Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Dual-Criterion Variable Inclusion Strategy for Improved Confounder Inclusion in Small-Sample Trauma Cohorts: An Empirical Methodological Study

Fuyao Yang
Second Hospital of Hebei Medical University
Shiqin Dong
Second Hospital of Hebei Medical University
Xiaoyu Li
Second Hospital of Hebei Medical University
Yayu Zhuo
Second Hospital of Hebei Medical University
Kai Tang
Second Hospital of Hebei Medical University
Jianping Sun
neuro.sjp@hebmu.edu.cn

Second Hospital of Hebei Medical University

Research Article

Keywords: Dual-criterion variable inclusion, Confounding control, Small-sample cohort, Empirical methodological study, Trauma, Logistic regression, Pulmonary infection

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10856829/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

Pulmonary infection is a common severe complication following decompressive craniectomy for cerebral contusion and laceration and is linked to poor prognosis in traumatic brain injury patients. Traditional single-threshold variable screening tends to exclude potential confounders in small-sample highly confounded cohorts, leading to incomplete covariate inclusion, biased effect estimates and unstable risk-factor identification. To optimize confounding adjustment in small-sample observational studies, we propose a dual-criterion variable-inclusion strategy to improve confounder completeness and the accuracy of risk-factor identification.

Methods

This empirical methodological study used a single-center retrospective cohort of 88 cerebral contusion patients receiving decompressive craniectomy. Candidate variables were screened via a dual-criterion union rule: baseline-imbalance testing or univariate-outcome testing, with conventional single-*P* value screening as a control. Ridge regression was conducted as a sensitivity analysis. Based on the dual-criterion strategy, a maximally-adjusted model and a parsimonious model were compared; calibration was assessed for the parsimonious model via bootstrap. A random-forest model was built as a comparator via leave-one-out cross-validation.

Results

The overall incidence of posttraumatic pulmonary infection was 58.43%. The single-criterion strategy omitted one key clinically critical confounder, subdural hematoma, whose omission caused substantial bias: false-positive age associations and masked neutrophil percentage effects. In contrast, the dual-criterion strategy included more confounders and produced different variable-selection results. Admission diastolic blood pressure and neutrophil percentage were stable independent risk-factors. The optimized model showed acceptable discrimination (apparent AUC = 0.712; bootstrap-corrected AUC = 0.693) and calibration (MAE = 0.033), whereas the random-forest model exhibited limited generalization performance (LOOCV AUC = 0.505).

Conclusions

The proposed dual-criterion strategy improves confounder inclusion completeness in small-sample trauma cohorts. It reduces the omission bias of conventional single-threshold screening and thereby yields a parsimonious prediction model with internally validated discrimination and calibration. This framework offers a reference for confounding adjustment and variable selection in similar small-sample

clinical observational studies. Our empirical comparison demonstrates that methodological choices in variable screening can substantially alter both risk-factor identification and model conclusions in small-sample settings. This framework highlights the importance of methodological transparency in variable selection for small-sample clinical prediction studies.

Background

Cerebral contusion and laceration (CCL) is a typical pathological subtype of traumatic brain injury (TBI). Direct external force can cause contusion, laceration, hemorrhage, and edema of brain tissue, which are often complicated by refractory intracranial hypertension [1]. For critically ill patients whose intracranial pressure cannot be controlled by medication, decompressive craniectomy (DC) is a life-saving procedure that expands the cranial cavity volume to alleviate brain swelling and reduce mortality. However, postoperative patients often experience prolonged bed rest, impaired consciousness, and compromised cough and swallowing reflexes; some patients require tracheostomy, making them highly susceptible to nosocomial pulmonary infection. The incidence of pulmonary infection after DC can reach 50–60%, which is higher than that in general neurosurgical patients. As a major complication, pulmonary infection prolongs hospital stays, worsens neurological outcomes, and increases mortality risk, posing a considerable. challenge to clinical management [2][3][4].

Current clinical strategies for the prevention and treatment of postoperative pulmonary infection (PPI) mainly include perioperative prophylactic antibiotics and enhanced airway-secretion clearance. Respiratory function monitoring is routinely performed to guide these interventions; nevertheless, the overall effectiveness of such measures remains suboptimal. The core challenge lies in the clinical lack of an early-risk identification tool capable of operating accurately and reliably in such small-sample, highly heterogeneous settings. Conventional predictive models select candidates for multivariate models based solely on the *P* value cutoff from univariate logistic regression. This method implicitly assumes that all potential confounders have a marginal association with the study outcome, and that the pre-specified *P* value cutoff of 0.1 is sufficient to detect all such factors during initial variable screening. However, in the complex clinical context of post-DC infections, baseline variables such as age and a history of drug or alcohol abuse are not only associated with infection outcomes but also strongly intertwined with TBI severity, physiological stress, and treatment intensity, forming a complex network of confounding factors [3][4]. Against this background, traditional univariate inclusion methods are prone to statistical bias: some indicators with only an apparent association may still appear to be significant factors in subsequent multivariate regression analyses and are consequently identified as independent risk factors. The key reason for this false significance is that variables with confounding effects are excluded because their *P* values in univariate analysis do not meet the inclusion criterion of 0.1. Insufficient adjustment for confounding factors fails to eliminate spurious associations, obscuring the true effects of each predictive factor. Thus, the models using this approach lack stability and cannot reliably support clinical decision-making.

The associations of admission diastolic blood pressure (DBP) and the peripheral-blood neutrophil percentage (NEUT) with PPIs remain controversial in the literature, partly because conventional screening strategies may have failed to adequately adjust for confounding factors. These two variables therefore provide a suitable test case for evaluating whether the proposed dual-criterion strategy can resolve such inconsistencies. Previous studies have suggested that abnormal fluctuations in hemodynamic parameters, such as DBP, may reflect decompensation of cerebrovascular autoregulation and excessive activation of the systemic stress response, suggesting a potential pathophysiological link with susceptibility to secondary infections [5]. Similarly, as the most rapidly responding inflammatory marker, elevated preoperative or postoperative NEUT has been shown in multiple studies to be associated with an increased risk of various types of postoperative surgical infections [6]. However, many of these studies failed to systematically consider and adjust for the confounding effects of baseline factors during the variable selection phase. For example, a patient's stress state at admission may simultaneously elevate both DBP and NEUT. Injury severity, such as subdural hematoma (typically excluded in the conventional single-criterion strategy but included in the dual-criterion strategy), also serves as a confounding factor linking abnormalities in these two indicators to adverse clinical outcomes. If such confounding effects are not corrected, the accuracy and reproducibility of the conclusions will inevitably be compromised. The innovative potential of this study lies in its methodological framework, which proposes a dual-criterion variable inclusion strategy to overcome the cognitive inertia of conventional single-criterion strategies. The optimized inclusion strategy allows for a more accurate assessment of the independent predictive value of DBP and NEUT for PPIs in highly confounded, small-sample datasets.

To address the above challenges, this study developed and applied a dual-criterion variable inclusion strategy combined with multimodel comparative validation. In terms of the methodological design, we adopted a manual variable categorization strategy. For continuous variables, we distinguished between normal and skewed distributions based on the mean-to-median ratio together with histograms and quantile-quantile (Q-Q) plots; for categorical variables, we classified them as ordered or unordered based on clinical judgment. Building on this foundation, a dual-criterion variable inclusion strategy was proposed and implemented. Rather than simply applying a single threshold, this strategy combines results from two criteria, significant differences in the baseline characteristic and univariate logistic regression screening, to form a union of variables for inclusion in multivariate models, aiming to more comprehensively account for potential confounders and predictors. We subsequently constructed three logistic regression models: one following the traditional inclusion approach, one deeply integrating the dual-criterion variable inclusion strategy as the main model, and another using the stepwise AIC algorithm to develop an automated screening model. To maximize the robustness of the results given the limited sample size, we further adopted Ridge regression for sensitivity analysis[7] and constructed a random forest model as a comparator for predictive performance, which was validated via leave-one-out cross-validation (LOOCV) [8]. The purpose of this analytical framework is to mitigate confounding bias and overfitting in small-sample studies.

We identified and quantified the independent risk effects of DBP and NEUT on PPI among high-risk, small-sample patients with CCL who underwent DC, based on the framework described above. Based on this, we constructed and evaluated a predictive model for the early identification of high-risk patients using these two predictors, which are routinely available on admission, to assess its preliminary screening utility. We also sought to use this clinical cohort to test the feasibility of the proposed methodological framework, a strategy that may offer a practical solution for identifying risk factors in highly confounded small-sample cohorts by improving the completeness of confounder inclusion. More broadly, the comparisons throughout both variable screening and model development show that methodological choices in variable screening can substantially alter risk-factor identification.

Methods

Study Design and Subjects

This was a single-center, retrospective cohort study that consecutively enrolled 99 patients with CCL who underwent DC at the Second Hospital of Hebei Medical University between April 28, 2019, and November 23, 2025. After excluding 11 patients whose data on pulmonary infection were incomplete, a total of 88 patients were ultimately included in the study and grouped based on the occurrence of PPI. Outcome and predictor variables were extracted retrospectively from electronic medical records. Blinding was not performed for either outcome or predictor assessment. The patient enrollment and exclusion procedures are summarized in Fig. 1.

