Predicting future occlusion or stenosis of lower extremity bypass grafts using artificial intelligence to simultaneously analyze all flow velocities collected in current and previous ultrasound examinations

Xiao Luo, PhD,[a] Fattah Muhammad Tahabi, BS,[a] Dave M. Rollins, RVT,[b] *and* Alan P. Sawchuk, MD,[c] *Indianapolis, IN*

ABSTRACT

Objective: Routine surveillance with duplex ultrasound (DUS) examination is recommended after femoral-popliteal and femoral-tibial-pedal vein bypass grafts with various intervals postoperatively. The presently used methodology to analyze bypass graft DUS examination does not use all the available data and has been shown to have a significant rate for missing impending bypass graft failure. The objective of this research is to investigate recurrent neural networks (RNNs) to predict future bypass graft occlusion or stenosis.

Methods: This study includes DUS examinations of 663 patients who had bypass graft operations done between January 2009 and June 2022. Only examinations without missing values were included. We developed two RNNs (a bidirectional long short-term memory unit and a bidirectional gated recurrent unit) to predict bypass graft occlusion and stenosis based on peak systolic velocities collected in the 2 to 5 previous DUS examinations. We excluded the examinations with missing values and split our data into training and test sets. Then, we applied 10-fold cross-validation on training to optimize the hyperparameters and compared models using the test data.

Results: The bidirectional long short-term memory unit model can gain an overall sensitivity of 0.939, specificity of 0.963, and area under the curve of 0.950 on the prediction of bypass graft occlusion, and an overall sensitivity of 0.915, specificity of 0.909, and area under the curve of 0.912 predicting the development of a future critical stenosis. The results on different bypass types show that the system performs differently on different types. The results on subcohorts based on gender, smoking status, and comorbidities show that the performance on current smokers is lower than the never smoker.

Conclusions: We found that RNNs can gain good sensitivity, specificity, and accuracy for the detection of impending bypass graft occlusion or the future development of a critical bypass graft stenosis using all the available peak systolic velocity data in the present and previous bypass graft DUS examinations. Integrating clinical data, including demographics, social determinants, medication, and other risk factors, together with the DUS examination may result in further improvements.

Clinical Relevance: Detecting bypass graft failure before it occurs is important clinically to prevent amputations, salvage limbs, and save lives. Current methods evaluating screening duplex ultrasound examinations have a significant failure rate for detecting a bypass graft at risk for failure. Artificial intelligence using recurrent neural networks has the potential to improve the detection of at-risk bypass graft before they fail. Additionally, artificial intelligence is in the news and is being applied to many fields. Vascular surgeons need to know its potential to improve vascular outcomes. (JVS—Vascular Science 2024;5:100192.)

Keywords: Bypass graft; Stenosis; Occlusion; Duplex ultrasound; Machine learning

Peripheral artery disease affects 8 to 12 million Americans. This results in approximately 185,000 amputations per year and approximately 2 million Americans living with limb loss.[1] The annual cost of amputation care in the United States is $13.7 billion.[2] Lower extremity revascularizations are done to alleviate disabling claudication, salvage limbs, and prevent amputations. Occlusion of a bypass graft causes recurrent limb ischemia with an increased risk for an amputation in patients with limb threatening ischemia. Unfortunately, lower extremity bypass grafts are prone to failure. The 3-year primary patency of a femoral to popliteal artery bypass above

From the School of Engineering and Technology, Indiana University Purdue University at Indianapolis[a]; IU Health[b]; and the Department of Surgery, Indiana University School of Medicine.[c]

Correspondence: Alan P. Sawchuk, MD, Indiana University School of Medicine, Indiana University Medical Center, Indianapolis, IN 46202 (e-mail: asawchuk@iupui.edu).

The editors and reviewers of this article have no relevant financial relationships to disclose per the Journal policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

2666-3503

https://doi.org/10.1016/j.jvssci.2024.100192

the knee for saphenous vein is 73%, and patency with a polytetrafluoroethylene graft is 66%. For infrapopliteal grafts the 3-year primary patency is 66% for saphenous vein and the 2-year patency with a polytetrafluoroethylene graft is 32%.[3] Bypass grafts remain at risk for failure over their lifetime. It is well-known that bypass graft patency and limb salvage are improved if a bypass graft can be repaired when it is failing but has not yet occluded.[3,4] To salvage limbs and prevent amputations, vascular physicians do bypass graft surveillance to look for signs that a bypass graft may fail so that an intervention can be done before the bypass graft occludes and threatens the limb. This surveillance consists of the clinical examination, including a lower extremity arterial bypass graft duplex ultrasound (DUS) examination, which examines how blood moves through the arteries and graft. It provides information about the speed of the blood flow, if there are any bypass graft stenoses, and, if so, where they are.

Unfortunately, present methods of bypass graft surveillance have been shown to not detect a significant number of impending bypass graft occlusions. The earliest method of bypass graft surveillance was done using the clinical examination and the ankle to brachial pressure ratio determined during a lower extremity Doppler study (LEAD) study. A LEAD study is generally considered suggestive of impending bypass graft failure if the ankle-brachial ratio decreases by >0.15.[5] Several studies have indicated that a LEAD study alone misses many impending bypass graft failures.[6,7] It was initially thought that bypass graft DUS examination, which can detect low flow in the bypass graft, or stenosis within the bypass graft, might improve bypass graft surveillance and mitigate bypass graft failure.[8,9] More recent studies, including meta-analysis, have indicated that the use of routine DUS surveillance does not result in a significant change in the primary, secondary, or assisted primary patency rates, or a statistically significant decrease in the amputation rate compared with an LEAD study and clinical examination.[10] Another study suggested that a prediction model might indicate which patients would benefit from periodic bypass graft DUS screening.[11] A prospective randomized trial of vein bypass graft surveillance also did not show any additional benefit to adding DUS to vein graft surveillance.[12] This finding indicates that the ability to predict bypass graft occlusion before it occurs needs improvement to salvage more bypass grafts and save more limbs.

