

# Development of a Machine Learning–Based Predictive Model for Arteriovenous Fistula Occlusion after Surgery: A Retrospective Cohort Study from 2015 to 2025 Predicting AV Fistula Occlusion Post-Surgery

[Jae Hoon Lee](#) and [Sang Gyu Kwak](#), \*

Posted Date: 14 November 2025

doi: 10.20944/preprints202511.1097.v1

Keywords: arteriovenous fistula; hemodialysis; machine learning; vascular access; occlusion



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Development of a Machine Learning–Based Predictive Model for Arteriovenous Fistula Occlusion After Surgery: A Retrospective Cohort Study from 2015 to 2025 Predicting AV Fistula Occlusion Post-Surgery

Jae Hoon Lee <sup>1</sup> and Sang Gyu Kwak <sup>2,\*</sup>

- <sup>1</sup> Division of Vascular and Endovascular Surgery, Department of Surgery, College of Medicine, Daegu Catholic University, Daegu, Republic of Korea.
- <sup>2</sup> Department of Medical Statistics, Daegu Catholic University School of Medicine, Daegu, Republic of Korea
- \* Correspondence: sgkwak@cu.ac.kr; Tel.: +82-53-650-4724; Fax: +82-53-621-4106

Abstract

**Background:** Arteriovenous fistula (AVF) occlusion remains a major cause of vascular access failure in hemodialysis patients. Early identification of high-risk patients may help prevent complications and improve outcomes. **Methods:** This retrospective cohort study included 1,498 adult patients who underwent AVF creation between 2015 and 2025 at Daegu Catholic University Medical Center. Clinical, surgical, and laboratory variables were used to develop machine learning (ML) models for predicting AVF occlusion. Five algorithms—LightGBM, CatBoost, XGBoost, Random Forest, and Logistic Regression—were trained and evaluated using stratified 5-fold cross-validation. Model performance was assessed using area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, and calibration. SHAP (Shapley Additive Explanations) analysis was used to interpret variable importance. **Results:** Among the 1,498 patients, 381 (25.4%) experienced AVF occlusion. LightGBM achieved the best performance (AUC = 0.887, accuracy = 0.858, specificity = 0.950), followed by CatBoost (AUC = 0.882) and XGBoost (AUC = 0.879). Calibration analysis demonstrated strong agreement between predicted and observed outcomes. SHAP analysis identified ferritin, hemoglobin, neutrophil percentage, and C-reactive protein as the most influential predictors, highlighting the role of inflammation and hematologic status in AVF failure. **Conclusions:** Gradient boosting–based ML models, particularly LightGBM and CatBoost, accurately predict AVF occlusion using routine clinical data. Explainable AI methods enhance interpretability, enabling early identification of high-risk patients and supporting precision vascular access management in hemodialysis care.

**Keywords:** arteriovenous fistula; hemodialysis; machine learning; vascular access; occlusion

Introduction

Arteriovenous fistula (AVF) is widely regarded as the optimal vascular access for patients undergoing maintenance hemodialysis due to its superior long-term patency and lower complication rates compared to arteriovenous grafts or central venous catheters. Despite its advantages, AVF occlusion remains a significant clinical problem that impairs dialysis efficiency and leads to repeated interventions such as angioplasty or revision surgery, thereby increasing patient morbidity and healthcare burden [1,2]. The occurrence of AVF occlusion is influenced by multiple factors including patient demographics (age, sex), comorbidities (diabetes, hypertension), vascular anatomy, and surgical techniques [3].

Previous studies have primarily investigated risk factors for AVF occlusion using traditional statistical methods such as univariate and multivariate logistic regression analyses [4–7]. Jin et al. [5] developed a risk prediction model for autogenous AVF thrombosis in maintenance hemodialysis patients; however, their model was limited by the relatively small number of variables and dataset size, restricting its generalizability.

