Am J Transl Res 2026;18(2):1431-1454
www.ajtr.org /ISSN:1943-8141/AJTR0169514

Original Article

Dynamic prediction model for early postoperative pain based on generalized estimating equations: a multi-timepoint longitudinal study

Yuanyuan Liao[1], Na Li[1], Juan Chen[2], Zhihong Tang[1]

[1]Department of Critical Care Medicine, West China Hospital, Sichuan University, No. 37, Wainan Guoxue Lane, Chengdu 610041, Sichuan, China; [2]Department of Emergency, The People's Hospital of Rugao, No. 278 Ninghai Road, Rucheng Town, Rugao 226500, Jiangsu, China

Received October 23, 2025; Accepted January 4, 2026; Epub February 15, 2026; Published February 28, 2026

Abstract: Objective: To identify the risk factors associated with early postoperative pain (Visual Analogue Scale [VAS] ≥3) one hour after extubation in surgical patients in the intensive care unit (ICU), and to establish a dynamic prediction model using generalized estimating equations (GEE) to support precise analgesic management. Methods: This retrospective longitudinal study was conducted in postoperative ICU patients of the West China Hospital (n=373). Patients were randomly divided into training (70%) and testing sets (30%) for model development and internal validation. An external validation cohort from The People's Hospital of Rugao (n=124) was used to assess generalizability. At 30 minutes, one hour, and two hours post-extubation, clinical, perioperative, and extubation variables were collected. Multivariable GEE modeling was performed based on significant factors identified from univariate analysis. Model performance was evaluated using AUC, accuracy, sensitivity, specificity, Hosmer-Lemeshow test, Brier score, and decision curve analysis (DCA). The discriminative performance of the model was compared with that of the Pain Catastrophizing Scale (PCS) using the DeLong test. Results: Independent risk factors for early postoperative pain included a higher Critical-Care Pain Observation Tool score at extubation, BMI, smoking history, older age, higher Acute Physiology and Chronic Health Evaluation II (APACHE-II) score, and intraoperative sedative use. Post-extubation analgesic pump use and time elapsed after extubation were protective factors. The model exhibited good discrimination with an AUC of 0.820 in the training set and 0.785 in the testing set. Stable performance was observed during external validation (AUC: 0.772). DCA demonstrated a significant net clinical benefit across a wide range of threshold probabilities. The model performed significantly better than the PCS ($p < 0.05$). Conclusions: Factors determining early post-extubation pain include patient-related, disease-related, sedation-related, and behavior-related factors. The GEE-based dynamic model provides robust discriminative ability and clinical utility for early identification of high-risk patients, supporting individualized analgesic interventions in the ICU.

Keywords: Intensive care unit, postoperative pain, extubation, generalized estimating equations, prediction model

Introduction

Postoperative pain is one of the most common and important complications arising after surgery. It has a high incidence rate and affects many bodily functions of patients physically, mentally, and socially [1]. According to statistics, up to 80% of surgical patients suffer from postoperative pain. Furthermore, it is estimated that around 75% of these patients report moderate to severe pain intensity [2]. Inadequate control of postoperative pain may delay recovery, prolong hospital stay, induce cardiovascular and respiratory complications, immunosuppression, anxiety, and depression, and may even lead to chronic pain syndrome [3]. According to Hong et al. [4], the intensity of acute pain 24 hours after surgery is an important predictor of chronic pain at 3 months. Literature indicates that the incidence of chronic postsurgical pain can be as high as 25-28%, severely impairing the quality of life of these patients [5]. Thus, accurate assessment and effective management of postoperative pain are of great significance for improving therapeutic outcomes and healthcare quality.

https://doi.org/10.62347/DDSO4916

The assessment and management of postoperative pain in clinical practice faces several challenges. Conventional pain assessment tools, such as the Visual Analogue Scale (VAS) and Numeric Rating Scale, are easy to administer but mainly provide information on momentary subjective experiences. They fail to track the temporal evolution of pain or its underlying mechanisms [6, 7]. In addition, assessments are usually performed after pain onset, indicating that this approach is reactive and inadequate for the "predictive analgesia" paradigm, which requires early identification and intervention of high-risk patients.

Although research on postoperative pain has increased in recent years, significant gaps remain. Existing studies often adopt one-off designs or assess only a single time point, failing to capture dynamic pain fluctuations at critical time points such as 30 minutes, 1 hour, and 2 hours postoperatively. Moreover, tissue injury and subsequent pain are time-dependent processes. Kagerer et al. [8] developed a machine learning model based on 70,000 patients with an AUC of 0.82. Most existing prediction models use classical techniques such as linear or logistic regression, which do not account for the correlation between repeated measurements from the same individual. Consequently, parameter estimation is less precise, and model instability may occur. Studies have shown that while some advanced machine learning algorithms perform well in predicting postoperative pain, postoperative pain is influenced by multiple factors, including baseline characteristics, surgical and anesthetic techniques, intraoperative medications, postoperative interventions, and early physiological parameters. Previous studies have used limited predictor sets and have not systematically explored multifactorial interactions, leading to low interpretability and predictive power of the resulting models. First of all, most existing tools perform suboptimally in terms of discrimination, calibration, and clinical utility, and their generalizability is often limited. Second, existing risk assessment formats are not intuitive and are not conducive to widespread adoption. Thirdly, it remains necessary to establish new prediction models that are more efficient and clinically valuable than the previously established Pain Catastrophizing Scale (PCS).

This study uses Generalized Estimating Equations (GEE) to develop a dynamic prediction model for early postoperative pain. GEE methods are valuable for longitudinal and repeated measures data, where multiple measurements are taken from the same subject. By incorporating the correlation of repeated measurements from the same subject into the estimation of fixed-effect parameters, more efficient and statistically robust estimates are obtained. Yang et al. [11] used a GEE-logistic model to account for data clustering, which performed well with an AUC of 0.944 in the training set and 0.836 in the validation set. Therefore, GEE is effective for modeling intra-individual correlations when there are multiple observation units per subject. The advantages of this method include directly modeling population-level means, accommodating both continuous and categorical variables, and supporting time-varying covariates - meaning it can adapt to the temporal characteristics of postoperative pain.

This study collected baseline, perioperative, and early postoperative physiological indicators. Univariate analysis and multivariable GEE modeling was performed to identify independent influencing factors and quantify their dynamic effects on pain incidence. Subsequently, a nomogram was constructed, and the model was evaluated for discrimination (AUC, accuracy, sensitivity, specificity), calibration (Hosmer-Lemeshow test, Brier score), and clinical net benefit (decision curve analysis, DCA). The DeLong test was used to compare the model's performance with PCS scores. The goal of this study is to provide a robust, accurate, and easily applicable dynamic prediction tool for early postoperative pain, enabling timely identification of high-risk patients and personalized analgesic strategies. This work facilitates the shift from reactive to predictive analgesia and may optimize patient outcomes.

Methods and materials

Sample size estimation

The Events Per Variable (EPV) ratio approach was used to calculate the sample size. Based on previous literature [4] and an estimated early postoperative pain incidence of approximately 40% at our center, we included around six potential risk factors in the multivariable model. According to the empirical rule of EPV ≥10, the number of events should be 60, corresponding to a total sample size of ≥150 patients. We further adopted EPV=15 to en-

hance model robustness, which requires 90 events and an overall sample size of ≥225 patients. A total of 373 patients were enrolled in this study, exceeding the minimum sample size requirement. This ultimately ensured the reliability of model development and validation. In addition, the sample size met the requirements for constructing the dynamic prediction model, as well as conducting internal and external validation.

