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.

# Development and Validation of a Nomogram for Predicting Postoperative ICU Admission in Patients Undergoing Emergency Abdominal Surgery: A Retrospective Cohort Study

**Chao Liu**
Second Central Hospital of Baoding
**Donghui Liu**
Second Central Hospital of Baoding
**Na Fan**
Second Central Hospital of Baoding
**Yongchang He**
Chengde Medical University
**Wangcheng Zhou**
Chengde Medical University
**Ruigang Feng**
fengruigang2021@163.com

Second Central Hospital of Baoding

**Research Article**

**Keywords:** acute abdomen, emergency surgery, ICU, nomogram, risk prediction, multiple imputation

**Posted Date:** November 23rd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11320812/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:

To construct and validate a nomogram for predicting postoperative ICU admission in patients undergoing emergency abdominal surgery.

# Methods:

This retrospective study used de-identified data from the Medical Informatics Operating Room Vitals and Events Repository (MOVER). Among 1,989 patients with acute abdomen-related discharge diagnoses, the primary analysis included the first eligible inpatient emergency abdominal surgery per patient (n = 1,557) to avoid clustering of repeated operations; all 3,699 inpatient episodes were analyzed in a sensitivity analysis using generalized estimating equations with the patient as the clustering unit. Data were split at the patient level into training (70%, n = 1,090) and validation (30%, n = 467) sets. Preoperative white blood cell count, hemoglobin and glucose, missing in approximately 64% of patients, were multiply imputed (m = 20) using models fitted in the training set only. Among 28 candidate predictors, LASSO regression with five-fold cross-validation selected variables retained across all imputed datasets, which entered a multivariable logistic regression model as 12 clinical predictors; a nomogram was constructed.

# Results:

In the primary cohort, 1,019 of 1,557 patients (65.4%) were admitted to the ICU; in the episode-level analysis, 2,815 of 3,699 (76.1%) involved ICU admission. The AUC was 0.818 (95% CI 0.792–0.844) in the training set and 0.770 (95% CI 0.726–0.814) in the validation set, with Brier scores of 0.160 and 0.180 (null model 0.227). At the optimal threshold of 0.548 (Youden index in the training set), validation sensitivity, specificity, PPV, NPV and accuracy were 82.9%, 52.1%, 76.4%, 62.0% and 72.2%. Calibration was good (slope 0.911; Hosmer-Lemeshow P = 0.302 and 0.415 in the training and validation sets), and decision curve analysis showed net benefit exceeding treat-all across thresholds of 0.2–0.5.

# Conclusions:

A 12-predictor nomogram based on readily available preoperative variables predicted postoperative ICU admission with good discrimination and adequate calibration, supporting preoperative risk evaluation and ICU resource allocation. External validation is required.

## Background

Emergency general surgery (EGS) poses a significant public health issue both in the United States and globally. A 10-year analysis of the Nationwide Inpatient Sample (2001–2010) identified more than 27 million EGS admissions, accounting for 7.1% of all hospitalizations in the United States; the population-adjusted admission rate reached 1,290 per 100,000 in 2010, exceeding the combined annual incidence of all new cancer diagnoses (650 per 100,000) [1]. Patients who need emergency abdominal surgery tend to be older, have more health issues, and face a much higher risk of complications and death than those having planned surgeries [2–4]. Emergency laparotomy has a 30-day mortality rate of 10% to 15% according to national audits [5], and unplanned admissions to critical care are frequently seen [6]. Identifying patients who might require ICU care after their operation is therefore essential for perioperative planning and ICU resource distribution [7].

Surgical risk scoring systems such as POSSUM, the ACS NSQIP Surgical Risk Calculator and the Surgical Outcome Risk Tool were developed mainly for surgical audit and the estimation of mortality or overall morbidity [8–12]. They are not designed to forecast postoperative ICU admission, and several depend on intraoperative variables that cannot be known before surgery.

First described in the 1940s, the ASA physical status classification was validated statistically in over 68,000 cases by Vacanti et al [13] and remains a leading tool for preoperative risk assessment. A meta-analysis of 77 studies with 165,705 patients confirmed that ASA classification moderately predicts postoperative mortality (AUC = 0.736) [14]. However, ASA classification alone does not capture disease-specific factors such as hemoglobin concentration, the extent of inflammation, or the complexity of the procedure. A systematic review of 44 studies on risk factors for unanticipated critical care admissions identified age, anemia, ASA classification, comorbidity burden, emergency operations, and duration of surgery as the most consistent independent predictors [7].

Recent models have begun to address ICU admission after surgery specifically. Xu and colleagues built a nomogram for non-cardiac surgery using VitalDB data, reaching an AUC of 0.917 [15]. Wang and colleagues developed a model for ICU admission after colorectal cancer surgery, achieving an AUC of 0.865 [16]. Martin et al. constructed a nomogram for surgical ICU readmission with an AUC of 0.71 [17], and nomogram approaches are increasingly used across surgical fields for tailored risk prediction [18]. However, these models either relied on intraoperative factors, focused on particular surgical populations, or were not developed for the broad emergency abdominal surgery group. Penalised regression methods such as LASSO are now used frequently for variable selection in clinical prediction models and reduce overfitting compared with stepwise selection [19].

Anemia before surgery is common in emergency abdominal operations and is linked to poorer outcomes after surgery. The PREVENTT randomized trial found that a substantial proportion of patients undergoing major abdominal surgery had preoperative anemia [20], and observational studies have associated anemia with more postoperative complications, increased transfusion needs, longer hospital stays, and unplanned ICU admissions [21–23]. A higher white blood cell count, indicating systemic inflammation, is likewise linked to poorer results after emergency abdominal surgery [24].

Despite this evidence, a nomogram for forecasting ICU admission after emergency abdominal surgery using only preoperative data is still lacking. This research aimed to design and internally validate such a nomogram from readily available preoperative variables.

## Methods

# Study Population and Cohort Construction

This was a retrospective cohort study based on de-identified data from the Medical Informatics Operating Room Vitals and Events Repository (MOVER), a publicly available perioperative database (EPIC cohort, University of California Irvine Medical Center; January 2018 to the date of data extraction) [25]. The workflow for patient selection is depicted in Fig. 1. Patients with acute abdomen-related principal discharge diagnoses were identified by keyword matching, including acute appendicitis, intestinal obstruction, acute cholecystitis, acute cholangitis, acute pancreatitis, gastrointestinal perforation, acute peritonitis, incarcerated or strangulated hernia, ruptured ectopic pregnancy, mesenteric ischemia, diverticulitis with perforation, volvulus, intussusception, intra-abdominal abscess, and peptic ulcer bleeding, yielding 1,989 unique patients. Non-abdominal emergencies, symptom-free biliary stones, uncomplicated ulcers, and chronic cholecystitis or pancreatitis were excluded. After matching to the operative record module, 432 patients without a matched inpatient surgery record were excluded, and 516 outpatient surgery records were removed, leaving 3,699 inpatient surgical episodes from 1,557 patients. Because repeated operations within the same patient are not independent observations, the primary analysis included only the first eligible inpatient emergency abdominal surgery per patient (n = 1,557), of whom 1,019 (65.4%) were admitted to the ICU postoperatively and 538 (34.6%) were not. All 3,699 inpatient episodes, of which 2,815 (76.1%) involved ICU admission, were retained for the sensitivity analysis.

