

Development and internal validation of a prediction model for postoperative vascular compromise following flap reconstruction in patients with diabetic foot ulcers

Xiaowen Zhang¹

Ningxia Medical University

Jing Yao^{2#}

Ningxia Medical University

Ying Cao³

Ningxia Medical University

Qing Shi⁴

Ningxia Medical University

Yanhua Ning²

nyh1_ning@126.com

Ningxia Medical University

Research Article

Keywords: diabetic foot ulcer, flap reconstruction, vascular compromise, prediction model, logistic regression, internal validation

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10783842/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Flap reconstruction for diabetic foot ulcers is often performed in the setting of coexisting ischemia, infection, and metabolic abnormalities. Postoperative vascular compromise can progress rapidly, whereas abnormalities detected by routine monitoring often become apparent only after perfusion has already deteriorated. This underscores the need to identify, by the end of surgery, patients who may require intensified postoperative surveillance. This study aimed to develop and internally validate a prediction model for postoperative vascular compromise following flap reconstruction in patients with diabetic foot ulcers.

Methods

We retrospectively collected data from patients with diabetic foot ulcers who were hospitalized at Ningxia Medical University General Hospital and underwent flap transfer or flap reconstruction between July 1, 2023, and July 1, 2026, yielding 200 surgical records. The prediction time point was defined as completion of flap reconstruction before the patient left the operating room, and the outcome was arterial, venous, or mixed vascular compromise occurring within 7 days after surgery. Logistic regression was used as the primary model, while a decision tree and random forest were included for exploratory comparison. Internal validation was performed using stratified 5-fold cross-validation grouped by clusters of highly similar records, with model performance assessed using the area under the receiver operating characteristic curve, the area under the precision-recall curve, the Brier score, calibration, and decision curve analysis.

Results

Vascular compromise was documented in 24 of 200 surgical records (12.0%). The final logistic regression model included preoperative fasting glucose and vascular pedicle length; for each 1 mmol/L increase in preoperative fasting glucose, the odds ratio for vascular compromise was 1.627 (95% CI 1.317–2.011, $P < 0.001$), whereas for each 1 cm increase in vascular pedicle length, the odds ratio was 0.026 (95% CI 0.006–0.113, $P < 0.001$). The internally validated AUCs for logistic regression, decision tree, and random forest were 0.859 (95% CI 0.774–0.927), 0.845 (95% CI 0.679–0.971), and 0.919 (95% CI 0.835–0.983), respectively, with corresponding Brier scores of 0.084, 0.068, and 0.066. The calibration intercept and slope of the logistic regression model were -0.038 and 1.138 , respectively.

Conclusions

Preoperative fasting glucose and vascular pedicle length may provide preliminary risk stratification before patients leave the operating room. For patients with a higher predicted risk, intensified early postoperative flap surveillance, prioritized assessment of pedicle course, dressing compression, and limb positioning, together with predefined escalation and surgical review procedures, may facilitate earlier recognition of vascular compromise and reduce delays in re-exploration. The model is intended to support allocation of monitoring resources rather than replace dynamic flap assessment, clinical diagnosis, or treatment decisions. Given the single-center design and limited number of vascular compromise events, the current findings do not support establishing a fixed risk score or clinical threshold.

Background

Diabetic foot ulcers are not merely skin defects but complex limb-threatening conditions arising from the interaction of peripheral neuropathy, peripheral artery disease, repetitive mechanical loading, and infection^[1–3]. Accordingly, treatment aims extend beyond wound closure to include infection control, restoration of adequate perfusion, preservation of foot structure and weight-bearing function, and reduction of recurrent ulceration and amputation whenever possible^[4–6]. In patients with persistent exposure of tendon, bone, or joint after adequate debridement, simple wound closure may be insufficient to meet the requirements of definitive reconstruction. In this setting, achieving stable, durable, and well-vascularized soft-tissue coverage is critical to subsequent rehabilitation and limb-salvage outcomes.

Flap reconstruction transfers well-vascularized tissue to complex wounds, providing durable soft-tissue coverage and creating a foundation for infection control and functional reconstruction^[7–9]. However, successful reconstruction depends not only on successful flap transfer but also on the maintenance of stable perfusion during the early postoperative period. Arterial insufficiency or impaired venous outflow can rapidly lead to ischemia or venous congestion and subsequently to irreversible tissue injury; failure to recognize and correct vascular compromise within an effective salvage window may result in flap necrosis, additional surgery, prolonged hospitalization, or even failure of limb salvage^[10]. Postoperative vascular compromise is therefore a highly time-sensitive complication, and its early recognition is directly relevant to flap salvage and limb outcomes.

Postoperative flap perfusion is commonly assessed by monitoring flap color, temperature, capillary refill, pinprick bleeding, and Doppler signals. These approaches remain essential components of postoperative flap assessment; however, abnormal findings often become apparent only after vascular impairment has already occurred and may be influenced by dressing coverage, limb positioning, observer experience, and the interval between assessments^[11–12]. In settings where nursing capacity and monitoring resources are limited, applying the same surveillance intensity to every patient may not represent the most efficient allocation of clinical resources. Risk stratification before the patient leaves the operating room could allow more frequent observation, reassessment of pedicle course, evaluation

of limb positioning and dressing compression, and expedited surgical review to be prioritized for patients at higher risk while maintaining standard monitoring for all patients.

Previous studies have largely investigated factors associated with free-flap failure, re-exploration, or vascular complications in heterogeneous populations involving different anatomical sites, etiologies, and flap types^[13]. The direct applicability of these findings to patients with diabetic foot ulcers remains limited because ischemia, infection, and metabolic abnormalities frequently coexist in this population, resulting in a risk profile that differs from that of other reconstructive populations. In addition, some studies have incorporated postoperative local findings or final flap outcomes as predictors, limiting their usefulness for risk assessment before vascular compromise occurs, while complex models developed from datasets with few outcome events may yield overly optimistic performance estimates. More importantly, models evaluated primarily on discrimination without adequate assessment of calibration and clinical net benefit still do not answer a practical clinical question: which patients require more intensive early monitoring immediately after surgery?

To address these gaps, we defined the prediction time point as completion of flap reconstruction before the patient left the operating room and included only preoperative and intraoperative variables available by that time to develop a model for vascular compromise occurring within 7 days after flap reconstruction in patients with diabetic foot ulcers. Logistic regression was selected as the primary model to preserve interpretability, while decision tree and random forest models were used for exploratory comparison. Model evaluation included not only discrimination but also calibration, overall prediction error, and decision curve analysis to provide a more comprehensive assessment of internal performance. The aim was not to replace dynamic postoperative flap monitoring or establish a threshold for re-exploration, but to provide a quantitative basis for early risk communication, stratified surveillance, and escalation pathways that can be further validated.