Sample source and ethics approval

This retrospective longitudinal study was conducted at two tertiary hospitals. The main derivation cohort came from West China Hospital, consisting of postoperative patients admitted to the ICU between January 2021 and April 2024. An independent external validation cohort was collected from The People's Hospital of Rugao during the same study period, using the same inclusion and exclusion criteria to evaluate the external applicability of the model. In total, 373 patients from West China Hospital and 124 patients from The People's Hospital of Rugao were analyzed. To ensure comparability, data collection methods, variable definitions, and pain assessment protocols were standardized across both centers. The study was approved by the Medical Ethics Committee of West China Hospital. Due to the retrospective design and the use of de-identified clinical records, informed consent was waived at both institutions.

Inclusion and exclusion criteria

Inclusion criteria: ① Patients aged ≥18 years; ② Patients who underwent surgery requiring postoperative ICU admission, endotracheal intubation, and mechanical ventilation; ③ Patients who were extubated and had available pain assessment data (including at least a VAS score) at 30 minutes, 1 hour, and 2 hours after extubation; ④ Patients with complete baseline clinical data, perioperative information, and key laboratory indicators.

Exclusion criteria: ① Pre-existing severe central nervous system disorders, consciousness disturbances, or communication barriers precluding accurate pain assessment; ② Chronic opioid use or substance dependence potentially affecting postoperative pain evaluation; ③ Perioperative reoperation, major complications, or death resulting in interrupted pain assessment; ④ Missing primary outcome data or key predictor variables that preclude analysis; ⑤ Pregnant or lactating patients.

Clinical data collection

All clinical data were extracted from the hospital electronic medical record system and ICU nursing documentation system. Two independent researchers extracted the data, with verification by a third researcher. Collected information included: demographic characteristics (age, gender, BMI, smoking history, medical history); preoperative status (American Society of Anesthesiologists [ASA] classification, preoperative pain VAS score [12]); surgical and anesthetic information (surgery type, operative duration, intraoperative blood loss, fluid volume, transfusion status, anesthetic modality, intraoperative use of sedatives and non-opioid analgesics); ICU-related parameters (APACHE-II score at ICU admission [13], duration of mechanical ventilation, time from medication cessation to extubation, duration of extubation procedure); physiological status at extubation (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, oxygen saturation); consciousness and pain assessment scales at extubation (Glasgow Coma Scale [GCS] score [14], CPOT score [15]); as well as post-extubation analgesic pump use, sore throat occurrence, and analgesia-related intervention records. Pain outcomes were evaluated using VAS scores to assess pain intensity. For patients unable to subjectively report pain, the CPOT was additionally used for accurate assessment. All data were uniformly processed before being entered into the database for further review.

Laboratory testing

All laboratory tests were performed by the clinical laboratory center of the institution. To evaluate routine blood parameters (including white blood cells, red blood cells, hemoglobin, and platelets), the Sysmex XN-1000 automated hematology analyzer (Sysmex Corporation, Japan) was used. The ABI 7500 fluorescence quantitative PCR system (Applied Biosystems, USA) was used to analyze total protein, albumin, globulin, albumin-globulin ratio, alanine aminotransferase (ALT), and aspartate aminotransferase (AST) (note: "photocopying" was

removed as it was logically irrelevant). All instruments were regularly calibrated, and internal quality control and external quality assessment were conducted. Laboratory data were extracted from the last test results before extubation, which helped reflect the early postoperative physiological and metabolic status of patients.

Follow-up

The bedside assessment and in-hospital electronic medical record review were used for follow-up in this study. After extubation, all patients underwent pain assessment at 30 minutes, 1 hour, and 2 hours. VAS or CPOT scores were recorded, along with pain-related analgesic interventions and adverse reactions. The follow-up endpoint was the completion of the 2-hour postoperative pain assessment, or the occurrence of events during the 2-hour period that precluded further assessment (e.g., patient transfer, deep sedation, acute adverse events). If necessary, additional records of in-hospital analgesic medication use, rescue analgesic strategies, and analgesia-related complications were documented to further analyze the progression of early postoperative pain and the efficacy of interventions.

Outcome measurement

The study used a longitudinal repeated-measures design, with early postoperative pain recorded at three standardized timepoints: 30 minutes, 1 hour, and 2 hours after extubation. Pain intensity was assessed using the VAS. A VAS score ≥3 was defined as clinically significant pain (denoted as "P"), while a VAS score <3 was defined as no pain (denoted as "N") [16]. For patients with limited communication ability, CPOT scoring was used in conjunction with VAS scoring. For GEE modeling, each patient contributed three repeated binary observations (pain vs. no pain) to characterize the dynamic changes of early postoperative pain over time. To classify patient risk and develop a nomogram, a composite outcome was defined: patients who experienced pain (VAS ≥3) at any timepoint within 2 hours after extubation were considered to have early postoperative pain. In addition, to assess whether the effects of specific risk factors varied by postoperative timepoint and patient subgroups, interaction terms (e.g., CPOT × time, BMI × intraoperative sedative use) were included in extended GEE models. Predicted probabilities under different interaction conditions were generated to characterize dynamic risk patterns. Secondary outcomes included dynamic trends in pain intensity, analgesic pump utilization, rescue analgesic interventions, and pain-associated physiological responses (e.g., heart rate and blood pressure variations).

Statistical analysis

All statistical analyses were conducted using R software (version 4.3.4) and SPSS software (version 27.0). Participants were screened according to predefined inclusion and exclusion criteria, and cases with incomplete clinical information were excluded prior to analysis; therefore, no missing data imputation was required. Continuous variables were tested for normality and presented as mean ± standard deviation or median (interquartile range), as appropriate. Group comparisons of continuous variables were performed using independent-samples t-tests or Wilcoxon rank-sum tests, while categorical variables were compared using χ^2 tests or Fisher's exact tests. To characterize the dynamic evolution of postoperative pain at 30 minutes, 1 hour, and 2 hours after extubation, GEE with a binomial distribution and logit link were applied. Three candidate working correlation structures (independence, exchangeable, and AR(1)) were compared using the quasi-likelihood under the independence model criterion (QIC). All models converged successfully, and the exchangeable structure yielded the lowest QIC (Table S1); thus, it was selected for the final GEE analyses. Candidate predictors were screened using univariate GEE analyses ($P<0.10$). Clinically relevant variables were prespecified based on prior literature and biological plausibility, including age, BMI, smoking history, APACHE-II score, intraoperative sedative use, and CPOT score at extubation. Multicollinearity was assessed using adjusted variance inflation factors (VIF_adj <2). A multivariable GEE model was then constructed, and regression coefficients, odds ratios (OR), and 95% confidence intervals (CI) were reported. All GEE models were fitted using the geepack package in R. Robust (sandwich) standard errors were used for all coefficient estimates to ensure valid inference even if the working correlation structure was misspecified. A nomogram was constructed based on the final model. Model performance was evaluated using

receiver operating characteristic (ROC) curves, area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. Calibration was assessed using the Hosmer-Lemeshow test, Brier score, and calibration plots. Clinical utility was evaluated using DCA. The DeLong test was used to compare the AUC of the proposed model with that of the PCS score. All statistical tests were two-sided.