Machine learning techniques have been applied recently to clinical data to predict various clinical outcomes of vascular surgery and other operations or diseases,[13-15] including the prediction of graft failure after liver transplantation,[16,17] predicting long-term mortality and graft failure in patients undergoing heart transplant,[18] and vein graft surveillance analysis.[19] The work on the vein graft surveillance analysis used a decision tree for prediction. The input to the decision tree included the results of one postoperative DUS examination and early clinical variables.[19] Prior relevant studies have not investigated the feasibility of applying deep learning methods on a set of DUS studies to predict potential bypass graft failure or the development of a high-grade stenosis before it occurring to promote limb salvage. Although a recent study shows substantial associations between DUS values and stenosis or occlusion,[20] there is no model developed using a large patient cohort.

The objective of this research was to investigate state-of-the-art deep learning algorithms on predicting bypass graft occlusion and high-level stenosis using a set of flow velocity values collected via DUS examination, which is done periodically during postoperative routine surveillance screening.

ARTICLE HIGHLIGHTS

- **Type of Research:** A global, multicenter, prospective, nonrandomized, single-arm, investigational device exemption study
- **Key Findings:** Treatment of symptomatic iliofemoral venous outflow obstruction using the Zilver Vena venous stent (Cook Ireland, Ltd, Limerick, Ireland) in 243 patients resulted in a 30-day freedom from major adverse events (MAE) rate of 96.7% and 12-month primary quantitative patency rate of 89.9%, which surpassed the corresponding performance goals. Also, significant improvement in clinical symptoms was demonstrated through 12 months.
- **Take Home Message:** The 12-month results indicate that the Zilver Vena venous stent is safe and effective.

METHODS

The study cohort consisted of 663 adults (aged ≥18 years) who underwent bypass graft operations between January 2009 to June 2022 and had two or more postoperative DUS surveillance screenings done in Indiana University Health. The study was approved by the Institutional Review Board of Indiana University.

All DUS examinations were performed in the hospitals by trained vascular ultrasound technologists in an Intersocietal Accreditation Commission accredited vascular laboratory. The bypass graft DUS measures peak systolic velocities (PSVs) and velocity ratios of adjacent PSVs within the graft and adjacent arteries. When performing DUS examinations, PSVs are obtained from the adjacent proximal inflow artery and distal outflow artery in addition to the full length of the graft conduit, including inflow artery (PSV_{ia}), proximal anastomosis (PSV_{pa}), proximal graft (PSV_{pg}), proximal/mid graft (PSV_{pmg}), mid graft (PSV_{mg}), mid/distal graft (PSV_{mg}), distal graft (PSV_{dg}), distal anastomosis (PSV_{da}), and outflow artery (PSV_{oa}). Quality assurance of these studies is done by comparison with computed tomography angiography (CTA) or angiography when available. Eleven percent of these bypass graft DUS examinations have a concurrent CTA or arteriography study. These radiographic studies are predominantly done in patients with abnormal findings, because an asymptomatic patient with a normal study would not have an indication for a more invasive study. The typical surveillance protocol included an examination 1 month after the procedure and at 3, 6, and 12 months after the procedure. Additional screenings are done on an annual or semiannual basis, depending on the findings. Indications for additional examinations included aberrant findings on DUS examination, impaired wound healing, aberrant pain, or other signs of ischemia.

Before the examination, a history is obtained from the referring physician, medical record, and the patient to determine the reason for the examination. The medical history focused on presenting symptoms, previous history of revascularization procedures, hypertension, diabetes, cerebrovascular disease, cardiac disease, tobacco use, and hyperlipidemia. The determination of an occlusion is based on the DUS examination when any PSV is measured as 0 or there is no flow in the graft. The high-level stenosis is input based on the clinical impression text on the report. The clinician diagnoses a >75% bypass graft stenosis if the ratios of the velocities between the stenotic segment and the adjacent nonstenotic segment is >3.5 or the PSV is >300 cm/s.[21] Additionally, a PSV of <40 cm/s is considered to be low flow in the graft with a risk for bypass graft failure. The inflow artery, all segments of the bypass graft, and the outflow artery were evaluated in these studies. If the clinical impression states that there is a >75% stenosis of the graft or any part of the artery, it is counted as a stenosis case. When patients had bypass grafts on both sides, each side was considered separately. One patient can have multiple occlusions or stenosis cases after operations or reinterventions. Each occlusion or stenosis is counted as an independent case. We excluded the DUS examinations if there are missing values. On average, 1.3 examinations of each patient are excluded.

In this research, we applied deep learning methods for the prediction of bypass graft failure and stenosis. Our method is to predict the probability of future bypass graft occlusion or stenosis based on the PSVs collected in all the previous DUS examinations. The prediction of future elevated PSVs can assist clinicians to determine which bypass grafts may develop stenosis or occlusion earlier. We applied stratified sampling on our data based on occlusion, stenosis, and number of examinations before the occlusion and stenosis and divided our data into training (80%) and test (20%) datasets. Then, we applied 10-fold cross-validation on the training data to identify the optimal parameters and compare the performances of two recurrent neural network (RNN) models. The optimal parameters were then used to build the models using training data only. Finally, the test data are used to produce the performances. Because the dataset is imbalanced with more normal cases, we applied an oversampling technique[22] to increase the occlusion and stenosis cases, so that the number of cases in each category is similar.