Variable Classification and Intergroup Comparisons

The variables were manually classified into four categories: normally distributed continuous variables, nonnormally distributed continuous variables, unordered categorical variables, and ordered categorical variables. Corresponding statistical tests were applied to conduct intergroup comparisons. For continuous variables, the mean-to-median ratio together with histograms and Q-Q plots was used to assess their distribution characteristics, thereby avoiding the limitation of relying merely on the Shapiro-Wilk test, which is overly sensitive to minor deviations in small-sample cohorts. For categorical variables, their ordered or unordered attributes were manually distinguished based on clinical logic, without relying on automatic classification by software. Only preoperative indicators were included in subsequent analyses, and all postoperative variables were excluded to avoid reverse causality.

For intergroup statistical comparisons, an independent-samples t test was adopted for normally distributed continuous variables, whereas a Mann-Whitney U test was used for continuous variables with skewed distributions. For unordered categorical variables, the chi-square test was used when all expected cell counts were ≥ 5 in 2 × 2 tables or when fewer than 20% of the cells had expected counts below 5 in larger tables. Otherwise, Fisher’s exact test was applied. Ordered categorical variables were compared via the Mann-Whitney U test to preserve hierarchical information of graded clinical variables.

In terms of data presentation, normally distributed continuous variables are described as the means ± standard deviations, nonnormally distributed continuous variables are reported as medians (interquartile ranges), and all categorical variables are expressed as numbers (percentages).

Variable Selection Strategy

Variables with $P < 0.05$ in the baseline analyses (Table 1) or $P < 0.1$ in the univariate logistic regression (Wald test) were selected and incorporated into the multivariate modeling pool following the dual-criterion variable inclusion strategy. Spearman's correlation matrix was calculated for these candidate variables, and highly correlated indicators were excluded to prevent multicollinearity prior to model construction. For example, the time from injury to admission and the time from injury to surgery were highly correlated; we retained the latter because it showed stronger baseline imbalance and a univariate P value below 0.1. The missing values of the remaining candidate variables were subsequently processed. The proportion of missing data was extremely low, with ≤ 2 missing cases per variable (< 3%). Continuous missing values were imputed using the median; categorical missing values were coded as a separate unknown category and included in the analysis.

Table 1
Baseline clinical characteristics of CCL patients receiving DC

Variables	Variable type	Noninfection group	Infection group	*P* value
Sex (Male)	Unordered categorical	12 (33.3%)	11 (21.2%)	0.302
Sex (Female)	Unordered categorical	24 (66.7%)	41 (78.8%)	0.302
Epidural hematoma (No)	Unordered categorical	30 (85.7%)	32 (61.5%)	0.028
Epidural hematoma (Yes)	Unordered categorical	5 (14.3%)	20 (38.5%)	0.028
Subdural hematoma (No)	Unordered categorical	0 (0.0%)	12 (23.1%)	0.001
Subdural hematoma (Yes)	Unordered categorical	36 (100.0%)	40 (76.9%)	0.001
Subarachnoid hemorrhage (No)	Unordered categorical	15 (41.7%)	10 (19.2%)	0.04
Subarachnoid hemorrhage (Yes)	Unordered categorical	21 (58.3%)	42 (80.8%)	0.04
Preoperative signs of brain herniation (No)	Unordered categorical	7 (19.4%)	23 (45.1%)	0.024
Preoperative signs of brain herniation (Yes)	Unordered categorical	29 (80.6%)	28 (54.9%)	0.024
Basal cistern compression (Normal)	Ordered categorical	2 (5.7%)	3 (5.8%)	0.169
Basal cistern compression (Compressed)	Ordered categorical	18 (51.4%)	35 (67.3%)	0.169
Basal cistern compression (Obliterated)	Ordered categorical	15 (42.9%)	14 (26.9%)	0.169
Age(years)	Normally distributed continuous	59.2 (14.1)	51.7 (13.8)	0.014
Right pupil diameter (mm)	Normally distributed continuous	4.1 (1.3)	3.2 (1.3)	0.002
Admission diastolic blood pressure (mmHg)	Normally distributed continuous	78.3 (13.7)	85.1 (14.0)	0.026
Admission body temperature (℃)	Normally distributed continuous	36.6 (0.3)	36.8 (0.5)	0.011

Variables	Variable type	Noninfection group	Infection group	*P* value
Peripheral-blood neutrophil percentage (%)	Normally distributed continuous	82.7 (14.0)	87.3 (4.5)	0.064
Activated Partial Thromboplastin Time (s)	Normally distributed continuous	31.0 (5.1)	27.9 (4.8)	0.005
Fibrinogen (g/L)	Normally distributed continuous	1.8 (1.1)	2.8 (2.9)	0.029
Serum potassium (mmol/L)	Normally distributed continuous	3.2 (0.4)	3.5 (0.5)	0.009
Intraoperative blood loss (mL)	Normally distributed continuous	990.3 (797.8)	681.7 (524.4)	0.047
Time from injury to admission (h)	Skewed continuous	4.0 (3.0–4.5)	4.5 (3.0–10.0)	0.046
Time from injury to surgery (h)	Skewed continuous	6.5 (4.8–9.0)	11.6 (6.0–30.1)	0.007

Construction and Validation of Multivariate Screening Models

After strict data preprocessing and quality control, three sets of multivariate logistic regression models were constructed in sequence:

Model 1 (traditional control model): only variables with $P < 0.1$ from the univariate logistic regression were included.

Model 2 (main model): candidate variables were determined via the dual-criterion variable inclusion strategy; the Akaike information criterion (AIC) of this model was calculated.

Model 3 (validation control model): stepwise regression was performed using the same variable pool as Model 2; its AIC value was also calculated; retained covariates and AIC values of Model 2 and Model 3 were compared to assess the stability of the dual-criterion screening strategy from an algorithmic perspective.

Sensitivity analyses were conducted on Model 1 and Model 2 via Ridge regression to validate the robustness of the findings. Multivariate analyses reported adjusted odds ratios (ORs) to reflect the independent risk effect of each variable.

Construction and Performance Evaluation of the Predictive Models

According to the events-per-variable (EPV) rule for clinical prediction models, the final parsimonious model included only two predictors, yielding an EPV of 25.5, which was well above the recommended

threshold of 20. This ensured stable estimation of the regression coefficients and reduced the risk of model overfitting.

A maximally-adjusted predictive model was constructed via all candidate predictors included in Model 2, whereas a parsimonious predictive model was built only with predictors that reached statistical significance in Model 2.

We assessed the discriminative ability of the two models via receiver operating characteristic (ROC) curves, the area under the curve (AUC), and accuracy (calculated based on the basis of the unified probability cutoff of 0.5). For the parsimonious model alone, a nomogram was generated, a calibration curve was plotted by bootstrapping (1000 resamples) on the fixed final predictors without repeated variable screening in each resample to avoid random inconsistency of predictor combinations, and the calibration goodness-of-fit was quantified via the mean absolute error (MAE).

A probability cutoff of 0.5 was uniformly adopted for the above two predictive models and the subsequent random forest model to standardize the evaluation criteria. For the parsimonious model, the Youden index was calculated over all probability thresholds on its ROC curve, with the resulting threshold used for patient risk stratification.

Machine learning comparative analysis

A random forest predictive model incorporating identical predictors as the parsimonious model was constructed concurrently. Given the small sample size, LOOCV was adopted to assess its generalization performance, and a parallel analysis was conducted with the parsimonious model.

All the statistical analyses in this study were performed via R software (version 4.5.2), and $P < 0.05$ was considered statistically significant. The 95% confidence intervals (95% CI) for diagnostic performance metrics were estimated via the Wilson score method.

Results

Baseline characteristics and incidence

A total of 88 patients were included in this study, with a PPI rate of 58.43%. A comparison of the baseline clinical data between the two groups is shown in Table 1.

Results of the multivariate logistic regression analysis

Compared with the single-criterion strategy, the dual-criterion approach yields a slightly expanded candidate variable pool. After collinear predictors were excluded and restricted to preoperative variables, the only additional variable retained in Model 2 was subdural hematoma (SDH).

Model 1: Constructed via a single-criterion inclusion strategy. SDH was not included in the candidate variables pool. After multivariate adjustment, only age and DBP had statistically significant effects, whereas NEUT had no statistically significant effect.

Model 2: Constructed via the dual-criterion inclusion strategy. SDH was included in addition to the candidate variables in Model 1. After adjustment, DBP (OR = 1.068, 95% CI, 1.009–1.142, *P* = 0.035) and NEUT (OR = 1.214, 95% CI, 1.033–1.505, P = 0.042) were identified as independent risk factors for PPI following DC. The quantitative results revealed that for each 1 mmHg increase in DBP and each 1% increase in NEUT, the odds of pulmonary infection increased by 6.8% and 21.4%, respectively, whereas age had no statistically significant effect.