Moreover, Hu et al. [6] compared post-thrombotic endovascular interventions with pre-emptive angioplasty and highlighted the procedural and anatomical factors associated with AVF dysfunction, whereas Li et al. [7] demonstrated that puncture-related thrombosis risk can be evaluated using structured risk assessment models. Although informative, these studies did not fully capture the complex interactions among clinical variables nor account for nonlinear relationships, underscoring the need for more sophisticated machine learning approaches.

Recent advances in medical big data and machine learning (ML) provide opportunities to overcome these limitations by integrating numerous clinical, laboratory, and procedural variables into predictive models with enhanced accuracy [8,9]. ML algorithms such as XGBoost, LightGBM, Random Forest, and Logistic Regression have demonstrated superior performance across diverse clinical prediction tasks by capturing complex patterns and interactions in data [10,11]. Furthermore, explainable AI techniques like SHAP (Shapley Additive Explanations) facilitate model interpretability, enabling clinicians to understand variable importance and effect directionality, thus improving clinical applicability.

Accordingly, this retrospective cohort study aims to develop and validate ML-based predictive models for postoperative AVF occlusion using comprehensive clinical data from adult patients who underwent AVF surgery at Daegu Catholic University Medical Center between 2015 and 2025. By leveraging a large-scale dataset encompassing patient demographics, comorbidities, surgical variables, and laboratory results, we seek to construct robust predictive models and identify key risk factors. The ultimate goal is to provide a reliable tool for early identification of high-risk patients to support timely preventive interventions and personalized management strategies, thereby improving patient outcomes and healthcare resource utilization.

## Methods

The study protocol was approved by the Institutional Review Board (IRB) of Daegu Catholic University Hospital (DCUMC 2025-11-012). Data from January 1, 2015, to September 30, 2024, were retrospectively collected and analyzed. As a retrospective study using de-identified medical records, individual patient consent was waived. All data were anonymized and securely stored in compliance with relevant privacy regulations.

### *Study Design and Population*

This retrospective cohort study included adult patients ( $\geq 18$  years) who underwent arteriovenous fistula (AVF) surgery for hemodialysis at Daegu Catholic University Medical Center from January 1, 2015, to September 30, 2025. Patients who died or were lost to follow-up within one month postoperatively, or whose records were incomplete or had excessive missing data, were excluded.

### *Data Collection and Preprocessing*

Clinical data, including demographic characteristics, comorbidities, laboratory results, vascular ultrasound findings, and dialysis history, were extracted from electronic medical records. Variables collected comprised age, sex, body mass index (BMI), presence of hyperlipidemia, hypertension, cardiovascular diseases, baseline laboratory values (hemoglobin, WBC, neutrophil, ferritin, CRP, ESR). Vascular ultrasound parameters included artery and vein diameters and presence of calcifications.

The primary outcome was AVF occlusion, defined as the occurrence of vascular access failure necessitating angiographic evaluation and intervention (e.g., percutaneous transluminal angioplasty or thrombectomy) during follow-up. Data were anonymized and missing values were analyzed to assess patterns. Missing variables were imputed using multiple imputation by chained equations (MICE) to minimize bias. Outliers were detected and handled based on interquartile range criteria and clinical plausibility. Categorical variables were encoded using label encoding or one-hot encoding as appropriate.

Model Development and Evaluation

We evaluated multiple machine learning algorithms for binary classification of AVF occlusion, including logistic regression, random forest, XGBoost, CatBoost and LightGBM. Hyperparameters for each model were optimized via grid search with 5-fold stratified cross-validation on the training set. Hyperparameters were set as follows:

- XGBoost: 300 estimators, learning rate 0.05, max depth 4, subsample 0.8, column sampling 0.8
- Random Forest: 300 trees, max depth 8
- Logistic Regression: maximum iterations 1000
- LightGBM: 300 estimators, learning rate 0.05
- CatBoost: 300 iterations, learning rate 0.05, depth 6

Predictive performance was evaluated on the testing dataset using multiple metrics: accuracy, area under the receiver operating characteristic curve (AUC), confusion matrix, precision, recall (sensitivity), specificity, and F1-score. To assess the stability of each model, 5-fold stratified cross-validation was performed on the training data, reporting mean and standard deviation of accuracy and AUC. Calibration curves were constructed to compare predicted probabilities of AVF occlusion with observed event rates across risk strata for each model. Receiver operating characteristic (ROC) curves were generated to visualize and compare discriminatory performance.