Results

Baseline characteristics of postoperative patients

A total of 373 patients were enrolled in this study. General demographic characteristics and perioperative clinical data are presented in **Table 1**. The incidence of pain was 48.53% at 30 minutes, 41.29% at 1 hour, and 35.66% at 2 hours after extubation. Overall, 181 patients (48.53%) reported early postoperative pain at any timepoint within the first 2 hours after extubation. The cohort had a predominance of male patients, most of whom were smokers. Most patients had no significant medical history. ASA physical status classification was predominantly Class II and III. The vast majority of patients underwent elective surgery, with low rates of preoperative analgesic use. Surgical types were primarily thoracic procedures, followed by abdominal and other surgeries. Most patients received intraoperative blood transfusion and sedative medications, while the utilization of non-opioid analgesics was relatively low. General anesthesia was the predominant anesthetic modality, with combined anesthesia used in a relatively small proportion of cases. Position changes during the perioperative period were infrequent. After extubation, most patients used analgesic pumps, and some experienced sore throat. The incidence of pain showed a progressive decline at 30 minutes, 1 hour, and 2 hours postoperatively. For continuous variables, the median patient age was in the intermediate range, with BMI levels within the normal range; preoperative pain scores (VAS) were low. Surgical duration, intraoperative blood loss, and intraoperative fluid volumes showed wide distributions. APACHE-II scores at ICU admission were moderate, with relatively short durations of mechanical ventilation and time from medication cessation to extubation. GCS scores at extubation were high, and CPOT scores indicated mild to moderate pain. Vital signs upon extubation were generally stable, including blood pressure, heart rate, respiratory rate, and oxygen saturation. Laboratory parameters (white blood cells, red blood cells, hemoglobin, platelets, total protein, and liver function-related indicators) were overall within normal physiological ranges (**Table 1** and **Figure 1**).

Comparison of baseline characteristics between training and testing groups

A total of 373 patients were enrolled, including 262 in the training group and 111 in the testing group. No statistically significant differences were observed between the two groups in categorical variables, including gender, smoking history, medical history, ASA classification, surgical urgency, preoperative analgesic use, surgical type, intraoperative transfusion, intraoperative non-opioid analgesic use, intraoperative sedative use, anesthetic modality, position changes, post-extubation analgesic pump use, and post-extubation sore throat occurrence (all $P>0.05$). Similarly, no statistically significant differences were noted in continuous variables, including age, BMI, preoperative VAS score, operative duration, intraoperative blood loss, intraoperative fluid volume, APACHE-II score at ICU admission, duration of mechanical ventilation, time from medication cessation to extubation in the ICU, GCS score at extubation, CPOT score, vital signs at extubation (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, oxygen saturation), duration of the extubation procedure, and all laboratory parameters (white blood cells, red blood cells, hemoglobin, platelets, total protein, albumin, globulin, albumin-globulin ratio, ALT, AST) (all $P>0.05$). Furthermore, standardized mean differences (SMD) for all variables were less than 0.2, indicating balanced distributions of baseline characteristics and good comparability between the training and testing groups (**Table 2**).

Variable types and assignment rules

Table 3 presents the variable types and assignment rules incorporated into the model analysis. All categorical variables were assigned binary or multicategory values based on clinical

Table 1. Baseline characteristics of patients

Variable	Category	Total (n=373)
Gender		
	Male	242 (64.88%)
	Female	131 (35.12%)
Smoking history		
	Yes	253 (67.83%)
	No	120 (32.17%)
Medical history		
	Yes	37 (9.92%)
	No	336 (90.08%)
ASA classification		
	2	82 (21.98%)
	3	284 (76.14%)
	4	7 (1.88%)
Surgical urgency		
	Emergency	21 (5.63%)
	Elective	352 (94.37%)
Preoperative analgesic use		
	Yes	20 (5.36%)
	No	353 (94.64%)
Surgery type		
	Thoracic	264 (70.78%)
	Abdominal	40 (10.72%)
	Other	69 (18.50%)
Intraoperative transfusion		
	Yes	241 (64.61%)
	No	132 (35.39%)
Intraoperative non-opioid analgesic use		
	Yes	75 (20.11%)
	No	298 (79.89%)
Intraoperative sedative use		
	Yes	324 (86.86%)
	No	49 (13.14%)
Anesthetic modality		
	General anesthesia	346 (92.76%)
	Combined anesthesia	27 (7.24%)
Position change		
	Yes	17 (4.56%)
	No	356 (95.44%)
Post-extubation analgesic pump use		
	Yes	288 (77.21%)
	No	85 (22.79%)
Post-extubation sore throat		
	Yes	79 (21.18%)
	No	294 (78.82%)
Pain at 30 min postoperatively		
	Yes	181 (48.53%)
	No	192 (51.47%)

Pain at 1 h postoperatively		
	Yes	154 (41.29%)
	No	219 (58.71%)
Pain at 2 h postoperatively		
	Yes	133 (35.66%)
	No	240 (64.34%)
Age		55.00 [42.00, 65.00]
BMI (kg/m^2)		23.29±3.74
Preoperative pain score (VAS)		1.00 [1.00, 1.00]
Operative duration (min)		212.00 [165.00, 308.00]
Intraoperative blood loss (mL)		50.00 [50.00, 120.00]
Intraoperative fluid volume (mL)		700.00 [600.00, 1100.00]
APACHE-II score at ICU admission		23.00 [17.00, 24.00]
Mechanical ventilation duration (h)		15.50 [6.00, 24.00]
Time from ICU medication cessation to extubation (h)		2.00 [1.00, 3.00]
GCS score at extubation		15.00 [14.00, 15.00]
CPOT score at extubation		2.00 [2.00, 4.00]
Systolic blood pressure at extubation (mmHg)		121.00 [115.00, 126.00]
Diastolic blood pressure at extubation (mmHg)		65.00 [63.00, 69.00]
Heart rate at extubation (bpm)		85.00 [76.00, 91.00]
Respiratory rate at extubation (bpm)		22.00 [20.00, 24.00]
Oxygen saturation at extubation (%)		100.00 [100.00, 100.00]
Extubation procedure duration (s)		2.00 [2.00, 3.00]
White blood cells (10^9/L)		4.84 [3.86, 6.23]
Red blood cells (10^9/L)		4.09±0.83
Hemoglobin (g/L)		145.00 [131.00, 161.00]
Platelets (10^9/L)		173.00 [147.00, 210.00]
Total protein (g/L)		53.21 [36.40, 65.60]
Albumin (g/L)		35.80 [29.80, 42.20]
Globulin (g/L)		24.60 [22.50, 27.10]
Albumin-globulin ratio (%)		1.49 [1.15, 1.79]
ALT (U/L)		21.00 [18.00, 25.00]
AST (U/L)		21.00 [18.00, 25.00]
PCS score		24.00 [22.00, 26.00]
Pain rate at 30 min post-surgery		181 (48.53%)
Pain rate at 1 h post-surgery		154 (41.29%)
Pain rate at 2 h post-surgery		133 (35.66%)

Note: ASA: American Society of Anesthesiologists, VAS: Visual Analogue Scale, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, GCS: Glasgow Coma Scale, CPOT: Critical-Care Pain Observation Tool, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BMI: Body Mass Index.

characteristics; dichotomous variables following the principle of "Yes =1, No =0", while multicategorical variables (e.g., ASA classification, surgical type) were assigned ordinal or categorical values. The original measured values of continuous variables were used for statistical analysis and model fitting without grouping or discretization. This assignment strategy ensures that all variables in the model are clearly defined and consistently applied. The GEE model used to assess the dynamic pattern of early postoperative pain is well defined.