Model 3: Based on the same pool of candidate variables as Model 2, Model 3 was constructed using the stepwise AIC (AIC = 78.38, lower than the 88.94 of Model 2). All the independent predictors identified in Model 2 were consistently retained in Model 3.

Sensitivity analyses using Ridge regression were conducted on Model 1 and Model 2. The OR values for each variable remained consistent with those in the initial models, suggesting that the conclusions of this study are robust.

A comparison of the ORs, 95% CIs, and *P* values of the core variables across the three models is presented in Table 2.

Table 2
OR, 95% CI and *P* values of the core variables across the three models

Variables	**Model 1 OR(95% CI) *P***	**Model 2 OR(95% CI) *P***	**Model 3 OR(95% CI) *P***
Age (years)	0.93(0.87–0.99) *P* = 0.043	0.95(0.88–1.01) *P* = 0.147	0.95(0.90–1.00) *P* = 0.051
Admission diastolic blood pressure (mmHg)	1.08(1.01–1.16) *P* = 0.026	1.07(1.01–1.14) *P* = 0.034	1.07(1.01–1.14) *P* = 0.022
peripheral-blood neutrophil percentage (%)	1.20(1.03–1.49) *P* = 0.053	1.21(1.03–1.50) *P* = 0.042	1.20(1.06–1.42) *P* = 0.012

Evaluation of Prediction Model Performance

Two prediction models were constructed based on Model 2. Their overall discrimination performance was as follows:

The maximally-adjusted model had an AUC = 0.943; with a probability threshold of 0.5, the model's overall accuracy was 84.09%.

The parsimonious model yielded an apparent AUC of 0.712 (95% CI, 0.600–0.824) in the training cohort. Bootstrap internal validation yielded an optimism-corrected AUC of 0.693, with a Dxy-based optimism of

0.039. With the same threshold of 0.5, the model calculated on the original training cohort was as follows: sensitivity (recall) of 86.54%, specificity of 41.67%, positive predictive value (precision, PPV) of 68.18%, negative predictive value (NPV) of 68.18%, and overall accuracy of 68.18%.

The calibration curves revealed that the predicted probabilities of the parsimonious model closely matched the actual infection outcomes, with an MAE of 0.033.

The optimal stratification threshold of this parsimonious model was further identified via the Youden index. The maximum Youden index corresponded to a cutoff value of 0.586. At this threshold, the model exhibited a sensitivity of 73.08%, specificity of 63.89%, PPV of 74.51%, NPV of 62.16%, and overall accuracy of 69.32%. The detailed diagnostic performance metrics under the two cutoff values are presented in Table 3.

Table 3
Diagnostic performance of the parsimonious model

Metric	**Youden index optimal cutoff (0.586)**	**Fixed threshold (0.5)**
TP	38	45
TN	23	15
FP	13	21
FN	14	7
Sensitivity (Recall)	73.08% (95% CI, 59.75–83.23%)	86.54% (95% CI, 74.73–93.32%)
Specificity	63.89% (95% CI, 47.58–77.52%)	41.67% (95% CI, 27.14–57.80%)
PPV	74.51% (95% CI, 61.13–84.45%)	68.18% (95% CI, 56.21–78.15%)
NPV	62.16% (95% CI, 46.10–75.94%)	68.18% (95% CI, 47.32–83.64%)
Accuracy	69.32%	68.18%

The final logistic regression prediction equation for the parsimonious model was as follows: $Logit(P) = -11.366 + 0.041 \times DBP + 0.098 \times NEUT$

The individual predicted probability of a PPI can be calculated by the following conversion formula:

$P = 1/(1 + \exp(-Logit(P)))$

The corresponding regression coefficients and 95% Wald CIs for the parsimonious model are summarized in Table 4.

Table 4
Regression coefficients for the parsimonious model

Variable	Coefficient	95% Wald CI	*P* value
Intercept	-11.366	(-21.128, -3.635)	0.0120
Admission diastolic blood pressure (mmHg)	0.041	(0.007, 0.078)	0.0225
Peripheral-blood neutrophil percentage (%)	0.098	(0.019, 0.197)	0.0355

A comparison of the ROC curves between the maximally-adjusted model and the parsimonious model is shown in Fig. 2; the predictive calibration performance of the parsimonious model is presented in Fig. 3, and the visualized nomogram is displayed in Fig. 4.

2.4 Comparison of Machine Learning Prediction Models

After being validated via LOOCV, the random forest model yielded an AUC of only 0.505. When this identical threshold of 0.5 was adopted, its precision (42.42%) and recall (38.89%) were close to random. Comparisons of these models suggested that the predictive performance of the random forest model was weaker than that of the Model 2 logistic regression model within this small-sample dataset. The ROC curve of the random forest model under LOOCV validation is presented in Fig. 5.

Discussion

Methodological Positioning of the Dual-Criterion Strategy

Early and accurate identification of high-risk patients for pulmonary infection after decompressive craniectomy remains a prominent challenge. This difficulty largely stems from the lack of systematic selection logic in traditional variable inclusion strategies. Some methods rely solely on univariate statistical test thresholds to screen variables for entry into multivariate models, while others depend primarily on researchers' clinical experience to subjectively decide which covariates to include. However, both strategies have inherent flaws. The former suffers from insufficient statistical power under small sample conditions, making it prone to omitting variables with weak marginal associations but genuine confounding effects, leading to inadequate confounding adjustment and biased effect estimates. The latter lacks standardized procedures and faces the dilemma of overinclusion or underinclusion: including too many highly correlated variables easily induces multicollinearity and model overfitting, whereas including too few variables may overlook important confounders. Moreover, in the context of incomplete clinical guidelines, experience-based selection itself lacks reproducibility and objectivity.

Despite their different forms, both strategies share a common flaw: they lack a systematic decision rule for determining which variables should enter the model. As a result, the selection of covariates becomes either statistically underpowered (single-P screening) or subjective (clinical experience). In both cases, variables with genuine confounding potential may be omitted before they can be adjusted for. This is

particularly problematic because multivariable logistic regression can only adjust for confounders that are actually included in the model. In small-sample, highly confounded cohorts, the variables most likely to be excluded by conventional single-*P* screening are those with weak univariate signals but strong confounding potential, such as SDH in this present cohort, which had an imbalanced baseline distribution (Fisher's test $P = 0.001$) yet a univariate *P* value = 0.987. If these variables are discarded before they enter the multivariable model, the adjustment capability of logistic regression cannot be fully exercised, and the resulting estimates remain biased. To address this methodological gap, we propose a dual-criterion variable inclusion strategy designed to retain, rather than discard, such variables, ensuring that the adjustment capability of multivariable regression can be fully applied. The following sections evaluate this strategy through multimodel comparisons, predictive performance assessments, and mechanistic interpretations.

Our results indicated that DBP and NEUT were key independent risk factors and that logistic regression with rigorous confounder adjustment provided more stable performance than did the random forest algorithm in this small-sample setting. These findings also help clarify previous uncertainties regarding traditional factors such as age and offer a methodological framework that can be adapted for confounding control and risk-factor identification in small-sample, highly confounded clinical data. Notably, several advanced approaches have been developed for small-sample clinical settings. From the modeling perspective, methods such as extreme gradient boosting with random effects [9] and Bayesian sparse regression with transfer learning [10] have shown promise, but they typically require multicenter hierarchical data or reliable external prior information. From the biomarker perspective, integrating circulating lncRNAs [11] or sEV-derived miRNAs with clinical markers [12] has achieved high predictive accuracy, yet these molecular assays require specialized laboratory techniques and often involve processing delays, while our two indicators are routinely available within minutes of admission. The least absolute shrinkage and selection operator (Lasso) regression has also become a widely used tool for variable shrinkage in prediction modeling, particularly in large-sample settings where the goal is to select a sparse set of variables with the strongest predictive signal. However, the role of Lasso is primarily prediction-driven: it removes variables that contribute little to predictive accuracy, regardless of their potential as confounders. In contrast, the core function of multivariable logistic regression is to adjust for confounding, but this adjustment can only occur if potential confounders are actually included in the model. Our dual-criterion strategy was therefore designed to retain, rather than discard, variables that may have weak predictive signals but strong confounding potential, such as SDH in the present cohort, ensuring that they remain available for adjustment in the multivariable stage. Thus Lasso, along with the other advanced modeling and biomarker approaches discussed above, and our dual-criterion strategy serve different research objectives and are complementary rather than competing approaches. Where these advanced methods are suited for scenarios with sufficient data, hierarchical structures, or specialized molecular assays, our strategy addresses the common but methodologically underexplored setting of confounder-driven risk-factor identification in small-sample cohorts.

Our strategy, through a transparent variable-inclusion rule, identified two routinely available clinical indicators (DBP and NEUT), offering a more accessible alternative that does not depend on either complex algorithms or molecular assays. Consistent with the study's stated objective, this investigation was designed specifically for the small-sample setting, not as a compromise but as the intended scope of inference. The proposed strategy is intended for use in small cohorts where conventional methods are prone to failure, and generalizability to large-sample or multicenter data is not the primary goal.