Statistical Analysis

Baseline characteristics were summarized using means with standard deviations for quantitative variables, and frequency with percentages for qualitative variables. Group comparison employed t-test, or chi-square test as appropriate. To enhance interpretability, Shapley Additive Explanations (SHAP) analysis was performed, primarily focusing on the best model due to its superior performance. Mean absolute SHAP values quantify the overall contribution of each predictor to the model’s output. Summary plots depicted both the magnitude and direction of each variable’s effect on predicted risk, with color coding representing high (red) and low (blue) feature values. This approach enabled identification of key risk factors influencing AVF occlusion. Statistical analyses were conducted using Python (version 3.11). A two-sided p-value <0.05 was considered statistically significant.

Results

Baseline Characteristics

A total of 1,498 patients who underwent AVF creation were included in the study, among whom 381 (25.4%) experienced AVF occlusion during the follow-up period, while 1,117 (74.6%) maintained AVF patency. The baseline clinical and laboratory characteristics according to AVF status are summarized in Table 1.

Table 1. Baseline Characteristics of Patients According to AVF Patency and Occlusion.

| Variable    |        | Patency (n=1,117) | Occlusion (n=381) | p-value |
|-------------|--------|-------------------|-------------------|---------|
| Age (years) |        | 64.925 ±13.722    | 63.493±13.227     | 0.071   |
| Sex         | Female | 465 (41.6)        | 171 (44.9)        | 0.294   |
|             | Male   | 652 (58.4)        | 210 (55.1)        | 0.294   |

|                            |            |                 |                   |        |
|----------------------------|------------|-----------------|-------------------|--------|
| BMI (kg/m <sup>2</sup> )   |            | 22.935±3.617    | 22.228±3.137      | <0.001 |
| AVF site                   | Left       | 955 (85.5)      | 271 (71.1)        | <0.001 |
|                            | Right      | 162 (14.5)      | 110 (28.9)        | <0.001 |
| Smoking                    | No         | 971 (86.9)      | 286 (75.1)        | <0.001 |
|                            | Yes        | 146 (13.1)      | 95 (24.9)         | <0.001 |
| Drinking                   | No         | 1,029 (93.0)    | 346 (91.1)        | 0.247  |
|                            | Yes        | 77 (7.0)        | 34 (8.9)          | 0.247  |
| Hyperlipidemia             | No         | 1,050 (94.0)    | 320 (84.0)        | <0.001 |
|                            | Yes        | 67 (6.0)        | 61 (16.0)         | <0.001 |
| Hypertension               | No         | 315 (28.2)      | 57 (15.0)         | <0.001 |
|                            | Yes        | 802 (71.8)      | 324 (85.0)        | <0.001 |
| COPD                       | No         | 1,108 (99.2)    | 381 (100.0)       | 0.170  |
|                            | Yes        | 9 (0.8)         | 0 (0.0)           | 0.170  |
| CVD                        | No         | 946 (84.7)      | 330 (86.6)        | 0.407  |
|                            | Yes        | 171 (15.3)      | 51 (13.4)         | 0.407  |
| CVA                        | No         | 838 (75.0)      | 322 (84.5)        | <0.001 |
|                            | Yes        | 279 (25.0)      | 59 (15.5)         | <0.001 |
| IAD (mm)                   | Pre-op     | 4.235±0.957     | 4.250±0.886       | 0.788  |
|                            | Post-op    | 4.352±0.946     | 4.417±0.914       | 0.239  |
|                            | Difference | 0.072 (0.268)   | 0.154±0.264       | <0.001 |
| DVD (mm)                   | Pre-op     | 3.903±1.831     | 4.469±1.862       | <0.001 |
|                            | Post-op    | 4.653±1.829     | 4.938±1.830       | 0.009  |
|                            | Difference | 0.679±0.710     | 0.632±0.558       | 0.185  |
| Hemoglobin (g/dL)          |            | 9.427±1.354     | 9.009±0.912       | <0.001 |
| WBC (x10 <sup>3</sup> /μL) |            | 7.869±3.110     | 8.400±2.545       | 0.001  |
| Neutrophil (%)             |            | 70.996±8.891    | 75.389±12.874     | <0.001 |
| Ferritin (ng/mL)           |            | 439.388±419.586 | 1,075.920±803.381 | <0.001 |
| CRP (mg/L)                 |            | 22.804±47.864   | 50.724±78.508     | <0.001 |
| ESR (mm/h)                 |            | 35.527±23.850   | 42.181±35.418     | 0.001  |