RNNs are a class of artificial neural networks that have been used in various predictive analysis tasks including those in the clinical domain, such as seizure detection or human brain state prediction using electroencephalographic data,[23,24] hospital visit prediction using electronic health records data,[25] and so on. RNNs are an extension of a conventional feed-forward neural network with a recurrent hidden state whose activation at each time is dependent on that of the previous time, shown as Equation 1 and 2, where x^t, h^t, and y^t are the input, hidden layer vector, and output at time t, $W, U,$ and b are the

$$h^t = g_h\left(W_h h^{t-1} + U_h x^t + b_h\right) \quad (1)$$

$$y^t = g_y\left(W_y h^t + b_y\right) \quad (2)$$

The advantage of RNNs is it uses the clinical time-series data in all previous time steps or, in this case, all previous bypass graft DUS examinations to predict the final outcomes. Deep RNNs are computationally expensive and may suffer from the problem of vanishing or exploding gradients. Hence, variations of RNNs, including bidirectional long short-term memory units (BiLSTMs)[26] and bidirectional gated recurrent units (BiGRUs)[27] were introduced to solve the vanishing gradient problem. In this study, we experiment with both BiLSTM and BiGRU models for bypass graft occlusion and stenosis prediction, since they are used in recent literature on other clinical predictive analysis.[28-31] The BiLSTM model is built on LSTM cells with two LSTM layers. The input sequence flows forward through a LSTM layer and backward in the additional LSTM layer. Then, we combine the outputs from both LSTM layers by concatenating them. Each LSTM cell acts like an RNN cell with the main difference being that LSTM cell has three gates: forget gate, input gate, and output gate. A GRU is a variation of an LSTM. The main difference between GRU and LSTM cells is that GRU combines the forget and input gates into an update gate and merges the cell state and hidden state. Based on the literature, BiGRU sometimes performs better than BiLSTM on prediction problems,[32,33] or vice versa.[34]

In our study, the input to the BiLSTM or BiGRU cells are the PSVs collected during DUS examinations at different time points ($DUS_0, DUS_1, \ldots, DUS_t$ shown as Fig 1). Each DUS examination includes all nine PSVs (PSV_{ia}, PSV_{pa}, PSV_{pg}, PSV_{pmg}, PSV_{mg}, PSV_{mg}, PSV_{dg}, PSV_{da}, and PSV_{oa}). The output of the hidden layer is sent to a classification layer for prediction. If a patient has $n\,(n \geq 3)$ DUS

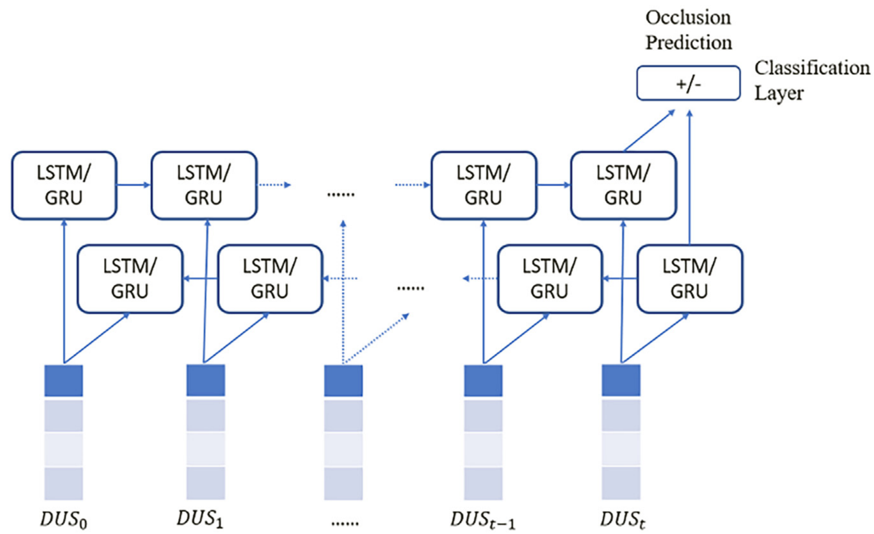

Fig 1. Recurrent neural networks (RNN) for the prediction of occlusion or stenosis. *DUS*, Doppler ultrasound examination; *GRU*, gated recurrent unit; *LSTM*, long short-term memory.

examinations, we generate $n-2$ sequences for prediction. The objective is to predict the examination results of $DUS_3, \ldots, DUS_n$. For example, if a patient has five DUS examinations, we generate three input sequences based on the first, second, third, and fourth DUS examinations, respectively. The objective is to predict the occlusion or stenosis status of the third, fourth, and fifth DUS examinations. The performances are calculated based on the sequences of the DUS examinations. We call each sequence a case in the following content.

Both RNN models were set to have 15 neurons and a dropout rate of 0.1; they were trained with 15 maximum epochs with a batch size of 500. The Adam optimizers were used with a learning rate of 0.01 and a weight decay of 0.05. For the loss function, categorical cross-entropy is used.

STUDY COHORT

Our study cohort has 663 unique patients in total. For patients for whom we had the original operative note, this included 172 femoropopliteal bypasses with vein (Current Procedural Terminology [CPT] code 35556), 122 femorotibial bypasses with vein (CPT code 35566), 73 femoropopliteal bypasses with polytetrafluoroethylene (CPT code 35656), 62 femoral-femoral bypasses with polytetrafluoroethylene (CPT code 35661), 25 popliteal-tibial bypasses with vein (CPT code 35571), and 24 axillary-femoral-femoral bypasses with polytetrafluoroethylene (CPT code 35654). The analyses of these individual types of graft were included in this research. After data preprocessing, there are 53 patients who have only occlusion, 101 patients who have only stenosis, 8 patients who have both occlusion and stenosis studies, 6 patients with more than one occlusion, 19 patients with multiple stenosis studies, and 501 patients in the control group. A patient was categorized in occlusion, stenosis, and control groups when we applied the univariate analysis. Table I shows the demographic and comorbidity characteristics of the study cohort. In univariate analysis, demographic and comorbidity data were compared between control patients and patients with occlusion or stenosis (independent vs nonindependent) using an independent *t* test for variables and a χ^2 test for categorical variables.