## Variable Definitions and Data Preprocessing

The main outcome was whether the patient was admitted to the ICU after surgery (ICU_ADMIN_FLAG = Yes). Potential predictors were: (1) demographics: age, sex, and BMI (weight was converted from ounces to kilograms using the conversion factor 1 oz = 0.02835 kg); (2) surgical considerations, encompassing ASA physical status, the duration of the surgery in hours, and the specific abdominal subtype; (3) laboratory values before surgery: the most recent measurements prior to incision; WBC, hemoglobin and glucose were retained for the primary analysis with multiple imputation, whereas the remaining laboratory variables (platelets, neutrophils, lymphocytes, NLR, creatinine, urea, sodium, potassium, bicarbonate, albumin, total bilirubin, AST, ALT, D-lactate, INR) were excluded because of > 60% missing data; (4) medical conditions present: hypertension, diabetes, coronary artery disease, heart failure, COPD, chronic kidney disease, cirrhosis, cerebrovascular disease, and atrial fibrillation; and (5) the dataset was divided at the patient level into a training set comprising 70% (n = 1,090 patients) and a

validation set comprising 30% (n = 467 patients), using a fixed random seed of 42, so that all surgical episodes of the same patient remained in the same set.

## Missing Data Handling

In the primary cohort (the first surgery per patient), demographic and surgical variables were nearly complete (missing rates 0.1%-10.8%; BMI 6.5%). Preoperative laboratory values at the time of the first surgery were missing in 63.8% of patients (WBC 63.8%, hemoglobin 63.5%, glucose 63.5%). We retained WBC, hemoglobin and glucose by multiple imputation by chained equations (MICE) instead of deleting them, using age, BMI, operation duration and the ICU outcome as auxiliary variables [19]. Missing laboratory values in the training set were imputed with MICE using predictive mean matching (PMM, m = 20); missing laboratory values in the validation set were completed with the corresponding training-set medians to prevent any leakage of validation-set information into the imputation or model-fitting steps. The remaining laboratory variables were excluded because more than 60% of their values were missing. For the small proportions of missing demographic and surgical continuous variables (0.1%-10.8%), training-set medians were used in both sets, and missing categorical values were coded as a separate 'unknown' category. In the sensitivity analysis of all inpatient episodes, laboratory values were missing in approximately 51% of episodes and the same MICE procedure was applied.

# Statistical Analysis

Candidate predictors comprised 28 variables: demographics (age, sex, BMI), ASA physical status class, acute abdomen subtype, nine comorbidities, and WBC, hemoglobin and glucose. Anesthesia type was not entered into the primary candidate set because of extreme class imbalance (12 of 1,557 patients) and was evaluated only in a sensitivity analysis. Variables were screened by LASSO logistic regression with five-fold cross-validation; the penalization level was selected by the 1 standard-error (1SE) rule, and variables selected in all twenty imputed datasets were retained for the final model. Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated from multivariable logistic regression fitted in the training set and pooled across the twenty imputed datasets using Rubin's rules, with ORs expressed per clinically meaningful unit (BMI per 1 kg/m², operation duration per 1 hour, WBC per $1\times10^9$/L, hemoglobin per 1 g/dL, glucose per 1 mmol/L). The ASA reference category was ASA class I-II combined (ASA class V was combined with class IV); the acute abdomen subtype reference was the aggregated 'other subtypes' group. Discrimination was assessed by the area under the ROC curve (AUC), calibration by calibration curves, the Hosmer-Lemeshow test, calibration slope and calibration-in-the-large, and overall accuracy by the Brier score (with the null-model Brier score as a reference). The optimal threshold was determined by maximizing the Youden index in the training set; sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy were calculated in both sets at this threshold. Decision curve analysis (DCA) was used to evaluate clinical utility, with net benefit quantified at thresholds of 0.2, 0.3, 0.4 and 0.5. Three model-level sensitivity analyses were performed: (S1) adding anesthesia type to the candidate set; (S2) excluding all laboratory variables; (S3) using intestinal

obstruction as the subtype reference. The non-independence of repeated operations was addressed in a GEE sensitivity analysis of all 3,699 inpatient episodes with an exchangeable correlation structure and robust standard errors, clustered by patient. Analyses were performed with R software (glmnet, rms, pROC, ResourceSelection, mice, geepack packages); two-sided $P < 0.05$ was considered statistically significant.

## Ethics Approval

The MOVER database includes only data that has been de-identified, with all patient identifiers removed prior to public release in compliance with relevant privacy laws [25]. Because this study was a secondary analysis of a publicly available, de-identified dataset, institutional review board approval and informed consent were not required. The data were accessed and analyzed under the repository's data usage agreement.

## Results

# Baseline Characteristics

The primary cohort comprised 1,557 patients who underwent their first eligible emergency abdominal surgery; 1,019 (65.4%) were admitted to the ICU after surgery, and 538 (34.6%) were not. The median age of the group was 54 years (IQR 37–66), and males made up 52.4% of the population. The ICU group consisted of a greater percentage of patients classified as ASA III or higher, had a shorter operation duration distribution toward lower BMI values, and included fewer patients with gynecologic emergencies or acute appendicitis. Table 1 provides a summary of the baseline characteristics. Full details of all candidate variables are provided in Supplementary Table S1.

Table 1
Baseline characteristics of the primary analysis cohort (1,557 patients, first eligible emergency abdominal surgery per patient), stratified by ICU admission

| Variable | All (n = 1,557) | ICU (n = 1,019) | Non-ICU (n = 538) | P |
|---|---|---|---|---|
| Age (years) | 54.0 (37.0–66.0) | 58.0 (44.0–69.0) | 44.0 (31.0–57.0) | < 0.001 |
| BMI (kg/m²) | 26.9 (23.0-31.3) | 26.5 (22.8–31.0) | 27.6 (23.4–32.2) | 0.001 |
| Operation duration (h) | 2.5 (1.6–3.9) | 2.8 (1.6–4.5) | 2.1 (1.6–3.2) | < 0.001 |
| WBC (×10⁹/L) | 10.3 (7.3–13.8) | 10.1 (7.2–14.2) | 10.4 (7.6–13.2) | 0.952 |
| Hemoglobin (g/dL) | 12.5 (10.5–14.0) | 11.8 (9.4–13.7) | 13.2 (11.9–14.4) | < 0.001 |
| Glucose (mmol/L) | 6.3 (5.4–7.6) | 6.6 (5.6–8.2) | 6.0 (5.3–7.1) | < 0.001 |
| Male sex | 816 (52.4%) | 568 (55.7%) | 248 (46.1%) | < 0.001 |
| ASA I-II | 550 (35.3%) | 224 (22.0%) | 326 (60.6%) | < 0.001 |
| ASA III | 659 (42.3%) | 491 (48.2%) | 168 (31.2%) | < 0.001 |
| ASA IV (incl. V) | 224 (14.4%) | 211 (20.7%) | 13 (2.42%) | < 0.001 |
| ASA unknown | 124 (7.96%) | 93 (9.13%) | 31 (5.76%) | 0.026 |
| Subtype: Other | 1274 (81.8%) | 935 (91.8%) | 339 (63.0%) | < 0.001 |
| Subtype: Gynecologic | 44 (2.83%) | 5 (0.491%) | 39 (7.25%) | < 0.001 |
| Subtype: Appendicitis | 239 (15.4%) | 79 (7.75%) | 160 (29.7%) | < 0.001 |
| Hypertension | 344 (22.1%) | 274 (26.9%) | 70 (13.0%) | < 0.001 |
| Chronic kidney disease | 162 (10.4%) | 143 (14.0%) | 19 (3.53%) | < 0.001 |
| Cerebrovascular disease | 402 (25.8%) | 317 (31.1%) | 85 (15.8%) | < 0.001 |