Methods

Study design and participants

This was a single-center retrospective prediction model study. We retrospectively collected data from patients with diabetic foot ulcers who were hospitalized at Ningxia Medical University General Hospital and underwent local, regional, or free flap transfer or flap reconstruction between July 1, 2023, and July 1, 2026. Each record corresponded to one flap procedure performed during a hospitalization and included postoperative follow-up information for the subsequent 7 days. Data were obtained from electronic medical records, operative reports, anesthesia records, nursing records, and the hospital laboratory information system.

Inclusion criteria

Records were eligible if they met all of the following criteria: a hospital diagnosis of diabetic foot ulcer consistent with the relevant International Working Group on the Diabetic Foot definitions^[1]; completion of flap transfer or flap reconstruction during the hospitalization; sufficient medical record information to determine whether vascular compromise occurred within 7 days after surgery; and availability of preoperative fasting glucose and vascular pedicle length for verification from the source records. Records were excluded from the final model if the primary outcome was missing, the same procedure had been entered more than once, key variables could not be verified, or apparently implausible values could not be confirmed from the source documentation. Wagner grade was obtained directly from the original medical record.

Data collection and variable definitions

Demographic and clinical variables included patient characteristics, duration of diabetes, Wagner grade, medical history, preoperative vital signs, and laboratory measurements. Procedure-related variables included flap type, donor site, vascular pedicle length, operative duration, intraoperative fluid administration, blood loss, flap area, and wound area. To avoid incorporating information obtained after outcome occurrence into the prediction model, postoperative flap findings, wound infection, and final flap survival were not considered candidate predictors.

Preoperative fasting glucose was defined as the most recent venous plasma glucose measurement obtained before surgery after at least 8 h without caloric intake^[14]; when multiple eligible measurements were available, the value closest to the start of surgery was used. Vascular pedicle length was extracted from operative reports in centimeters. Because the source records did not consistently specify the measurement start and end points, potential measurement variability was acknowledged as a study limitation. All continuous variables were retained on their original measurement scales.

Outcome definition

The primary outcome was vascular compromise occurring within 7 days after flap reconstruction. Clinical signs of arterial compromise could include flap pallor, decreased temperature, prolonged capillary refill, reduced bleeding on pinprick testing, and attenuation of the arterial Doppler signal. Venous compromise could present with dark red or cyanotic discoloration of the flap, swelling, shortened capillary refill, dark red bleeding on pinprick testing, and impaired venous outflow; cases showing features of both arterial and venous compromise were classified as mixed vascular compromise^[10–11]. In the source data, arterial, venous, and mixed vascular compromise were coded as 1, whereas absence of vascular compromise was coded as 0.

Candidate predictors and model complexity

Candidate predictors were identified from the reconstructive surgery clinical data form, relevant literature, and two rounds of expert consultation. Expert consultation was used primarily to assess the clinical relevance of each variable and its availability before the predefined prediction time point. Expert opinion served as one component of candidate predictor selection and did not replace verification against the source records or subsequent statistical evaluation.

Before model development, variables obtained after the occurrence of vascular compromise or variables that substantially overlapped with the outcome definition were excluded. Continuous variables were then assessed for their distributions, outliers, linearity, and multicollinearity. Because only 24 vascular compromise events were observed, the number of parameters in the primary model was deliberately restricted to reduce the risk of overfitting. Preoperative fasting glucose and vascular pedicle length were ultimately retained as predictors.

Data cleaning, model development, and internal validation

Variable names, coding schemes, and measurement units were standardized, and apparent data-entry errors were checked against the source records. Flap area and wound area were calculated as length × width, and the midpoint was used for continuous variables recorded as ranges.

Clinical fields other than the record identifier and outcome were compared; records with completely identical clinical fields were flagged as potential duplicates, whereas records with $\geq 95\%$ concordance were grouped into clusters of highly similar cases. Similar records were not removed automatically. Because the de-identified data structure did not allow reliable distinction between repeated procedures in the same patient and duplicate entries, internal validation was grouped by case cluster so that records from the same cluster were assigned to the same fold, thereby reducing potential information leakage and overestimation of model performance due to within-patient correlation.

The same two predictors, preoperative fasting glucose and vascular pedicle length, were used in all three models. Logistic regression served as the primary model, with regression coefficients, odds ratios, and 95% confidence intervals reported. The decision tree was configured with a maximum depth of 3 and a minimum of 8 observations per leaf node, whereas the random forest consisted of 500 trees with a maximum depth of 4 and a minimum of 5 observations per leaf node. Model performance was evaluated using established approaches for clinical prediction models, calibration, and decision curve analysis^[17–19]. Because previous studies have not consistently demonstrated superior performance of more complex algorithms over logistic regression in low-dimensional clinical data, the decision tree and random forest were included only for exploratory comparison^[20–21].

Internal validation was performed using stratified 5-fold cross-validation grouped by case cluster, with all records from the same cluster assigned to a single fold. Out-of-fold predicted probabilities were obtained for each record. Model discrimination was assessed using the area under the receiver operating characteristic curve (AUC) and the area under the precision-recall curve (PR-AUC), while overall

prediction error was assessed using the Brier score; 95% confidence intervals for the AUC were estimated using 2,000 bootstrap resamples with the case cluster as the resampling unit. Calibration analysis and decision curve analysis were both based on out-of-fold predicted probabilities^[17–19]. The maximum Youden index was used only to compare internal classification performance across models and was not interpreted as a clinical decision threshold.

Statistical analyses were performed using Python version 3.13.5, with the principal packages including scikit-learn 1.8.0, statsmodels 0.14.6, NumPy 2.3.5, and pandas 2.2.3. The present study included internal validation only, and no external validation was performed. The sample size for future external validation should be determined according to the expected event rate in the target population and the precision required for estimates of calibration and discrimination^[22].

Quality control

A variable dictionary was established to standardize variable definitions, measurement units, and data extraction time points. After data entry, particular attention was given to verification of outcome coding, preoperative fasting glucose, vascular pedicle length, intraoperative blood loss, and flap-related variables. For each data correction, the original value, corrected value, supporting medical record documentation, and date of correction were retained to ensure traceability of the data management process. All analytical data were de-identified before statistical analysis.

Results

Study population and vascular compromise

Between July 1, 2023, and July 1, 2026, a total of 200 surgical records were collected from patients with diabetic foot ulcers who underwent flap transfer or flap reconstruction. Vascular compromise occurred within 7 days after surgery in 24 records, corresponding to a record-level proportion of 12.0%, while no vascular compromise was documented in the remaining 176 records. According to the original medical record coding, the 24 vascular compromise events comprised 8 arterial and 16 venous events, with no mixed vascular compromise recorded. Because the de-identified data structure did not allow complete exclusion of repeated procedures in the same patient or duplicate records, vascular compromise subtypes are described according to the original medical record coding.

