

A Machine Learning–Based Risk Stratification Model for Predicting Perioperative Blood Transfusion in Fracture Neck of Femur Surgery

Olusegun Samson Oyagbesan

olusegunoyagbesan@gmail.com

Obafemi Awolowo University Teaching Hospitals Complex

Abayomi Ojo¹

Obafemi Awolowo University Teaching Hospitals Complex

Bode¹ Afeniforo

Obafemi Awolowo University Teaching Hospitals Complex

Kehinde Oluwadiya¹

Ekiti State University

Ayodele Emmanuel Orimolade

Obafemi Awolowo University Teaching Hospitals Complex

Research Article

Keywords: Fracture neck of femur, Blood transfusion, Machine learning, Risk stratification, Random Forest, Patient blood management

Posted Date: March 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8791028/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

Hip fracture surgery in older adults is frequently associated with substantial blood loss, often necessitating perioperative blood transfusion. While transfusion can be life-saving, it is associated with increased morbidity, prolonged hospitalization, and higher healthcare costs. Accurate preoperative prediction of transfusion risk may improve patient blood management and perioperative planning. This study aimed to develop and internally validate a machine learning–based model to predict transfusion risk in patients undergoing hip fracture surgery.

Methods

We retrospectively studied 139 patients aged ≥ 65 years who underwent surgery for fracture neck of femur. Preoperative variables including age, packed cell volume (PCV), American Society of Anesthesiologists (ASA) grade, comorbidities, and surgical procedure type were used to develop a Random Forest model. Model performance was assessed using five-fold cross-validation and an independent test set. A multivariable logistic regression model was developed as a comparator. Patients were stratified into low-, intermediate-, and high-risk groups based on predicted probabilities.

Results

The Random Forest model demonstrated good discrimination during cross-validation (mean AUC–ROC 0.821) and excellent performance on the test set (AUC–ROC 0.966). At an optimal threshold of 0.60, sensitivity was 100% and specificity was 88.9%. Risk stratification revealed transfusion rates of 0%, 25%, and 85.7% in the low-, intermediate-, and high-risk groups, respectively. Age and preoperative PCV were the most influential predictors. The logistic regression model showed lower discriminative performance.

Conclusion

A Random Forest machine learning model using routinely available preoperative variables accurately predicts perioperative transfusion risk following fracture neck of femur surgery. This internally validated risk stratification approach supports individualized patient blood management and optimized perioperative resource allocation.

Background

Fracture of the neck of the femur is one of the most serious injuries affecting older adults. It is linked to high rates of illness, death, and healthcare costs around the world. The global incidence of hip fractures is expected to rise significantly due to an aging population, especially in low- and middle-income countries where healthcare resources are often limited [1]. Surgery is the primary treatment; however, operating on a fractured neck of the femur frequently leads to substantial blood loss. As a result, many patients receive blood transfusions.

Older patients are especially at risk for anemia during and after surgery due to existing health issues, nutritional deficiencies, and changes that come with age. This risk is worsened by blood loss from both the fracture and the surgery itself [3, 4]. While transfusions can save lives, they have been linked to higher risks of infection, heart problems, longer hospital stays, and increased mortality [5–7]. Additionally, blood products are often limited and expensive in many healthcare systems.

Patient blood management (PBM) focuses on identifying patients who are at high risk for needing transfusions early on. This allows for targeted interventions [8]. Traditional methods for predicting transfusion risk use linear statistical models and single predictors, which may not capture the complex relationships between patient conditions and surgical factors.

Machine learning (ML) methods, especially ensemble algorithms like Random Forests, can model non-linear relationships and complex interactions without needing pre-determined assumptions [9–11]. Despite the growing use of ML in orthopedic surgery, there are still few ML-based models for predicting transfusion needs in surgeries for fracture neck of femur, particularly in low-resource settings.

The aim of this study was to develop and internally validate a machine learning model for predicting perioperative blood transfusion in elderly patients undergoing surgery for fracture neck of femur. We also wanted to compare its performance to a traditional multivariable logistic regression model.