Empirical Evidence: Confounding Control and Bias Correction

As noted in the background, the baseline variables in this cohort form a complex confounding network: SDH reflects injury severity, DBP and NEUT reflect physiological and immune-inflammatory status, and age is intertwined with both severity and host response. Within this network, the exclusion of the SDH, a key node, can distort the estimated effects of other interconnected variables, as illustrated below. Notably, in this small-sample cohort, Model 1 exhibited marked selection bias and model distortion. The model failed to identify the NEUT as a statistically significant independent risk factor but erroneously indicated a significant association between age and the PPI. This occurred because, under small-sample conditions, the Wald *P* value from univariable regression cannot adequately reflect true confounding effects, thereby leading to biased variable selection. In this cohort, SDH was a critical potential confounder (no SDH: 0 patients in the noninfected control group versus 12 patients in the pulmonary infection group). Owing to the presence of zero cells in the contingency table and insufficient expected frequencies in certain cells, Fisher's exact test was employed for between-group comparisons according to the prespecified statistical rules, yielding P = 0.001, which clearly indicated a significant difference in the SDH distribution between the two groups. However, in univariable logistic regression, the Wald test was distorted by sparse data, resulting in an inflated *P* value and an extreme deviation in the OR output. The same variable thus produced two entirely contradictory statistical signals: Fisher's test indicated a strong baseline imbalance with confounding risk, whereas Wald univariable regression suggested no statistical association. The conventional screening strategy that relies solely on univariable regression *P* values adopted the latter, leading to the exclusion of SDH from the multivariable model because $P > 0.1$, although this variable should have been incorporated as a key confounder for adjustment.

This omission introduced residual confounding, masking the independent association between the NEUT and pulmonary infection and rendering the NEUT statistically nonsignificant in Model 1. Moreover, age, whose baseline distribution was imbalanced ($P = 0.014$), absorbed the surrogate effects of the other excluded covariates, producing a spuriously significant association. This issue is not unique to the present study. A systematic review by Xu et al. of prognostic prediction models for acute exacerbation of chronic obstructive pulmonary disease revealed that 73.6% of published prognostic models employed univariable analysis for predictor selection. Using the prediction-model risk-of-bias assessment tool, that review reported that this analytical strategy can introduce systematic error and compromise internal validity and is classified as having a high risk of bias in the analysis domain [13]. However, that review did not elucidate the pathway through which such bias arises in small-sample, sparse cohorts. The present

study, which is based on a real clinical cohort, concretely demonstrates how this screening procedure omits key confounders and induces spurious statistical associations. The dual-criterion variable selection strategy was specifically designed to address this limitation: it simultaneously integrates evidence from baseline between-group comparisons and univariable regression, retaining any variable that satisfies either between-group comparison $P < 0.05$ or univariable regression $P < 0.1$. After incorporating SDH into Model 2, the significant association of age disappeared, DBP remained significant, and NEUT transitioned from nonsignificant to significant. This correction was achieved with only one additional variable and did not increase the model burden, supporting the specificity rather than the breadth of the dual-criterion approach. These results suggest that the age effect observed in Model 1 was not a causal association but rather a composite manifestation of indirect effects and confounding bias mediated jointly by SDH, DBP, NEUT, and other covariates. In summary, the methodological contribution of this study lies in systematically capturing confounding signals arising from baseline distribution imbalances during the variable selection stage, thereby addressing the inherent limitations of traditional variable selection strategies in small-sample studies, which are prone to insufficient power and the omission of important confounding variables.

Model Performance and Threshold Selection

Model 2 and Model 3 differed only in the variable reduction strategies. Although Model 3 (AIC = 78.38) showed better fit than Model 2 (AIC = 88.94), it relied entirely on the data-driven stepwise AIC algorithm, which tends to produce unstable variable combinations and weak generalizability in small samples. Model 3 was therefore used only as an auxiliary check rather than the primary model. Notably, both models consistently retained the two core independent predictors, DBP and NEUT, confirming the robustness of the dual-criterion screening results.

Ridge regression, rather than Lasso, was chosen as the sensitivity analysis tool. Lasso performs L1-based premodeling variable screening that removes weak-effect variables before they can enter the multivariable model, a process that could eliminate important confounders such as SDH and is therefore incompatible with our study's primary goal of maximizing confounder inclusion. In contrast, Ridge regression applies L2 shrinkage after the variable set has been determined, retaining all the candidate variables while only shrinking their coefficients. The Ridge was therefore used as a postestimation sensitivity check rather than a premodeling filter, allowing us to assess the stability of our findings without prematurely discarding potential confounders.

The maximally-adjusted model, which included all the candidate variables from Model 2, was also constructed as a reference. It achieved an apparent AUC of 0.943, representing the theoretical upper limit of predictive performance for this dataset after comprehensive confounder adjustment. However, this model has insufficient EPV under small-sample conditions and carries a potential risk of overfitting. It was therefore used only for performance comparison, not for clinical application. Two independently significant predictors screened from Model 2 were selected to establish the parsimonious model. Although its apparent AUC (0.712) was lower than that of the maximally-adjusted model, bootstrap internal validation yielded an optimism-corrected AUC of 0.693 and an MAE of 0.033, indicating

acceptable discrimination and calibration. The parsimonious model was not intended as a standalone clinical tool but rather as an application-level validation of the dual-criterion screening strategy to test whether the selected predictors retained meaningful predictive signals after confounding adjustment. In a small-sample cohort of 88 cases, a corrected AUC close to 0.7 with only two variables, both routinely available and mechanistically supported, was achieved, suggesting that the dual-criterion strategy successfully identified clinically meaningful predictors while avoiding the overfitting inherent in more complex models. This suggests that the parsimonious model is a transparent and interpretable demonstration that the screening strategy outputs have practical predictive utility. The performance discrepancy between these two models illustrates the inherent trade-off between comprehensive confounder adjustment and clinical feasibility in small-sample settings, a reflection of the bias-variance trade-off. Compared with the complex maximally-adjusted model, the parsimonious model contains only two easily accessible variables and is more suitable for bedside risk assessment. The parsimonious model achieved a recall rate of 86.54% at the unified 0.5 cutoff, meaning that fewer than 14% of truly high-risk patients were overlooked during screening for PPIs. Given the severe consequences of PPI after DC and the risk of irreversible progression if the early intervention window is missed, a higher false-positive rate is clinically acceptable to minimize missed diagnoses at the screening stage. However, this threshold yielded a low specificity (0.417), generating numerous false positives. To balance sensitivity and specificity for formal risk stratification, we calculated the Youden index, which corresponded to an optimal cutoff of 0.586. At this threshold, the sensitivity was 73.08%, and the specificity was 63.89%. The 0.5 threshold thus fits the demands of perioperative emergency screening, whereas the Youden-derived cutoff is more suitable for individualized long-term risk stratification. The MAE of 0.033 further indicates that the model yields well-calibrated predicted probabilities across risk strata, supporting its potential utility in clinical risk communication pending external validation.

The random forest model built in this study achieved an AUC of only 0.505 after LOOCV, with a predictive performance close to random guessing. This contrasted with the stable performance of logistic regression in this small-sample setting. The main reason is that the random forest requires a large sample size and high signal-to-noise ratio to perform well. With only 88 cases, the feature space became extremely sparse, and the split nodes of each tree lacked sufficient sample support, making them vulnerable to random noise and leading to poor generalizability [14]. Even when LOOCV was adopted for our random forest model, it failed to bridge the fundamental gap between model capacity and the limited effective information of our cohort. In contrast, Ridge regression effectively suppressed variance inflation in the high-dimensional variable space, reducing the risk of overfitting without excessively sacrificing bias. These empirical findings reinforce the practical principle that, for small-sample highly confounded datasets, controlling confounding factors prior to modeling and prioritizing interpretable regularized generalized linear models is more appropriate than directly applying complex black-box algorithms. Similar findings have been reported in other clinical contexts, where regression models with appropriate variable selection outperformed machine learning alternatives [15]. Importantly, this does not imply that machine learning approaches are inferior. Rather, their advantages, particularly in capturing complex interactions and achieving high predictive accuracy, are contingent on a sufficient sample size

and signal-to-noise ratio. In small-sample settings, the choice of modeling strategy should prioritize confounder control and interpretability over algorithmic complexity. This observation aligns with the broader methodological principle that data availability should inform model selection, rather than assuming that more complex algorithms are universally superior.

