

Risk Prediction Models for Autogenous Arteriovenous Fistula Maturation Failure in Patients with End-Stage Renal Disease: A Systematic Review and Meta-Analysis

Jinhong Miao^{*}, Jiajia Jia, Yujie Zhu, Juan Du, Yutian Lei, Qingyang Zhang

Hemodialysis Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

ABSTRACT

Background: Autogenous Arteriovenous Fistula (AVF) is the preferred vascular access for maintenance hemodialysis in patients with End-Stage Renal Disease (ESRD), but its clinical utility is limited by high rates of maturation failure. Although multiple predictive models have been developed, their reliability and methodological quality remain uncertain.

Objective: To systematically review and evaluate the predictive performance and risk of bias of existing models for AVF maturation failure in ESRD patients.

Methods: A systematic search of nine databases (including PubMed, Embase, Web of Science, and Cochrane Library) was conducted from inception to November 14, 2024. Studies involving adult ESRD patients undergoing new AVF creation and reporting prediction models for AVF maturation failure were included. Data extraction followed the CHARMS framework, and methodological quality was assessed using the PROBAST tool. Meta-analysis of model performance was conducted using AUC values, with heterogeneity assessed via I^2 statistics and Cochran's Q test.

Results: Eighteen studies comprising 18 prediction models (sample sizes 66–14,892) across eight countries were included. The incidence of AVF maturation failure ranged from 11.0% to 69.7%. Common predictors included preoperative vein diameter, age, sex, diabetes, and peripheral vascular disease. Model performance showed moderate to good discrimination, with AUC values ranging from 0.629 to 0.886. However, only three studies conducted external validation, and most relied on internal validation methods. Seventeen models were assessed as having a high risk of bias, primarily due to limitations in study design, predictor measurement, outcome assessment, and statistical analysis.

Conclusions: Existing prediction models for AVF maturation failure demonstrate variable performance and are generally limited by high risk of bias and insufficient external validation. Future research should focus on improving methodological rigor, standardizing predictor and outcome definitions, and conducting robust external validations to enhance clinical applicability.

Keywords: AVF maturation; Outcome assessment; Statistical analysis; Diameter

INTRODUCTION

Autogenous Arteriovenous Fistula (AVF) is the preferred vascular access for Maintenance Hemodialysis (MHD) in patients With End-Stage Renal Disease (ESRD) due to its

superior longevity, high blood flow, low maintenance costs, and reduced infection risk compared to arteriovenous grafts and hemodialysis catheters. According to the Dialysis Outcomes and Practice Patterns Study (DOPPS), the rates of forearm AVF establishment vary significantly by region: 84% in Japan, 54% in

Correspondence to: Jinhong Miao, Hemodialysis Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China; E-mail: jiajia@gs.zzu.edu.cn

Received: 17-Dec-2024, Manuscript No. AOA-24-35940; **Editor assigned:** 20-Dec-2024, PreQC No. AOA-24-35940 (PQ); **Reviewed:** 03-Jan-2025, QC No. AOA-24-35940; **Revised:** 09-Feb-2026, Manuscript No. AOA-24-35940 (R); **Published:** 16-Feb-2026, DOI: 10.35841/2329-9495.26.14.564

Citation: Miao J, Jia J, Zhu Y, Du J, Lei Y, Zhang Q (2026) Risk Prediction Models for Autogenous Arteriovenous Fistula Maturation Failure in Patients with End-Stage Renal Disease: A Systematic Review and Meta-Analysis. Angiol Open Access. 14:564.

Copyright: © 2026 Miao J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Europe/Australia-New Zealand, and 32% in the United States. However, they are also associated with a low maturation rate, which results in an increased reliance on first-time dialysis catheters, prolonged catheterization, and a heightened risk of potential infection and mortality in patients with ESRD. Research indicates that the incidence of AVF maturation failure varies between 20% and 54%, with rates further increasing in individuals of advanced age, comorbid diabetes mellitus, elevated cholesterol levels, and a history of cervical cannulation. Risk prediction, accurate identification, and timely intervention for AVF maturation failure are essential for improving AVF maturation rates [1].