There were no significant differences in age or sex distribution between the occlusion and patency groups. The mean age was 63.5 ± 13.2 years in the occlusion group and 64.9 ± 13.7 years in the patency group (p = 0.071). The proportion of males was slightly higher in the patency group (58.4%) than in the occlusion group (55.1%), but this difference was not statistically significant (p = 0.294).

Significant differences were observed in several clinical parameters. The occlusion group had a lower mean body mass index (BMI) compared to the patency group (22.2 ± 3.1 vs. 22.9 ± 3.6, p < 0.001). AVF site distribution differed significantly, with a higher proportion of right-sided AVFs in the occlusion group (28.9% vs. 14.5%, p < 0.001). Smoking prevalence was higher in the occlusion group (24.9% vs. 13.1%, p < 0.001), as was the prevalence of hyperlipidemia (16.0% vs. 6.0%, p < 0.001) and hypertension (85.0% vs. 71.8%, p < 0.001).

Vascular measurements revealed that the preoperative draining vein diameter (DVD) was larger in the occlusion group (4.47 ± 1.86 mm) compared to the patency group (3.90 ± 1.83 mm, p < 0.001), while differences in inflow artery diameter (IAD) were not statistically significant preoperatively. However, the postoperative increase in IAD was greater in the occlusion group (0.154 ± 0.264 mm vs. 0.072 ± 0.268 mm, p < 0.001).

Laboratory values also differed significantly: hemoglobin levels were lower in the occlusion group (9.01 ± 0.91 g/dL vs. 9.43 ± 1.35 g/dL, p < 0.001), whereas inflammatory markers such as white blood cell count (WBC), neutrophil percentage, ferritin, C-reactive protein (CRP), and erythrocyte sedimentation rate (ESR) were elevated (all p-values < 0.01), indicating a higher inflammatory state in patients experiencing AVF occlusion.

Predictive Model Performance

Five machine learning models—LightGBM, CatBoost, XGBoost, Random Forest, and Logistic Regression—were developed to predict AVF occlusion. Table 2 presents a comprehensive

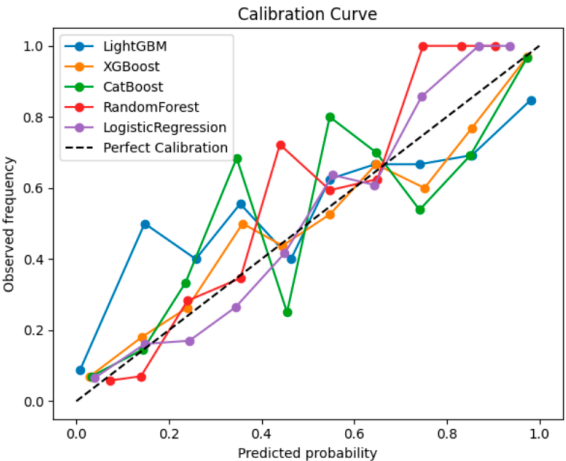
comparison of model performance metrics. LightGBM outperformed other models, achieving the highest area under the ROC curve (AUC) of 0.887, accuracy of 85.8%, and F1-score of 0.68. CatBoost and XGBoost demonstrated comparable performance with AUCs of 0.882 and 0.879, respectively. Random Forest showed slightly lower predictive ability (AUC 0.865), while Logistic Regression exhibited the lowest performance (AUC 0.792). While gradient boosting models demonstrated high specificity (93–95%), their sensitivities ranged from 0.57 to 0.60, indicating moderate ability to detect AVF occlusion cases. Logistic Regression showed lower sensitivity (0.43) and precision (0.62), underscoring the superiority of ensemble machine learning techniques in this context.