Table 1

Preoperative and procedure-related characteristics in records with and without vascular compromise

Variable	Overall (n = 200)	Vascular compromise (n = 24)	No vascular compromise (n = 176)	<i>P</i> value
Age (years)	58.00 (55.00, 63.25)	61.00 (29.75, 71.50)	58.00 (55.00, 63.00)	0.332
Duration of diabetes (years)	13.00 (5.00, 20.00)	9.50 (3.00, 28.00)	13.00 (5.00, 20.00)	0.710
Preoperative fasting glucose (mmol/L)	7.35 (6.50, 9.80)	9.80 (9.30, 11.80)	7.20 (6.50, 9.80)	< 0.001
Ankle-brachial index	1.10 (1.00, 1.20)	1.10 (1.00, 1.30)	1.10 (1.00, 1.20)	0.238
Operative duration (min)	59.50 (50.00, 81.00)	88.00 (39.00, 100.00)	59.00 (50.75, 75.00)	0.008
Intraoperative blood loss (mL)	30.00 (10.00, 50.00)	25.00 (20.00, 170.00)	30.00 (10.00, 50.00)	0.036
Vascular pedicle length (cm)	2.00 (2.00, 2.00)	1.50 (1.00, 2.00)	2.00 (2.00, 2.00)	< 0.001
Male, n (%)	152 (76.0)	24 (100.0)	128 (72.7)	0.001
Wagner grade ≥ 3 , n (%)	168 (84.0)	24 (100.0)	144 (81.8)	< 0.001

Continuous variables are presented as median (first quartile, third quartile), and categorical variables as n (%). *P* values are provided only to describe sample composition and were not used for predictor selection.

Data quality assessment and predictor determination

Data quality assessment identified two pairs of records with completely identical clinical fields and 84 pairs with $\geq 95\%$ concordance across clinical fields; the related records were grouped into 131 case clusters. To reduce potential information leakage from highly similar records being assigned to both training and validation folds, internal validation was performed at the case-cluster level, with records from the same cluster restricted to a single fold. Postoperative flap temperature, capillary refill, swelling, pain, wound infection, and final flap survival were either obtained after the occurrence of vascular compromise or substantially overlapped with the outcome definition and were therefore excluded from the prediction model.

All 24 vascular compromise events occurred in male patients, and all were observed in records with Wagner grade 3 or higher. Complete or near-complete separation was also present for anesthesia type, donor site, flap type, and some Wagner grade categories, resulting in unstable parameter estimates for these variables. Intraoperative blood loss was sensitive to a small number of extreme values; after excluding the eight records with blood loss ≥ 100 mL, the direction of its regression coefficient changed, and it was therefore retained only for supplementary observation. Preoperative fasting glucose and vascular pedicle length were both available before the occurrence of vascular compromise, and the variance inflation factor was 1.014 for each variable, indicating no substantial multicollinearity; these two variables were ultimately retained in the primary model.

Logistic regression model

Preoperative fasting glucose and vascular pedicle length were entered simultaneously into the logistic regression model. For each 1 mmol/L increase in preoperative fasting glucose, the odds of vascular compromise within 7 days after surgery increased by a factor of 1.627 (95% CI 1.317–2.011, $P < 0.001$), whereas each 1 cm increase in vascular pedicle length was associated with an odds ratio of 0.026 (95% CI 0.006–0.113, $P < 0.001$). The final prediction equation was $\text{logit}(P) = -0.107 + 0.487 \times \text{preoperative fasting glucose} - 3.668 \times \text{vascular pedicle length}$, where P represents the predicted probability of vascular compromise within 7 days after surgery.

Table 2
Two-predictor logistic regression model for postoperative vascular compromise

Variable	β	Standard error	Wald χ^2	<i>P</i> value	OR (95% CI)
Intercept	-0.107	1.000	0.012	0.915	0.898 (0.127–6.373)
Preoperative fasting glucose, per 1 mmol/L increase	0.487	0.108	20.343	< 0.001	1.627 (1.317–2.011)
Vascular pedicle length, per 1 cm increase	-3.668	0.761	23.231	< 0.001	0.026 (0.006–0.113)

Model internal validation and comparison

Grouped 5-fold cross-validation by case cluster generated out-of-fold predicted probabilities for each record. The AUCs for logistic regression, decision tree, and random forest were 0.859, 0.845, and 0.919, respectively; the corresponding PR-AUCs were 0.472, 0.643, and 0.715, and the Brier scores were 0.084, 0.068, and 0.066. Random forest showed the highest discrimination and the lowest overall prediction error in the current dataset; however, all three models used only the same two predictors, preoperative fasting glucose and vascular pedicle length. Therefore, the apparent internal advantage of the random forest may still reflect characteristics of the single-center data distribution and sample structure rather than superior clinical generalizability.

Table 3
Discrimination and overall prediction error of the three prediction models

Model	AUC (95% CI)	PR-AUC	Brier score
LR	0.859 (0.774–0.927)	0.472	0.084
DT	0.845 (0.679–0.971)	0.643	0.068
RF	0.919 (0.835–0.983)	0.715	0.066

At the internal cutoff corresponding to the maximum Youden index, the logistic regression model had a sensitivity of 1.000 and a specificity of 0.693. The corresponding sensitivity and specificity were 0.833 and 0.801 for the decision tree, and 1.000 and 0.727 for the random forest. Across the three models, positive predictive values ranged from 0.308 to 0.364, while negative predictive values ranged from 0.972 to 1.000. These classification metrics were used only for internal model comparison in the current dataset and should not be interpreted as estimates of diagnostic or predictive performance in external populations.

Table 4
Classification performance of the three prediction models

Model	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score
LR	1.000	0.693	0.308	1.000	0.730	0.471
DT	0.833	0.801	0.364	0.972	0.805	0.506
RF	1.000	0.727	0.333	1.000	0.760	0.500

Model calibration, decision curve analysis, and variable importance

The calibration intercept of the logistic regression model was -0.038 (95% CI -0.743 to 0.666), and the calibration slope was 1.138 (95% CI 0.692 to 1.584). The respective confidence intervals included the ideal values of 0 and 1, suggesting no obvious systematic calibration deviation; however, the limited number of vascular compromise events resulted in considerable uncertainty around these estimates.

Decision curve analysis showed that all three models generally provided a positive net benefit across threshold probabilities from 0.02 to 0.50 and outperformed the treat-all and treat-none strategies at most thresholds. However, these findings were based on internal validation and do not demonstrate that use of the models improves clinical outcomes.

In the random forest model, the relative variable importance values for preoperative fasting glucose and vascular pedicle length were 0.674 and 0.326, respectively. These values reflect only the relative predictive contribution of the variables within the model.

Discussion

Main findings and clinical interpretation

The principal contribution of this study is not simply the identification of statistical associations between preoperative fasting glucose, vascular pedicle length, and vascular compromise, but the shift of the prediction time point to the end of flap reconstruction, before routine postoperative monitoring has begun, using only information already available at that stage. This temporal framework reduces the risk of incorporating post-outcome information, such as flap color, temperature, capillary refill, or wound infection, into the prediction process and therefore allows the model to address a clinically relevant question: which patients may require more intensive monitoring during the early postoperative period?