Univariate analysis results of early postoperative pain influencing factors based on GEE

Univariate GEE analysis of factors influencing early postoperative pain revealed significant

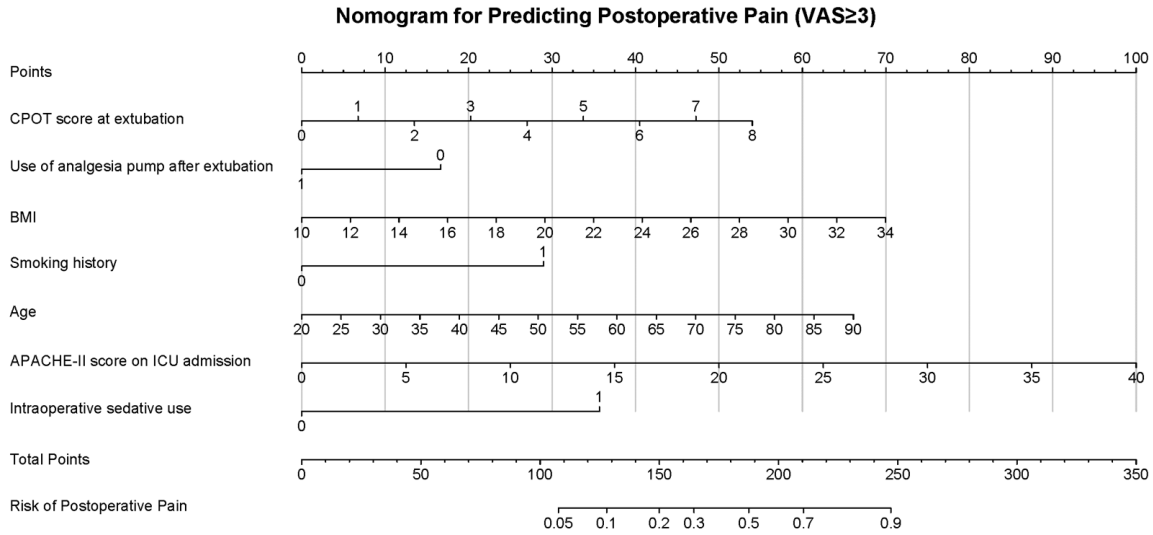

Figure 1. Nomogram for early postoperative pain risk prediction based on multivariable GEE model (VAS ≥3). Note: GEE: Generalized Estimating Equation, CPOT: Critical-Care Pain Observation Tool, VAS: Visual Analogue Scale, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, BMI: Body Mass Index.

Table 2. Comparison of baseline characteristics between training and testing groups

Variable	Training Group (n=262)	Testing Group (n=111)	t/Z/χ^2	*P*-value	SMD
Gender			0.501	0.479	0.080
Male	167 (63.74%)	75 (67.57%)			
Female	95 (36.26%)	36 (32.43%)			
Smoking history			0.308	0.579	0.063
Yes	180 (68.70%)	73 (65.77%)			
No	82 (31.30%)	38 (34.23%)			
Medical history			<0.001	0.997	<0.001
Yes	26 (9.92%)	11 (9.91%)			
No	236 (90.08%)	100 (90.09%)			
ASA classification			1.020	0.600	0.102
2	59 (22.52%)	23 (20.72%)			
3	197 (75.19%)	87 (78.38%)			
4	6 (2.29%)	1 (0.90%)			
Surgical urgency			0.015	0.903	0.014
Emergency	15 (5.73%)	6 (5.41%)			
Elective	247 (94.27%)	105 (94.59%)			
Preoperative analgesic use			0.278	0.598	0.060
Yes	13 (4.96%)	7 (6.31%)			
No	249 (95.04%)	104 (93.69%)			
Surgery type			1.434	0.488	0.120
Thoracic	185 (70.61%)	79 (71.17%)			
Abdominal	31 (11.83%)	9 (8.11%)			
Other	46 (17.56%)	23 (20.72%)			
Intraoperative transfusion			0.166	0.684	0.046
Yes	171 (65.27%)	70 (63.06%)			
No	91 (34.73%)	41 (36.94%)			
Intraoperative non-opioid analgesic use			0.880	0.348	0.106
Yes	56 (21.37%)	19 (17.12%)			
No	206 (78.63%)	92 (82.88%)			
Intraoperative sedative use			0.657	0.418	0.092
Yes	230 (87.79%)	94 (84.68%)			
No	32 (12.21%)	17 (15.32%)			

Anesthetic modality			0.738	0.390	0.097
General anesthesia	245 (93.51%)	101 (90.99%)			
Combined anesthesia	17 (6.49%)	10 (9.01%)			
Position change			0.261	0.609	0.058
Yes	11 (4.20%)	6 (5.41%)			
No	251 (95.80%)	105 (94.59%)			
Post-extubation analgesic pump use			0.533	0.465	0.083
Yes	205 (78.24%)	83 (74.77%)			
No	57 (21.76%)	28 (25.23%)			
Post-extubation sore throat			0.936	0.333	0.110
Yes	52 (19.85%)	27 (24.32%)			
No	210 (80.15%)	84 (75.68%)			
Age	55.00 [41.25, 63.75]	54.00 [42.50, 66.00]	0.180	0.857	0.007
BMI (kg/m^2)	23.15±3.72	23.62±3.81	-1.109	0.268	0.126
Preoperative pain score (VAS)	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	0.058	0.954	0.008
Operative duration (min)	217.80 [165.00, 308.00]	204.60 [157.00, 274.00]	0.565	0.572	0.087
Intraoperative blood loss (mL)	50.00 [50.00, 200.00]	50.00 [50.00, 100.00]	0.680	0.497	0.148
Intraoperative fluid volume (mL)	700.00 [600.00, 1100.00]	700.00 [600.00, 1100.00]	0.034	0.973	0.081
APACHE-II score at ICU admission	23.00 [17.00, 24.75]	21.00 [17.50, 24.00]	0.713	0.476	0.085
Mechanical ventilation duration (h)	18.00 [6.00, 24.00]	10.00 [5.00, 23.00]	1.929	0.054	0.133
Time from ICU medication cessation to extubation (h)	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	1.001	0.317	0.050
GCS score at extubation	15.00 [14.00, 15.00]	14.00 [14.00, 15.00]	1.231	0.219	0.128
CPOT score at extubation	2.00 [2.00, 4.00]	2.00 [2.00, 3.00]	0.399	0.690	0.039
Systolic blood pressure at extubation (mmHg)	121.00 [115.00, 126.00]	121.00 [115.00, 125.00]	0.188	0.851	0.031
Diastolic blood pressure at extubation (mmHg)	65.00 [63.00, 69.00]	65.00 [63.00, 69.00]	0.15	0.881	0.008
Heart rate at extubation (bpm)	85.00 [76.00, 90.75]	85.00 [76.50, 93.00]	0.01	0.992	0.042
Respiratory rate at extubation (bpm)	22.00 [20.00, 24.00]	22.00 [20.00, 24.00]	0.813	0.416	0.125
Oxygen saturation at extubation (%)	100.00 [100.00, 100.00]	100.00 [100.00, 100.00]	0.623	0.533	0.071
Extubation procedure duration (s)	2.00 [2.00, 3.00]	2.00 [2.00, 3.00]	0.486	0.627	0.077
White blood cells (10^9/L)	4.79 [3.79, 5.91]	5.00 [4.09, 6.57]	1.514	0.13	0.048
Red blood cells (10^9/L)	4.08±0.81	4.12±0.89	-0.445	0.657	0.05
Hemoglobin (g/L)	144.50 [130.00, 159.00]	147.00 [132.00, 166.00]	1.116	0.264	0.038
Platelets (10^9/L)	174.00 [145.00, 211.00]	167.00 [149.50, 207.50]	0.272	0.786	0.014
Total protein (g/L)	52.81 [37.38, 65.30]	53.47 [35.75, 66.85]	0.300	0.764	0.003
Albumin (g/L)	35.79 [28.95, 42.10]	35.80 [31.20, 43.30]	0.200	0.842	0.012
Globulin (g/L)	24.54 [22.50, 26.80]	24.90 [22.48, 27.20]	0.806	0.42	0.072
Albumin-globulin ratio (%)	1.49 [1.15, 1.79]	1.52 [1.17, 1.76]	0.184	0.854	0.050
ALT (U/L)	20.00 [17.00, 24.00]	21.00 [18.00, 25.00]	1.115	0.265	0.041
AST (U/L)	21.00 [18.00, 25.00]	21.00 [19.00, 24.50]	0.256	0.798	0.029
PCS score	24.00 [22.00, 26.00]	24.00 [22.00, 26.00]	0.811	0.417	