**Table 2.** Performance of Machine Learning Models for AVF Occlusion Prediction.

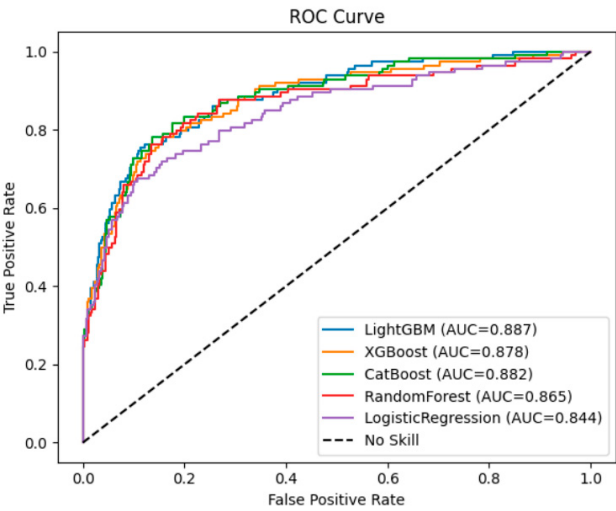
| Model       | LightGBM | CatBoost | XGBoost | Random Forest | Logistic Regression |
|-------------|----------|----------|---------|---------------|---------------------|
| Precision   | 0.790    | 0.790    | 0.760   | 0.730         | 0.620               |
| Sensitivity | 0.600    | 0.570    | 0.600   | 0.540         | 0.430               |
| Specificity | 0.950    | 0.950    | 0.930   | 0.930         | 0.910               |
| F1-score    | 0.680    | 0.660    | 0.670   | 0.620         | 0.510               |
| Accuracy    | 0.858    | 0.853    | 0.849   | 0.833         | 0.789               |
| AUC         | 0.887    | 0.882    | 0.879   | 0.865         | 0.792               |

Calibration and Discrimination

Figure 1 shows the calibration curves comparing the predicted probabilities of AVF occlusion with the observed event frequencies across the five machine learning models. Among the models evaluated, LightGBM and CatBoost demonstrated the best calibration, with their curves lying closest to the reference diagonal line representing perfect calibration. This indicates that these models’ predicted probabilities closely matched the actual observed risk of AVF occlusion, suggesting well-calibrated probability estimates suitable for clinical interpretation. Overall, gradient boosting–based algorithms (LightGBM, CatBoost, and XGBoost) provided more reliable and clinically interpretable probability predictions than the conventional Logistic Regression model. In addition, to evaluate discriminative performance, Figure 2 presents the receiver operating characteristic (ROC) curves for the five machine learning models used to predict AVF occlusion. All models demonstrated strong discriminative ability, with areas under the curve (AUC) exceeding 0.84. Among them, LightGBM achieved the highest AUC (0.887), followed by CatBoost (0.882), XGBoost (0.878), Random Forest (0.865), and Logistic Regression (0.844).



**Figure 1.** Calibration curves compare the predicted probabilities of AVF occlusion versus observed outcomes for different machine learning models.