Because vascular compromise was relatively infrequent in the current dataset, the models were more likely to yield high negative predictive values; however, a low predicted risk should not be interpreted as justification for reducing standard monitoring. Given the potentially severe consequences and need for rapid intervention, the model is better viewed as a tool for risk prioritization and allocation of monitoring resources rather than as a substitute for clinical diagnosis. The limited number of vascular compromise events also warrants caution when interpreting model performance. With only 24 events, estimates of the AUC, calibration slope, and threshold-dependent performance measures may vary substantially with the composition of the sample and may be sensitive to data partitioning. Accordingly, the internally validated performance observed in this study should be regarded as evidence of model feasibility at the development stage rather than as evidence of a stable clinical score ready for direct application ^[15–18].

Preoperative hyperglycemia as a marker of microcirculatory vulnerability

Preoperative fasting glucose was positively associated with postoperative vascular compromise in this study, but previous microsurgical reconstruction studies have reported inconsistent findings regarding whether hyperglycemia directly increases the risk of flap thrombosis, re-exploration, or flap failure. More consistent associations have been observed between hyperglycemia and postoperative infection or overall complications, whereas larger studies have not uniformly demonstrated that diabetes itself substantially increases the risk of free-flap failure ^[23–30]. These findings suggest that, in the present study, preoperative fasting glucose may be better interpreted as an integrated marker of perioperative physiological reserve and microcirculatory vulnerability rather than as an isolated determinant of thrombosis.

This interpretation is biologically plausible. Hyperglycemia may impair neutrophil function, endothelial vasodilation, and local microcirculation, while also increasing platelet reactivity and impairing tissue repair ^[31–34]. In the setting of low-flow conditions, tissue edema, and local inflammation after microvascular anastomosis, these changes may reduce the tolerance of the flap to transient fluctuations in perfusion. However, infection, pain, and surgical stress can themselves increase blood glucose levels, meaning that a single preoperative fasting glucose measurement cannot distinguish chronic poor glycemic control from an acute stress response. Clinical interpretation should therefore incorporate

HbA1c, infection burden, renal function, nutritional status, and perioperative glucose trends rather than relying on a single preoperative glucose value to define thresholds for cancellation of surgery or intensified glucose-lowering therapy.

Vascular pedicle length as a marker of surgical configuration rather than a simple length effect

Vascular pedicle length showed a strong inverse association with vascular compromise in this study, but this finding should not be interpreted simply as indicating that a longer pedicle is safer. Pedicle length is influenced by multiple procedural factors, including flap design, recipient vessel selection, anastomotic site, transfer distance, and operative technique; an excessively short pedicle may increase tension at the anastomosis, whereas excessive length may result in redundancy, folding, or twisting. Microvascular thrombosis is generally related to a combination of endothelial injury, low-flow conditions, and mechanical compression, while vessel kinking and the configuration of venous outflow may also affect the stability of flap perfusion ^[35–38]. Pedicle length may therefore be better regarded as a surrogate marker of surgical configuration and vascular course rather than as an independent linear protective factor.

Vascular pedicle length in this study was obtained from operative reports, but the source records did not consistently specify the start and end points of measurement, and the observed values were concentrated within a relatively narrow range. The large effect estimate may therefore have been influenced by measurement variability, center-specific operative practices, or a small number of extreme records. Future studies should standardize the measurement of pedicle length prospectively and simultaneously record pedicle diameter, course angle, sites of compression, recipient vessel quality, anastomotic technique, the number of arterial and venous anastomoses, and postoperative positioning. Only with more comprehensive technical information will it be possible to determine whether pedicle length retains independent predictive value.

Balancing model complexity and generalizability

Although the random forest achieved a higher AUC and a lower Brier score during internal validation, this does not necessarily indicate that a more complex algorithm is clinically superior to logistic regression. All three models used only two predictors, preoperative fasting glucose and vascular pedicle length, and in such low-dimensional data with few outcome events, a random forest may more readily fit center-specific data patterns and may also be more sensitive to duplicate records, concentrated predictor distributions, and data partitioning. Previous methodological research has likewise not consistently demonstrated that machine-learning approaches outperform well-specified logistic regression in common clinical prediction tasks ^[20]. In contrast, logistic regression provides directly interpretable effect directions, predictor scales, and individual predicted probabilities and is more amenable to external validation and recalibration; it was therefore considered the more appropriate primary model, with the decision tree and random forest retained for exploratory comparison.

Grouping cross-validation by clusters of highly similar records reduced the risk of overestimating model performance caused by similar records appearing in both the training and validation sets, but this approach could not fully resolve the issue of record independence. Because the de-identified data structure did not allow reliable distinction between repeated procedures in the same patient and duplicate entries, some records may still have been correlated within patients. If individual patients contributed more than one surgical record, such within-patient correlation could lead to underestimated standard errors and overly optimistic estimates of model performance. The robustness of the current internal validation findings therefore requires further evaluation in an independent cohort with clearly identifiable patients and confirmed record independence.

Clinical implementation of risk-stratified monitoring

The model is better suited for integration into existing postoperative flap-monitoring pathways than for use as a separate diagnostic pathway. For patients with a higher predicted risk, early postoperative care could prioritize reassessment of pedicle course, dressing compression, and limb positioning, together with more frequent evaluation of flap color, temperature, capillary refill, pinprick bleeding, and Doppler signals. Predefined escalation and surgical review procedures may also help reduce delays between recognition of suspected perfusion abnormalities and clinical assessment. Importantly, routine clinical surveillance remains central to the detection of vascular compromise, and model-based risk estimates should never delay direct evaluation or necessary re-exploration when perfusion impairment is suspected ^[10–12].

Future evaluation of the model should extend beyond improvements in AUC and examine whether its use can shorten the time to recognition of vascular compromise and reduce delays to re-exploration. Measures such as false alarms, nursing workload, flap salvage, length of hospital stay, and limb-salvage outcomes should also be systematically assessed to determine whether risk stratification translates into meaningful clinical benefit. Lower-extremity reconstruction in patients with diabetic foot ulcers often requires coordinated input from plastic surgery, vascular surgery, endocrinology, and nursing teams. The model will have practical clinical value only if risk stratification improves interdisciplinary coordination without introducing unnecessary interventions or additional resource burden ^[39–42].

Strengths, limitations, and directions for validation

A major strength of this study was the explicit definition of the prediction time point, together with exclusion of information obtained after vascular compromise had occurred. The three models were compared using the same predictors, and internal performance was evaluated across multiple domains, including discrimination, calibration, overall prediction error, and decision curve analysis. Several important limitations should nevertheless be considered. First, the single-center retrospective design and the small number of vascular compromise events limited the stability of model estimates and the generalizability of the findings; in addition, the de-identified data structure did not allow complete confirmation of record independence. Outcome classification relied on existing medical record documentation, and the start and end points used to measure vascular pedicle length were not

standardized, both of which may have introduced measurement error. Furthermore, all 24 vascular compromise events occurred in male patients, preventing evaluation of model performance in women and limiting assessment of generalizability across sex subgroups. The current cross-validation procedure was conducted after the two predictors and model parameters had been fixed, while variable processing, selection, and parameter specification were not fully repeated within each training fold; internal performance estimates may therefore remain somewhat optimistic. Finally, potentially relevant information, including HbA1c, continuous glucose measurements, severity of peripheral artery disease, recipient vessel quality, anastomotic technique, surgeon-related factors, and postoperative positioning, was not available.