Although risk factors for AVF maturation failure are known, the effectiveness of predictive models in predicting AVF maturation failure remains uncertain. A systematic review revealed that models predicting AVF dysfunction in MHD patients are biased due to varied data sources and analysis reporting. While machine learning and risk scoring systems demonstrate strong predictive accuracy within their development cohorts, they are not exempt from bias due to under-reporting and the absence of external validation. The emergence of new studies on AVF maturation failure models underscores the necessity for ongoing research and a comprehensive reassessment of these predictive tools. This study aims to systematically review the predictive performance and potential bias in existing risk prediction models for AVF maturation failure in ESRD patients. The findings will provide valuable references for clinical practice and future research [2].

MATERIALS AND METHODS

The study protocol was registered on PROSPERO (CRD42024610923).

Search strategy

China National Knowledge Infrastructure (CNKI), Wanfang Database, China Science and Technology Journal Database (VIP), SinoMed, PubMed, Web of Science, The Cochrane Library, Embase, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) and were searched from inception to November 14, 2024. The retrieval method employed in this review combined subject terms and free-text searches, while also manually tracing the references of the included literature for further supplement [3].

To conduct the systematic review, we adhered to the PICOTS framework, endorsed by the Cochrane prognostic methodology group for a comprehensive and structured approach. The key components of our systematic review are outlined as follows: P: ESRD patients undergoing new Arteriovenous Fistula (AVF) surgery; I: Prediction models assessing the risk of AVF maturation failure; C: No competing models; O: The risk of AVF maturation failure; T: 12 weeks post-AVF surgery to evaluate whether maturation is sufficient for puncture use; S: Nephrology and hemodialysis centers.

Selection criteria

Inclusion criteria: (1) The study population consisted of ESRD patients aged 18 years or older who were undergoing the creation of a new AVF; (2) The study focused on the construction and/or validation of a predictive model assessing the risk of AVF maturation failure; (3) The primary outcome was defined as AVF maturation failure, characterized by difficulties in puncture and/or inadequate blood flow, which necessitated functional exercise or balloon-assisted maturation interventions; (4) The study design was case-control, cross-sectional, or cohort; (5) The language of the study was either Chinese or English.

Exclusion criteria: (1) Studies with fewer than two model predictors; (2) Modeling based on reviews or systematic evaluations; (3) Conference abstracts or duplicate publications; (4) Literature with inaccessible full text, incorrect original data, duplicates, or unextractable [4].

Study screening and data extraction

All literature retrieved from 9 databases was imported into EndNote (version 20.0, Clarivate Analytics, Philadelphia, PA, USA) for literature management according to the search strategy. Subsequent screening process and data extraction were carried out independently by 2 reviewers, and any discrepancies were resolved through discussion or by a third reviewer. Firstly, duplicates were removed. Secondly, the remaining studies were assessed for eligibility based on titles and abstracts. Then the full texts were reviewed based on inclusion and exclusion criteria, and reviews and reference lists of all eligible studies were examined to identify potentially relevant studies.

Data extraction was performed following CHARMS (critical appraisal and data extraction for systematic reviews of prediction model studies). Extracted data included (1) Basic characteristics of the studies: First author, year of publication, country, study design, study population, data source, definition of outcome and incidence; (2) Specifics of the prediction model construction: Sample size, model development method, method of screening for predictor variables, treatment of missing values, model predictive performance and model presentation; (3) Model validation: Internal and external validation [5].

Methodological quality assessment

The quality of the included literature was evaluated by two researchers independently using the PROBAST (Prediction Model Risk of Bias Assessment Tool) developed by Wolff et al. PROBAST is a tool specifically designed to critically assess the risk of bias and applicability for predictive model construction and validation. The risk of bias assessment consists of 20 items in four domains: Research object, predictor, outcome, and statistical analysis. Each signaling question can be answered as "yes", "probably yes", "no", "probably no" or "no information". If at least one signaling question in a domain is answered as "no" or "probably no", that domain should be considered at high risk of bias. Only when all domains are judged as low risk of bias, the overall bias can be considered low risk. The assessment of applicability includes the first three domains, and the assessment