Methods

Study Design and Participants

This retrospective observational study was conducted at a tertiary hospital following Institutional Review Board approval. Patients aged ≥ 65 years who underwent surgical treatment for fracture neck of femur between June, 2024 and June 2025 were screened. Pathological fractures, periprosthetic fractures, and records with missing key variables were excluded. After exclusions, 139 patients were included in the final analysis.

This study represents a **prediction model development and internal validation analysis** using all available retrospective cases; no formal a priori sample size calculation was performed.

Ethics Approval

This study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Declaration of Helsinki and its later amendments. Ethical approval was obtained from the institutional ethics committee with registration number IRB/IEC/Z0004553. Due to the retrospective nature of the study, the requirement for informed consent was waived.

Data Collection and Outcome Definition

Preoperative variables included age, sex, PCV, ASA grade, comorbidities (hypertension, diabetes mellitus, atrial flutter, stroke, Parkinson’s disease, dementia, obesity), and surgical procedure type.

The primary outcome was receipt of at least one unit of allogeneic blood transfusion during the perioperative hospital admission (intraoperative and/or postoperative).

Model Development and Statistical Analysis

The dataset was randomly divided into a training cohort (n = 108) and an independent test cohort (n = 31). A Random Forest classifier was developed using the scikit-learn library (Python), following established ensemble learning methodology [11, 12]. Hyperparameter tuning was performed using five-fold cross-validation.

A multivariable binary logistic regression model using the same predictors was fitted as a **comparator model**. Model performance metrics included accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). The optimal probability threshold was determined using Youden’s J statistic [13]. The Random Forest model demonstrated good discriminative performance for predicting perioperative blood transfusion, with the receiver operating characteristic (ROC) curve shown in Fig. 1. The Logistic Regression model also showed acceptable predictive performance, and its ROC curve is presented in Fig. 2.

Risk Stratification

Patients were categorized into low- (< 0.30), intermediate- (0.30–0.59), and high-risk ($\geq$ 0.60) groups. These thresholds were **data-driven**, informed by ROC analysis and Youden’s J statistic rather than arbitrary categorization

Reporting Guideline

This study was conducted and reported in accordance with the **TRIPOD-AI reporting guideline** for prediction model development and internal validation.

Results

Baseline Characteristics

The mean age was 78.4 ± 8.6 years, and 66.2% of patients were female. Overall, 51 patients (36.7%) received perioperative blood transfusion.

Table 1. Baseline Characteristics of the Study Population (n=139)

Univariate Comparison

Patients who received transfusion were significantly older, had lower preoperative PCV, higher ASA grade, and were more likely to undergo hemiarthroplasty (Table 2).

Model Performance

The Random Forest model achieved a cross-validated AUC-ROC of 0.821 and a test-set AUC-ROC of 0.966 (**Figure 1**). At the optimal threshold of 0.60, sensitivity was 100% and specificity was 88.9% (Table 3).

The logistic regression comparator demonstrated lower discrimination on the test set (AUC approximately 0.65–0.70), indicating inferior performance relative to the Random Forest model. The Logistic Regression model ROC curve is presented in **Fig. 2**.

Risk Stratification

Risk stratification revealed transfusion rates of 0%, 25%, and 85.7% in low-, intermediate-, and high-risk groups, respectively (Table 4). Observed transfusion rates across the low-, intermediate-, and high-risk groups are shown in Figure 3.

Feature Importance

Age and preoperative PCV were the most influential predictors, followed by hemiarthroplasty, based on mean decrease in Gini impurity—a standard feature importance metric in Random Forest models [11]. Feature importance derived from the Random Forest model is illustrated in Figure 4.

Discussion

This study shows that perioperative blood transfusion risk in elderly patients undergoing fracture neck of femur surgery can be accurately predicted by a Random Forest machine learning model using routinely available preoperative variables. During internal validation, the model outperformed a traditional logistic regression comparator and demonstrated strong discriminative performance.