Mechanistic interpretability of the identified risk factors

After we identified independent risk factors, we observed that each 1 mmHg increase in DBP was associated with a 6.8% relative increase in the odds of PPI following DC (OR = 1.068, P = 0.035). The underlying pathophysiological mechanism of this association has not been fully clarified, and the prevailing hypothesis centers on the overactivated sympathetic nervous system secondary to TBI. CCL directly or indirectly stimulates the sympatho-adrenomedullary axis, leading to massive release of catecholamines into the circulation, which further induces marked systemic vasoconstriction and redistribution of the volume load. Given its characteristics of low pressure and high capacity, pulmonary circulation is particularly sensitive to increased venous return. The clinical presentation corresponds to the prodromal phase of neurogenic pulmonary edema, with elevated pulmonary capillary hydrostatic pressure and increased permeability of the alveolar-capillary barrier [16][17]. Within this microenvironment, protein-rich exudate in the alveolar space serves as an ideal culture medium for microbial colonization. Meanwhile, ventilation-perfusion mismatch impairs local host defense. Collectively, these alterations substantially increase susceptibility to pulmonary infection. Furthermore, elevated DBP may reflect heterogeneity in the magnitude of individual stress responses. Patients presenting more dramatic hemodynamic fluctuations after equivalent TBI may inherently remain in a hyperreactive state characterized by dysregulated immune and inflammatory modulation. Accordingly, DBP not only serves as a quantitative indicator of systemic circulation but also functions as a comprehensive clinical surrogate marker integrating the degree of neuroendocrine-immune axis dysregulation. The above-described pathophysiological cascade supports the mechanistic rationale for DBP as an independent predictive factor.

After correction via the dual-criterion inclusion strategy, NEUT was identified as another independent risk factor (OR = 1.214, P = 0.042). Each 1% increase in NEUT corresponded to a 21.4% increase in the odds of PPI. As first-line effector cells of innate immunity, elevated NEUT may not merely reflect a secondary response to established infection; instead, they could suggest a preactivated inflammatory state prior to the onset of infection. Massive tissue necrosis and the release of cellular contents induced by cerebral contusions generate abundant damage-associated molecular patterns (DAMPs). These mediators can activate circulating neutrophils via pattern recognition receptors (e.g., Toll-like receptors) and drive them into a functionally polarized state, characterized by an increased propensity to synthesize, prestore, and then release adhesion molecules, reactive oxygen species, and proteolytic enzymes[18][19]. When these preactivated neutrophils traverse the pulmonary circulation, the unique anatomical features of the pulmonary microvasculature (extensive capillary beds, sluggish blood flow, and high deformability requirements) render them prone to mechanical retention and transendothelial migration, triggering local degranulation and premature release of neutrophil extracellular traps (NETs), a pathological process

termed neutrophil extracellular trap formation. While NETs exert antibacterial effects, their excessive production can directly injure the alveolar epithelium and endothelium, disrupting the integrity of the alveolar-capillary barrier. The resultant elevated permeability further facilitates bacterial translocation and the establishment of PPIs. Thus, this mechanism supports that an elevated NEUT predicts new-onset postoperative pulmonary infection rather than reflecting a preexisting infection.

For neuroimmune crosstalk, the two independent risk factors identified in this study, DBP and NEUT, jointly point to the core pathophysiological cascade of the systemic inflammation-sympathetic storm following CCL. Elevated DBP suggests hyperactive central autonomic outflow induced by CCL. Specifically, compression or ischemic stimulation of the medullary vasomotor center triggers excessive release of norepinephrine and epinephrine from sympathetic nerve terminals and the adrenal medulla. These catecholamines not only mediate vasoconstriction in the systemic and pulmonary circulation via α_1 receptors but also directly regulate immune cell function: macrophages and dendritic cells express adrenergic receptors on their surface, and persistent catecholamine stimulation induces their polarization toward the M2-tolerogenic phenotype, impairing the capacity of innate immunity to eliminate pathogens [16]. Meanwhile, the elevation and preactivation of circulating NEUT represent a parallel arm of immune responses triggered by DAMPs released by injured tissues. PPI after CCL is not driven by a single pathway but arises from the intersection of two axes: neurogenic hemodynamic disturbance and immune inflammatory activation. Elevated DBP deteriorates the hydrodynamic microenvironment of pulmonary circulation, whereas immune preactivation provides cytotoxic effector molecules. Together, they create a “second hit” at the alveolar-capillary interface, jointly breaking the local defense threshold. This systemic perspective is consistent with previous findings that clinical parameters reflecting overall host status often outweigh local immune features in determining postoperative outcomes [20][21]. This explains why several conventional infectious biomarkers, such as C-reactive protein and procalcitonin, failed to emerge as independent predictors in our analysis. These markers predominantly reflect established infection or nonspecific acute-phase reactions triggered by surgical trauma rather than sustained latent signals of the preinfection susceptible state. In addition, surgical procedures alone can drastically increase C-reactive protein levels, severely diluting its discriminative power as a preoperative risk stratification tool. This mechanistic interpretation further underscores the methodological merit of the dual-criterion inclusion strategy in isolating indicators that truly reflect long-term pathophysiological susceptibility.

Limitations and Future Directions

Although this study established a robust analytical framework combining dual-criterion variable inclusion and hierarchical multimodel sensitivity analysis, several limitations should be noted.

First, the sample size (n = 88; 51 events) was sufficient for the planned multivariable modeling (EPV = 25.5), but it lacked external validation or extensive subgroup analyses, an issue that warrants attention in future larger-scale studies. The single-center retrospective design further limits generalizability. Full nested bootstrap with repeated variable selection was not performed for calibration assessment

because re-running variable screening within each bootstrap resampling would introduce random fluctuations in predictor selection and destabilize the final model. Instead, bootstrap calibration was performed on the fixed final predictor set, as reported in the Results. Ridge regression was used to mitigate multicollinearity-induced parameter fluctuations, but it cannot eliminate all potential selection bias from the preliminary screening stage.

Second, the retrospective design cannot fully eliminate residual confounding factors. Although the dual-criterion strategy expanded the candidate variable pool to correct baseline imbalance, unrecorded confounders, such as preoperative antibiotic use, intraoperative details, and postoperative airway management, may persist. Moreover, this study included only patients receiving DC for CCL, and the applicability of this framework to other surgical populations remains to be tested.

Third, while we assumed linear relationships for DBP and NEUT based on the observed data distribution, we did not formally test for nonlinearity using restricted cubic splines. Given the small sample size, such an analysis would have limited statistical power, but it could be considered in future external validation cohorts. Similarly, we did not perform decision curve analysis to assess clinical net benefit, as decision curve analysis requires stable probability stratification, which is difficult to achieve with 88 cases and would likely produce unstable estimates; this could be revisited if the model is validated in larger independent datasets.

Despite these limitations, the multi-model comparison provided solid internal evidence for the robustness of the proposed small-sample analytical strategy. Model 2 consistently outperformed Model 1 in confounding adjustment, and Model 3 confirmed the stability of the key predictors. These layered sensitivity analyses, although conducted within a single dataset, further support the methodological validity and targeted superiority of the dual-criterion approach for heavily confounded small-sample surgical cohorts.

Conclusion

This study constructed a comprehensive analytical framework for small-sample surgical cohorts with severe confounding, centered on a dual-criterion variable inclusion strategy. Multimodel sensitivity analyses, covering Model 1, Model 2, and Model 3 with Ridge regression verification, evaluated the stability of the variable screening and confounding adjustment. Based on Model 2, we established a maximally-adjusted model and a parsimonious model, with random forest as a comparator. Two cutoffs were set for the parsimonious model: the universal 0.5 threshold for cross-model comparison and rapid screening and the Youden-derived 0.586 threshold for individualized risk stratification.

The dual-criterion strategy improves confounder inclusion, avoids omission bias from single-threshold screening, and identifies predictors that are both mechanistically supported and clinically accessible. The resulting parsimonious model demonstrated acceptable discrimination and calibration, confirming that the screening strategy captures meaningful predictive signals. This analytical pipeline offers a

structured and replicable template for variable selection and confounding control in small-sample clinical studies.

Further validation in multi-center prospective cohorts is required to assess the generalizability of this framework.

Abbreviations

AIC, Akaike information criterion;

AUC, area under the curve;

CCL, cerebral contusion and laceration;

95% CI, 95% confidence interval;

DAMPs, damage-associated molecular patterns;

DBP, admission diastolic blood pressure;

DC, decompressive craniectomy;

EPV, events-per-variable;

Lasso, least absolute shrinkage and selection operator;

LOOCV, leave-one-out cross-validation;

MAE, mean absolute error;

NETs, neutrophil extracellular traps;

NEUT, peripheral-blood neutrophil percentage;

NPV, negative predictive value;

PPV, positive predictive value;

Q-Q plot, quantile-quantile plot;

OR, odds ratio;

PPI, postoperative pulmonary infection;

PPV, positive predictive value;

ROC, receiver operating characteristic;

SDH, subdural hematoma;

TBI, traumatic brain injury.

Declarations

Ethics approval and consent to participate

The Institutional Review Board at The Second Hospital of Hebei Medical University approved this retrospective analysis of anonymized data and waived informed consent. The study adhered to the Declaration of Helsinki.

Clinical trial number: not applicable.

Consent for publication

Not applicable, as no individual-level identifiable participant data are presented within this manuscript.

Availability of data and materials

The datasets analyzed during the present study are not publicly available due to hospital-related patient-data protection regulations. The datasets are available from the corresponding author upon reasonable formal request.

Competing interests

The authors declare that they have no competing interests.