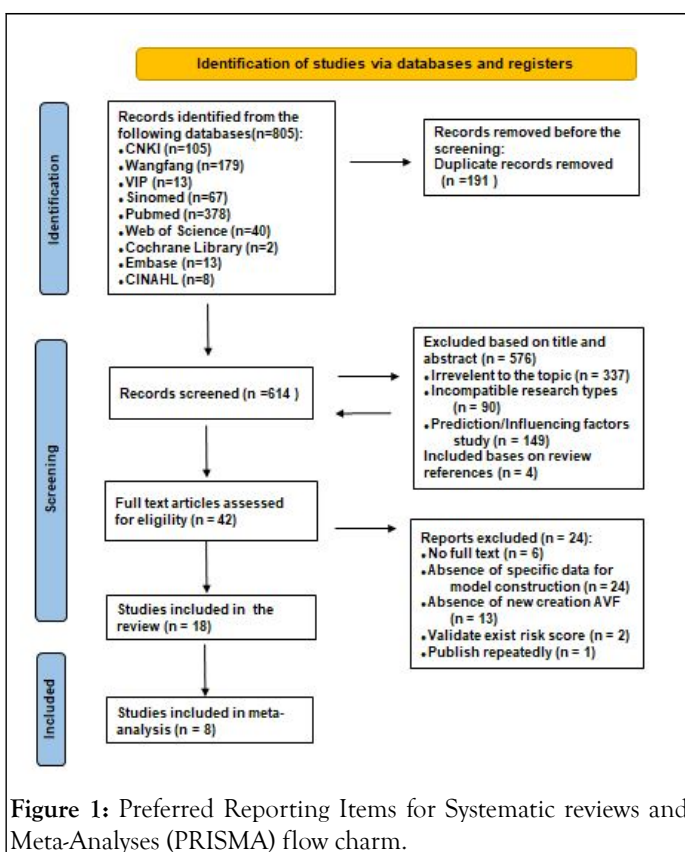
process and criteria are similar to the assessment of risk of bias.

Statistical analysis

A meta-analysis of the Area Under the Curve (AUC) values of the validated models was performed using R software (version 4.4.2, meta package, The R Project for Statistical Computing, Auckland, New Zealand). Heterogeneity was assessed using the I^2 statistic and Cochran's Q test. The fixed effects model was selected if the data heterogeneity was small ($P \geq 0.1$, $I^2 \leq 50\%$); the random effects model was determined if the data heterogeneity was significant ($P < 0.1$ or $I^2 > 50\%$). Egger's test was applied to evaluate publication bias, with $P > 0.05$ suggesting a low risk of such bias [6].

RESULTS

A total of 805 articles were obtained through database search, and 18 articles were finally included according to the inclusion criteria through gradual screening (Figure 1).



Studies characteristics

18 studies encompassing 18 predictive models and involving 66 to 14,892 subjects were published from 2003 to 2024, with 6 studies appearing in the last 5 years. The studies were conducted in 8 countries, including USA, UK, China, Canada, Thailand, India, Netherlands and Singapore. 15 studies encompassed model development and validation, while 3 concentrated on development alone. The studies were evenly split between retrospective and prospective analyses, with 9 of each. Beyond

single-center research, 4 studies utilized multi-center data, while another 4 drew from extensive renal registry databases [7].

Predicted time points were 6 weeks, 3 months and 6 months postoperatively, with a follow-up of 1-2 years. Predicted outcomes were based on the clinical AVF puncture use and guidelines (AVF diameter, flow, and depth). The incidence of AVF maturation failure ranged from 11.0% to 69.7%. For studies focusing on AVF patency, Twine et al.; Heindel et al. and Zhang et al. predicted AVF primary patency (30.3-69.0%); Wongmahisorn et al. and Bosanquet et al. predicted AVF secondary patency at 1 and 2 years (53.5-86.2%), respectively [8].

Model predictors and assignments

The number of predictors ultimately included in the prediction models ranged from 2 to 13, with the most frequently occurring predictors including preoperative vein diameter, age, gender, comorbid diabetes, and peripheral vascular disease. Predictors assigned higher values included preoperative vein diameter, peripheral vascular disease, gender, race, and age. 6 models categorized the risk of AVF maturation failure into low, medium, high, and very high-risk groups based on predictor scores. In the machine learning-based models, predictor screening, handling of missing values, and model presentation were excluded from the analysis [9].