*Continuous variables are presented as median (IQR); categorical variables as n (%) within each column. P values: Wilcoxon rank-sum test (continuous) and chi-square test (categorical). ASA V (moribund) was combined with ASA IV. Laboratory values are based on observed measurements before imputation; WBC in ×10⁹/L, hemoglobin in g/dL, glucose in mmol/L.*

# Variable Selection

Of the 28 candidate predictors, LASSO regression with five-fold cross-validation and the 1SE rule retained 14 variable indicators, identical across all twenty imputed datasets; because categorical predictors (ASA classification and acute abdomen subtype) were expanded into indicator (dummy) variables, these correspond to 12 clinical predictors: age, BMI, operation duration, sex, WBC, hemoglobin,

glucose, ASA classification, acute abdomen subtype, hypertension, chronic kidney disease and cerebrovascular disease (Fig. 2). Anesthesia type was not selected into the primary model.

# Multivariable Logistic Regression

The multivariable logistic regression analysis (Table 2) showed that hemoglobin was inversely associated with ICU admission (OR = 0.81 per 1 g/dL, 95% CI 0.70–0.94, P = 0.004). Each 1-hour increase in operation duration increased the odds of ICU admission by 15% (OR = 1.15, 95% CI 1.05–1.25, P = 0.002). Relative to ASA class I-II, ASA class IV (OR = 6.26, 95% CI 2.97–13.21) and ASA class III (OR = 2.09, 95% CI 1.39–3.14) were associated with markedly higher risk, as was ASA unknown (OR = 2.37, 95% CI 1.23–4.57, P = 0.010). Among the acute abdominal subtypes, both gynecologic emergency (OR = 0.13, 95% CI 0.04–0.44) and acute appendicitis (OR = 0.51, 95% CI 0.31–0.84) were associated with significantly lower risk of ICU admission compared with the other-subtype reference group. BMI showed a modest inverse association (OR = 0.98 per 1 kg/m², 95% CI 0.96-1.00, P = 0.118). WBC (OR = 1.03 per $1\times10^9$/L, 95% CI 0.98–1.09, P = 0.216) was retained by LASSO but remained non-significant after multivariable adjustment; because it is clinically plausible and was consistently selected across all twenty imputed datasets, it was kept in the final model rather than over-pruning the penalised selection [19]. Preoperative glucose showed no significant association with ICU admission after multivariable adjustment (OR = 0.99 per 1 mmol/L, 95% CI 0.91–1.09, P = 0.874). Figure 3 shows the forest plot of the pooled model.

**Table 2. Multivariable logistic regression for postoperative ICU admission (training set, n = 1,090; estimates pooled over MICE m = 20 imputed datasets)**

| Variable | OR | 95% CI | P |
|---|---|---|---|
| Age (years) | 1.008 | 0.998–1.019 | 0.123 |
| BMI (kg/m2) | 0.981 | 0.958–1.005 | 0.118 |
| Operation duration (h) | 1.145 | 1.052–1.246 | 0.002 |
| Male sex | 1.303 | 0.886–1.917 | 0.179 |
| WBC (10^9/L) | 1.033 | 0.981–1.087 | 0.216 |
| Hemoglobin (g/dL) | 0.810 | 0.700–0.936 | 0.004 |
| Glucose (mmol/L) | 0.993 | 0.905–1.088 | 0.874 |
| ASA III (vs I-II) | 2.088 | 1.388–3.139 | <0.001 |
| ASA IV (vs I-II) | 6.263 | 2.970–13.207 | <0.001 |
| ASA unknown | 2.372 | 1.232–4.568 | 0.010 |
| Gynecologic (vs Other) | 0.125 | 0.035–0.444 | 0.001 |
| Appendicitis (vs Other) | 0.512 | 0.311–0.843 | 0.009 |
| Hypertension | 1.247 | 0.743–2.094 | 0.404 |
| Chronic kidney disease | 1.570 | 0.753–3.272 | 0.229 |
| Cerebrovascular disease | 1.511 | 0.970–2.353 | 0.068 |

**Model fit statistics**

| Statistic | Value |
|---|---|
| Likelihood ratio chi2 | 327.7 |
| df | 15 |
| Likelihood ratio P | <2e-16 |
| Nagelkerke R2 | 0.359 (range 0.330-0.391) |
| AIC | 1107.6 (range 1072.3-1137.5) |

*OR, odds ratio; CI, confidence interval. ASA was modeled as a categorical variable with I–II (including ASA V combined with IV) as the reference; subtype was modeled with other acute abdomen subtypes as the reference. Estimates pooled over MICE m = 20 imputed datasets. WBC in ×10⁹/L, hemoglobin in g/dL, operation duration in hours, BMI in kg/m², age in years, glucose in mmol/L. Model fit statistics are averaged across the 20 imputed datasets: likelihood ratio chi-square test of the full model against the null model; Nagelkerke pseudo-R² (range across imputations); AIC, Akaike information criterion (range across imputations). Brier score and Hosmer–Lemeshow calibration are reported in* Table 3.

# Model Performance

The nomogram incorporated the 12 clinical predictors (Fig. 4A). As shown in Table 3, the AUC was 0.818 (95% CI 0.792–0.844) in the training set and 0.770 (95% CI 0.726–0.814) in the validation set (Fig. 4D), with Brier scores of 0.160 and 0.180, respectively (null-model Brier score 0.227). The optimal threshold of 0.548 was determined by the Youden index in the training set; at this threshold, sensitivity, specificity and accuracy were 85.2%, 62.9% and 77.5% in the training set, and sensitivity, specificity, PPV, NPV and accuracy were 82.9%, 52.1%, 76.4%, 62.0% and 72.2% in the validation set. The Hosmer-Lemeshow test indicated good calibration (training $\chi^2 = 9.5$, P = 0.302; validation $\chi^2 = 8.2$, P = 0.415). The validation calibration slope was 0.911 and the calibration-in-the-large (intercept) was 0.031, indicating no systematic miscalibration; the calibration curve is shown in Fig. 4B. In DCA (Fig. 4C), the model provided net benefit exceeding treating all patients across thresholds of 0.2–0.5 (net benefit at 0.2/0.3/0.4/0.5: 0.566/0.514/0.448/0.375 vs. 0.564/0.501/0.418/0.302 for treating all), supporting clinical utility across the clinically relevant threshold range.