There are no significant differences between the age of control, occlusion, or stenosis patients. There are more female patients in the occlusion group ($P < .01$). Regarding smoking status, the control group has more patients who never smoked ($P < .01$). The collected comorbidities that are often associated with bypass failure or stenosis patients are also listed in Table I. Compared with the control group, the occlusion group has more patients with hypertension ($P < .01$), and the stenosis group has more patients with hyperlipidemia ($P < .05$).

RESULTS

There are 69 studies that indicated occlusion, 133 studies that indicated a >75% stenosis, and 3198 normal studies. On quality assurance data, the interpreting physicians had an overall accuracy of 92% matching the gold standard of CTA or arteriography for determining whether a bypass graft had <50%, 50% to 74%, or 74% to 99% stenosis matching the location of the stenosis when compared with their DUS interpretation when CTA or arteriographic imaging was available. Through applying the paired *t* test on the area under the curve (AUC) values gained using the 10-fold cross-validation on training, we found that BiLSTM works better on stenosis prediction ($P < .05$), whereas they performed competitively on occlusion prediction. Based on the performances shown in Tables II and III, the

Table I. Summary of the study cohort

	Total	Control	Occlusion only	*P* value occlusion	Stenosis only	*P* value stenosis
No. of patients	663	501	53		101	
Age, years	73.95 ± 13.47	74.12 ± 13.83	72.47 ± 13.74	.42	74.79 ± 11.48	.77
Sex						
Male	407 (61.4)	314 (62.7)	25 (47.2)	**<.01**	63 (62.4)	.84
Female	254 (38.3)	186 (37.1)	28 (52.8)	**<.01**	37 (36.6)	.84
Unknown	2 (0.3)	1 (0.2)	0 (0.0)	-	1 (1.0)	-
Smoking status						
Never smoker	250 (37.7)	196 (39.1)	14 (24.4)	**<.01**	38 (37.6)	.82
Previous smoker	208 (31.4)	155 (30.9)	19 (35.8)	.68	33 (32.7)	.83
Current smoker	205 (30.9)	150 (29.9)	20 (37.7)	.36	30 (29.7)	.76
Comorbidity						
Diabetes	279 (42.1)	201 (40.1)	25 (47.2)	.51	49 (48.5)	.07
CAD	228 (34.4)	171 (34.1)	19 (35.8)	.83	33 (32.7)	.82
Hypertension	517 (78.0)	380 (75.8)	47 (88.7)	**<.01**	82 (82.2)	.37
Hyperlipidemia	466 (70.3)	341 (68.1)	40 (75.5)	.41	77 (76.2)	**<.05**

CAD, Coronary artery disease.
Values are man ± standard deviation or number (%). Boldface entries indicate statistical significance.

Table II. Confusion Matrix of bidirectional long short-term memory units (BiLSTM) model on Test data

	Actual		
Predicted	Normal	Occlusion	Stenosis
Normal	1525	68	163
Occlusion	2	31	0
Stenosis	5	0	54

Table III. Confusion Matrix of bidirectional gated recurrent units (BiGRU) model on Test data

	Actual		
Predicted	Normal	Occlusion	Stenosis
Normal	1377	180	199
Occlusion	2	31	0
Stenosis	6	1	52

performances of BiLSTM on the test set had an overall sensitivity of 0.939, specificity of 0.963, and receiver operating characteristic (ROC)-AUC of 0.950 for predicting impending bypass graft failure before occlusion in bypass grafts that occluded. Based on the confusion matrices shown in Tables II and III, both models detected 31 of the 33 bypasses that occluded in test sets, leaving only 2 bypasses presenting occlusion without studies indicating an impending graft failure. The BiLSTM model on the test set had an overall sensitivity of 0.915, specificity of 0.909, and ROC-AUC of 0.912 for detecting a bypass graft that would proceed to develop a >75% stenosis before a standard bypass graft DUS examination would classify the graft at risk. The confusion matrices shown in Tables I and II show that BiLSTM detected 54 of the 59 bypass stenosis cases, whereas BiGRU detected 52 of 59 bypass stenosis cases.

The results of both BiLSTM and BiGRU in Tables IV and V show that, when the input for the number of DUS examinations is three or more, the performance of BiLSTM can gain high sensitivity, specificity, and AUC-ROC values. When at most two consecutive DUS examinations are used for occlusion prediction, BiLSTM can gain an AUC-ROC of 0.902. With the increase of the number of DUS examinations, the prediction performance increases. The overall performance on stenosis prediction is approximately 3.8% less than the occlusion prediction based on the AUC-ROC value on test data when BiLSTM is used.

We have also included the performances on different bypasses in Tables VI and VII. There are no occlusion cases of the popliteal-tibial bypasses and only two occlusion cases of axillary-femoral-femoral bypasses in the study cohort. There is no stenosis case of the axillary-femoral-femoral bypass cases in the study cohort. Thus, the results of these two types are not included. The results show that both models perform better on the femoropopliteal bypass occlusion prediction. However, the BiLSTM model cannot predict the only one occlusion case of femoral-femoral bypass in the test data (in Table VI, the sensitivity is 0 and the AUC value cannot be calculated), whereas the BiGRU model predicted it correctly. Both models do not perform as well on the stenosis cases that happened with the femorotibial bypass.

