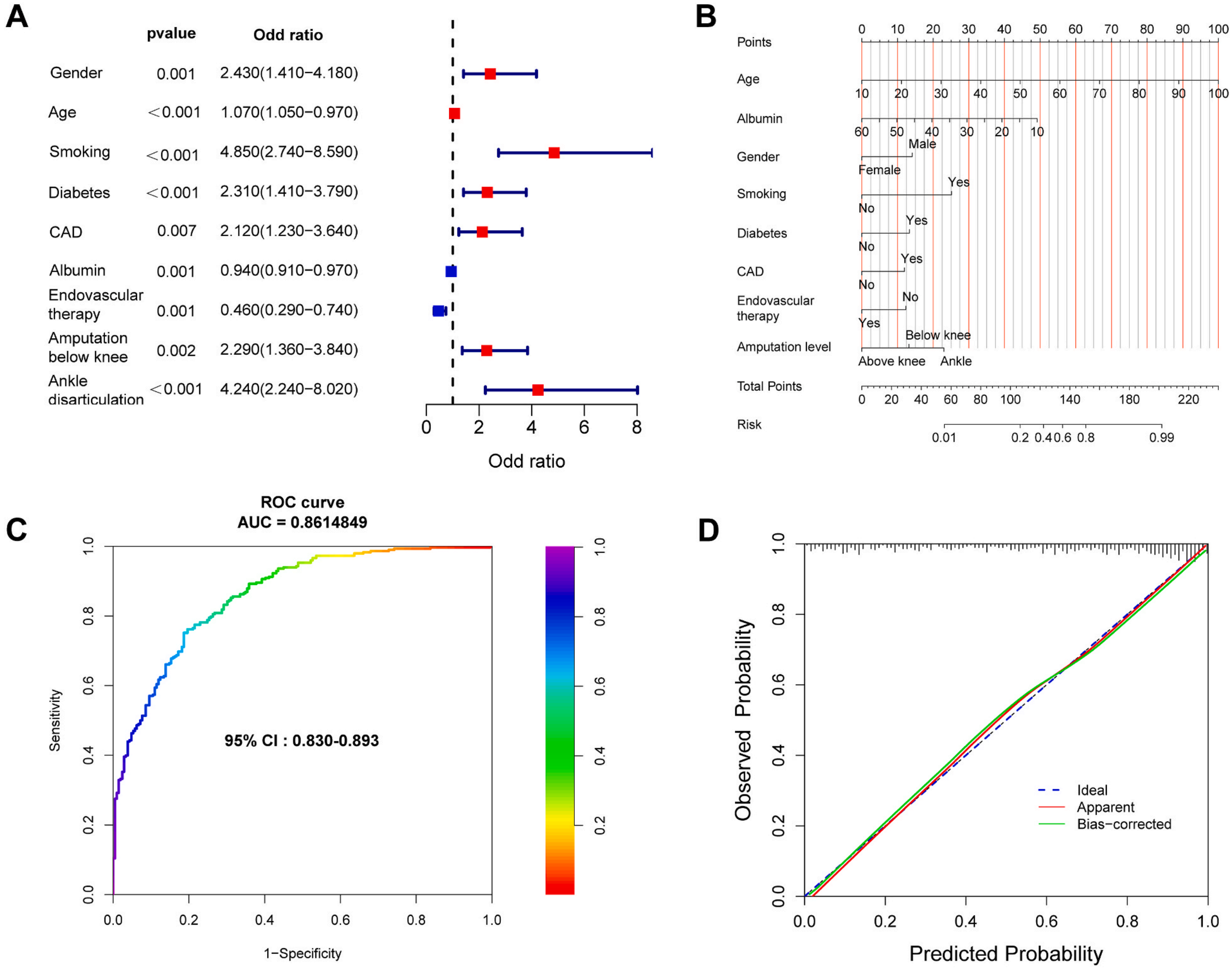

Fig. 2. Construction and validation of the predictive model.
(A)Forest plot of the pre-operative variables. (B) Nomogram for prediction model. (C) AUC of the nomogram. (D) Calibration curves for predictive model.
AUC: Area under curve; CI: Confidence interval; CAD: Coronary artery disease; ROC: Receiver operating characteristic.

longitudinal trajectories using unsupervised machine learning among healthcare workers; the findings of that study provided insights into the disease progression of non-hospitalized patients [18]. Similarly, Demanse et al. used the UMLA to identify clinical phenotypes in an osteoarthritis initiative database. Using this method, they understood and categorized patients more effectively [19]. Our study contributes to suggesting the effect of machine learning on personalized medicine. By leveraging advanced analytical techniques, we can enhance risk stratification and guide tailored interventions for patients with PAD undergoing amputation. This approach holds promise for improving patient outcomes and optimizing resource allocation in healthcare settings.