Note: ASA: American Society of Anesthesiologists, VAS: Visual Analogue Scale, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, GCS: Glasgow Coma Scale, CPOT: Critical-Care Pain Observation Tool, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BMI: Body Mass Index, SMD: Standardized Mean Difference.

associations with smoking history (OR=3.745, 95% CI: 2.258-6.211, P<0.001), ASA classification (OR=0.314, 95% CI: 0.181-0.544, P<0.001), preoperative analgesic use (OR=3.279, 95% CI: 1.090-9.863, P=0.035), intraoperative non-opioid analgesic use (OR=1.805, 95% CI: 1.038-3.138, P=0.036), intraoperative sedative use (OR=2.728, 95% CI: 1.305-5.704, P=0.008), post-extubation analgesic pump use (OR=0.192, 95% CI: 0.105-0.351, P<0.001), age (OR=1.040, 95% CI: 1.023-1.058, P<0.001), BMI (OR=1.194, 95% CI: 1.116-1.277, P<0.001), APACHE-II score at ICU admission (OR=1.105, 95% CI: 1.056-1.156, P<0.001), time from medication cessation to extubation in the ICU (OR=1.377, 95% CI: 1.193-1.590, P<0.001), CPOT score at extubation (OR=1.520, 95% CI: 1.343-1.721, P<0.001), systolic blood pressure at extubation (OR=0.963, 95% CI: 0.940-0.985, P=0.001), and diastolic blood pressure at extubation (OR=0.962, 95% CI: 0.934-0.990, P=0.008). Among these factors, smoking history, preoperative analgesic use, intraoperative non-opioid analgesic use, intra-

Table 3. Variable types and assignment rules

Variable Name	Variable Type	Assignment Rule
Gender	Categorical	Male =1, Female =0
Smoking history	Categorical	Yes =1, No =0
Medical history	Categorical	Yes =1, No =0
ASA classification	Categorical	2=1, 3=2, 4=3
Surgical urgency	Categorical	Emergency =1, Elective =0
Preoperative analgesic use	Categorical	Yes =1, No =0
Surgery type	Categorical	Thoracic =1, Abdominal =2, Other =3
Intraoperative transfusion	Categorical	Yes =1, No =0
Intraoperative non-opioid analgesic use	Categorical	Yes =1, No =0
Intraoperative sedative use	Categorical	Yes =1, No =0
Anesthetic modality	Categorical	General anesthesia =1, Combined anesthesia =0
Position change	Categorical	Yes =1, No =0
Post-extubation analgesic pump use	Categorical	Yes =1, No =0
Post-extubation sore throat	Categorical	Yes =1, No =0
Age	Continuous	Original data used for analysis
BMI (kg/m^2)	Continuous	Original data used for analysis
Preoperative pain score (VAS)	Continuous	Original data used for analysis
Operative duration (min)	Continuous	Original data used for analysis
Intraoperative blood loss (mL)	Continuous	Original data used for analysis
Intraoperative fluid volume (mL)	Continuous	Original data used for analysis
APACHE-II score at ICU admission	Continuous	Original data used for analysis
Mechanical ventilation duration (h)	Continuous	Original data used for analysis
Time from ICU medication cessation to extubation (h)	Continuous	Original data used for analysis
GCS score at extubation	Continuous	Original data used for analysis
CPOT score at extubation	Continuous	Original data used for analysis
Systolic blood pressure at extubation (mmHg)	Continuous	Original data used for analysis
Diastolic blood pressure at extubation (mmHg)	Continuous	Original data used for analysis
Heart rate at extubation (bpm)	Continuous	Original data used for analysis
Respiratory rate at extubation (bpm)	Continuous	Original data used for analysis
Oxygen saturation at extubation (%)	Continuous	Original data used for analysis
Extubation procedure duration (s)	Continuous	Original data used for analysis
White blood cells ($10^9/L$)	Continuous	Original data used for analysis
Red blood cells ($10^9/L$)	Continuous	Original data used for analysis
Hemoglobin (g/L)	Continuous	Original data used for analysis
Platelets ($10^9/L$)	Continuous	Original data used for analysis
Total protein (g/L)	Continuous	Original data used for analysis
Albumin (g/L)	Continuous	Original data used for analysis
Globulin (g/L)	Continuous	Original data used for analysis
Albumin-globulin ratio (%)	Continuous	Original data used for analysis
ALT (U/L)	Continuous	Original data used for analysis
AST (U/L)	Continuous	Original data used for analysis

Note: ASA: American Society of Anesthesiologists, VAS: Visual Analogue Scale, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, GCS: Glasgow Coma Scale, CPOT: Critical-Care Pain Observation Tool, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BMI: Body Mass Index.

operative sedative use, age, BMI, APACHE-II score, time from medication cessation to extubation in the ICU, and CPOT score at extubation were identified as risk factors. In contrast, ASA

classification, post-extubation analgesic pump use, and systolic and diastolic blood pressure at extubation were protective factors. Other variables (including gender, medical history, surgical urgency, surgical type, intraoperative transfusion, anesthetic modality, vital signs, and hematological parameters) showed no statistically significant associations with early postoperative pain (all P>0.05) (**Table 4**).

Multicollinearity diagnostic results before entering variables into the GEE model for early postoperative pain

Multicollinearity testing revealed that VIF_adj values for all included independent variables were <2, indicating no significant collinearity issues. VIF_adj values for systolic and diastolic blood pressure at extubation approached but did not exceed 2, while VIF_adj values for the remaining variables (including CPOT score, post-extubation analgesic pump use, preoperative analgesic use, ASA classification, age, smoking history, BMI, intraoperative sedative use, APACHE-II score, time from ICU medication cessation to extubation, and intraoperative non-opioid analgesic use) were well below 2 (**Table 5**).

Multivariable analysis results of early postoperative pain influencing factors based on GEE

Building upon univariate analysis and collinearity diagnostics, variables with statistical significance or clinical relevance were entered into the multivariable GEE model. Analysis results identified time (time_hours) (OR=0.609, 95% CI: 0.509-0.728, P<0.001) and post-extubation analgesic pump use (OR=0.358, 95% CI: 0.187-0.682, P=0.002) as independent protective factors for early postoperative pain. Independent risk factors included CPOT score at extubation (OR=1.247, 95% CI: 1.067-1.457, P=0.005), BMI (OR=1.134, 95% CI: 1.060-1.213, P<0.001), smoking history (OR=2.612, 95% CI: 1.475-4.624, P=0.001), APACHE-II score at ICU admission (OR=1.075, 95% CI: 1.027-1.126, P=0.002), age (OR=1.036, 95% CI: 1.019-1.053, P<0.001), and intraoperative sedative use (OR=4.113, 95% CI: 1.709-9.900, P=0.002). Additionally, time from ICU medication cessation to extubation exhibited borderline significance (OR=1.177, 95% CI: 1.000-1.385, P=0.050). Other variables, including ASA classification, systolic and diastolic blood pressure at extubation, preoperative analgesic use, intraoperative non-opioid analgesic use, and gender, showed no statistically significant associations and were excluded from the final model (see **Table 6**).