Conclusions

This study moved assessment of vascular compromise risk to the end of flap reconstruction, before the patient left the operating room, and suggests that preoperative fasting glucose and vascular pedicle length may be useful for preliminary risk prioritization. For patients with a higher predicted risk, more intensive early postoperative flap surveillance and more rapid escalation pathways may facilitate earlier recognition of potentially reversible perfusion impairment and reduce delays in surgical review and re-exploration; however, a low predicted risk should not be used to justify reducing standard monitoring. At the departmental level, future implementation could explore risk-stratified monitoring schedules, structured checks of pedicle course, dressing compression, and limb positioning, and multidisciplinary rapid-response pathways, with clinical utility evaluated using time to recognition of vascular compromise, timeliness of re-exploration, flap salvage, limb-salvage outcomes, length of hospital stay, and nursing workload. Given the single-center design, limited number of vascular compromise events, uncertainty regarding complete record independence, and lack of standardized measurements, the model should not replace dynamic flap assessment or be used to define diagnostic or management thresholds. Until multicenter external validation and prospective evaluation of clinical utility are completed, the model is better regarded as a basis for optimizing monitoring workflows and generating hypotheses for further validation.

Abbreviations

AUC

area under the receiver operating characteristic curve

CI

confidence interval

DCA

decision curve analysis

DT

decision tree

IWGDF

International Working Group on the Diabetic Foot

LR
logistic regression
NPV
negative predictive value
PPV
positive predictive value
PR-AUC
area under the precision-recall curve
RF
random forest
ROC
receiver operating characteristic curve
VIF
variance inflation factor.

Declarations

Abbreviations

Not applicable. (No abbreviations are used in this manuscript.)

Ethics approval and consent to participate

The study was approved by the Ethics Committee of Ningxia Medical University General Hospital (approval No. KYLL-2025-0237). Given the retrospective nature of the study and the use of de-identified medical records, the Ethics Committee waived the requirement for informed consent. All data were anonymized prior to analysis.

Consent for publication

Not applicable. The manuscript does not contain any identifiable patient information, images, or case details.

Availability of data and materials

Individual-level clinical data cannot be made publicly available because of patient privacy considerations. Subject to approval by the ethics committee and the relevant institution, de-identified variable documentation, model code, and aggregated results may be made available by the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Ningxia Natural Science Foundation (grant No. 2025AAC030778). The funder had no role in study design, data analysis, interpretation of the results, or preparation of the manuscript.

Clinical trial number: Not applicable. (This study is not a clinical trial.)

Authors' contributions

XWZ was responsible for the study conceptualization, design, and overall project planning. XWZ and JY wrote the main manuscript text. JY prepared the figures and tables. YC and QS collected and curated the data. YHN supervised the project and provided critical guidance throughout the study. All authors reviewed the manuscript and approved the final version.

Acknowledgements

The authors thank all medical and nursing staff in the Department of Burn and Plastic Surgery and the Department of Hand, Foot and Ankle Surgery at the General Hospital of Ningxia Medical University for their support and assistance.

References

1. Van Netten JJ, Bus SA, Apelqvist J, et al. Definitions and criteria for diabetes-related foot disease (IWGDF 2023 update). *Diabetes Metab Res Rev.* 2024;40(3):e3654. 10.1002/dmrr.3654.
2. Schaper NC, Van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetes-related foot disease (IWGDF 2023 update). *Diabetes Metab Res Rev.* 2024;40(3):e3657. 10.1002/dmrr.3657.
3. Armstrong DG, Boulton AJM, Bus SA. Diabetic foot ulcers and their recurrence. *N Engl J Med.* 2017;376(24):2367–75. 10.1056/NEJMr1615439.
4. Fitridge R, Chuter V, Mills J, et al. The intersocietal IWGDF, ESVS, SVS guidelines on peripheral artery disease in people with diabetes mellitus and a foot ulcer. *Diabetes Metab Res Rev.* 2024;40(3):e3686. 10.1002/dmrr.3686.
5. Senneville É, Albalawi Z, Van Asten SA, et al. IWGDF/IDSA guidelines on the diagnosis and treatment of diabetes-related foot infections. *Diabetes Metab Res Rev.* 2024;40(3):e3687. 10.1002/dmrr.3687.
6. Monteiro-Soares M, Hamilton EJ, Russell DA, et al. Guidelines on the classification of foot ulcers in people with diabetes. *Diabetes Metab Res Rev.* 2024;40(3):e3648. 10.1002/dmrr.3648.
7. Reed AJM, Lim NTY, Yip SWL, et al. Outcomes of flap reconstruction for diabetic foot ulcers: a systematic review and meta-analysis. *Plast Reconstr Surg.* 2024;154(5):1118–30. 10.1097/PRS.00000000000011231.

8. Bhat S, Chia B, Barry IP, et al. Free tissue transfer in diabetic foot ulcers: a systematic review and meta-analysis. *Eur J Vasc Endovasc Surg.* 2023;66(5):670–7. 10.1016/j.ejvs.2023.07.031.
9. Tan CH, Wu Y, Satkunanantham M. Flaps for lower limb diabetic wound reconstruction: a systematic review and meta-analysis. *J Plast Reconstr Aesthet Surg.* 2023;79:74–86. 10.1016/j.bjps.2023.01.032.
10. Suh HP, Park CJ, Hong JP. Special considerations for diabetic foot reconstruction. *J Reconstr Microsurg.* 2021;37(1):12–6. 10.1055/s-0040-1714431.
11. Shen AY, Lonie S, Lim K, et al. Free flap monitoring, salvage, and failure timing: a systematic review. *J Reconstr Microsurg.* 2021;37(3):300–8. 10.1055/s-0040-1722182.
12. Kwasnicki RM, Noakes AJ, Banhidly N, et al. Quantifying the limitations of clinical and technology-based flap monitoring strategies. *Plast Reconstr Surg Glob Open.* 2021;9(7):e3663. 10.1097/GOX.0000000000003663.
13. Lese I, Biedermann R, Constantinescu M, et al. Predicting risk factors that lead to free flap failure and vascular compromise. *J Plast Reconstr Aesthet Surg.* 2021;74(3):512–22. 10.1016/j.bjps.2020.08.126.
14. American Diabetes Association Professional Practice Committee. Diagnosis and classification of diabetes: Standards of Care in Diabetes 2026. *Diabetes Care.* 2026;49(Suppl 1):S27–49. 10.2337/dc26-S002.
15. Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: binary and time-to-event outcomes. *Stat Med.* 2019;38(7):1276–96. 10.1002/sim.7992.
16. Van Smeden M, Moons KGM, De Groot JAH, et al. Sample size for binary logistic prediction models. *Stat Methods Med Res.* 2019;28(8):2455–74. 10.1177/0962280218784726.
17. Steyerberg EW. *Clinical Prediction Models.* 2nd ed. Cham: Springer; 2019. 10.1007/978-3-030-16399-0.
18. Van Calster B, McLernon DJ, Van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17:230. 10.1186/s12916-019-1466-7.
19. Vickers AJ, Elkin EB. Decision curve analysis. *Med Decis Mak.* 2006;26(6):565–74. 10.1177/0272989X06295361.
20. Christodoulou E, Ma J, Collins GS, et al. No performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12–22. 10.1016/j.jclinepi.2019.02.004.
21. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32. 10.1023/A:1010933404324.
22. Riley RD, Debray TPA, Collins GS, et al. Minimum sample size for external validation of a clinical prediction model. *Stat Med.* 2021;40(19):4230–51. 10.1002/sim.9025.
23. Bollig CA, Spradling CS, Dooley LM, et al. Impact of perioperative hyperglycemia in microvascular reconstruction. *Head Neck.* 2018;40(6):1196–206. 10.1002/hed.25097.