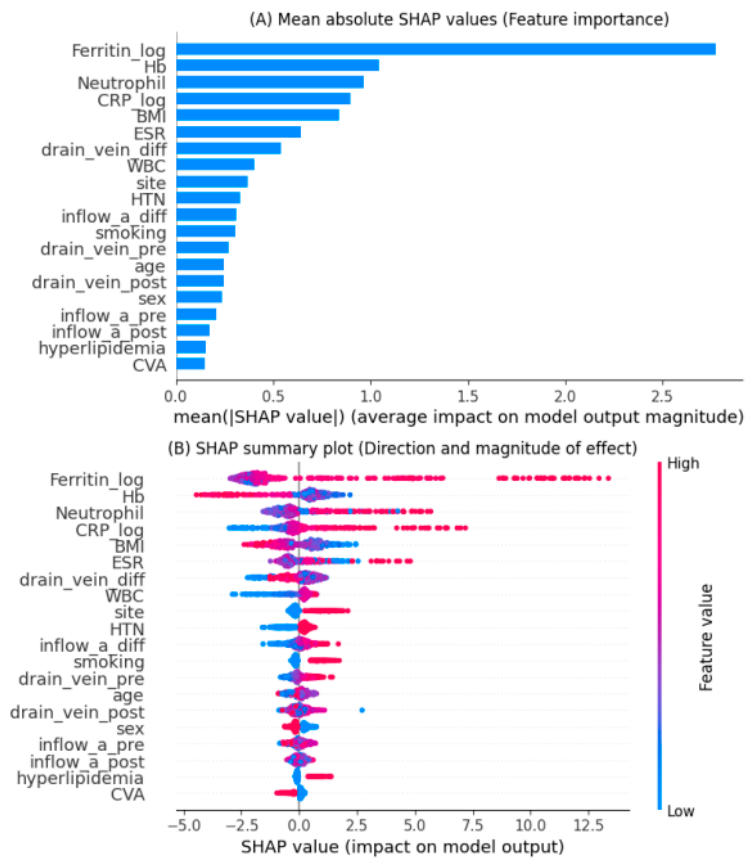
**Figure 2.** Receiver operating characteristic (ROC) curves illustrating the predictive performance of various machine learning models for AVF occlusion.

The ROC curves of the three gradient boosting-based models (LightGBM, CatBoost, XGBoost) were positioned consistently above Logistic Regression model, confirming superior classification performance and robustness. The LightGBM model showed particularly high true-positive rates across a wide range of false-positive rates, suggesting more accurate discrimination between patients who experienced AVF occlusion and those who maintained patency.

Together with calibration curve results (Figure 1), these findings indicate that LightGBM and CatBoost not only provided accurate probability estimation but also offered the best balance between sensitivity and specificity for clinical risk prediction.

*Feature Importance and Model Interpretability*

Feature importance was evaluated using Shapley Additive Explanations (SHAP) analysis applied to the LightGBM model, which demonstrated the best predictive performance. Figure 3 summarizes both the magnitude and direction of the contribution of each feature to AVF occlusion prediction.



**Figure 3.** Feature importance analysis based on SHAP values. (A) The bar plot shows the mean absolute SHAP value for each feature, representing its overall contribution to the LightGBM model predicting AVF occlusion. (B) The summary plot illustrates the direction and magnitude of the effect of each variable. Variables with higher SHAP values indicate stronger influence on the predicted risk; red dots indicate higher feature values, while blue dots indicate lower feature values.

Panel (A) shows the mean absolute SHAP values, indicating the overall importance of each variable in the LightGBM model. The most influential predictors were serum ferritin (log-transformed), hemoglobin, neutrophil percentage, and C-reactive protein (CRP, log-transformed). These variables were followed by body mass index (BMI), erythrocyte sedimentation rate (ESR), and the difference in draining vein diameter before and after surgery, suggesting that both inflammatory markers and vascular characteristics significantly contributed to model predictions.

Panel (B) illustrates the SHAP summary plot, which depicts both the magnitude and direction of each variable's effect. High ferritin, neutrophil, and CRP levels (represented in red) were associated with an increased predicted risk of AVF occlusion, whereas higher hemoglobin and BMI values (blue) were linked to a lower risk. The results indicate that systemic inflammation and impaired hematologic status are major contributors to AVF dysfunction.