Model development, performance and validation

Model development: In the included studies, modeling methods were logistic regression, Cox regression, and machine learning models (random forests, artificial neural networks, decision trees, etc.). 10 studies categorized the continuous variables in the predictor variables and 6 studies screened the predictor variables by backward stepwise selection. 7 studies reported missing data and treatments, of which only 3 reported specific missing values or proportions of 25.0-43.0%, 0.3-26.3%, and 5 cases. In terms of model presentation, 5 studies used scoring equations, 6 studies used scoring systems, 3 studies used line diagrams, and 2 studies used web calculators. The remaining 2 studies, Kordzadeh et al. and Heindel et al. used machine learning to construct the model and did not present the model.

Model performance and validation: 16 studies reported model predictive performance with an area under the receiver operating characteristic curve (AUC) of 0.629-0.886. Calibration was reported in 11 models using the Hosmer-Lemeshow (HL) test, calibration curve, and decision curve methods. 13 models performed internal validation of predictive models using bootstrap resampling and sample splitting, while 3 models performed external validation across time, geography, and populations. Only Lok et al. combined internal and external validation methods for model validation.

Literature quality assessment

Risk of bias assessment: 17 risk prediction models were at high risk of bias overall: 9 models in the subject domain, 4 in the predictor domain, 13 in the outcome domain, and 1 in the

statistical analysis domain. Firstly, retrospective cohort studies rely on existing case data, which may lead to large discrepancies between predicted and actual outcomes, resulting in biased data sources. Subjective bias was deemed unclear due to the lack of clarity in the inclusion and exclusion criteria of the model. Secondly, 4 studies that used multicenter clinical data for modeling did not indicate whether the predictor measurements were uniform across centers. 9 retrospective studies faced potential bias due to confounding in predictor measurement and unclear blinding, leading to an unclear bias risk rating. Thirdly, inconsistencies in the timing of outcome event measurements and the lack of blinding for predictors during outcome event measurements in retrospective studies increase the risk of bias in the outcome domain. Lastly, variable conversion from continuous to categorical, reporting of missing values, incomplete screening of predictors based on a single factor, and incomplete assessment of the predictive performance of models cause bias in statistical analyses.

Applicability assessment: 9 studies were evaluated as high applicability risk overall. To start with, 3 models had a high risk of applicability in terms of the study population, namely AVF and AVG, AVF in patients with diabetic nephropathy, and newly constructed AVF after CVC in 67 years or older. In addition, 3 models had a high risk of applicability in terms of predictors. Specifically, Tng et al., included preoperative venous mapping charts as a predictive factor, yet the study was constrained by the practical limitations of clinical examinations, affecting the accuracy of preoperative venous mapping for each AVF. Kanchanasuttirak et al., included data from physical examination and ultrasound at 2 and 6 weeks postoperatively, and Heindel et al., included ultrasound data from visits at 4-6 weeks postoperatively, which may have resulted in the inability to routinely perform ultrasound in the same time due to ultrasound accessibility. Besides, 5 models had a high risk of applicability due to endpoint discrepancies. The studies' measurements at 6 weeks post-AVF surgery and 1-year patency rates differ from the outcomes defined in this review.

Meta-analysis results

Inclusion was limited by insufficient literature on the model development process, with 8 studies meeting the inclusion criteria. Using the random effects model, the combined AUC was estimated to be 0.77 (95% CI: 0.70-0.85), indicating good predictive accuracy. High heterogeneity ($I^2=98.1\%$, $P<0.01$) and no significant publication bias ($P=0.299$) among studies was observed in the analysis (Figure 2).

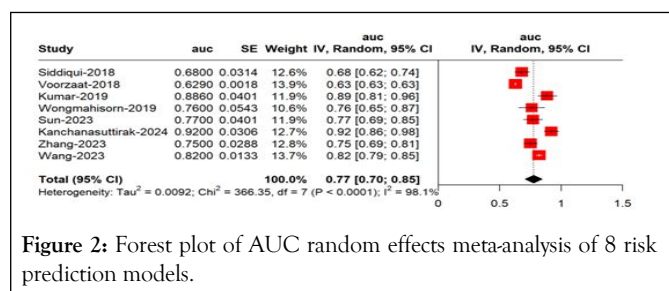


Figure 2: Forest plot of AUC random effects meta-analysis of 8 risk prediction models.