Table 3
Performance of the ICU admission prediction model in the training (n = 1,090) and validation (n = 467) sets

| Metric | Training set | Validation set |
|---|---|---|
| AUC (95% CI) | 0.818 (0.792–0.844) | 0.770 (0.726–0.814) |
| Optimal threshold (Youden, training set) | 0.548 | – |
| Sensitivity | 0.852 | 0.829 |
| Specificity | 0.629 | 0.521 |
| PPV | 0.814 | 0.764 |
| NPV | 0.69 | 0.62 |
| Accuracy | 0.775 | 0.722 |
| Brier score | 0.1602 | 0.1804 |
| Hosmer-Lemeshow $\chi^2$ (P) | 9.5 (0.302) | 8.19 (0.415) |
| Calibration slope (95% CI, validation set) | – | 0.911 (0.730–1.114) |
| Calibration intercept (95% CI, validation set) | – | 0.031 (− 0.232–0.279) |

*AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value. The optimal threshold was determined by the Youden index in the training set. Hosmer-Lemeshow test: P > 0.05 indicates good calibration. Calibration slope and intercept were estimated in the validation set with 95% CIs derived from 1,000 bootstrap resamples. Null-model Brier score: 0.227.*

# Sensitivity Analysis

Three model-level sensitivity analyses were performed. S1: adding anesthesia type to the candidate set yielded a validation AUC of 0.772, essentially unchanged from the primary model (0.770); epidural anesthesia was used in only 12 patients (0.8%) and its effect estimate was unstable (OR = 0.16, 95% CI 0.04–0.66), so it was not included in the primary model. S2: excluding all laboratory variables reduced the training AUC from 0.818 to 0.794, whereas the validation AUC was essentially unchanged (0.774 vs. 0.770). The modest increase in the validation AUC is likely attributable to estimation variability in the smaller validation sample (n = 467) rather than to a genuine loss of prognostic information; the consistent decline in the training set supports the incremental value of the imputed laboratory markers. S3: using intestinal obstruction as the subtype reference gave ORs for gynecologic emergency (0.11, 95% CI 0.03–0.42) and acute appendicitis (0.48, 95% CI 0.28–0.83), consistent with the primary analysis. In the GEE analysis of all 3,699 inpatient episodes clustered by patient, the directions of the effects were consistent with the primary analysis: operation duration (OR = 1.11 per 1 hour, 95% CI 1.07–1.15), ASA class IV (OR = 5.28, 95% CI 3.99–6.98), gynecologic emergency (OR = 0.14, 95% CI 0.05–0.38) and acute appendicitis (OR = 0.36, 95% CI 0.27–0.49) remained significantly associated with ICU admission, whereas BMI was not significantly associated (OR = 0.99, 95% CI 0.98–1.01, P = 0.501). When WBC and hemoglobin were added in a second GEE model among 1,814 episodes with complete laboratory data, both retained independent associations with ICU admission (WBC OR = 1.06 per $1\times10^9$/L, 95% CI 1.04–1.09; hemoglobin OR = 0.79 per 1 g/dL, 95% CI 0.75–0.85), corroborating the prognostic value of laboratory variables at the episode level.

# Discussion

We developed a 12-predictor nomogram for ICU admission after the first emergency abdominal surgery in 1,557 patients. The model performed consistently across all sensitivity analyses, including an episode-level GEE analysis of 3,699 operations, and decision curve analysis supported its clinical use. We are not aware of a previous model developed specifically to predict ICU admission after emergency abdominal surgery across a broad range of acute abdomen diagnoses using only preoperative variables. Although the Emergency Surgery Score (ESS) and the PIP score have been associated with ICU utilization in emergency laparotomy cohorts, the ESS was designed primarily to predict mortality and relies on nine laboratory variables that may not all be available in the emergency setting, while the PIP score incorporated intraoperative transfusion status and was developed in a small single-center cohort of 108 patients without external validation.

The ICU admission rate was 65.4% in the primary cohort and 76.1% across all inpatient episodes, considerably higher than the 15.6% reported by Xu et al. [15] and higher than usual rates for planned abdominal surgeries [23]. The high rate most likely reflects our exclusive focus on emergency abdominal surgery, where both disease severity [2, 3, 9, 26] and local ICU admission practice [27] differ from elective settings.

ASA class was the strongest predictor. Compared with ASA I-II, ASA IV carried a six-fold higher odds of ICU admission (OR 6.26), and ASA III a two-fold higher odds (OR 2.09). We also found that patients with unknown ASA had higher risk (OR 2.37). This is probably not a biological effect; rather, in the most urgent

cases the ASA assessment is often incomplete, and the missing label itself marks a sicker patient. The association between ASA class and postoperative mortality has been established since Vacanti et al.'s validation in 68,388 cases [13] and confirmed in a meta-analysis of 77 studies (AUC 0.736) [14]; ASA class was also a leading predictor of mortality after emergency laparotomy in the NELA risk model [5].

Each additional hour of surgery increased ICU admission odds by 15% (OR 1.15), in line with previous emergency laparotomy audits [5, 7].We initially expected obesity to increase the need for ICU care, but higher BMI was associated with lower risk of ICU admission (OR = 0.98 per 1 kg/m²). A more likely explanationis the 'obesity paradox' reported in several emergency and critical care populations: in acutely ill surgical patients, lower BMI could reflect chronic disease burden, frailty and reduced nutritional reserve. Because BMI was measured at emergency admission, fluid shifts and acute illness may also distort the measurement. The remaining predictors-age, male sex, hypertension, chronic kidney disease and cerebrovascular disease-were selected by LASSO but none reached statistical significance after multivariable adjustment, suggesting that their independent contribution to ICU admission is limited in this cohort. Epidural anesthesia was used in only 12 patients and was therefore not included in the primary model; its apparent protective effect in the sensitivity analysis (OR 0.16) most likely reflects selection of less severely ill patients.

The gynecologic emergency subtype carried notably lower risk of ICU admission (OR = 0.13) than the reference group. This probably indicates that these patients are generally younger and healthier, and that the procedures are shorter with less physiological stress [33]. Acute appendicitis also carried lower risk (OR = 0.51): appendicitis typically presents earlier with less severe physiological disturbance than perforation, obstruction or ischemia, so the surgery is simpler and shorter [32, 34].

Retaining laboratory variables despite high missingness was a key feature of the primary model. Lower preoperative hemoglobin was independently associated with ICU admission (OR = 0.81 per 1 g/dL), which agrees with randomized and observational evidence linking preoperative anemia to worse outcomes after major abdominal surgery [20–23, 28–31]. WBC and glucose illustrate an instructive contrast between the patient-level and episode-level analyses. Both were selected by LASSO in every imputed dataset but neither reached statistical significance after multivariable adjustment (WBC OR = 1.03, P = 0.216; glucose OR = 0.99, P = 0.874). A likely explanation is that at the patient level WBC correlates with other markers of severity and its signal is absorbed by ASA class and comorbidity, whereas in the episode-level GEE analysis-which uses all 3,699 operations-WBC became independently significant (OR = 1.06, 95% CI 1.04–1.09). This contrast suggests that laboratory values carry prognostic information beyond clinical variables alone, and supports routine preoperative blood tests in emergency abdominal surgery.