Conversely, vascular features such as postoperative changes in draining vein and inflow artery diameters also influenced risk estimation, reflecting the impact of surgical and anatomical factors.

Overall, the SHAP analysis provided an interpretable framework linking biological plausibility with model outputs, enabling identification of clinically relevant predictors and supporting the potential use of the model for individualized risk stratification.

Discussion

This study developed and validated machine learning-based predictive models for postoperative arteriovenous fistula (AVF) occlusion using a large cohort of 1,498 patients who underwent AVF creation between 2015 and 2025. Among the five models tested, gradient boosting-based algorithms—particularly LightGBM and CatBoost—demonstrated superior predictive performance, achieving the highest accuracy, discrimination, and calibration compared to traditional Logistic Regression and Random Forest models. The results highlight the feasibility and clinical potential of applying explainable artificial intelligence (AI) techniques to predict AVF dysfunction in hemodialysis patients.

Previous studies have primarily employed traditional regression-based approaches to identify risk factors for AVF dysfunction[4,5]. Although these models have provided valuable insights, they were limited by small sample sizes and an inability to account for complex, nonlinear relationships among clinical variables. In contrast, our study incorporated diverse preoperative and perioperative features, including demographic, vascular, and inflammatory parameters, allowing the machine learning models to capture multidimensional interactions. The superior discrimination of LightGBM (AUC 0.887) and CatBoost (AUC 0.882) confirms the strength of ensemble learning in handling heterogeneous clinical data.

SHAP analysis revealed that elevated ferritin, neutrophil percentage, and C-reactive protein (CRP) were the most influential predictors of AVF occlusion, whereas higher hemoglobin and body mass index (BMI) were associated with lower risk. These findings are biologically plausible and consistent with prior reports indicating that systemic inflammation and oxidative stress contribute to neointimal hyperplasia and thrombosis in AVFs[4,12]. High ferritin levels may reflect chronic inflammation and iron overload, both of which promote vascular injury and impaired endothelial function. Similarly, CRP and neutrophil activation have been linked to endothelial dysfunction and hypercoagulability, predisposing to vascular access failure. In addition, vascular parameters such as changes in inflow artery and draining vein diameters also contributed meaningfully to the predictive model. These anatomical factors likely reflect surgical technique and vessel adaptability, which are critical determinants of postoperative patency. Therefore, the integration of both biological and anatomical variables underscores the importance of a holistic approach to AVF outcome prediction.

A major advantage of this study lies in the use of explainable machine learning through SHAP analysis, which allows for transparent interpretation of model predictions. The ability to quantify each feature's contribution enables clinicians to identify high-risk patients early and implement targeted interventions—such as intensified monitoring, preemptive angioplasty, or aggressive management of inflammation and anemia. The calibration analysis further confirmed that LightGBM and CatBoost produced well-calibrated probability estimates, making them suitable for clinical decision-support applications.

The strengths of this study include the use of a large real-world dataset spanning 10 years, multiple machine learning algorithms for comparative analysis, and explainable AI for model interpretation. However, several limitations should be acknowledged. First, this was a single-center retrospective study, which may limit generalizability to other populations or surgical settings. Second, external validation with data from different institutions is needed to confirm the robustness of the developed models. Third, some potential predictors—such as detailed intraoperative parameters and postoperative hemodynamic measurements—were not available in this dataset.

Future research should focus on prospective validation and dynamic modeling that incorporates longitudinal variables such as flow rate changes and repeated laboratory measurements. Integration of these models into electronic health record systems could enable real-time risk prediction and individualized clinical decision-making.

## Conclusions

In summary, this study demonstrates that machine learning, particularly gradient boosting algorithms such as LightGBM and CatBoost, can effectively predict AVF occlusion using routinely collected clinical and laboratory data. The incorporation of interpretable AI methods enhances

transparency and clinical trust, providing a foundation for precision vascular access management and improved hemodialysis outcomes.