The application of our artificial intelligence (AI)-driven model in clinical practice might aid in the management of PAD. For PAD amputees at high risk, our model suggests that a tailored strategy involving intensive monitoring and stronger clinical interventions needs to be implemented. The personalized approach aims to delay disease progression and mitigate the risk of severe complications, such as infections, ulcers, and secondary amputation. This targeted intervention approach supports the principles of precision medicine, which emphasizes individualized care based on specific risk profiles.

4.2. Risk factors related to dismal prognostic outcome following amputation for PAD

We identified male gender, advanced age, smoking history, a medical history of diabetes and coronary artery disease, low pre-operative albumin levels, below-knee amputation, ankle disarticulation, and a lack of pre-operative endovascular intervention to open occluded blood vessels as high-risk factors related to dismal prognostic outcome of PAD patients.

PAD shows an increasing incidence with age, and those who are 70 years old or older have a significantly greater incidence than those under 70 years old. The probability of poor prognosis in patients ≥70 years increases by 2.206 folds relative to those <70 years [20]. Some researchers have suggested that diabetics are associated with an increased PAD risk. The probability of intermittent claudication in diabetic patients is 2–3 times greater than that in non-diabetic individuals [21]. Diabetes primarily affects arteries below the knees and increases the risks of lower limb ischemic ulcers, amputation, or even death; therefore, diabetes deteriorates outcomes for PAD patients [22]. Smokers have a two-fold increased PAD risk compared with non-smokers [23]. Although quitting smoking reduces AD risk, according to one community-based cohort study, the risk for ex-smokers can reach the level of risk for non-smokers after 30 years [24].

Several studies on PAD and multiple vascular diseases have been conducted. Multiple vascular diseases, including PAD, occur in atherosclerotic patients and affect diverse vascular beds. According to certain studies evaluating novel antithrombotic or lipid-lowering therapeutic techniques for preventing adverse cardiovascular events, patients with multi-vascular diseases had a greater risk than those without such diseases. The combined risk of PAD and myocardial infarction for major adverse cardiovascular events was the highest, and the 2.5-year risk was 14.9 %. PAD patients without myocardial infarction were associated with a greater major adverse cardiovascular event risk (10.3 %) compared with non-PAD myocardial infarction cases (7.6 %) [25]. Patients with PAD and concomitant conditions like coronary artery disease have an increased risk of developing systemic arterial sclerosis, which significantly increases the cardiovascular event risk. Coronary artery disease can increase amputation risk due to both intraoperative and post-operative complications. Thus, it has severe adverse effects on patient prognosis [26].

With advancements in medical expertise and equipment, the application of endovascular interventions for PAD patients has become popular because these interventions are minimally invasive, safe, and efficacious and facilitate rapid recovery [27]. Thrombolysis, balloon dilation, stent implantation, intraluminal rotational cutting, arterial dilation, and reconstruction represent common endovascular intervention methods [28]. Endovascular intervention can be used for efficiently expanding and reconstructing narrowed or occluded vessels. Thus, it can accelerate blood circulation, greatly enhance vascular patency, and considerably decrease the mortality and amputation rates for patients [29]. Due to improvements in technology and equipment, interventional treatment for PAD has advanced significantly; specifically, post-operative success rates and short-term efficacy have greatly improved [30].

To summarize, in this study, we showed that factors like age, sex, smoking history, diabetes and coronary artery disease history, pre-operative ALB levels, amputation level, and whether endovascular intervention to open occluded blood vessels is conducted significantly affect the prognosis of PAD amputees. Our findings highlighted the need to comprehensively consider these risk factors while developing individualized treatment plans for improving disease outcome and reducing complications in patients.

Currently, the association between the amputation level and post-amputation prognosis remains controversial. A notable finding in our study was the significantly worse prognosis observed in patients who underwent below-knee amputation and ankle disarticulation than in those who underwent above-knee amputation. Patients with higher amputation levels often experience a lack of tissue viability, necessitating more extensive amputation procedures. Conversely, those who undergo lower-level amputations may face an increased risk of non-healing complications stemming from inadequate blood perfusion in distal diseased arteries [31]. Despite the common practice of lowering amputation levels to enhance the post-amputation quality of life, our results underscore the critical importance of considering vascular patency. Vascular occlusion directly influences post-operative healing in amputees, with compromised blood flow leading to delayed wound healing, ulceration, infection, and the need for secondary amputation [32]. Thus, while striving to optimize the quality of life for amputees, clinicians must carefully weigh the benefits of lower amputation levels against the risk of vascular complications to mitigate adverse outcomes.