The training and validation AUCs (0.818 and 0.770) did not reach statistical significance (DeLong test P = 0.064; bootstrap 95% CI for the difference − 0.003 to 0.098). Given the limited size of the validation set (n = 467), this difference is borderline and should be interpreted cautiously; the modest drop in discrimination appears compatible with sampling variation. Our validation AUC of 0.770 is close to, but

somewhat lower than, the AUCs reported for previous ICU admission models (0.917 by Xu et al. [15] and 0.865 by Wang et al. [16]). The difference probably reflects three things. Those models included intraoperative variables such as crystalloid administration, transfusion and surgical approach, which are unavailable preoperatively and inherently improve discrimination; our model was deliberately restricted to preoperative variables to support preoperative triage. Our cohort also covers a wider range of disease severities and procedures, which increases outcome heterogeneity. Also, most patients are concentrated in the moderate-to-high predicted risk range (baseline event rate 65.4%), which narrows the spread of predicted probabilities and may modestly limit the ceiling of achievable discrimination; absolute risk estimates should therefore be recalibrated before use in centers with different ICU admission thresholds.

The Hosmer-Lemeshow test was not significant in either set (training P = 0.302; validation P = 0.415), the calibration slope was 0.911 (95% CI 0.730–1.114) and the calibration-in-the-large was 0.031 (95% CI -0.232-0.279), indicating that predicted probabilities closely matched observed event rates across the risk range (Fig. 4B). Because triage decisions depend on absolute risk, the good calibration we observed in the validation set is encouraging [35, 36]; it should nonetheless be re-examined in external cohorts. In practice, the nomogram can be completed from routine preoperative data within minutes: a predicted probability above the training-derived threshold (0.548) would support early ICU bed planning, and lower scores would allow ward-based monitoring.

Our study has limitations. It was retrospective and single-center, and we validated the model only internally; external validation in independent cohorts is needed before clinical use. The acute abdomen cohort is clinically heterogeneous, and although subtype categories were included as predictors, model performance within specific subtypes may vary [26]. Preoperative laboratory values were missing in 63.8% of patients; we used multiple imputation (m = 20), but residual bias cannot be fully excluded. For the validation set, missing laboratory values were completed with training-set medians rather than multiple imputation. This choice was necessary to avoid data leakage, but single-value imputation does not reflect imputation uncertainty and may therefore underestimate the uncertainty of validation-set predictions; a full multiple-imputation procedure should be applied in future external validation to quantify this effect.

We deliberately excluded intraoperative variables such as blood loss, fluid requirements and transfusion to keep the model preoperative, which may limit its performance [7, 16]. The optimal threshold was derived from the training set by the Youden index; it remains cohort-specific, and decision curve analysis should guide threshold selection in practice. Finally, the high ICU admission rate at our center (65.4%) may not generalize to centers with different ICU admission thresholds.

# Conclusions

We developed a preoperative nomogram for ICU admission after emergency abdominal surgery using routinely available variables. The model performed consistently across sensitivity analyses, including an

episode-level GEE analysis. External validation is now needed to determine whether it can support ICU triage in other centers.

## Declarations

# Clinical trial number

Not applicable.

# Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

# Author Contribution

RF: study conception and design, statistical analysis, manuscript drafting and revision. CL: study design, data curation, manuscript drafting. DL: clinical interpretation of intensive care outcomes, manuscript review and editing. YH and WZ: data collection and literature search. NF: data collection and administrative support. All authors read and approved the final manuscript.

## References

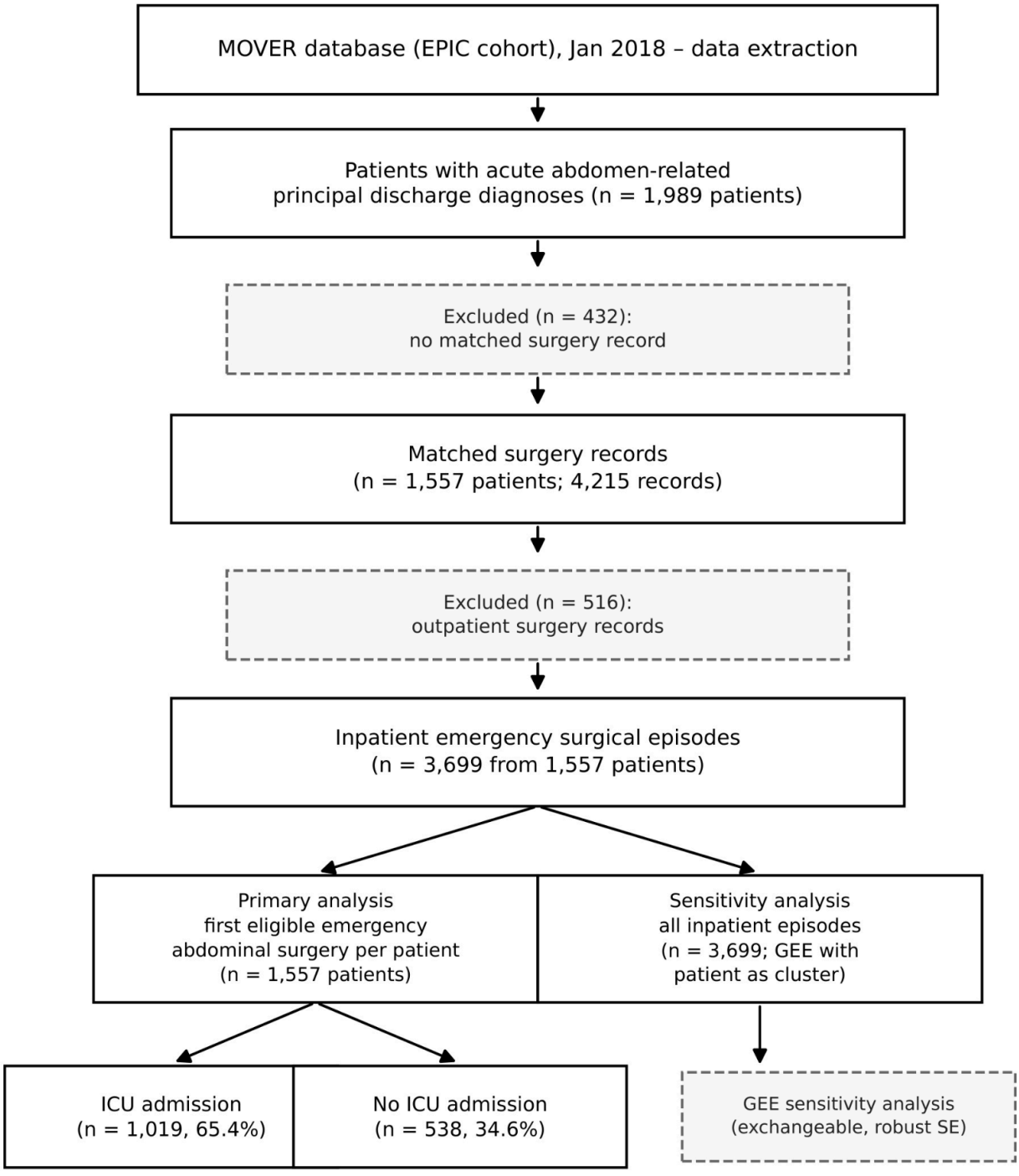
1. Gale SC, Shafi S, Dombrovskiy VY, Arumugam D, Crystal JS. The public health burden of emergency general surgery in the United States: A 10-year analysis of the Nationwide Inpatient Sample–2001 to 2010. J Trauma Acute Care Surg. 2014;77(2):202–8. 10.1097/TA.0000000000000362.
2. Havens JM, Columbus AB, Seshadri AJ, Brown CVR, Tominaga GT, Mowery NT, et al. Risk stratification tools in emergency general surgery. Trauma Surg Acute Care Open. 2018;3(1):e000160. 10.1136/tsaco-2017-000160.
3. Kokkinakis S, Kritsotakis EI, Lasithiotakis K. Artificial Intelligence in Surgical Risk Prediction. J Clin Med. 2023;12(12). 10.3390/jcm12124016.
4. Scotton G, La Greca A, Lirusso C, Mariani D, Zago M, Chiarugi M, et al. Can the American College of Surgeons NSQIP Surgical Risk Calculator Accurately Predict Adverse Postoperative Outcomes in Emergency Abdominal Surgery? An Italian Multicenter Analysis. J Am Coll Surg. 2023;236(2):387–98. 10.1097/XCS.0000000000000445.
5. Eugene N, Kuryba A, Martin P, Oliver CM, Berry M, Moppett IK, et al. Development and validation of a prognostic model for death 30 days after adult emergency laparotomy. Anaesthesia. 2023;78(10):1262–71. 10.1111/anae.16096.