All values were presented by mean  $\pm$  standard deviation or frequency(percent). P-values were obtained by two sample t-test for quantitative variables or chi-square test for qualitative variables. AVF: Arteriovenous Fistula; BMI: Body Mass Index; COPD: Chronic Obstructive Pulmonary Disease; CRP: C-Reactive Protein; CVA: Cerebrovascular Accident; CVD: Cardiovascular Disease; DVD: Draining Vein Diameter; ESR: Erythrocyte Sedimentation Rate; IAD: Inflow Artery Diameter.

**Author Contributions:** Conceptualization: Jae Hoon Lee, Sang Gyu Kwak; Data curation: Jae Hoon Lee, Sang Gyu Kwak; Formal analysis: Sang Gyu Kwak; Methodology: Sang Gyu Kwak; Writing - original draft: Sang Gyu Kwak; Writing - review & editing: Jae Hoon Lee. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study did not receive any specific grants from funding agencies in the public, commercial, or non-profit sectors.

**Institutional Review Board Statement:** The study was approved by the Institutional Review Board of Daegu Catholic University Medical Center (DCUMC 2025-11-012). Patient consent was waived due to the retrospective nature of the study and the use of de-identified data.

**Informed Consent Statement:** Patient consent was waived due to the retrospective design and the use of de-identified clinical data.

**Data Availability Statement:** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Conflicts of Interest:** No potential conflict of interest relevant to this article was reported.

## References

1. Nordyke, R.J., et al., Costs Attributable to Arteriovenous Fistula and Arteriovenous Graft Placements in Hemodialysis Patients with Medicare coverage. *Am J Nephrol*, 2019. **50**(4): p. 320-328.
2. Thamer, M., et al., Medicare Costs Associated With Arteriovenous Fistulas Among US Hemodialysis Patients. *Am J Kidney Dis*, 2018. **72**(1): p. 10-18.
3. Lok, C.E. and R. Foley, *Vascular access morbidity and mortality: trends of the last decade*. *Clin J Am Soc Nephrol*, 2013. **8**(7): p. 1213-9.
4. Viecelli, A.K., et al., The pathogenesis of hemodialysis vascular access failure and systemic therapies for its prevention: Optimism unfulfilled. *Semin Dial*, 2018. **31**(3): p. 244-257.
5. Jin, X., et al., Construction of Risk-Prediction Models for Autogenous Arteriovenous Fistula Thrombosis in Patients on Maintenance Hemodialysis. *Blood Purif*, 2024. **53**(10): p. 813-823.
6. Hu, X., et al., Hemodialysis Arteriovenous Fistula Dysfunction: Retrospective Comparison of Post-thrombotic Percutaneous Endovascular Interventions with Pre-emptive Angioplasty. *Ann Vasc Surg*, 2022. **84**: p. 286-297.
7. Li, M., C. Sun, and X. Du, Application Value and Relevance Analysis of the Risk Evaluation System for Arteriovenous Fistula Puncture in Thrombosis after Puncture. *J Healthc Eng*, 2021. **2021**: p. 6919979.
8. Heindel, P., et al., Predicting radiocephalic arteriovenous fistula success with machine learning. *npj Digital Medicine*, 2022. **5**(1): p. 160.
9. Meng, L. and P. Ho, A systematic review of prediction models on arteriovenous fistula: Risk scores and machine learning approaches. *J Vasc Access*, 2025. **26**(3): p. 735-746.
10. Caro Acevedo, P., et al., A systematic follow-up protocol achieving a low hemodialysis graft thrombosis rate. *J Vasc Access*, 2019. **20**(6): p. 683-690.

11. Broeck, G.V.d., et al., *On the Tractability of SHAP Explanations*. J. Artif. Intell. Res., 2022. **74**: p. 36.
12. Girerd, S., et al., Arteriovenous fistula thrombosis is associated with increased all-cause and cardiovascular mortality in haemodialysis patients from the AURORA trial. Clin Kidney J, 2020. **13**(1): p. 116-122.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.