Table IV. Comparison of performances for occlusion prediction

No. of DUS examinations	2	3	4	5	Overall
Support	16	7	8	2	33
BiLSTM					
Sensitivity	0.875	1.000	1.000	1.000	0.939
Specificity	0.929	0.979	0.992	0.979	0.963
AUC	0.902	0.989	0.996	0.989	0.950
BiGRU					
Sensitivity	0.938	1.000	0.875	1.000	0.939
Specificity	0.829	0.921	0.951	0.983	0.901
AUC	0.883	0.961	0.913	0.991	0.920

AUC, Area under the curve; *BIGRU*, bidirectional gated recurrent units; BiLSTM, bidirectional long short-term memory units; DUS, Doppler ultrasound.

Table V. Comparison of performances for stenosis prediction

No. of DUS examinations	2	3	4	5	Overall
Support	23	22	7	7	59
BILSTM					
Sensitivity	0.957	0.864	0.857	1.000	0.915
Specificity	0.850	0.935	0.943	0.974	0.909
AUC	0.903	0.899	0.900	0.987	0.912
BIGRU					
Sensitivity	0.783	0.955	0.857	1.000	0.881
Specificity	0.799	0.921	0.962	0.970	0.889
AUC	0.791	0.938	0.909	0.985	0.885

AUC, Area under the curve; *BIGRU*, bidirectional gated recurrent units; *BiLSTM*, bidirectional long short-term memory units; *DUS*, Doppler ultrasound.

Table VI. Performances on different bypass types for occlusion prediction

CPT codes	35,556	35,566	35,656	35,661
Support	10	8	3	1
BILSTM				
Sensitivity	1	0.875	1	0
Specificity	0.973	0.969	0.966	0.975
AUC	0.986	0.922	0.983	-
BIGRU				
Sensitivity	1	0.875	1	1
Specificity	0.923	0.948	0.816	0.884
AUC	0.962	0.911	0.908	0.942

AUC, Area under the curve; *BIGRU*, bidirectional gated recurrent units; *BiLSTM*, bidirectional long short-term memory units; CPT, Current Procedural Terminology.

Because the study cohort has more male patients, and most of the patients have various chronic conditions, we investigated the performances of BiLSTM model performance on each of the subcohorts defined by the gender, smoking status, and comorbidities. AUC-ROC curves, which are analyses of accuracy for imbalanced data, are shown in Figs 2 and 3. For occlusion prediction, shown as Fig 2, there is no significant difference in prediction when patients have hypertension. However, the AUC is slightly higher for male patients or without diabetes. If the patient is a current smoker, the returned AUC is lower.

For the prediction of a bypass graft developing a stenosis, shown as Fig 3, the performance in female patients is higher than that in male patients. Like occlusion prediction, if the patient is a current smoker, the returned AUC

Table VII. Performances on different bypass types for stenosis prediction

CPT codes	35,556	35,566	35,656	35,661	35,571
Support	16	26	8	4	5
BILSTM					
Sensitivity	1.000	0.846	1.000	1.000	1.000
Specificity	0.937	0.885	0.905	0.975	0.857
AUC	0.969	0.865	0.953	0.987	0.922
BIGRU					
Sensitivity	1.000	0.885	1.000	0.500	1.000
Specificity	0.863	0.863	0.841	0.915	0.870
AUC	0.932	0.874	0.920	0.708	0.870

AUC, Area under the curve; *BIGRU*, bidirectional gated recurrent units; *BiLSTM*, bidirectional long short-term memory units; *CPT*, Current Procedural Terminology.

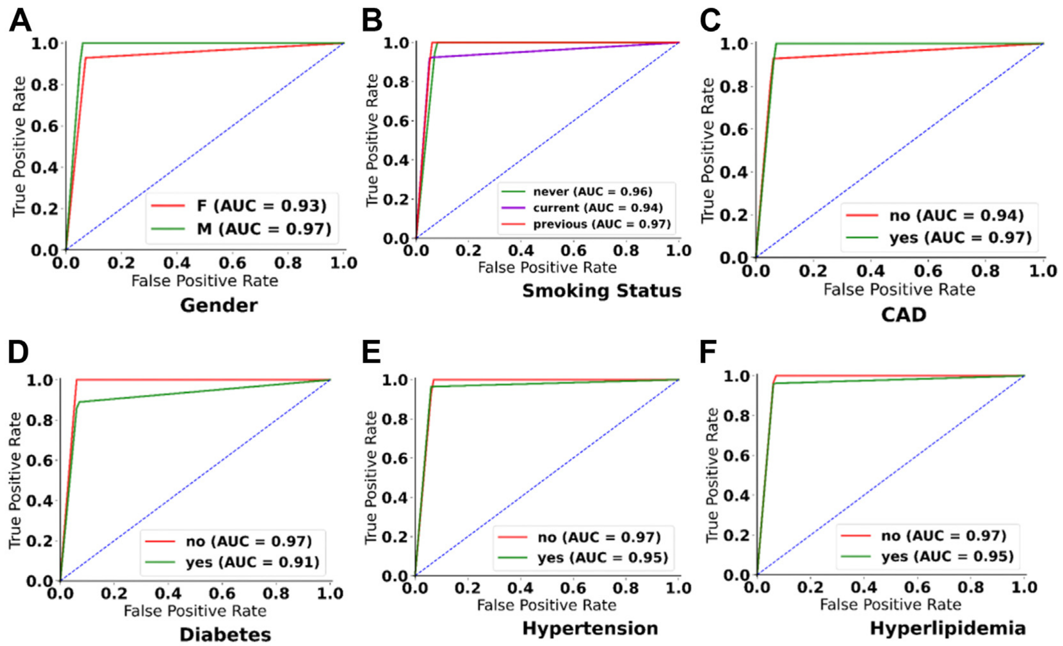

Fig 2. Typical receiver operator cure area under the curves (*AUC*) showing the accuracy of using bidirectional long short-term memory units (BiLSTMs) for occlusion prediction on test data.

is slightly lower. If the patient has coronary artery disease or diabetes, the prediction accuracy is lower. If the patient has hyperlipidemia, the prediction accuracy is higher. Hypertension makes no significant difference in stenosis prediction.