6. Li L, He H, Xiang L, Wang Y. A warning model for predicting patient admissions to the intensive care unit (ICU) following surgery. Perioper Med (Lond). 2025;14(1):60. 10.1186/s13741-025-00544-6.
7. Onwochei DN, Fabes J, Walker D, Kumar G, Moonesinghe SR. Critical care after major surgery: a systematic review of risk factors for unplanned admission. Anaesthesia. 2020;75(Suppl 1):e62–74. 10.1111/anae.14793.
8. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg. 1991;78(3):355–60. 10.1002/bjs.1800780327.
9. Jones HJ, de Cossart L. Risk scoring in surgical patients. Br J Surg. 1999;86(2):149–57. 10.1046/j.1365-2168.1999.01006.x.
10. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg. 2013;217(5):833–42. e831-833.
11. Protopapa KL, Simpson JC, Smith NC, Moonesinghe SR. Development and validation of the Surgical Outcome Risk Tool (SORT). Br J Surg. 2014;101(13):1774–83. 10.1002/bjs.9638.
12. Oakland K, Cosentino D, Cross T, Bucknall C, Dorudi S, Walker D. External validation of the Surgical Outcome Risk Tool (SORT) in 3305 abdominal surgery patients in the independent sector in the UK. Perioper Med (Lond). 2021;10(1):4. 10.1186/s13741-020-00173-1.
13. Vacanti CJ, VanHouten RJ, Hill RC. A statistical analysis of the relationship of physical status to postoperative mortality in 68,388 cases. Anesth Analg. 1970;49(4):564–6.
14. Koo CY, Hyder JA, Wanderer JP, Eikermann M, Ramachandran SK. A meta-analysis of the predictive accuracy of postoperative mortality using the American Society of Anesthesiologists' physical status classification system. World J Surg. 2015;39(1):88–103. 10.1007/s00268-014-2783-9.
15. Xu Z, Yao S, Jiang Z, Hu L, Huang Z, Zeng Q, et al. Development and validation of a prediction model for postoperative intensive care unit admission in patients with non-cardiac surgery. Heart Lung. 2023;62:207–14. 10.1016/j.hrtlng.2023.08.001.
16. Wang L, Wu Y, Deng L, Tian X, Ma J. Construction and validation of a risk prediction model for postoperative ICU admission in patients with colorectal cancer: clinical prediction model study. BMC Anesthesiol. 2024;24(1):222. 10.1186/s12871-024-02598-3.
17. Martin LA, Kilpatrick JA, Al-Dulaimi R, Mone MC, Tonna JE, Barton RG, et al. Predicting ICU readmission among surgical ICU patients: Development and validation of a clinical nomogram. Surgery. 2019;165(2):373–80. 10.1016/j.surg.2018.06.053.
18. Eom BW, Joo J, Kim YW, Park B, Yoon HM, Ryu KW, et al. Nomogram Estimating the Probability of Intraabdominal Abscesses after Gastrectomy in Patients with Gastric Cancer. J Gastric Cancer. 2015;15(4):262–9. 10.5230/jgc.2015.15.4.262.
19. Pavlou M, Ambler G, Seaman S, De Iorio M, Omar RZ. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. Stat Med. 2016;35(7):1159–77. 10.1002/sim.6782.

20. Richards T, Baikady RR, Clevenger B, Butcher A, Abeysiri S, Chau M, et al. Preoperative intravenous iron to treat anaemia before major abdominal surgery (PREVENTT): a randomised, double-blind, controlled trial. Lancet. 2020;396(10259):1353–61. 10.1016/S0140-6736(20)31539-7.
21. Buhl H, Nørgaard A, Otkjaer A, Jørgensen LN, Jørgensen HL. The impact of anemia and blood transfusion on mortality after open abdominal surgery in the elderly. Langenbecks Arch Surg. 2023;408(1):421. 10.1007/s00423-023-03122-w.
22. Schack A, Ekeloef S, Ostrowski SR, Gögenur I, Burcharth J. Blood transfusion in major emergency abdominal surgery. Eur J Trauma Emerg Surg. 2022;48(1):121–31. 10.1007/s00068-020-01562-3.
23. Myles PS, Richards T, Klein A, Wood EM, Wallace S, Shulman MA, et al. Postoperative anaemia and patient-centred outcomes after major abdominal surgery: a retrospective cohort study. Br J Anaesth. 2022;129(3):346–54. 10.1016/j.bja.2022.06.014.
24. Falor AE, Zobel M, Kaji A, Neville A, De Virgilio C. Admission variables predictive of gangrenous cholecystitis. Am Surg. 2012;78(10):1075–8.
25. Samad M, Angel M, Rinehart J, Kanomata Y, Baldi P, Cannesson M. Medical Informatics Operating Room Vitals and Events Repository (MOVER): a public-access operating room database. JAMIA Open. 2023;6(4):ooad084. 10.1093/jamiaopen/ooad084.
26. Akinbami F, Askari R, Steinberg J, Panizales M, Rogers SO Jr.. Factors affecting morbidity in emergency general surgery. Am J Surg. 2011;201(4):456–62. 10.1016/j.amjsurg.2010.11.007.
27. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, et al. Mortality after surgery in Europe: a 7 day cohort study. Lancet. 2012;380(9847):1059–65. 10.1016/S0140-6736(12)61148-9.
28. Au Yong PS, Ke Y, Kok EJY, Tan BPY, Kadir HA, Abdullah HR. Preoperative anemia in older individuals undergoing major abdominal surgery is associated with early postoperative morbidity: a prospective observational study. Can J Anaesth. 2024;71(3):353–66. 10.1007/s12630-023-02676-z.
29. van Kooten RT, Bahadoer RR, Peeters K, Hoeksema JHL, Steyerberg EW, Hartgrink HH, et al. Preoperative risk factors for major postoperative complications after complex gastrointestinal cancer surgery: A systematic review. Eur J Surg Oncol. 2021;47(12):3049–58. 10.1016/j.ejso.2021.07.021.
30. Moon T, Smith A, Pak T, Park BH, Beutler SS, Brown T, et al. Preoperative Anemia Treatment with Intravenous Iron Therapy in Patients Undergoing Abdominal Surgery: A Systematic Review. Adv Ther. 2021;38(3):1447–69. 10.1007/s12325-021-01628-7.
31. Macpherson KJ. Postoperative anaemia increases unplanned readmission: an international prospective cohort study of patients undergoing major abdominal surgery. Br J Surg. 2024;111(7). 10.1093/bjs/znae158.
32. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004;91(1):28–37. 10.1002/bjs.4464.
33. Dafna L, Ganer Herman H, Ginath S, Shalev A, Kerner R, Keidar R, et al. Comparison of Emergent and Elective Surgeries of Endometriomas: Is Adnexal Torsion a Concern? Isr Med Assoc J. 2019;21(4):251–4.