24. Offodile AC, Chou HY, Lin JAJ, et al. Hyperglycemia and adverse outcomes following microvascular reconstruction. *Oral Oncol.* 2018;79:15–9. 10.1016/j.oraloncology.2018.02.009.
25. De Vries FEE, Gans SL, Solomkin JS, et al. Perioperative blood glucose targets and surgical-site infection. *Br J Surg.* 2017;104(2):e95–105. 10.1002/bjs.10424.
26. Kantar RS, Rifkin WJ, David JA, et al. Diabetes and free flap failure in 6030 patients. *Microsurgery.* 2019;39(1):14–23. 10.1002/micr.30332.
27. Veith J, Donato D, Holoyda K, et al. Postoperative complications in lower extremity free flap reconstruction. *Microsurgery.* 2019;39(7):621–8. 10.1002/micr.30502.
28. Brady JS, Govindan A, Crippen MM, et al. Impact of diabetes on free flap surgery. *Microsurgery.* 2018;38(5):504–11. 10.1002/micr.30276.
29. Lee ZH, Daar DA, Stranix JT, et al. Free-flap reconstruction for diabetic lower extremity limb salvage. *J Surg Res.* 2020;248:165–70. 10.1016/j.jss.2019.12.008.
30. Sato T, Ohura N, Seike S, et al. Risk factors for free flap failure in diabetic foot ulcers. *J Plast Surg Hand Surg.* 2025;60:204–8. 10.2340/jphs.v60.44803.
31. Turina M, Fry DE, Polk HC Jr. Acute hyperglycemia and the innate immune system. *Crit Care Med.* 2005;33(7):1624–33. 10.1097/01.CCM.0000170106.61978.D8.
32. Kolluru GK, Bir SC, Kevil CG. Endothelial dysfunction and diabetes. *Int J Vasc Med.* 2012;2012:918267. 10.1155/2012/918267.
33. Holt RIG, Cockram CS, Ma RCW, et al. Diabetes and infection. *Diabetologia.* 2024;67(7):1168–80. 10.1007/s00125-024-06102-x.
34. Kaur R, Kaur M, Singh J. Endothelial dysfunction and platelet hyperactivity in type 2 diabetes. *Cardiovasc Diabetol.* 2018;17:121. 10.1186/s12933-018-0763-3.
35. Williams JG, French RJ, Lalonde DH. Why do free flap vessels thrombose? *Can J Plast Surg.* 2004;12(1):23–6. 10.1177/229255030401200112.
36. Liu Y, Zhao YF, Huang JT, et al. Venous compromise in radial forearm free flaps. *Int J Oral Maxillofac Surg.* 2012;41(4):448–52. 10.1016/j.ijom.2012.01.016.
37. Horng SY, Chen CK, Lee CH, et al. Vascular kinking and twisting. *Ann Vasc Surg.* 2010;24(8):1154–5. 10.1016/j.avsg.2010.07.023.
38. Dodd SX, Morzycki A, Nickel KJ, et al. One or two venous pedicles for lower extremity free flaps. *Microsurgery.* 2021;41(8):792–801. 10.1002/micr.30811.
39. Bigdeli AK, Gazyakan E, Schmidt VJ, et al. Long-term outcome after lower extremity free flap salvage. *J Reconstr Microsurg.* 2019;35(4):263–9. 10.1055/s-0038-1675146.
40. Pu LLQ. Free flaps in lower extremity reconstruction. *Clin Plast Surg.* 2021;48(2):201–14. 10.1016/j.cps.2020.12.002.
41. Azoury SC, Stranix JT, Kovach SJ, et al. Principles of orthoplastic surgery for lower extremity reconstruction. *J Reconstr Microsurg.* 2021;37(1):42–50. 10.1055/s-0039-1695753.

42. Piwnica-Worms W, Stranix JT, Othman S, et al. Risk factors for lower extremity amputation following attempted free flap limb salvage. J Reconstr Microsurg. 2020;36(7):528–33. 10.1055/s-0040-1710358.