DISCUSSION

RNN models can be built to predict impending bypass graft occlusion and stenosis based on a set of PSVs measured in past DUS examinations. The occlusion and stenosis predictions are more accurate when there are three or more DUS examinations in medical history. Even with PSVs collected in two DUS examinations, the sensitivity, specificity, and AUC values can exceed >0.9 when the BiLSTM model is used. We believe this model can assist physicians to identify possible occlusion and stenosis cases before they turn into occlusions or stenosis in high-risk patients. If a patient has a low risk of stenosis or occlusions, the follow-up visit can be postponed. If the patient is at risk of developing stenosis or occlusion, intervention or further diagnostic imaging can be planned. This patient-centered method is desirable for this delicate patient population. Nonetheless, before being implemented in clinical practice, the model needs to undergo further validation. Our approach is based on two or more DUS examinations. There are cases when occlusion or stenosis happens after one DUS examination. To detect those situations, we think it is important to integrate other data, such as symptoms, patient physical activities, and other clinical data for prediction.

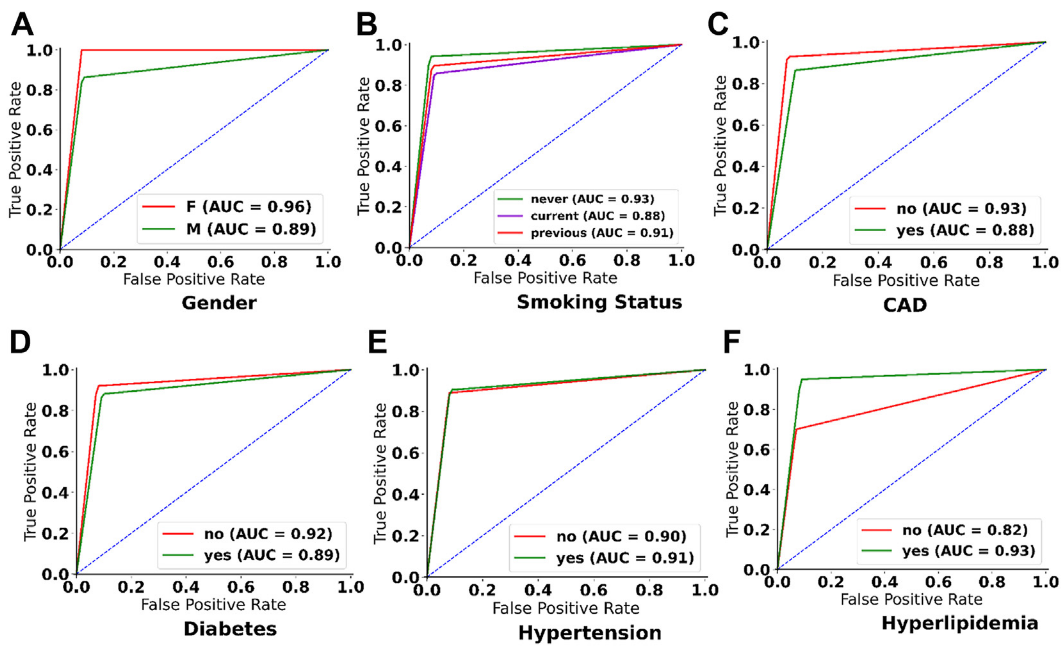

Fig 3. Typical receiver operator cure area under the curves (*AUC*) showing the accuracy of using bidirectional long short-term memory units (BiLSTMs) for stenosis prediction on test data.

The occlusion and stenosis prediction on the subcohorts based on the gender, smoking status, and comorbidities show that the prediction accuracy varies based on different patient populations. The impact of the medications taken by the patient in different patient populations may also be relevant to investigate. The research in the literature shows that the risk factors for failure of lower extremity revascularization procedures include age,[35] gender,[36,37] smoking status,[38] hypertension,[35] hyperlipidemia,[35] and diabetes.[35,39] This finding implies that using demographic information, social determinants of health, medication, and comorbidities can build a personalized prediction system. The future design of the RNN models can be modified to consider the dynamic clinical features such as diagnoses, medications, laboratory test results, and symptoms at the time when the DUS examinations are done. The future system can also consider static clinical features, such as gender and age, before the classification layer.

The performance of the artificial intelligence RNN model was better on femoropopliteal bypasses than femorotibial bypasses. We hypothesize that femorotibial bypass failure with smaller outflow vessels may be more dependent on changes in these outflow vessels, which are not fully evaluated in a standard bypass graft DUS examination. Future studies should evaluate more segments of the outflow tibial vessel to see if this enhances test accuracy.

The inclusion of the proposed bypass graft surveillance predictive models into the clinical workflow needs to consider the integration of all DUS examination results into the typical EHR systems and model fine-tuning needs to be done periodically. In addition, model interpretation should be included in the future to provide details on the risk factors of individuals that drive the prediction result.