DISCUSSION

Factors predicting AVF poor maturation

Predictive models incorporating demographic data, preoperative vascular attributes, and medical history can accurately forecast AVF maturation failure. Anastomotic vein diameter is a key predictor, with patients having veins under 3 mm (OR=0.28, 95% CI 0.14-0.53) and those with veins under 2 mm (HR=0.69, 95% CI 0.50-0.96) exhibiting lower maturation rates. Therefore, meticulous preoperative evaluation is essential for vessels under 2 mm in diameter to ensure optimal AVF quality. Elderly patients, particularly those aged 65 and above, are at an increased risk of AVF maturation failure, with an odds ratio of 0.23 (95% CI 1.25-3.96) and those over 73 years exhibit an even higher risk, with a hazard ratio of 1.56 (95% CI 1.02-2.40). Healthcare providers should be alert to the potential for AVF dysfunction in this demographic. Studies including those by Siddiqui et al.; Voorzaat et al., and Sun et al. have consistently identified female gender as an independent risk factor for AVF maturation failure. Diabetes and a history of peripheral vascular disease are indeed common risk factors for AVF maturation failure. In diabetic patients, compromised glucose and lipid metabolism, along with impaired endothelial function, contribute to atherosclerosis and vascular calcification. These conditions can disrupt vasoconstriction and diastolic function, leading to altered hemodynamics and, ultimately, poor vascular remodeling, which is associated with AVF maturation failure.

Our primary analysis focused on established risk factors, but the diversity across studies suggests that additional variables, such as nutritional status, inflammatory markers, and lifestyle habits, may also play a role. A subset of studies has underscored the significance of these factors, indicating a need for further investigation. Future research should aim to standardize data collection to gain a more holistic understanding of AVF outcomes.

Good accuracy and predictive performance existed in the models

The meta-analysis of 8 models revealed a pooled AUC of 0.77 (95% CI: 0.70-0.85), signifying high overall predictive accuracy. The performance of these models is gauged by two principal metrics: Differentiation and calibration. Differentiation is commonly evaluated through the AUC value or C index, whereas calibration is assessed *via* goodness-of-fit tests, Brier scores, and calibration curves. In our systematic review, 16 studies assessed model differentiation, with 11 also evaluating calibration. Collectively, the models demonstrated robust predictive performance and a degree of practical applicability.

High bias found in the development and validation of the models

The majority of studies were deemed to be at high risk of bias by the PROBAST assessment, which diminishes the practical applicability of the predictive models. This significant heterogeneity among the models can likely be attributed to differences in population demographics, predictor variables, and methodological approaches.

Model development: The majority of studies relied on retrospective data, which may omit essential information. For future model development or optimization, it is recommended to use prospective cohort, case-control, or nested case-control studies to minimize information bias in model construction. In terms of variable selection, most studies utilized univariate and multivariate analyses. LASSO regression, which helps to reduce multi-collinearity among numerous predictors, is particularly advantageous. However, only studies by Sun et al. and Heindel et al. employed this method. It is suggested that multiple imputations for missing values and avoiding early categorization of continuous variables in model development should be adopted. While 9 studies did not report on missing value handling, and 10 studies prematurely categorized continuous variables. Machine learning methods tend to be more accurate than traditional logistic regression in terms of sample size, handling of continuous variables, and predictor selection. Therefore, future research should provide a robust and interpretable variable selection mechanism in machine learning-based methods and then tailor these to score-generating regression models.

Model validation and results reporting: Most studies lack multi-center data and external validation, with some models showing poor external validation results ($AUC < 0.7$), which limits their predictive power and applicability. Larger, multi-center studies are needed to validate these models and ensure their broad clinical utility and accuracy. Additionally, the majority of studies did not follow the TRIPOD standards, leading to a lack of transparency in model construction and affecting the quality assessment of the literature. Future studies should adhere to TRIPOD guidelines to enhance the clarity of the modeling process.

Implications for clinical practice

The high failure rate of AVF maturation has substantial clinical implications, emphasizing the need for precise and timely risk assessment and intervention in ESRD patients to reduce adverse outcomes. In our study, out of 18 prediction models, 9 demonstrated AUC values above 0.7, indicating good predictive performance. The number of predictors in these models varied from 2 to 10, and most predictors were objectively measurable and easily collectible, enhancing their clinical utility compared to traditional rating scales. Furthermore, 6 models were risk-scored, and 2 provided a free online calculator to visualize AVF maturation failure risk prediction models, allowing clinicians to stratify patients into low, moderate, and high-risk groups and to implement targeted preventive and therapeutic strategies based on individual risk levels.