Figures

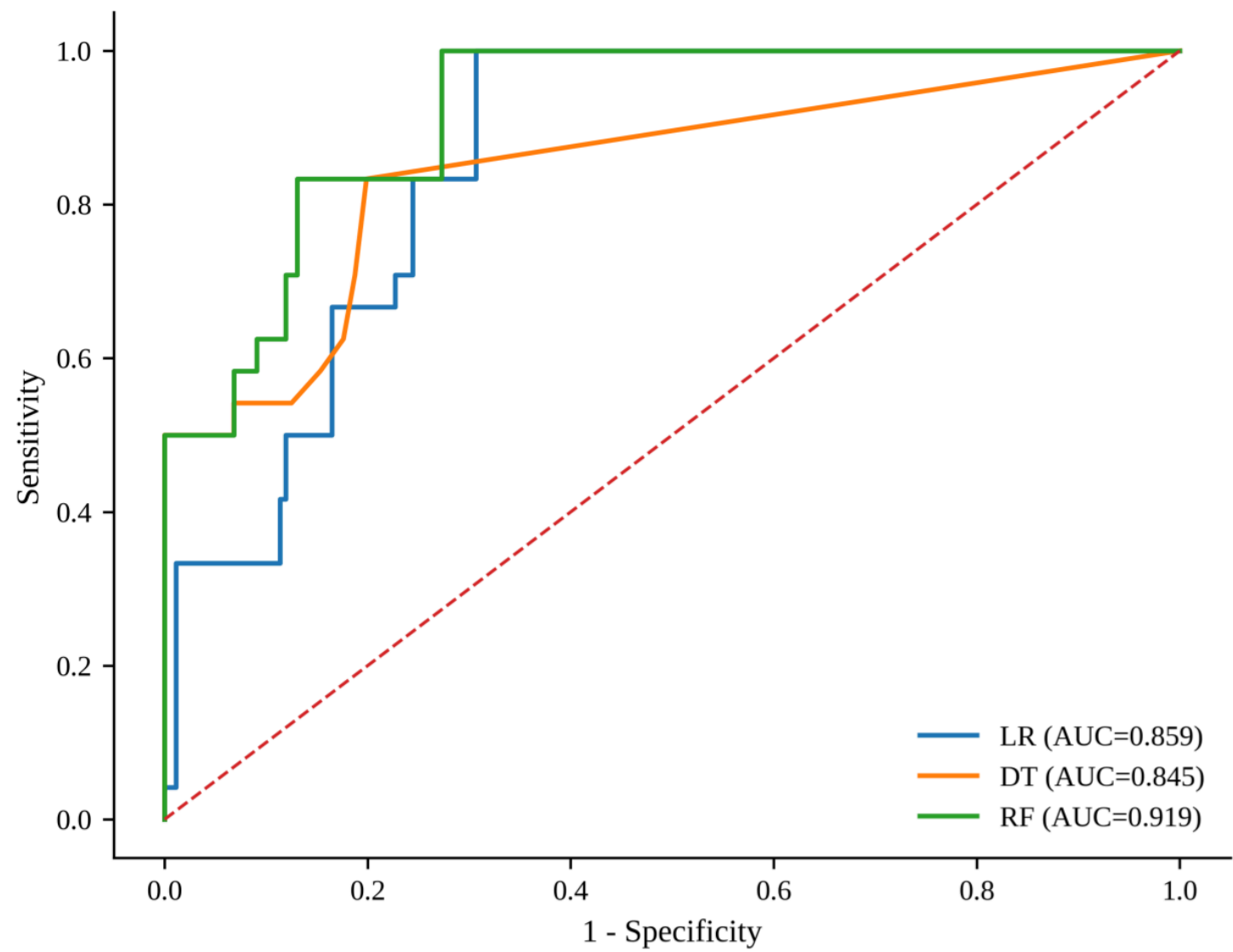


Figure 1

Receiver operating characteristic curves for the three prediction models

Note: LR, logistic regression; DT, decision tree; RF, random forest; AUC, area under the receiver operating characteristic curve.

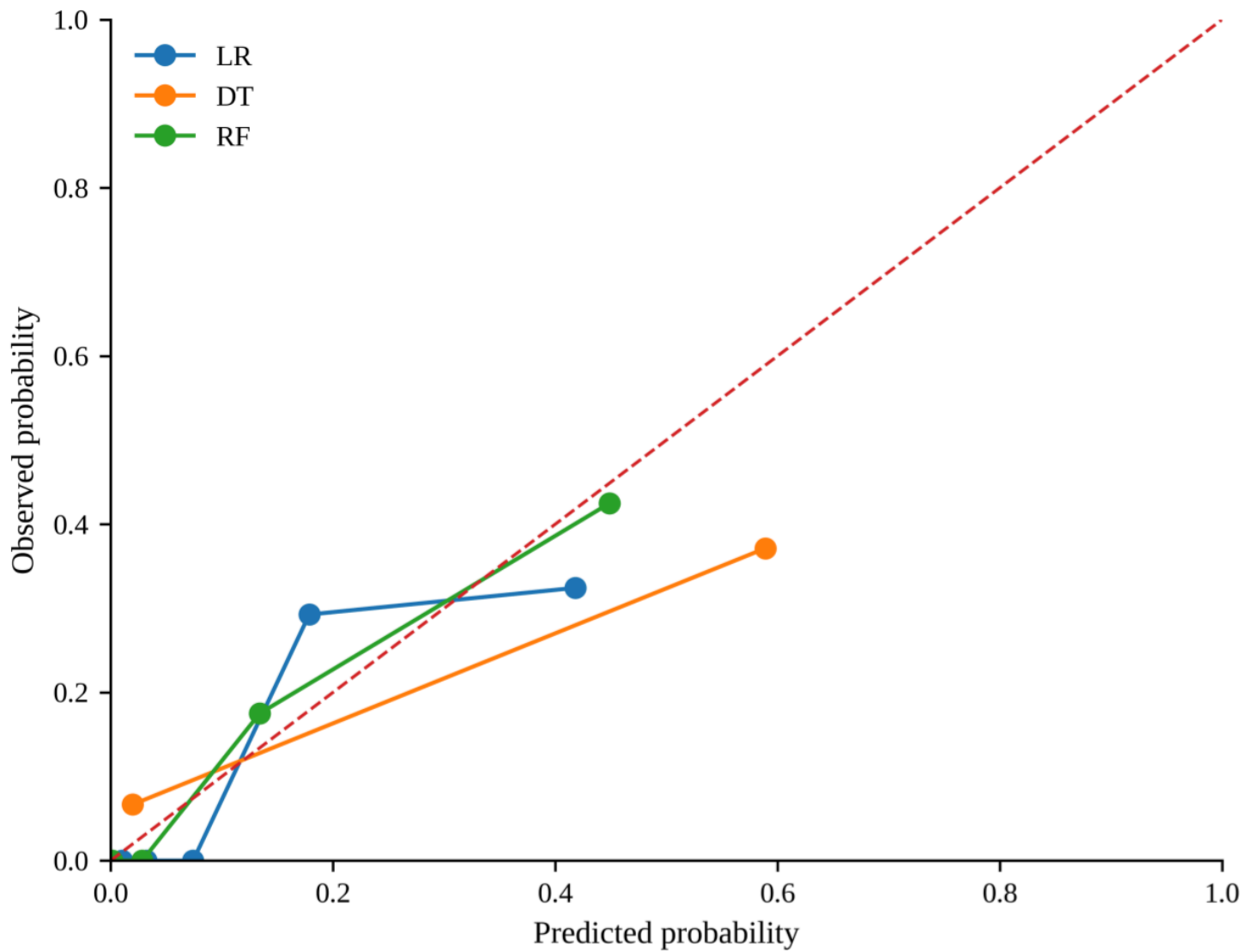


Figure 2

Calibration curves for the three prediction models

Note: The dashed line represents perfect calibration; LR, logistic regression; DT, decision tree; RF, random forest.