Funding

Supported by Hebei Natural Science Foundation (H202520660) and Hebei Provincial Health Commission Medical Research Project (20211355).

Authors' contributions

FYY: Conceptualization, Methodology, Software, Formal analysis, Validation, Visualization, Investigation, Data curation, Writing – original draft, Writing – review & editing.

SQD: Investigation, Data curation, Writing – review & editing.

XYL: Investigation, Data curation, Writing – review & editing.

YYZ, KT: Investigation, Writing – review & editing.

JPS: Supervision, Project administration, Funding acquisition, Resources, Writing – review & editing.

All authors read and approved the final manuscript.

Acknowledgments

We thank President Wu Yuming and Professor Teng Xu of Hebei Medical University for their valuable guidance and support.

References

1. JIRLOW U, HOSSAIN I, KORHONEN O, et al. Cerebral contusions - Pathomechanism, predictive factors for progression and historical and current management. Brain Spine. 2024. 10.1016/j.bas.2024.103329.
2. QURESHI M S S, SHAD Z S, SHOAIB F et al. Early Versus Late Tracheostomy After Decompressive Craniectomy. Cureus, 2018[2026-06-01]. https://www.cureus.com/articles/15699-early-versus-late-tracheostomy-after-decompressive-craniectomy. 10.7759/cureus.3699
3. KUMAR R, ZENIAN M S, MAENG T Y, et al. Patient Outcomes following Immediate Tracheostomy and Emergency Decompressive Craniectomy in the Same Setting. Int J Environ Res Public Health. 2022;19(23):15746. 10.3390/ijerph192315746.
4. ROBBA C, REBORA P. Incidence, Risk Factors, and Effects on Outcome of Ventilator-Associated Pneumonia in Patients With Traumatic Brain Injury. Chest. 2020;158(6):2292–303. 10.1016/j.chest.2020.06.064.
5. EVANS L, RHODES A, ALHAZZANI W, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock 2021. Crit Care Med. 2021;49(11):e1063–143. 10.1097/CCM.0000000000005337.
6. ZHANG H, WANG Y, QU M, et al. Neutrophil, neutrophil extracellular traps and endothelial cell dysfunction in sepsis. Clin Translational Med. 2023;13(1):e1170. 10.1002/ctm2.1170.
7. ŠINKOVEC H, HEINZE G, BLAGUS R, et al. To tune or not to tune, a case study of ridge logistic regression in small or sparse datasets. BMC Med Res Methodol. 2021;21(1):199. 10.1186/s12874-021-01374-y.
8. LIU X, CHEN B, ZHANG H, et al. Enhancing differentiation between unipolar and bipolar depression through integration of machine learning and electroencephalogram analysis. J Affect Disord. 2025;389:119599. 10.1016/j.jad.2025.119599.
9. DONG T, ORONTI I B SINHAS, et al. Enhancing Cardiovascular Risk Prediction: Development of an Advanced Xgboost Model with Hospital-Level Random Effects. Bioengineering. 2024;11(10):1039. 10.3390/bioengineering11101039.
10. LI K M, REPS J M, NISHIMURA A, et al. Transfer-learning on federated observational healthcare data for prediction models using Bayesian sparse logistic regression with informed priors. J Am Med Inform Assoc. 2026;33(2):409–23. 10.1093/jamia/ocaf146.

11. WANG N, YAO C, LUO C, et al. Integrated plasma and exosome long noncoding RNA profiling is promising for diagnosing non-small cell lung cancer. Clin Chem Lab Med (CCLM). 2023;61(12):2216–28. 10.1515/cclm-2023-0291.
12. ZHOU X, LIU M. Circulating small extracellular vesicles microRNAs plus CA-125 for treatment stratification in advanced ovarian cancer. J Translational Med. 2023;21(1):927. 10.1186/s12967-023-04774-4.
13. XU Z, LI F, XIN Y, et al. Prognostic risk prediction model for patients with acute exacerbation of chronic obstructive pulmonary disease (AECOPD): a systematic review and meta-analysis. Respir Res. 2024;25(1):410. 10.1186/s12931-024-03033-4.
14. BRIEUC M S O, WATERS C D, DRINAN D P, et al. A practical introduction to Random Forest for genetic association studies in ecology and evolution. Mol Ecol Resour. 2018;18(4):755–66. 10.1111/1755-0998.12773.
15. GUO Y, LI L, ZHENG K, et al. Development and validation of a survival prediction model for patients with advanced non-small cell lung cancer based on LASSO regression. Front Immunol. 2024;15:1431150. 10.3389/fimmu.2024.1431150.
16. ZIAKA M. Pathophysiology of acute lung injury in patients with acute brain injury: the triple-hit hypothesis. Crit Care. 2024;28(1):71. 10.1186/s13054-024-04855-w.
17. RODRÍGUEZ A, CLAVERIAS L, MARTÍN-LOECHES I, et al. Ventilator-Associated Pneumonia (VAP) in Neurocritical Patients: The Hidden Dialog of Brain and Infection. Biomedicines. 2025;13(12):3112. 10.3390/biomedicines13123112.
18. CASTELLANOS-MOLINA A, BRETHEAU F, BOISVERT A, Brain, et al. Behav Immun. 2024;122:583–95. 10.1016/j.bbi.2024.07.047.
19. LIU Y W, LI S, DAI SS. Neutrophils in traumatic brain injury (TBI): friend or foe? J Neuroinflamm. 2018;15(1):146. 10.1186/s12974-018-1173-x.
20. LIN H, HUA J, WANG Y, et al. Prognostic and predictive values of a multimodal nomogram incorporating tumor and peritumor morphology with immune status in resectable lung adenocarcinoma. J Immunother Cancer. 2025;13(3):e010723. 10.1136/jitc-2024-010723.
21. MAO Y, WANG A, GE X. PD-1/PD-L1 inhibitors monotherapy vs. combination therapy in elderly advanced NSCLC: a real-world study and nomogram for survival prognosis. BMC Pulmonary Med. 2025,;25(1):350. 10.1186/s12890-025-03791-x. :.

Figures

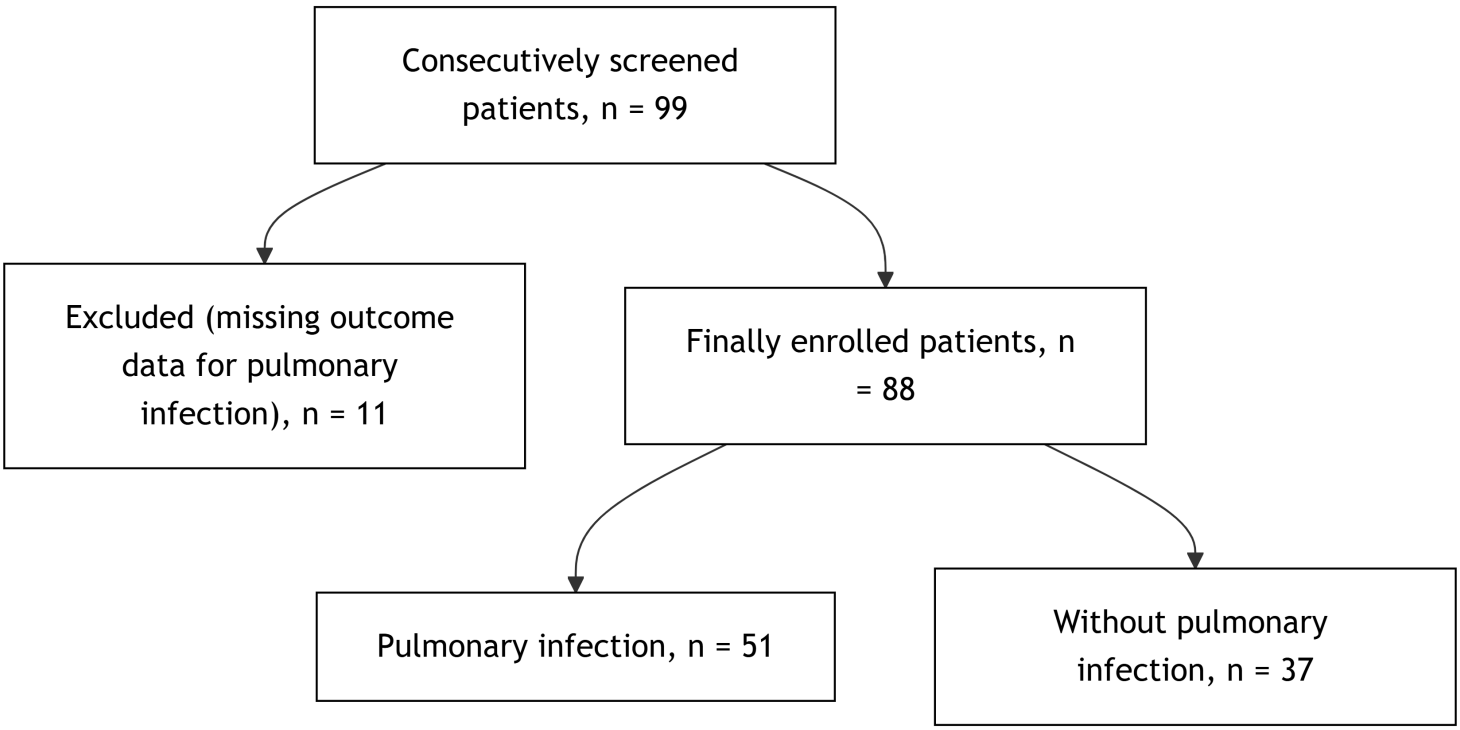


Figure 1

Patient enrollment and exclusion flowchart.