4.3. Limitations

This study had certain limitations. First, as the study had a retrospective design, we could not eliminate selection bias. Second, all data were obtained from one individual center; consequently, multicenter studies are needed for external validation and for optimizing the prediction model.

5. Conclusion

UMLA was used to identify new PAD phenotypes among patients showing different risks of amputation. Our constructed AI-driven predictive model for PAD subtypes showed that it can be used for risk stratification and clinical management with high accuracy and reliability.

Ethics statement

Our study protocols gained approval from Ethics Committee of The First Affiliated Hospital of Guangxi Medical University (approval number: 2023-E314-01). The present work was carried out following ethical principles in Declaration of Helsinki. The Ethics Committee of The First Affiliated Hospital of Guangxi Medical University exempted the need of informed consent in this retrospective study.

Funding

The present study was funded by the National Natural Science Foundation of China [grant numbers: 81960091].

Data availability statement

The original contributions displayed in the present work can be obtained from article/supplementary material. Further inquiries are directed to corresponding authors.

CRediT authorship contribution statement

Yuanliang Ma: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Lin Zhang:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Data curation, Conceptualization. **Que Li:** Writing – original draft, Validation, Resources, Methodology, Investigation, Data curation. **Xiao Qin:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Our thanks should go to Prof. Xiao Qin (Vascular Surgery Ward, The First Affiliated Hospital of Guangxi Medical University) for their assistance in the present work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e34602.

References

[1] Y.H. Zheng, X.T. Song, [Progress and prospect of the treatment of lower extremity arteriosclerosis obliterans], Zhonghua Wai Ke Za Zhi 59 (12) (2021) 961–964, https://doi.org/10.3760/cma.j.cn112139-20210923-00453.

[2] W. Lian, et al., Clinical significance of endothelin-1 and C reaction protein in restenosis after the intervention of lower extremity arteriosclerosis obliterans, J. Invest. Surg. 34 (7) (2021) 765–770, https://doi.org/10.1080/08941939.2019.1690600.

[3] E. Martelli, et al., Sex-related differences and factors associated with peri-procedural and 1 Year mortality in chronic limb-threatening ischemia patients from the CLIMATE Italian registry, J Pers Med 13 (2) (2023), https://doi.org/10.3390/jpm13020316.

[4] W. Yao, et al., Effects of valsartan on restenosis in patients with arteriosclerosis obliterans of the lower extremities undergoing interventional therapy: a prospective, randomized, single-blind trial, Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. 26 (2020) e919977, https://doi.org/10.12659/msm.919977.

[5] M. Ye, et al., Neutrophil-lymphocyte ratio and platelet-lymphocyte ratio predict severity and prognosis of lower limb arteriosclerosis obliterans, Ann. Vasc. Surg. 64 (2020) 221–227, https://doi.org/10.1016/j.avsg.2019.09.005.

[6] K. Claesson, T. Kölbel, S. Acosta, Role of endovascular intervention in patients with diabetic foot ulcer and concomitant peripheral arterial disease, Int. Angiol. 30 (4) (2011) 349–358.

[7] V.A. Badtieva, D.N. Voroshilova, N.V. Sichinava, [Use of enhanced external counterpulsation in the treatment and rehabilitation of patients with atherosclerosis obliterans of the lower extremity], Vopr. Kurortol. Fizioter. Lech. Fiz. Kul't. 96 (4) (2019) 5–11, https://doi.org/10.17116/kurort2019960415.

[8] J. Dong, et al., Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care, Crit. Care 25 (1) (2021) 288, https://doi.org/10.1186/s13054-021-03724-0.

[9] J. Pan, et al., Unsupervised machine learning identifies predictive progression markers of IPF, Eur. Radiol. 33 (2) (2023) 925–935, https://doi.org/10.1007/s00330-022-09101-x.

[10] C.M. Eckhardt, et al., Unsupervised machine learning methods and emerging applications in healthcare, Knee Surg. Sports Traumatol. Arthrosc. 31 (2) (2023) 376–381, https://doi.org/10.1007/s00167-022-07233-7.

[11] F. Chantraine, et al., Classification of stiff-knee gait kinematic severity after stroke using retrospective k-means clustering algorithm, J. Clin. Med. 11 (21) (2022), https://doi.org/10.3390/jcm11216270.

[12] S. Wu, et al., Genome-wide identification of immune-related alternative splicing and splicing regulators involved in abdominal aortic aneurysm, Front. Genet. 13 (2022) 816035, https://doi.org/10.3389/fgene.2022.816035.

[13] A. Sebastian, et al., Association of snoring characteristics with predominant site of collapse of upper airway in obstructive sleep apnea patients, Sleep 44 (12) (2021), https://doi.org/10.1093/sleep/zsab176.

[14] E. Ahlqvist, et al., Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables, Lancet Diabetes Endocrinol. 6 (5) (2018) 361–369, https://doi.org/10.1016/s2213-8587(18)30051-2.