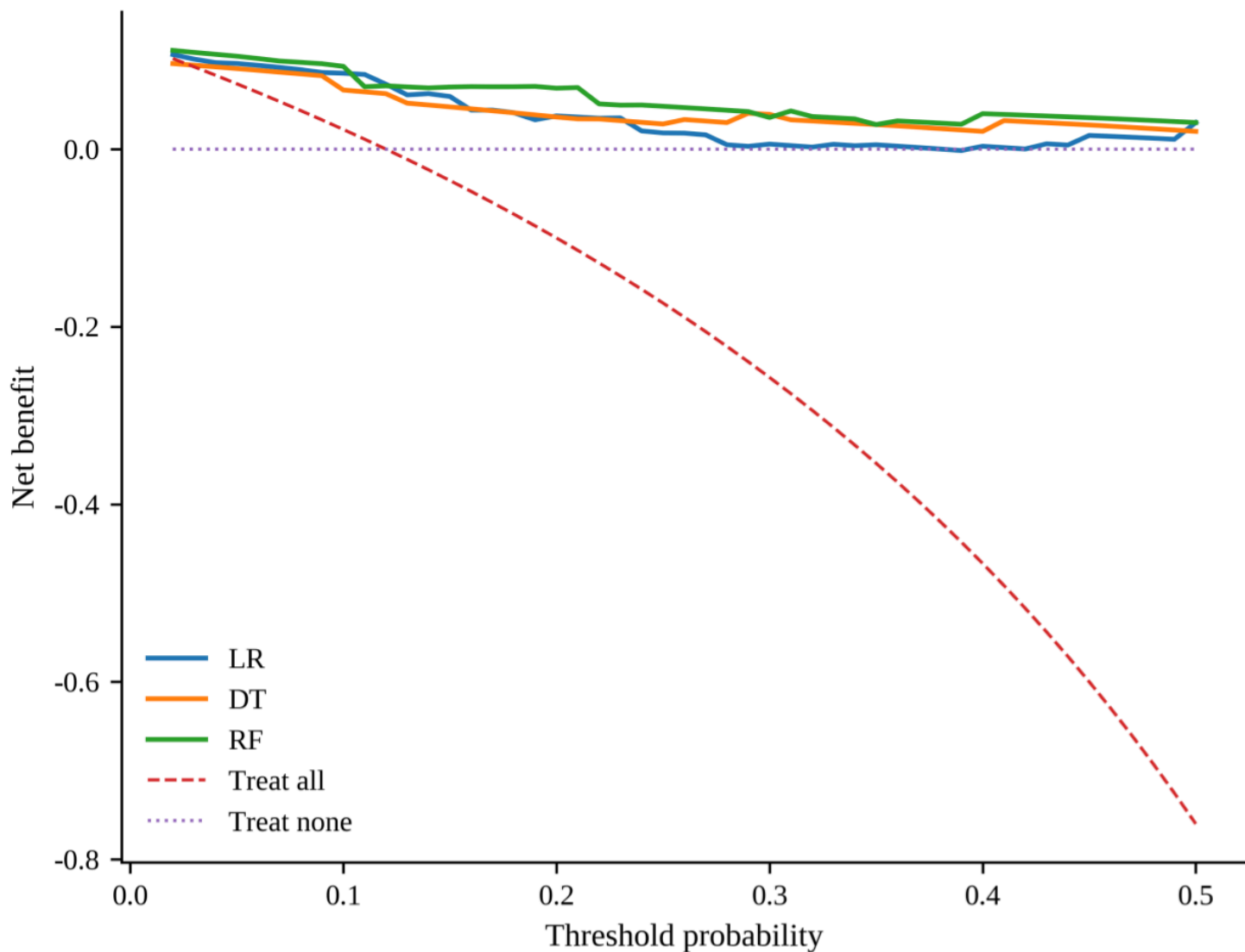


Figure 3

Decision curves for the three prediction models

Note: Treat all indicates intensified monitoring for all patients; Treat none indicates no intensified monitoring for any patient; LR, logistic regression; DT, decision tree; RF, random forest.

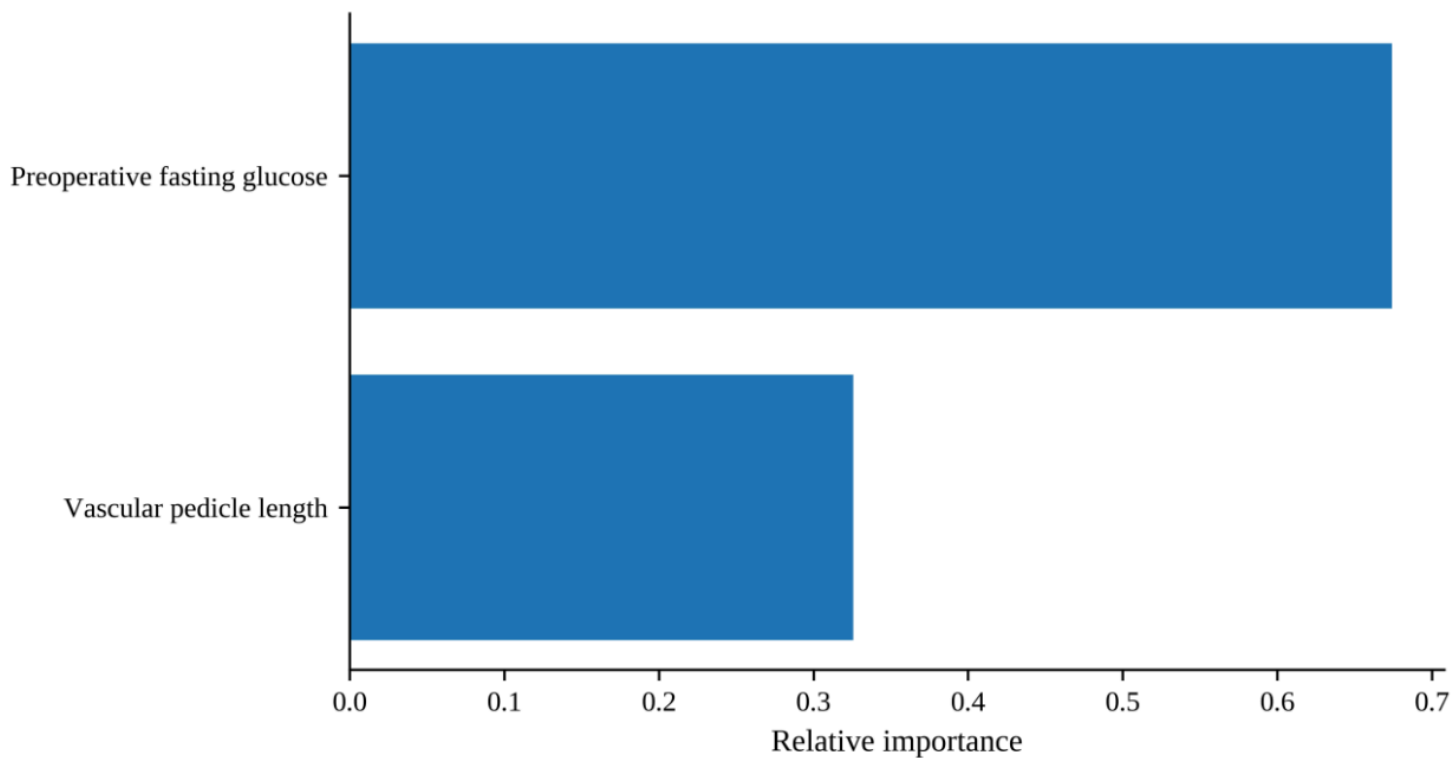


Figure 4

Relative importance of the two predictors in the random forest model

Note: Variable importance reflects relative predictive contribution within the random forest model and does not represent the magnitude or direction of a causal effect.