This flowchart shows the screening, enrollment, and exclusion procedures for the retrospective cohort of 99 patients with cerebral contusion and laceration who underwent decompressive craniectomy. After 11 patients with incomplete data on pulmonary infection were excluded, 88 patients were included in the final analysis.

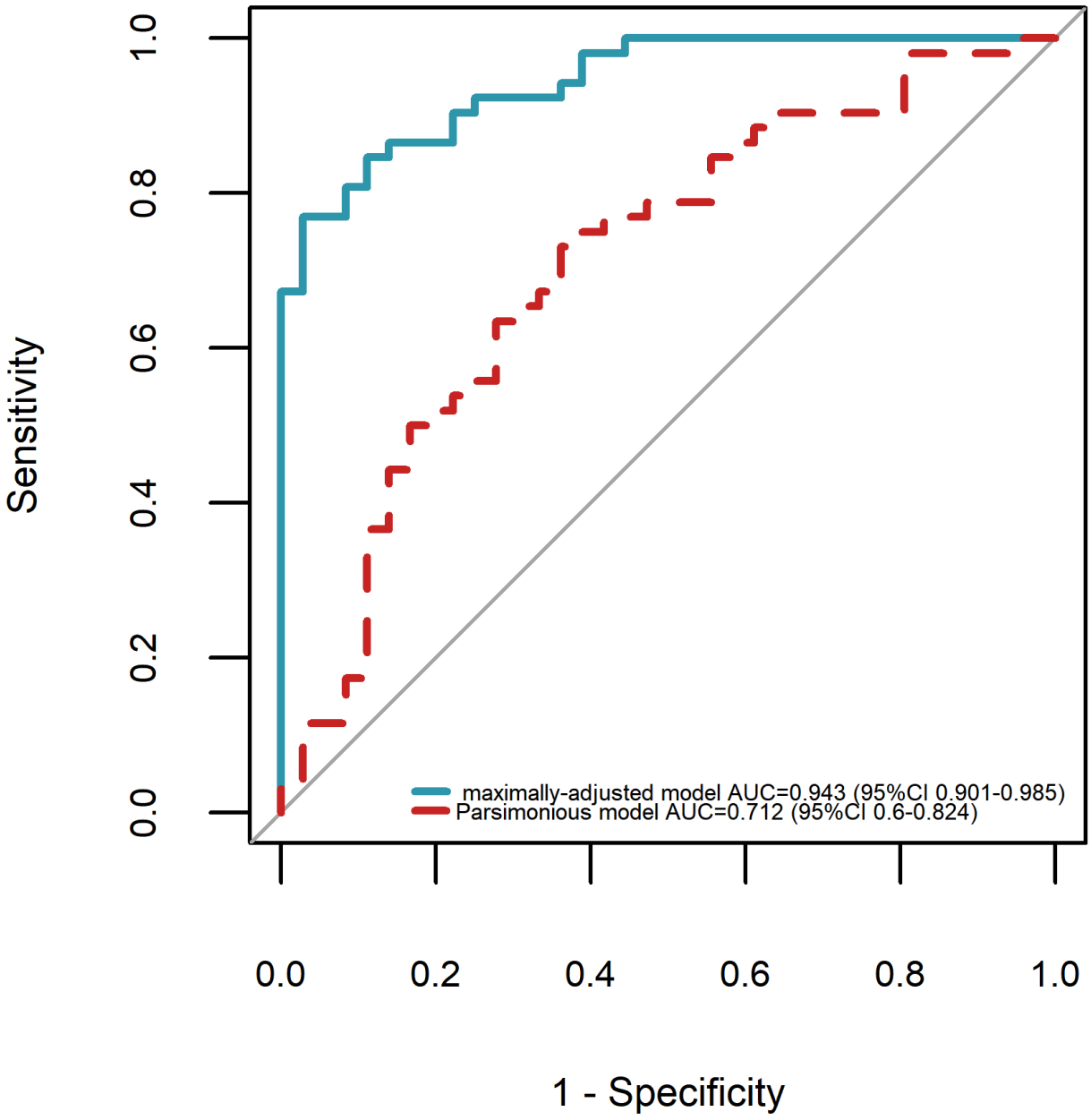


Figure 2

ROC curves comparing the maximally-adjusted model and the parsimonious model.

ROC curves comparing the discriminative performance of the maximally-adjusted model and the parsimonious model. Sensitivity is plotted against 1-specificity. The maximally-adjusted model incorporates all candidate variables identified by the dual-criterion strategy; the parsimonious model includes only DBP and NEUT. The maximally-adjusted model achieved an AUC of 0.943, while the parsimonious model achieved an AUC of 0.712 (95% CI, 0.600–0.824). The AUC values and confidence

intervals are displayed within the figure and detailed in the legend. DBP, admission diastolic blood pressure; NEUT, peripheral-blood neutrophil percentage; AUC, area under the curve.

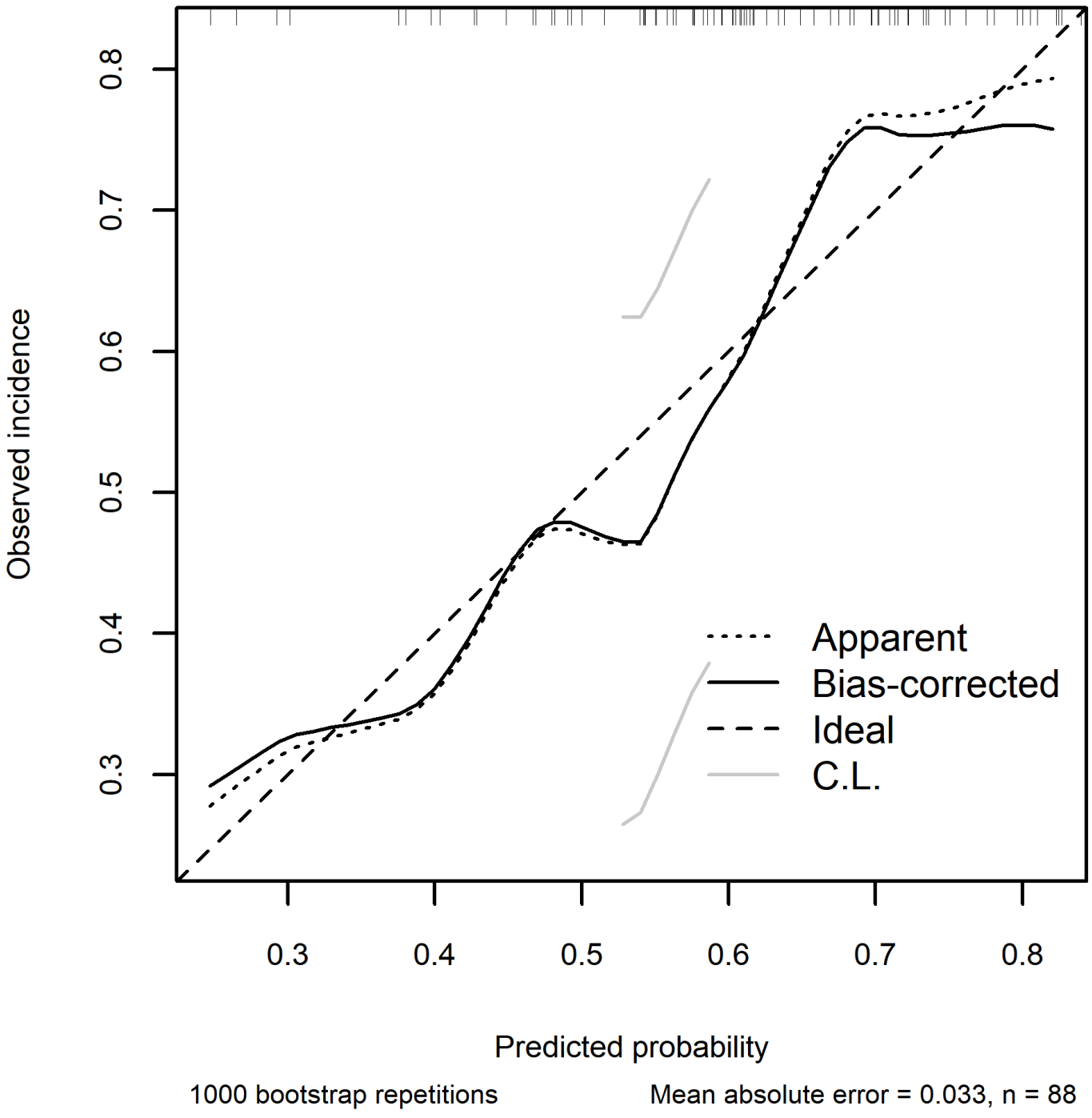


Figure 3

Calibration curve of the parsimonious model.