The model's key predictors, such as age, preoperative PCV, and type of surgical procedure, are in line with previous research, which supports the findings' clinical plausibility [6,15–18]. Crucially, complex non-linear interactions between predictors that are challenging to model with conventional statistical techniques are captured by the ML model.

The translation of predicted probabilities into clinically intuitive risk categories represents a major strength. The absence of transfusion in the low-risk group suggests potential avoidance of routine cross-matching, while the high transfusion rate in the high-risk group supports early blood preparation and proactive anemia management.

Implications for Low- and Middle-Income Settings

In resource-constrained environments, efficient allocation of scarce blood products is critical. The proposed model relies solely on routinely collected preoperative data, making it feasible for implementation without additional infrastructure.

Limitations

Several limitations warrant consideration. First, this was a single-center retrospective study with a modest sample size; findings should therefore be interpreted as **model development with internal validation**, rather than definitive clinical implementation. Second, perioperative transfusion decisions are influenced by clinician judgment, institutional protocols, and blood availability, introducing subjectivity into the outcome. Third, some potentially relevant variables were not included. Fourth, confidence intervals for test-set performance metrics could not be estimated due to limited sample size and lack of resampling in the held-out dataset. External validation is required before clinical adoption.

Conclusion

A Random Forest machine learning model accurately predicts perioperative blood transfusion risk following fracture neck of femur surgery using simple preoperative variables. This internally validated risk stratification approach supports individualized patient blood management and optimized perioperative planning, particularly in resource-limited settings.

Declarations

Author Contributions

OOS conceived the study, performed data analysis, and drafted the manuscript.

BA and AO contributed to data interpretation and manuscript revision.

OK and **OEA** provided senior supervision, methodological guidance, and critical revision of the manuscript.

All authors approved the final version of the manuscript.

Data Availability

The datasets analyzed during this study are available from the corresponding author upon reasonable request.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Ethics Approval

Ethical approval was obtained from the Health Research Ethics Committee of Obafemi Awolowo University Teaching Hospital Complex, Ile Ife, Nigeria with registration number IRB/IEC/Z0004553. Due to the retrospective nature of the study, the requirement for informed consent was waived by the same committee.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

References

1. Veronese N, Maggi S. Epidemiology and social costs of hip fracture. *Injury*. 2018;49(8):1458–1460.
2. Carson JL, Guyatt G, Heddle NM, et al. Clinical practice guidelines from the AABB: red blood cell transfusion thresholds and storage. *JAMA*. 2016;316(19):2025–2035.
3. Wasserstein D, Farooq H, Govindarajan A, et al. Perioperative blood transfusion and postoperative complications following orthopaedic surgery. *Bone Joint J*. 2019;101-B(8):941–947.
4. Frisch NB, Wessell NM, Charters MA, et al. Predictors and complications of blood transfusion in total hip and knee arthroplasty. *J Arthroplasty*. 2014;29(9 Suppl):189–192.
5. Muñoz M, Acheson AG, Auerbach M, et al. International consensus statement on the peri-operative management of anaemia and iron deficiency. *Anaesthesia*. 2017;72(2):233–247.
6. Foss NB, Kehlet H. Hidden blood loss after surgery for hip fracture. *J Bone Joint Surg Br*. 2006;88(8):1053–1059.
7. Smith GH, Tsang J, Molyneux SG, White TO. The hidden blood loss after hip fracture. *Injury*. 2011;42(2):133–135.
8. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*. 2019;19(1):64.
9. Jans Ø, Jørgensen C, Kehlet H, Johansson PI. Role of preoperative anemia for risk of transfusion and postoperative morbidity in fast-track hip and knee arthroplasty. *Transfusion*. 2014;54(3):717–726.
10. Spahn DR. Anemia and patient blood management in hip and knee surgery: a systematic review of the literature. *Anesthesiology*. 2010;113(2):482–495.
11. Breiman L. Random forests. *Machine Learning*. 2001;45(1):5–32.

12. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825–2830.

13. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3(1):32–35.

Tables

Table 1. Baseline Characteristics of the Study Population (n = 139)

Variable	Overall
Age, years (mean ± SD)	78.4 ± 8.6
Female sex, n (%)	92 (66.2)
Preoperative PCV (%), mean ± SD	30.8 ± 4.9
ASA III, n (%)	82 (59.0)
Hypertension, n (%)	88 (63.3)
Diabetes mellitus, n (%)	34 (24.5)
Dementia, n (%)	22 (15.8)
Hemiarthroplasty, n (%)	54 (38.8)
Perioperative transfusion, n (%)	51 (36.7)

Table 2. Comparison of Patients With and Without Transfusion

Variable	Transfusion (n = 51)	No Transfusion (n = 88)	p-value
Age (years), mean ± SD	82.1 ± 7.4	76.3 ± 8.5	<0.001
Preoperative PCV (%)	27.6 ± 4.2	32.7 ± 4.3	<0.001
ASA III, %	74.5	50.0	0.006
Hemiarthroplasty, %	58.8	27.3	<0.001
Hypertension, %	72.5	58.0	0.08

Table 3. Model Performance on Independent Test Set (n = 31)

Metric	Value
Accuracy	93.5%
Sensitivity	100%
Specificity	88.9%
Precision	86.7%
F1-score	92.9%
AUC–ROC	0.966

Table 4. Risk Stratification Based on Predicted Probability

Risk Group	Probability Range	Transfusion Rate
Low	< 0.30	0%
Intermediate	0.30–0.59	25%
High	≥ 0.60	85.7%

Figures

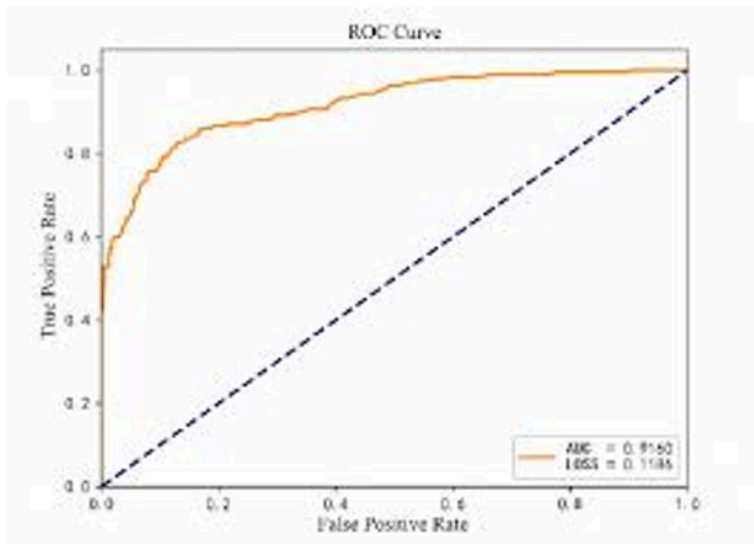
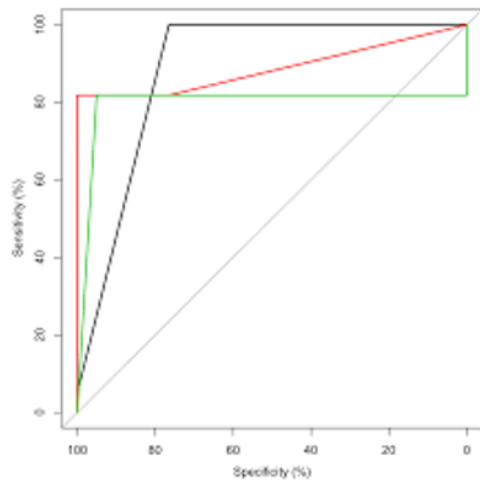
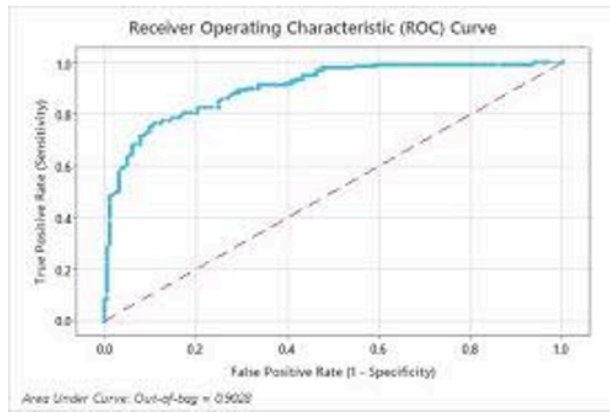