We attempted to do a retrospective comparison between artificial intelligence analysis of bypass graft DUS examinations vs physician interpretation. Upon reviewing the chart, we discovered several intricacies that would make it difficult to make a straightforward comparison. In 2 of 69 patients with a bypass graft occlusion the interpreting physician detected a >75% bypass graft stenosis, but the referring physician did not choose to do further imaging or treatment. In 7 of 69 patients, there were single findings suggestive of low bypass graft flow or segmental stenosis, but there were other findings, such as no significant change from previous studies or bypass graft tortuosity, that led the interpreting physician to conclude that abnormal findings were not due to a critical stenosis. In 60 of 69 bypass graft occlusions, there was no evidence of impending bypass graft occlusion using standard criteria for physician interpretation. However, in a retrospective review of these patients' charts, there were some instances of severe intercurrent illnesses, such as cancer or patient noncompliance with the examination schedule, where the physician may have picked up signs of impending bypass graft occlusion if those intercurrent events had not interfered with routine patient care. In a few cases, the patient underwent an intervention based on an abnormal DUS examination, a subsequent normal DUS examination, and an occlusion on the subsequent follow-up examination. In those few cases, the abnormal DUS examination before the intervention may have contributed to artificial intelligence using RNNs predicting a future occlusion.

This study demonstrates that artificial intelligence using RNNs can automate the interpretation of bypass graft DUS examinations with good sensitivity, specificity, and accuracy. It suggests that an artificial intelligence analysis may improve the predictive value of impending bypass graft failure over currently used methodology. However, the RNNs are time-series analysis models that need multiple ultrasound examinations to analyze the trend in the PSVs for the prediction. When there is only one examination, RNNs cannot be applied. From these results, we can tell that, when there are only two examinations, the systems do not perform as well as the situations that have three or more examinations. This result means that the system can assist physicians in making decisions only when the patients have multiple ultrasound examinations. In contrast, we believe that conclusion requires a prospective comparison between artificial intelligence and current interpretation techniques, which we would like to conduct in the future. The artificial intelligence model can be automated to assist community-based physicians in detecting when a bypass graft has an increased risk of failure so that patient care can be modified to try to preserve the bypass. A decision could be made to decrease the time between screening intervals, perform further imaging studies such as a CTA, or proceed with arteriography and possible intervention.

There are limitations to this study. We have excluded the examinations with missing values. Imputation methods can be explored in the future to improve the robustness of the method. The DUS measurements are operator dependent on positioning of the ultrasound probe, the angle of the ultrasound probe and the possibility that there were stenotic segments of the graft missed by the ultrasonographer, although in our quality assurance 92% of the ultrasound examinations matched CTA or digital subtraction angiography when those studies were available for comparison. It is worth mentioning that other advanced RNN models can also be applied to improve the performance of the models, such as an RNN with an attention layer, and so on. This study should be verified with further prospective studies, including a prospective comparison between the sensitivity, specificity, and accuracy for predicting bypass graft stenosis or occlusion should be made between using artificial intelligence with RNNs and standard physician interpretation.

CONCLUSIONS

We collected DUS examination data from a group of patients undergoing open lower extremity revascularization for peripheral artery disease. We developed RNNs to predict the impending occlusion and stenosis after two or more DUS examinations are done. We found that the model can gain high performance on occlusion and stenosis prediction. There are disparities in patients with different categories of smoking and comorbidities. These findings lay the groundwork for research on integrating full clinical data including demographics, social determinants of health, medications, and other risk factors into the predictive analysis. The sensitivity, specificity, and accuracy are excellent when multiple ultrasound examinations are available for analysis, but the error rate is higher when there is only one DUS examination. Integrating other clinical data may improve the performance of this model when there are limited examinations. Future work should focus on developing a comprehensive predictive analysis system to include all relevant risk factors and provide immediate feedback to the physician when an occlusion or stenosis case is predicted.

AUTHOR CONTRIBUTIONS

Conception and design: XL, AS
Analysis and interpretation: XL, FM, DM, AS
Data collection: XL, DM, AS
Writing the article: XL, AS
Critical revision of the article: XL, FM, DM, AS
Final approval of the article: XL, FM, DM, AS
Statistical analysis: XL, AS
Obtained funding: Not applicable
Overall responsibility: AS

REFERENCES

1. Ma VY, Chan L, Carruthers KJ. Incidence, prevalence, costs, and impact on disability of common conditions requiring rehabilitation in the United States: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. *Arch Phys Med Rehabil*. 2014;95:986–995.e1.
2. Al-Thani H, Sathian B, El-Menyar A. Assessment of healthcare costs of amputation and prosthesis for upper and lower extremities in a Qatari healthcare institution: a retrospective cohort study. *BMJ open*. 2019;9:e024963.
3. Green R, Ouriel K, Ricotta J, DeWeese J. Revision of failed infrainguinal bypass graft: principles of management. *Surg*. 1986;100:646–654.
4. Baldwin ZK, Pearce BJ, Curi MA, et al. Limb salvage after infrainguinal bypass graft failure. *J vasc surg*. 2004;39:951–957.
5. Chang JB. *Modern Vascular Surgery*. 5. Springer Science & Business Media; 2012.
6. O'Donnell TF, Deery SE, Schermerhorn ML, et al. The impact of perioperative ankle-brachial index and clinical status on outcomes following lower extremity bypass. *Ann vasc surg*. 2018;53:139–147.
7. Green RM, McNamara J, Ouriel K, DeWeese JA. Comparison of infrainguinal graft surveillance techniques. *J vasc surg*. 1990;11:207–215.
8. Beidle T, Brom-Ferral R, Letourneau J. Surveillance of infrainguinal vein grafts with duplex sonography. *AJR Am J Roentgenol*. 1994;162:443–448.
9. Fasih T, Rudol G, Ashour H, Mudawi A, Bhattacharya V. Surveillance versus nonsurveillance for femoro-popliteal bypass grafts. *Angiology*. 2004;55:251–256.
10. Mohammed K, Farah W, Haydour Q, et al. Systematic review and meta-analysis of duplex ultrasound surveillance for infrainguinal vein bypass grafts. *Journal of vascular surgery*. 2017;66:1885–1891. e8.