[15] M.L.S. Mationg, et al., Determining the impact of a school-based health education package for prevention of intestinal worm infections in the Philippines: protocol for a cluster randomized intervention trial, JMIR Res Protoc 9 (6) (2020) e18419, https://doi.org/10.2196/18419.

[16] L. Lovmar, et al., Silhouette scores for assessment of SNP genotype clusters, BMC Genom. 6 (2005) 35, https://doi.org/10.1186/1471-2164-6-35.

[17] H. Habehh, S. Gohel, Machine learning in healthcare, Curr. Genom. 22 (4) (2021) 291–300, https://doi.org/10.2174/1389202922666210705124359.

[18] V. Gupta, et al., Unsupervised machine learning to investigate trajectory patterns of COVID-19 symptoms and physical activity measured via the MyHeart Counts App and smart devices, NPJ Digit Med 6 (1) (2023) 239, https://doi.org/10.1038/s41746-023-00974-w.

[19] D. Demanse, et al., Unsupervised machine-learning algorithms for the identification of clinical phenotypes in the osteoarthritis initiative database, Semin. Arthritis Rheum. 58 (2023) 152140, https://doi.org/10.1016/j.semarthrit.2022.152140.

[20] Y.B. Yang, et al., A risk predictor of restenosis after superficial femoral artery stent implantation: relevance of mean platelet volume, BMC Cardiovasc. Disord. 20 (1) (2020) 361, https://doi.org/10.1186/s12872-020-01633-8.

[21] F.N. Brand, R.D. Abbott, W.B. Kannel, Diabetes, intermittent claudication, and risk of cardiovascular events. The Framingham Study, Diabetes 38 (4) (1989) 504–509, https://doi.org/10.2337/diab.38.4.504.

[22] N. Ding, et al., Traditional and nontraditional glycemic markers and risk of peripheral artery disease: the Atherosclerosis Risk in Communities (ARIC) study, Atherosclerosis 274 (2018) 86–93, https://doi.org/10.1016/j.atherosclerosis.2018.04.042.
[23] E.M. Willigendael, et al., Influence of smoking on incidence and prevalence of peripheral arterial disease, J. Vasc. Surg. 40 (6) (2004) 1158–1165, https://doi.org/10.1016/j.jvs.2004.08.049.
[24] N. Ding, et al., Cigarette smoking, smoking cessation, and long-term risk of 3 major atherosclerotic diseases, J. Am. Coll. Cardiol. 74 (4) (2019) 498–507, https://doi.org/10.1016/j.jacc.2019.05.049.
[25] M.P. Bonaca, et al., Low-density lipoprotein cholesterol lowering with evolocumab and outcomes in patients with peripheral artery disease: insights from the FOURIER trial (further cardiovascular outcomes research with PCSK9 inhibition in subjects with elevated risk), Circulation 137 (4) (2018) 338–350, https://doi.org/10.1161/circulationaha.117.032235.
[26] J. Steffel, et al., The COMPASS trial: net clinical benefit of low-dose rivaroxaban plus aspirin as compared with aspirin in patients with chronic vascular disease, Circulation 142 (1) (2020) 40–48, https://doi.org/10.1161/circulationaha.120.046048.
[27] W. Ye, et al., [Endovascular revascularization for lower extremity atherosclerosis obliterans in elderly patients:a report of 86 cases], Zhonghua Wai Ke Za Zhi 48 (4) (2010) 261–264.
[28] D.V. Kosaev, [Early outcomes of therapy and indirect revascularization surgery in patients with critical ischemia of lower extremities], Khirurgiia (8) (2020) 55–60, https://doi.org/10.17116/hirurgia202008155.
[29] H.Y. Duan, et al., Special issues in endovascular treatment of lower extremities arteriosclerosis obliterans], Zhonghua Yixue Zazhi 93 (13) (2013) 1008–1011.
[30] H. Lu, P. Guo, [Plasma heparin cofactor II activity correlates with the incidence of in-stent restenosis after the intervention of arteriosclerosis obliterans in lower extremity], Zhong Nan Da Xue Xue Bao Yi Xue Ban 40 (2) (2015) 177–181, https://doi.org/10.11817/j.issn.1672-7347.2015.02.010.
[31] J.A. Beckman, et al., Microvascular disease, peripheral artery disease, and amputation, Circulation 140 (6) (2019) 449–458, https://doi.org/10.1161/circulationaha.119.040672.
[32] N. Narula, et al., Pathology of peripheral artery disease in patients with critical limb ischemia, J. Am. Coll. Cardiol. 72 (18) (2018) 2152–2163, https://doi.org/10.1016/j.jacc.2018.08.002.