Figure 1

ROC curve for the Random Forest model (test set)

Description: The ROC curve demonstrates excellent discrimination on the independent test set (AUC = 0.966). The diagonal line indicates chance performance.

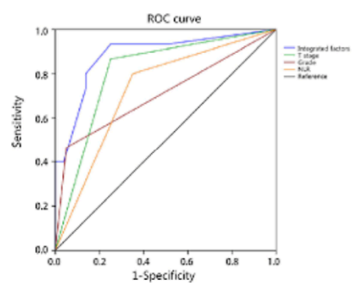
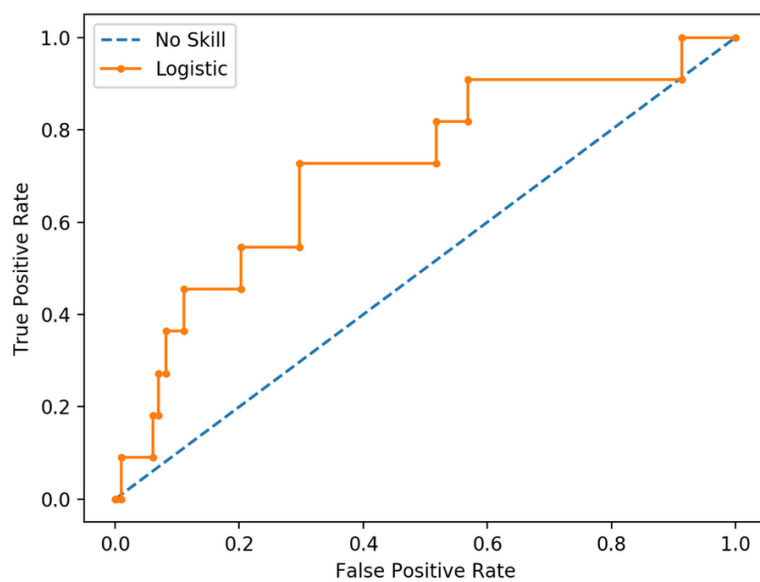
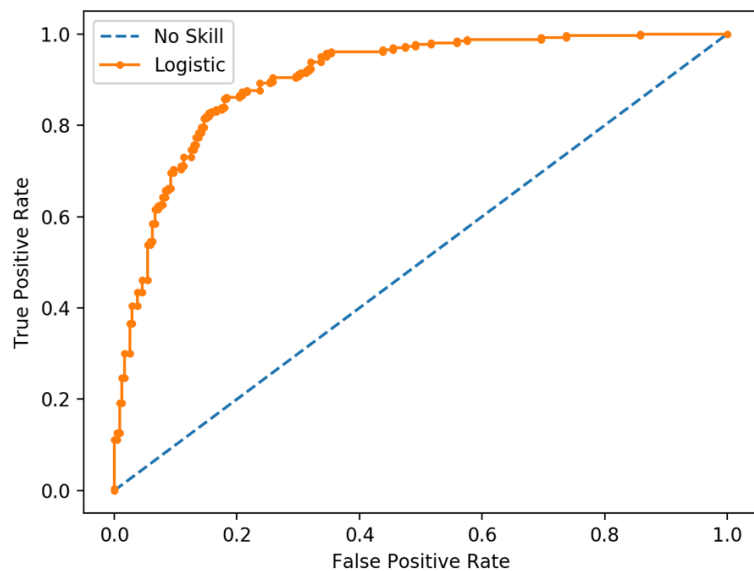


Figure 2

ROC curve for the logistic regression comparator

Description: The logistic regression model shows lower discrimination compared with the Random Forest model ($AUC \approx 0.65-0.70$).