Bootstrap-corrected calibration curve for the parsimonious model based on 1000 resamples (n = 88). The solid line represents the bias-corrected calibration, and the dashed line represents the ideal

calibration (where the predicted probabilities equal the observed outcomes). The MAE is 0.033.

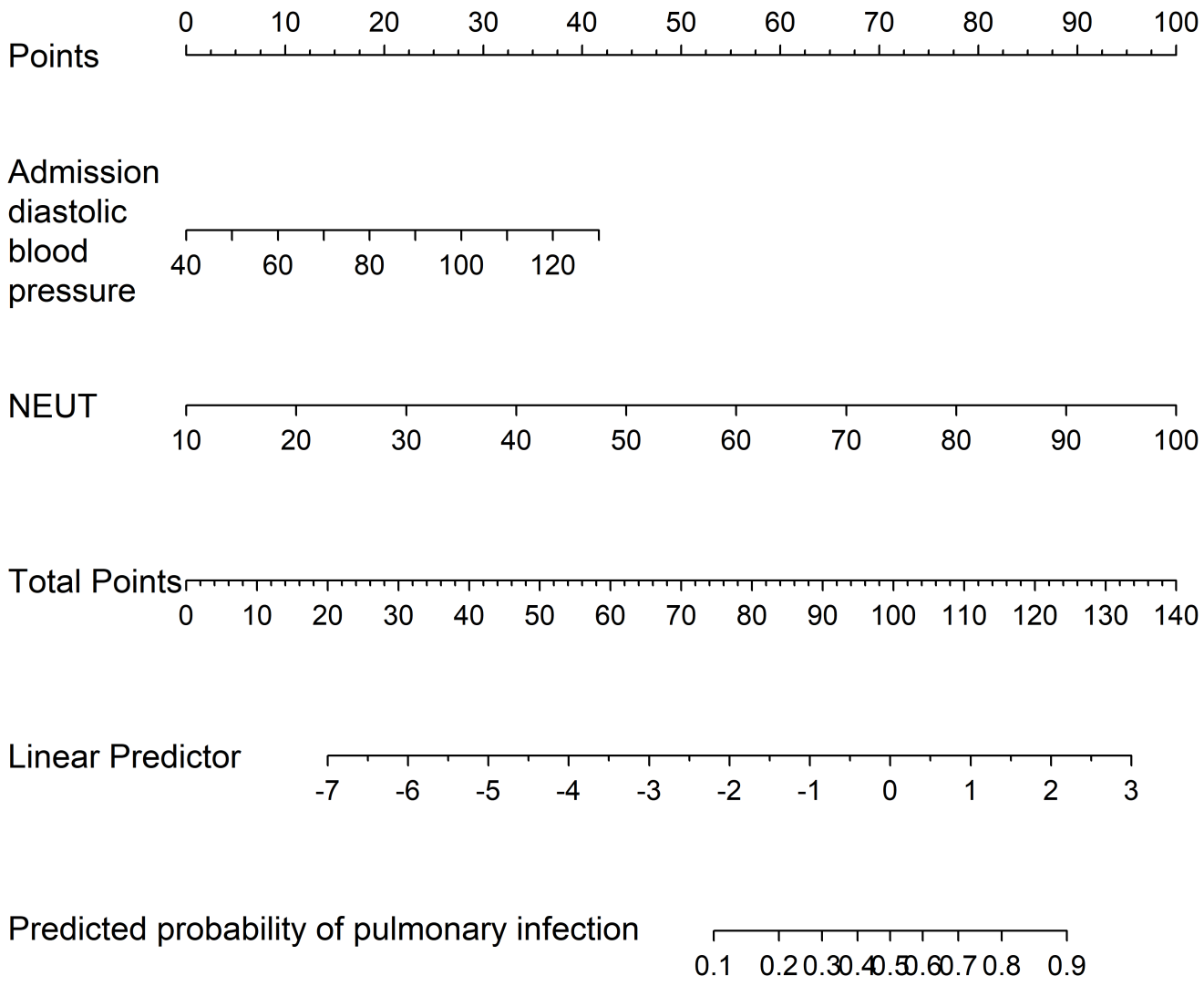


Figure 4

Nomogram for predicting postoperative pulmonary infection via the parsimonious model.

The nomogram visualizes the parsimonious prediction model incorporating DBP and NEUT. To estimate an individual patient's risk, the values of each predictor on their respective axis are located, a vertical line is drawn to the points axis to obtain a score for each variable, the scores are summed, the total score on the total points axis is located, and a vertical line to the risk axis is drawn to obtain the predicted probability of postoperative pulmonary infection.

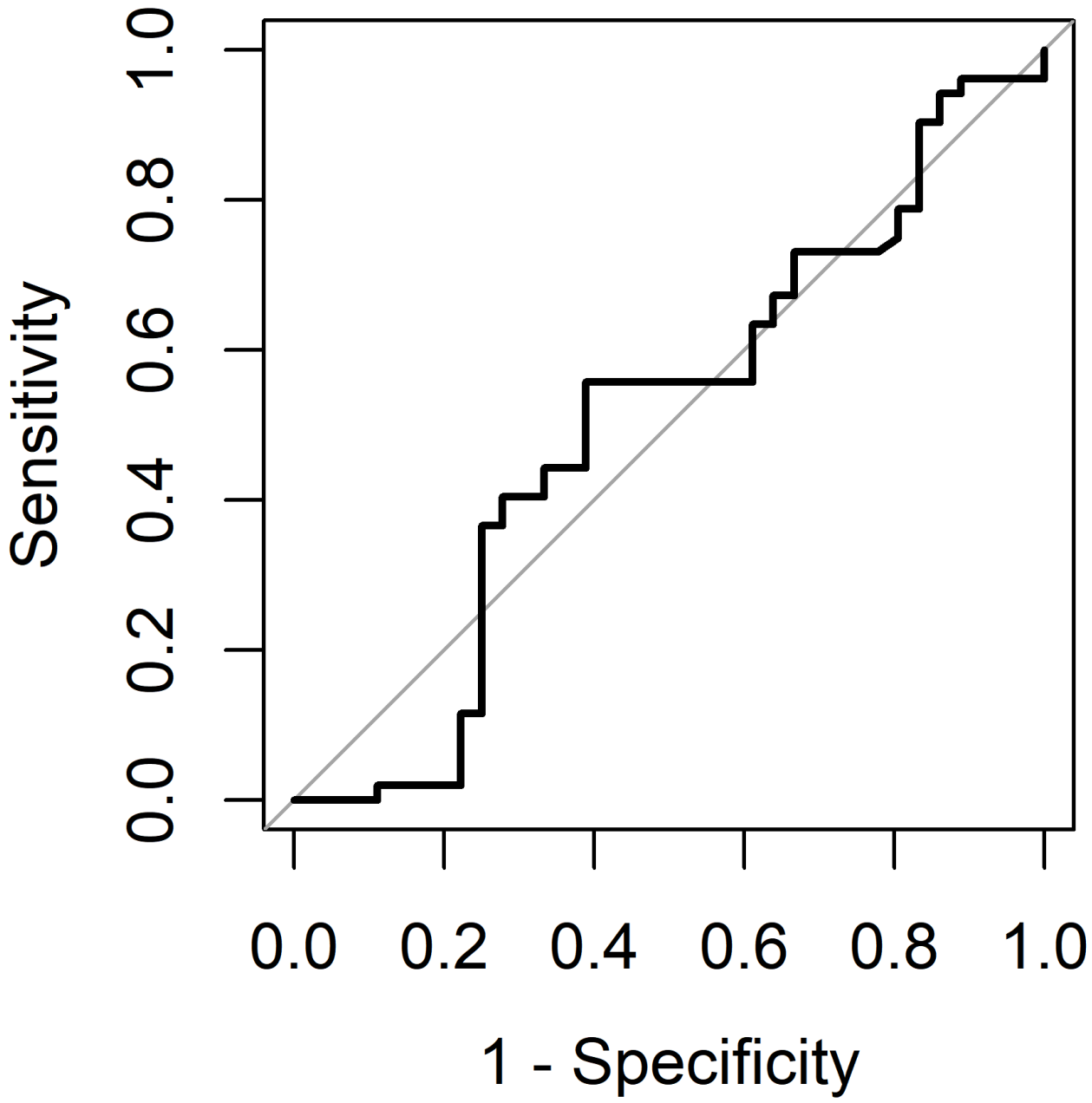


Figure 5

ROC curve of the random-forest model validated by leave-one-out cross-validation.

ROC curve for the random forest prediction model using the same two predictors (DBP and NEUT) as the parsimonious model. The model was validated by LOOCV, with sensitivity plotted against 1-specificity. The LOOCV AUC is 0.505