Selection of the optimal working correlation structure

Three candidate working correlation structures (independence, exchangeable, and AR(1)) were compared. All models converged successfully. As shown in Table S1, the exchangeable structure had the lowest QIC value (831.52), slightly lower than that of the independence structure (831.60) and notably lower than that of the AR(1) structure (835.02). Therefore, the exchangeable structure was selected for the final GEE model.

Nomogram for early postoperative pain risk prediction based on multivariable GEE model

Based on multivariable GEE analysis results, independent influencing factors (including CPOT score at extubation, post-extubation analgesic pump use, BMI, smoking history, APACHE-II score at ICU admission, age, and intraoperative sedative use) were incorporated into a nomogram model for predicting early postoperative pain risk (defined as VAS ≥3). As illustrated in **Figure 1**, each predictor variable corresponds to a specific risk score; the predicted probability of postoperative pain can be calculated by summing the total scores. Smoking history, CPOT score, BMI, APACHE-II score, age, and intraoperative sedative use were associated with an increased risk of postoperative pain, whereas post-extubation analgesic pump use significantly reduced this risk. The constructed nomogram intuitively reflects the combined impact of various factors on postoperative pain risk, which can facilitate the development of individualized clinical analgesic strategies (**Figure 1**).

Comparison of key risk factor variables across the three cohorts

To compare the derivation and validation datasets, key risk factor variables were compared across the training (n=262), testing (n=111) and external validation (n=124) cohorts. As shown in Table S2, no significant between-group differences were observed in categorical

Table 4. Univariate analysis results based on generalized estimating equations (GEE)

Variable	Coefficient	Standard Error	*P*-value	OR (95% CI)
Gender	-0.427	0.230	0.063	0.653 (0.416-1.024)
Smoking history	1.320	0.258	<0.001	3.745 (2.258-6.211)
Medical history	0.125	0.362	0.730	1.133 (0.557-2.304)
ASA classification	-1.159	0.281	<0.001	0.314 (0.181-0.544)
Surgical urgency	-0.192	0.455	0.674	0.826 (0.338-2.014)
Preoperative analgesic use	1.187	0.562	0.035	3.279 (1.090-9.863)
Surgery type	-0.309	0.349	0.377	0.735 (0.370-1.457)
Intraoperative transfusion	0.301	0.234	0.198	1.352 (0.854-2.139)
Intraoperative non-opioid analgesic use	0.591	0.282	0.036	1.805 (1.038-3.138)
Intraoperative sedative use	1.004	0.376	0.008	2.728 (1.305-5.704)
Anesthetic modality	0.217	0.408	0.594	1.242 (0.559-2.761)
Position change	0.396	0.564	0.483	1.485 (0.492-4.483)
Post-extubation analgesic pump use	-1.649	0.308	<0.001	0.192 (0.105-0.351)
Post-extubation sore throat	-0.023	0.273	0.933	0.977 (0.572-1.669)
Age	0.039	0.009	<0.001	1.040 (1.023-1.058)
BMI (kg/m^2)	0.177	0.034	<0.001	1.194 (1.116-1.277)
Preoperative pain score (VAS)	0.305	0.304	0.316	1.356 (0.748-2.461)
Operative duration (min)	<0.001	0.001	0.705	1.000 (0.998-1.001)
Intraoperative blood loss (mL)	<0.001	0.001	0.603	1.000 (0.998-1.001)
Intraoperative fluid volume (mL)	<0.001	<0.001	0.178	1.000 (0.999-1.000)
APACHE-II score at ICU admission	0.100	0.023	<0.001	1.105 (1.056-1.156)
Mechanical ventilation duration (h)	-0.007	0.006	0.195	0.993 (0.982-1.004)
Time from ICU medication cessation to extubation (h)	0.32	0.073	<0.001	1.377 (1.193-1.590)
GCS score at extubation	0.046	0.154	0.763	1.047 (0.775-1.416)
CPOT score at extubation	0.419	0.063	<0.001	1.520 (1.343-1.721)
Systolic blood pressure at extubation (mmHg)	-0.038	0.012	0.001	0.963 (0.940-0.985)
Diastolic blood pressure at extubation (mmHg)	-0.039	0.015	0.008	0.962 (0.934-0.990)
Heart rate at extubation (bpm)	0.011	0.008	0.195	1.011 (0.995-1.027)
Respiratory rate at extubation (bpm)	-0.028	0.037	0.445	0.972 (0.905-1.045)
Oxygen saturation at extubation (%)	-0.182	0.315	0.564	0.834 (0.450-1.545)
Extubation procedure duration (s)	-0.079	0.069	0.257	0.924 (0.807-1.059)
White blood cells ($10^9/L$)	-0.035	0.039	0.369	0.965 (0.893-1.043)
Red blood cells ($10^9/L$)	-0.082	0.129	0.525	0.921 (0.715-1.186)
Hemoglobin (g/L)	0.002	0.004	0.585	1.002 (0.994-1.011)
Platelets ($10^9/L$)	0.001	0.002	0.762	1.001 (0.997-1.004)
Total protein (g/L)	-0.006	0.007	0.412	0.994 (0.980-1.008)
Albumin (g/L)	0.002	0.014	0.868	1.002 (0.976-1.029)
Globulin (g/L)	-0.021	0.027	0.428	0.979 (0.928-1.032)
Albumin-globulin ratio (%)	0.137	0.237	0.564	1.147 (0.720-1.826)
ALT (U/L)	<0.001	0.003	0.890	1.000 (0.994-1.007)
AST (U/L)	-0.001	0.004	0.792	0.999 (0.992-1.006)

Note: GEE: Generalized Estimating Equation, ASA: American Society of Anesthesiologists, VAS: Visual Analogue Scale, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, GCS: Glasgow Coma Scale, CPOT: Critical-Care Pain Observation Tool, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BMI: Body Mass Index, OR: Odds Ratio, CI: Confidence Interval.

Table 5. Multicollinearity diagnostic results for each variable

Variable	Df	VIF_adj	Collinearity Level
Systolic blood pressure at extubation	1	1.892	No collinearity
Diastolic blood pressure at extubation	1	1.873	No collinearity
CPOT score at extubation	1	1.12	No collinearity
Post-extubation analgesic pump use	1	1.075	No collinearity
Preoperative analgesic use	1	1.072	No collinearity
ASA classification	2	1.062	No collinearity
Age	1	1.054	No collinearity
Smoking history	1	1.049	No collinearity
BMI	1	1.044	No collinearity
Intraoperative sedative use	1	1.039	No collinearity
APACHE-II score at ICU admission	1	1.036	No collinearity
Time from ICU medication cessation to extubation (h)	1	1.028	No collinearity
Intraoperative non-opioid analgesic use	1	1.025	No collinearity
Gender	1	1.022	No collinearity

Note: Df: Degrees of freedom, VIF_adj: Variance Inflation Factor adjusted, GEE: Generalized Estimating Equation, CPOT: Critical-Care Pain Observation Tool, APACHE-II: Acute Physiology and Chronic Health Evaluation II, BMI: Body Mass Index, ASA: American Society of Anesthesiologists.