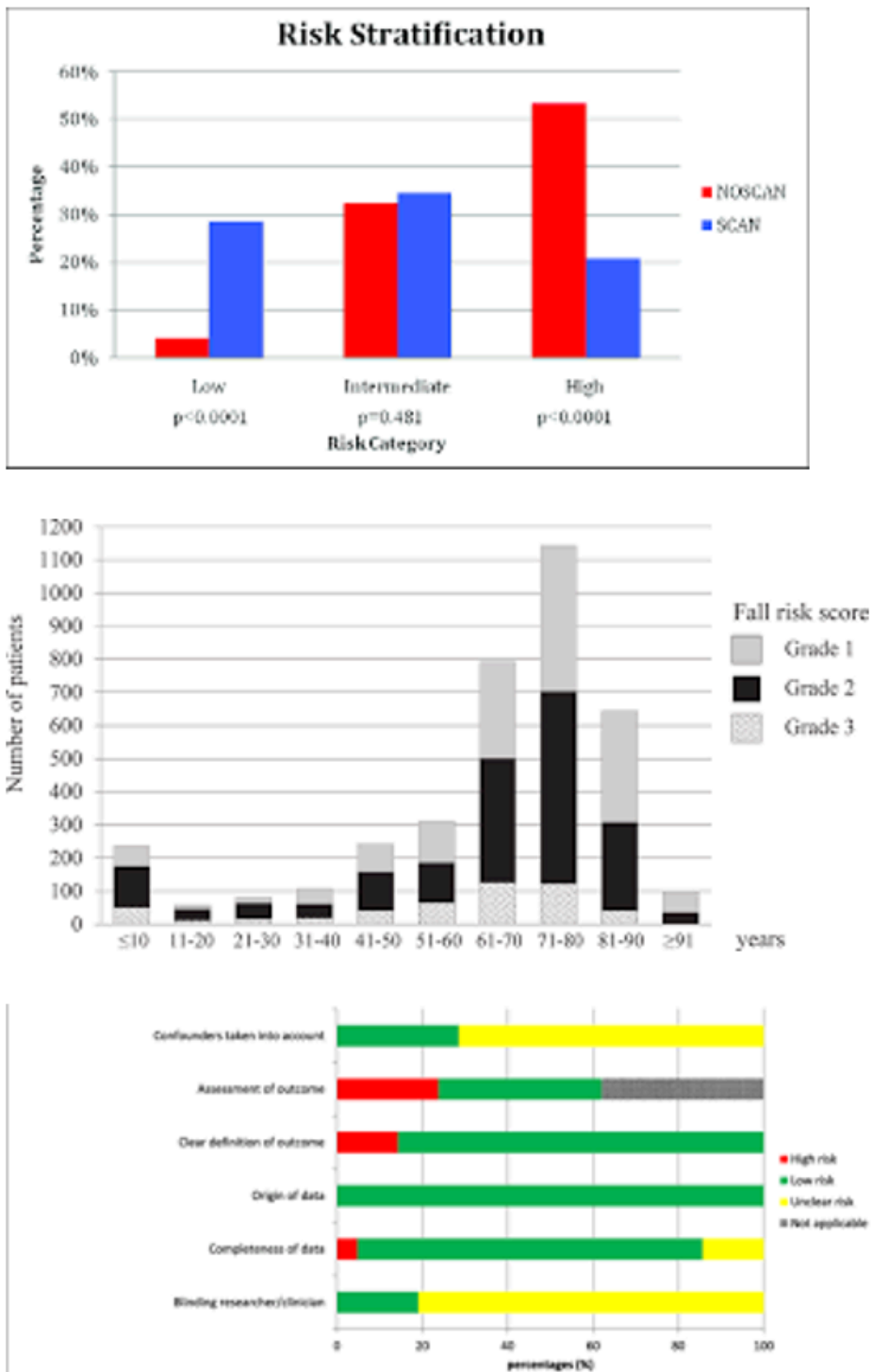


Figure 3

Risk stratification by predicted transfusion probability

Description: Observed transfusion rates across low-, intermediate-, and high-risk groups derived from data-driven probability thresholds.