CONCLUSION

This study encompassed 18 AVF maturation failure risk prediction models, which exhibited good discriminative ability, calibration, and applicability. However, there was a notable risk of heterogeneity and bias across the models. The study identified key predictors of AVF maturation failure, including anastomotic vein diameter, age, gender, diabetes, and peripheral vascular disease. Future research should integrate traditional statistical methods with machine learning techniques and adhere to stringent reporting standards to develop models with robust design and multi-center validation. By doing so, these risk prediction models can effectively identify patients at high risk for AVF maturation failure, enabling early detection, prevention, and intervention, and ultimately improving the quality of life for ESRD patients.

LIMITATIONS OF THIS STUDY

This study only included studies published in English and Chinese, there may be limitations related to language barriers. Studies based on retrospective data and included inconsistent predictors may have excluded key factors such as nutritional status, inflammatory factors, and dialysis data, thus affecting the accuracy of our risk prediction models. The heterogeneity among the included studies posed challenges in conducting a meta-analysis to quantitatively synthesize the results.

DATA AVAILABILITY

The data used to support the findings of this study can be found in the appendix. Further information can be directed to the corresponding author.

CONFLICTS OF INTEREST

All authors have no conflict of interests.

AUTHORS' CONTRIBUTIONS

Jinhong Miao: Conceptualization; literature search; integration of the structure; methodology; project administration; paper review and editing.

Jiajia Jia: Literature search; literature screening; data extraction; quality evaluation; methodology; writing the original draft.

Yujie Zhu: Literature screening; data extraction; quality evaluation.

Juan Du: Data integration; quality evaluation.

Yutian Lei: Research content review and integration.

Qingyang Zhang: Software; research content review and integration.

FUNDING

This study was funded by "The First Affiliated Hospital of Zhengzhou University Special Nursing Care Project in 2023:

Construction and application of a risk pre-control model for vascular access for dialysis" (NO: HLKY2023003).

REFERENCES

1. Zhao X, Niu Q, Gan L, Hou FF, Liang X, Ni Z, et al. Baseline data report of the China Dialysis Outcomes and Practice Patterns Study (DOPPS). *Sci Rep.* 2021;11(1):873.
2. Pisoni RL, Zepel L, Zhao J, Burke S, Lok CE, Woodside KJ, et al. International comparisons of native arteriovenous fistula patency and time to becoming catheter-free: Findings from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Am J Kidney Dis.* 2021;77(2): 245–254.
3. Voorzaat BM, van der Bogt KEA, Janmaat CJ, van Schaik J, Dekker FW, Rotmans JL. Arteriovenous fistula maturation failure in a large cohort of hemodialysis patients in the Netherlands. *World J Surg.* 2018;42(6):1895–1903.
4. Bylsma LC, Gage SM, Reichert H, Dahl SLM, Lawson JH. Arteriovenous fistulae for haemodialysis: A systematic review and meta-analysis of efficacy and safety outcomes. *Eur J Vasc Endovasc Surg.* 2017;54(4):513–522.
5. Chytilova E, Jemcov T, Malik J, Pajek J, Fila B, Kavan J. Role of Doppler ultrasonography in the evaluation of hemodialysis arteriovenous access maturation and influencing factors. *J Vasc Access.* 2021;22(Suppl 1):42–55.
6. Robbin ML, Greene T, Allon M, Dember LM, Imrey PB. Prediction of arteriovenous fistula clinical maturation from postoperative ultrasound measurements: Findings from the hemodialysis fistula maturation study. *J Am Soc Nephrol.* 2018;29(11):2735–2744.
7. Meng L, Ho P. A systematic review of prediction models on arteriovenous fistula: Risk scores and machine learning approaches. *J Vasc Access.* 2024;25.
8. Han M, Zhao Q, Zhao J, Xue X, Wu H. Risk prediction models for autogenous arteriovenous fistula failure in maintenance hemodialysis patients: A systematic review and meta-analysis. *World J Surg.* 2024;48(10):2526–2542.
9. Lok CE, Huber TS, Lee T, Shenoy S, Yevzlin AS, Abreo K, et al. KDOQI clinical practice guideline for vascular access: 2019 update. *Am J Kidney Dis.* 2020;75(4 Suppl 2):S1–S164.