Table 6. Multivariable analysis results of early postoperative pain influencing factors based on GEE

Variable	Coefficient	Standard Error	*P*-value	OR (95% CI)
(Intercept)	-6.591	2.231	0.003	0.001 (<0.001-0.109)
time_hours	-0.496	0.091	<0.001	0.609 (0.509-0.728)
CPOT score at extubation	0.221	0.079	0.005	1.247 (1.067-1.457)
Post-extubation analgesic pump use	-1.028	0.33	0.002	0.358 (0.187-0.682)
BMI	0.126	0.034	<0.001	1.134 (1.060-1.213)
Smoking history	0.96	0.291	0.001	2.612 (1.475-4.624)
Age	0.036	0.008	<0.001	1.036 (1.019-1.053)
Time from ICU medication cessation to extubation (h)	0.163	0.083	0.05	1.177 (1.000-1.385)
APACHE-II score at ICU admission	0.073	0.023	0.002	1.075 (1.027-1.126)
ASA classification (2)	-0.189	0.342	0.58	0.827 (0.423-1.618)
ASA classification (3)	0.407	0.695	0.559	1.502 (0.384-5.867)
Systolic blood pressure at extubation	-0.002	0.025	0.95	0.998 (0.952-1.048)
Intraoperative sedative use	1.414	0.448	0.002	4.113 (1.709-9.900)
Diastolic blood pressure at extubation	-0.018	0.029	0.533	0.982 (0.929-1.039)
Preoperative analgesic use	0.184	0.61	0.762	1.203 (0.364-3.978)
Intraoperative non-opioid analgesic use	0.551	0.309	0.074	1.736 (0.947-3.181)
Gender	-0.517	0.278	0.063	0.596 (0.345-1.029)

Note: GEE: Generalized Estimating Equation, CPOT: Critical-Care Pain Observation Tool, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, BMI: Body Mass Index, ASA: American Society of Anesthesiologists, VAS: Visual Analogue Scale, OR: Odds Ratio, CI: Confidence Interval, All standard errors (SE) reported are robust (sandwich) standard errors.

variables including smoking history (P=0.806), preoperative analgesic use (P=0.588), and post-extubation analgesic pump use (P=0.977). Similarly, continuous variables including age (P=0.940), BMI (P=0.697), APACHE-II score at ICU admission (P=0.910), and CPOT score at extubation (P=0.950) were comparable across the three cohorts. These results confirm that the three cohorts were well balanced in terms of key clinical risk factors, supporting the valid-

ity and generalizability of subsequent model training and external validation (Table S2).

Model discrimination, calibration, and clinical net benefit evaluation based on the training set

In the training dataset, the model exhibited good discriminative ability, with an AUC of 0.820. Model performance metrics included an accuracy of 0.767, sensitivity of 0.762, specificity of 0.774, positive predictive value (PPV) of 0.812, negative predictive value (NPV) of 0.718, and an optimal cutoff value of 0.569 (see **Table 7** and **Figure 2A**). A Hosmer-Lemeshow test *P* value of 0.381 and a Brier score of 0.165 indicated good calibration, with predicted probabilities consistent with actual outcomes (**Figure 2B**). DCA demonstrated that the model yielded a greater net clinical benefit than both the "treat-all" and "treat-none" strategies across threshold probabilities of 0.05-0.06 and 0.11-0.99. At threshold probabilities of 0.01-0.04, the net benefit of the model was comparable to that of the "treat-all" strategy. The maximum net benefit (approximately 0.538) was achieved at a threshold probability of approximately 0.05 (**Figure 2C**). Collectively, the model exhibited stable discriminative ability, good calibration, and substantial clinical utility across a wide range of threshold probabilities in the training dataset (see **Figure 2** and **Table 7**).

Model discrimination, calibration, and clinical benefit analysis based on testing set

The testing dataset was an internal validation cohort derived from a random split of the same-center population. In this dataset, the model maintained stable internal discriminative performance, with an AUC of 0.785 (95% CI: 0.702-0.869) (**Figure 3A**). The optimal cutoff probability was determined to be 0.420, which achieved the best balance between event probability and the minimum sum of false negative and false positive rates (**Table 7**; **Figure 3A**). Calibration assessment showed a Hosmer-Lemeshow test *P* value of 0.175 and a Brier score of 0.205, indicating acceptable agreement between predicted probabilities and observed outcomes (**Figure 3B**). DCA revealed that the model provided a greater net clinical benefit than the "treat-all" and "treat-none" strategies across multiple threshold probability ranges, including 0.04-0.09, 0.14, 0.18-0.19, 0.21-0.28, and 0.32-0.98. At two very high threshold probabilities (0.93 and 0.97) and select intermediate ranges (particularly 0.75-0.79, 0.82, and 0.89-0.93), the model outperformed the "treat-all" strategy. The highest net benefit (approximately 0.560) was observed at a threshold probability of around 0.04 (**Figure 3C**). Overall, these results indicate that the model achieved acceptable discriminative ability, good calibration, and consistent clinical utility during internal validation (see **Figure 3** and **Table 7**).

External validation of model performance

The external validation cohort comprised 124 patients in total. The model exhibited stable predictive performance in this independent dataset, with an AUC of 0.772 (95% CI: 0.690-0.853) on the ROC curve (**Figure 4A**). At the cutoff value of 0.597, the model achieved an accuracy of 0.702, sensitivity of 0.653, specificity of 0.769, positive predictive value (PPV) of 0.797, and negative predictive value (NPV) of 0.615, demonstrating acceptable discriminative power (**Table 7**; **Figure 4A**). Based on the Hosmer-Lemeshow goodness-of-fit test (P=0.451) and Brier score (0.192), the model was well calibrated, with good agreement between predicted probabilities and observed outcomes. The calibration curve also closely approximated the reference line, indicating stable calibration performance in the external cohort (**Figure 4B**). DCA further confirmed the model's clinical utility: the model yielded a higher net benefit than both the "treat-all" and "treat-none" strategies across a broad range of threshold probabilities (0.01-0.13, 0.16-0.19, 0.21-0.28, 0.32-0.48, and 0.52-0.90). The maximum net benefit (approximately 0.576) was achieved at threshold probabilities of 0.01-0.02 (**Figure 4C** and **Table 7**), confirming consistent and clinically meaningful applicability in real-world settings.

Model discriminative performance comparison based on Delong test

To further evaluate the discriminative ability of the proposed prediction model, DeLong tests were performed to compare the area under the ROC curves (AUCs) of the prediction model with that of the Pain Catastrophizing Scale (PCS) in the training set, testing set, and external validation set. In the training set, the AUC of the model was 0.8073 (95% CI: 0.756-0.859) and was significantly higher than the AUC of the

Table 7. Model discriminative performance metrics in training and testing sets

Metric	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV	Optimal Cutoff	Youden index
Training set	0.820	0.770-0.870	0.767	0.762	0.774	0.812	0.718	0.569	0.536
Testing set	0.785	0.702-0.869	0.739	0.797	0.66	0.761	0.705	0.42	0.457
External validation set	0.772	0.690-0.853	0.702	0.653	0.769	0.797	0.615	0.597	0.422