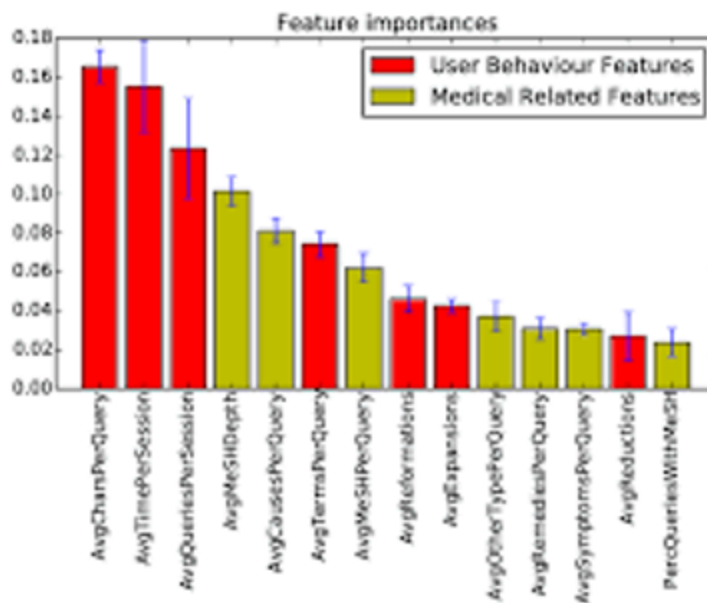
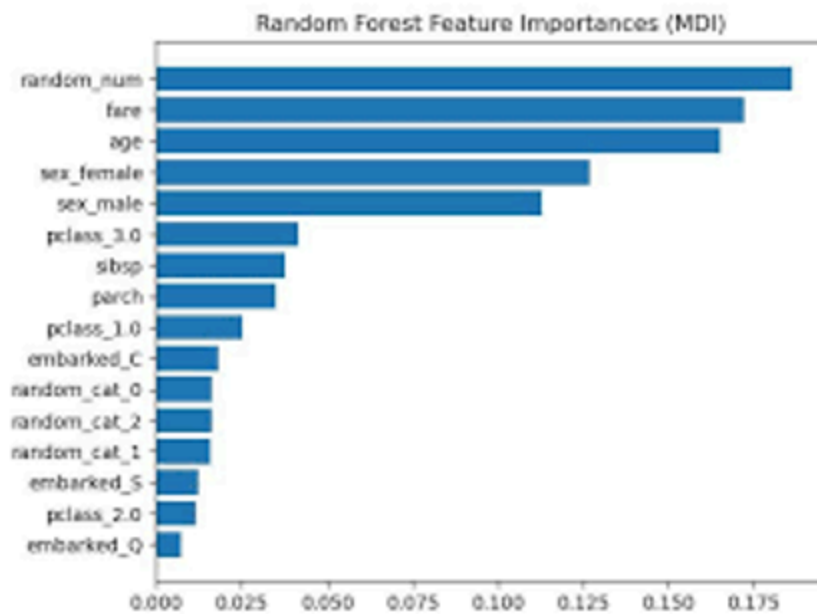


Figure 4

Feature importance for the Random Forest model

Description: Feature importance based on mean decrease in Gini impurity, highlighting age and preoperative PCV as dominant predictors.