34. Zhu C, Zhang R, Song X, Shang K, Lu Y, Wang ML, et al. A foundation model for acute abdomen diagnosis stratification and triage on noncontrast computed tomography. Nat Commun. 2026;17:76634. 10.1038/s41467-026-76634-w.
35. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ. 2016;352:i6. 10.1136/bmj.i6.
36. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26(6):565–74. 10.1177/0272989X06295361.

# Figures




ICU, intensive care unit; GEE, generalized estimating equations; SE, standard error

**Figure 1**

Flow diagram of patient selection from the MOVER database (EPIC cohort). ICU, intensive care unit; GEE, generalized estimating equation.

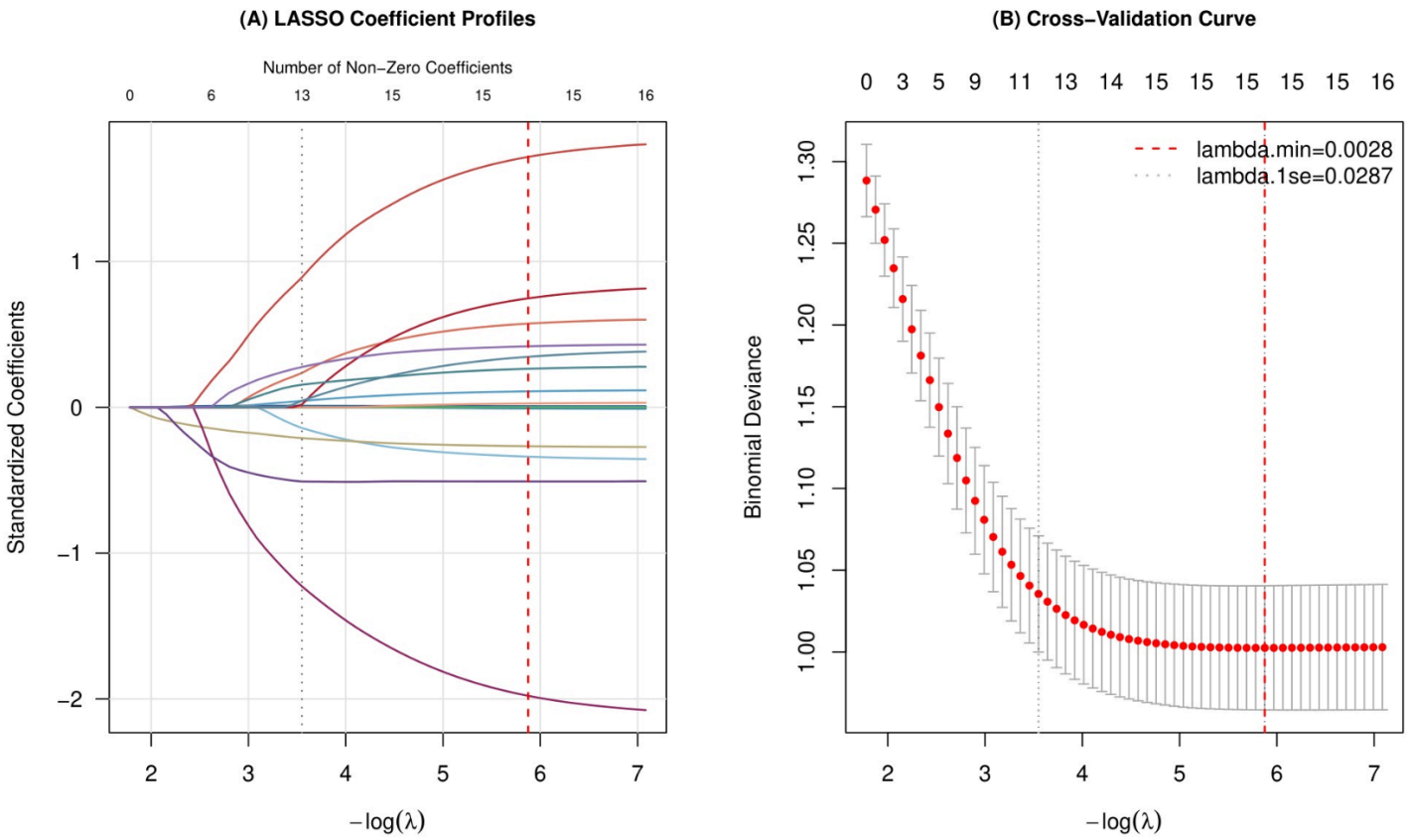



**Figure 2**

LASSO regression for variable selection. (A) Coefficient profiles of the candidate predictors; (B) cross-validation deviance curve of the LASSO regression. Red dashed and gray dotted vertical lines indicate lambda.min and lambda.1se, respectively. Coefficient profiles are shown for the first imputed dataset; 14 variable indicators were consistently selected across all twenty imputed datasets. These 14 indicators correspond to 12 clinical predictors after combining dummy variables for ASA class and acute abdomen subtype.