11. Tinder CN, Chavanpun JP, Bandyk DF, et al. Efficacy of duplex ultrasound surveillance after infrainguinal vein bypass may be enhanced by identification of characteristics predictive of graft stenosis development. *J Vasc Surg*. 2008;48:613—618.
12. Davies A, Hawdon A, Sydes M, Thompson S. Is duplex surveillance of value after leg vein bypass grafting? Principal results of the Vein Graft Surveillance Randomised Trial (VGST). *Circulation*. 2005;112:1985—1991.
13. Guo S, Cui J, Zhao Y, et al. Machine learning—based operation skills assessment with vascular difficulty index for vascular intervention surgery. *Med Biol Eng Comput*. 2020;58:1707—1721.
14. Rossi L, Varriale A, Deleo G, Segramora VM. Machine learning and neural networks in vascular surgery. *Eur J Vasc Endovasc Surg*. 2019;58:e786—e787.
15. Raffort J, Adam C, Carrier M, Lareyre F. Fundamentals in artificial intelligence for vascular surgeons. *Ann Vasc Surg*. 2020;65:254—260.
16. Lau L, Kankanige Y, Rubinstein B, et al. Machine-learning algorithms predict graft failure after liver transplantation. *Transplantation*. 2017;101:e125.
17. Senanayake S, White N, Graves N, Healy H, Baboolal K, Kularatna S. Machine learning in predicting graft failure following kidney transplantation: a systematic review of published predictive models. *Int J Med Inform*. 2019;130:103957.
18. Agasthi P, Buras MR, Smith SD, et al. Machine learning helps predict long-term mortality and graft failure in patients undergoing heart transplant. *Gen Thorac Cardiovasc Surg*. 2020;68:1369—1376.
19. Mofidi R, McBride OM, Green BR, Gatenby T, Walker P, Milburn S. Validation of a decision tree to streamline infrainguinal vein graft surveillance. *Ann Vasc Surg*. 2017;40:216—222.
20. Huizing E, Schreve MA, Kum S, et al. Development of a prediction model for the occurrence of stenosis or occlusion after percutaneous deep venous arterialization. *Diagnostics*. 2021;11:1008.
21. Bandyk DF. Surveillance after lower extremity arterial bypass. *Perspect Vasc Surg Endovasc Ther*. 2007;19:376—383.
22. Provost F. *Machine learning from imbalanced data sets 101*. AAAI Press; 2000:1—3.
23. Petrosian A, Prokhorov D, Homan R, Dasheiff R, Wunsch D. Recurrent neural network based prediction of epileptic seizures in intra-and extracranial EEG. *Neurocomputing*. 2000;30:201—218.
24. Daoud H, Bayoumi MA. Efficient epileptic seizure prediction based on deep learning. *IEEE Trans Biomed Circuits Syst*. 2019;13:804—813.
25. Rasmy L, Wu Y, Wang N, et al. A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set. *J Biomed Inform*. 2018;84:11—16.
26. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural networks*. 2005;18:602—610.
27. Dey R, Salem FM. *Gate-variants of gated recurrent unit (GRU) neural networks*. IEEE; 2017:1597—1600.
28. Golbus JR, Pescatore NA, Nallamothu BK, Shah N, Kheterpal S. Wearable device signals and home blood pressure data across age, sex, race, ethnicity, and clinical phenotypes in the Michigan Predictive Activity & Clinical Trajectories in Health (MIPACT) study: a prospective, community-based observational study. *Lancet Digit Health*. 2021;3:e707—e715.
29. Jeong S, Cheon W, Cho S, Han Y. Clinical applicability of deep learning-based respiratory signal prediction models for four-dimensional radiation therapy. *Plos One*. 2022;17:e0275719.
30. He J, Guo Z, Shao Z, Zhao J, Dan G. An LSTM-based prediction method for lower limb intention perception by integrative analysis of Kinect visual signal. *J Healthc Eng*. 2020;2020:8024789.
31. D. Perna and A. Tagarelli, Deep auscultation: Predicting respiratory anomalies and diseases via recurrent neural networks, 2019. 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain. pp. 50-55, https://doi.org/10.1109/CBMS.2019.00020
32. Fu R, Zhang Z, Li L. *Using LSTM and GRU neural network methods for traffic flow prediction*. 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China; 2016:324—328. https://doi.org/10.1109/YAC.2016.7804912.
33. Yamak PT, Yujian L, Gadosey PK. *A comparison between arima, lstm, and gru for time series forecasting*. 2019:49—55.
34. Yang S, Yu X, Zhou Y. Lstm and gru neural network performance comparison study: taking yelp review dataset as an example. *IEEE*. 2020:98—101.
35. Owens CD, Ho KJ, Conte MS. Risk factors for failure of lower-extremity revascularization procedures: are they different for bypass and percutaneous procedures? *Semin Vasc Surg,*. 2008;21(3):143—153.
36. Jain AK, Velazquez-Ramirez G, Goodney PP, Edwards MS, Corriere MA. Gender-based analysis of perioperative outcomes associated with lower extremity bypass. *Am Surg*. 2011;77:844—849.
37. Seymour KA, Sadowitz B, Amankwah KS, Gahtan V. Outcome of lower extremity revascularization for peripheral artery occlusive disease: is there a difference between men and women? *Vasc*. 2011;19:59—67.
38. Young JC, Paul NJ, Karatas TB, et al. Cigarette smoking intensity informs outcomes after open revascularization for peripheral artery disease. *J Vasc Surg*. 2019;70:1973—1983.e5.
39. Paraskevas KI, Baker DM, Pompella A, Mikhailidis DP. Does diabetes mellitus play a role in restenosis and patency rates following lower extremity peripheral arterial revascularization? A critical overview. *Annals Vasc Surg*. 2008;22:481—491.

Submitted May 31, 2023; accepted Nov 26, 2023.