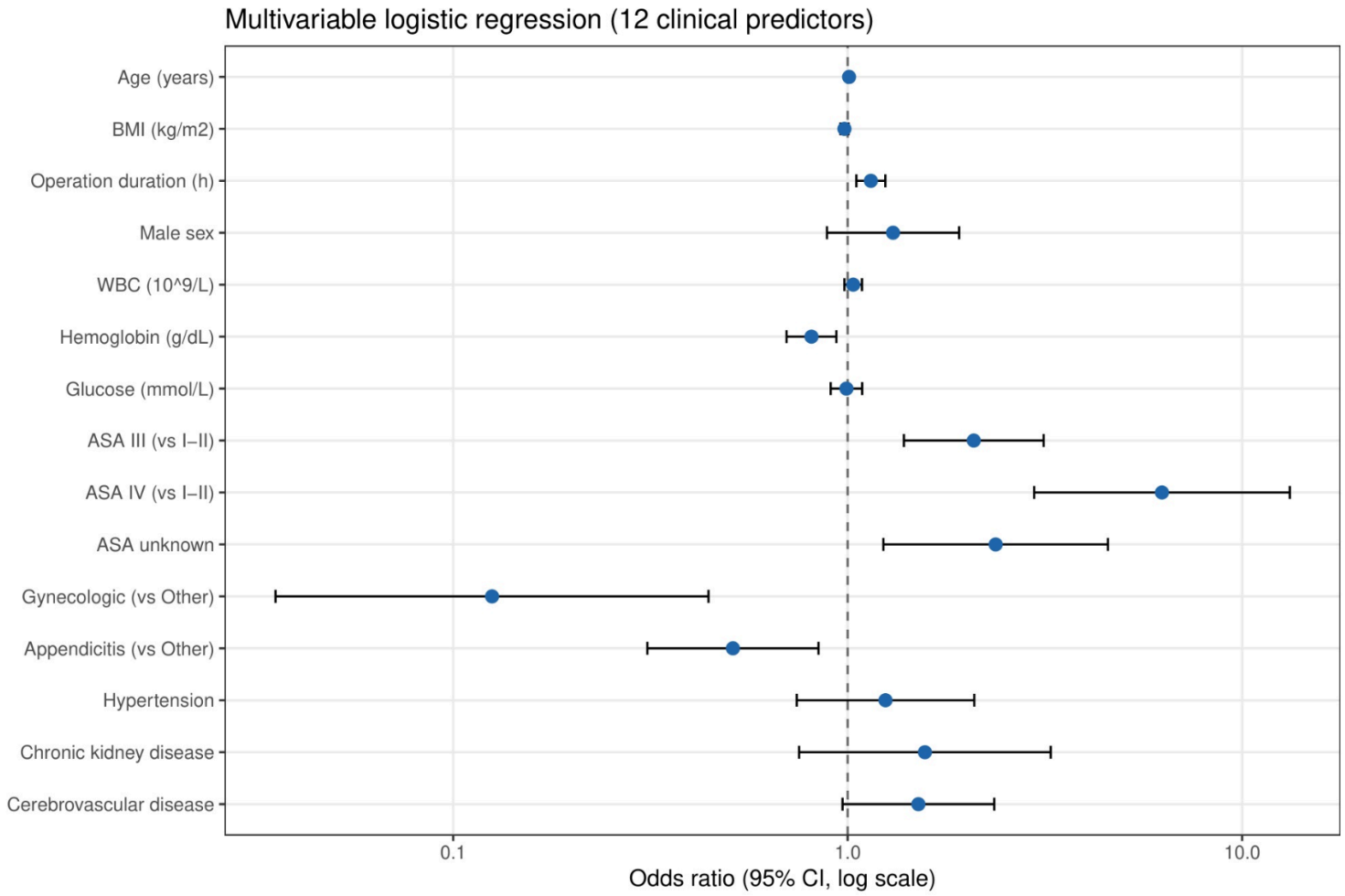



**Figure 3**

Forest plot of the multivariable logistic regression model showing adjusted odds ratios and 95% confidence intervals for postoperative ICU admission, fitted in the training set (n = 1,090) with estimates pooled over 20 imputed datasets (Rubin's rules). ASA, American Society of Anesthesiologists; BMI, body mass index; CI, confidence interval; ICU, intensive care unit; OR, odds ratio.

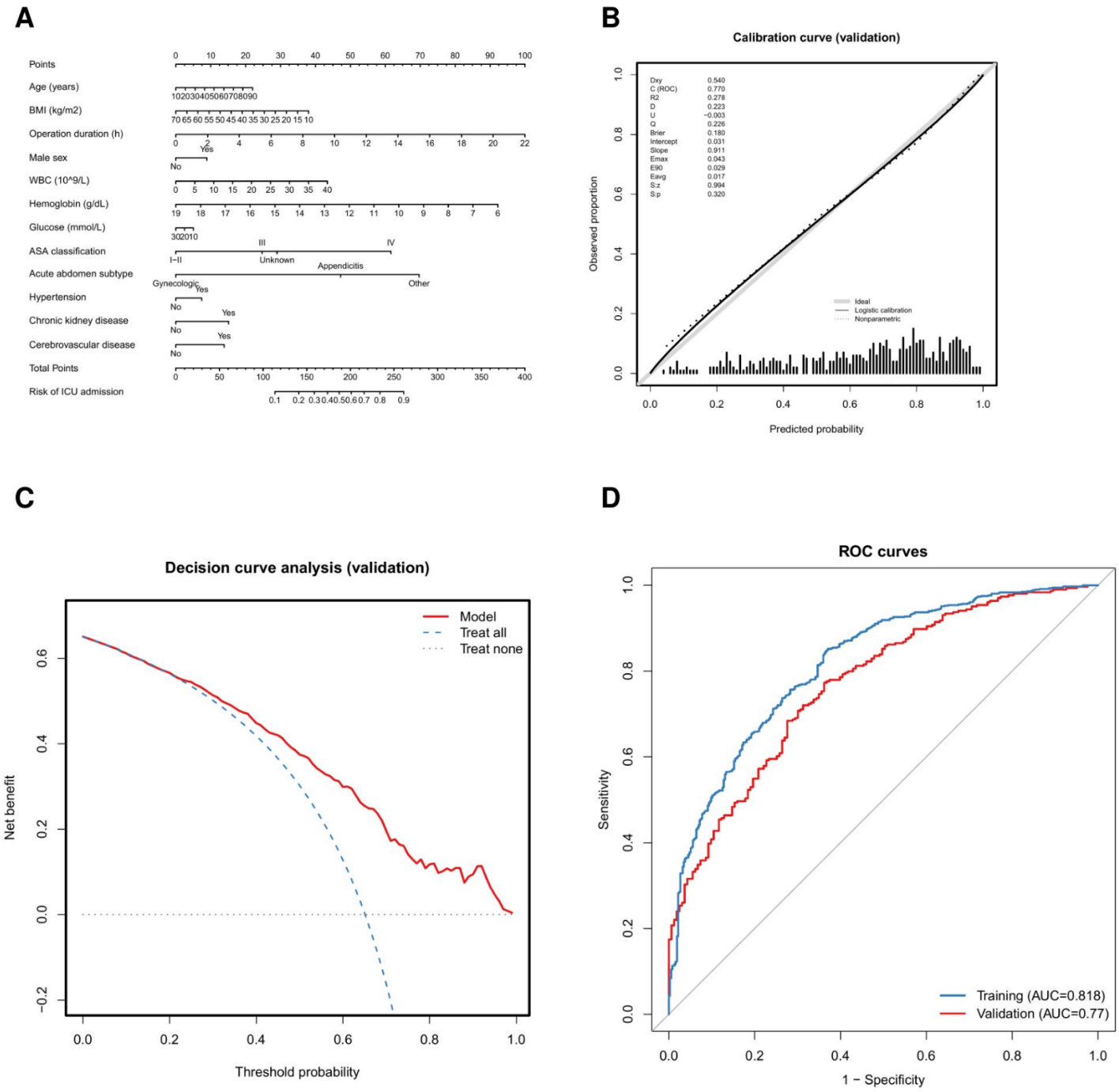



**Figure 4**

Performance evaluation of the nomogram for predicting postoperative ICU admission. (A) nomogram for predicting postoperative ICU admission; (B) calibration curve in the validation set; (C) decision curve analysis in the validation set; (D) receiver operating characteristic (ROC) curves of the model in the training (AUC 0.818) and validation (AUC 0.770) sets. AUC, area under the ROC curve; ICU, intensive care unit; ROC, receiver operating characteristic.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.docx