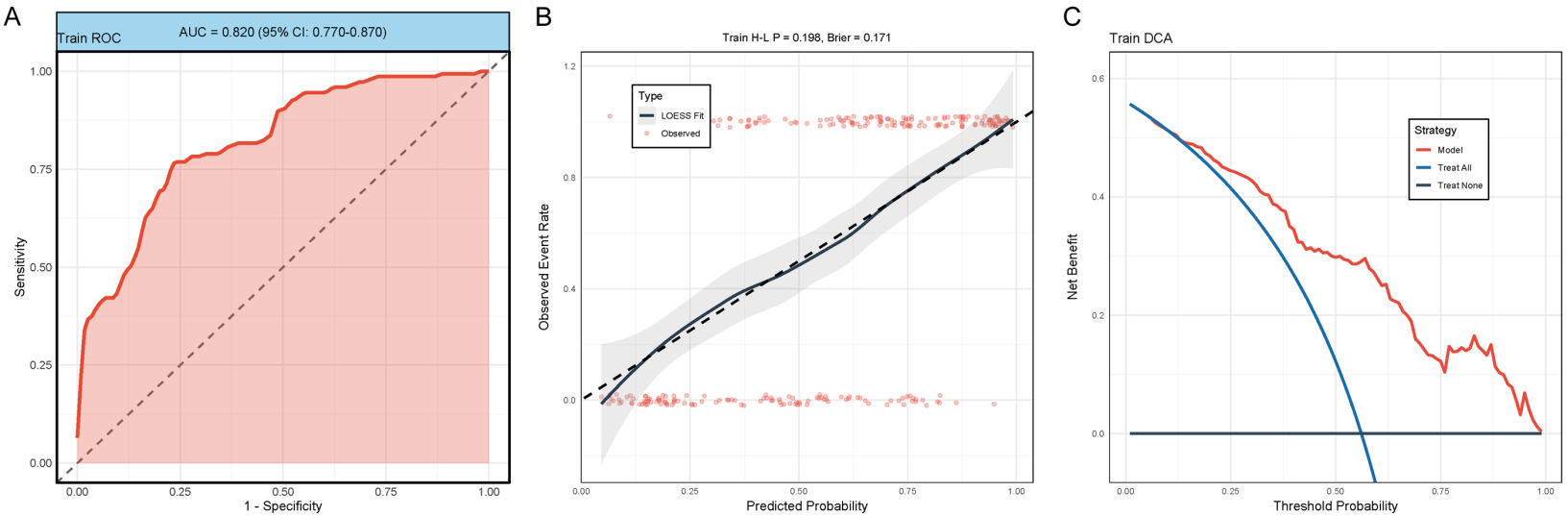

Figure 2. Model performance evaluation based on training set. A. ROC curve of the training set; B. Calibration curve of the training set; C. DCA of the training set. Note: ROC: Receiver Operating Characteristic, DCA: Decision Curve Analysis.

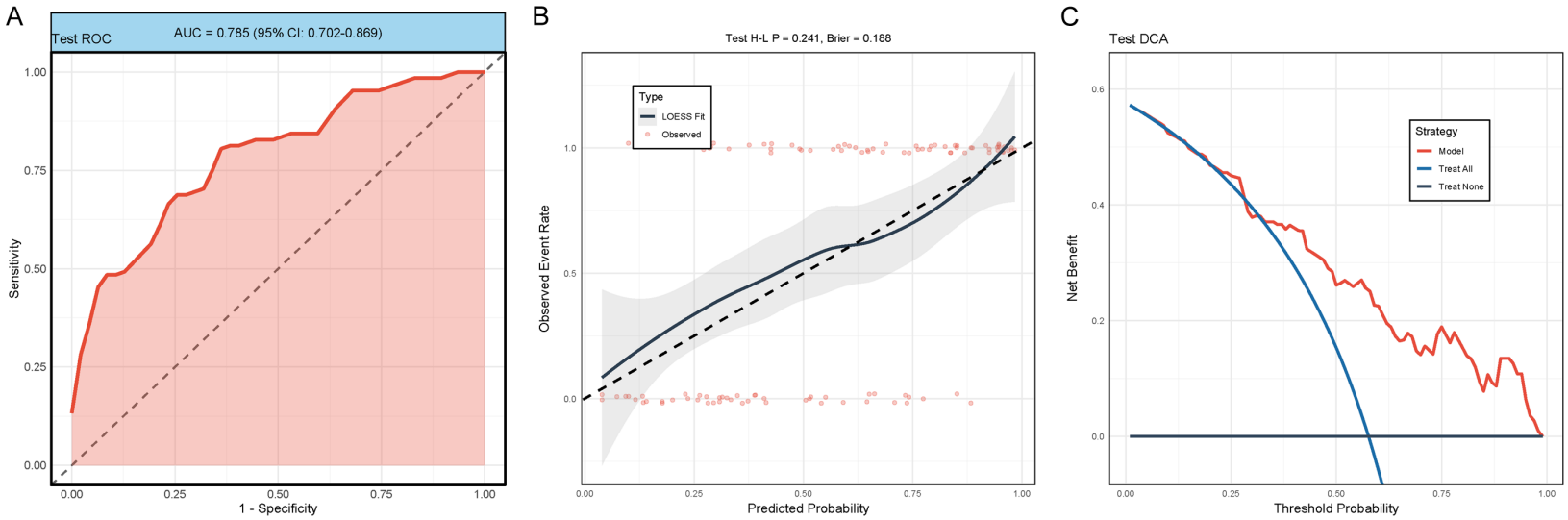

Figure 3. Model performance evaluation based on testing set. A. ROC curve of the testing set; B. Calibration curve of the testing set; C. DCA of the testing set. Note: ROC: Receiver Operating Characteristic, DCA: Decision Curve Analysis.

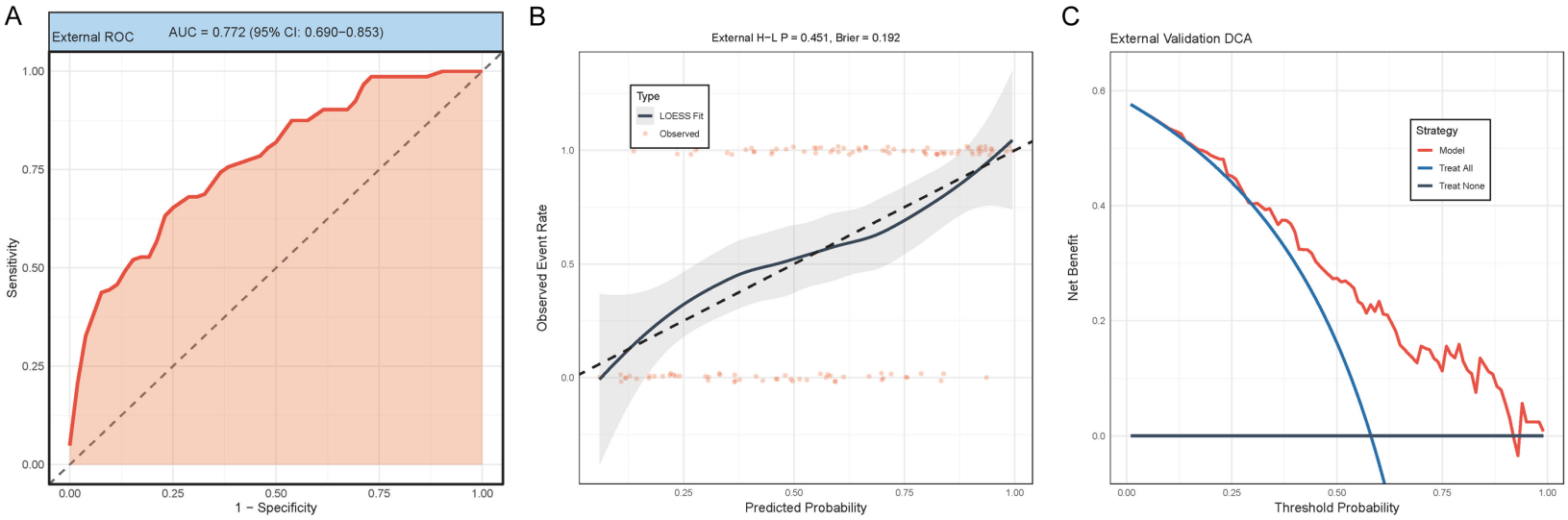

Figure 4. Model performance evaluation in the external validation set. A. ROC curve of the external validation set; B. Calibration curve of the external validation set; C. DCA of the external validation set. Note: ROC: Receiver Operating Characteristic, DCA: Decision Curve Analysis.

PCS score (0.6313, 95% CI: 0.563-0.700) (Delong test, P=0.0001) (**Figure 5A**). In the testing set, the model also showed superior discriminative performance, with an AUC of 0.8301 (95% CI: 0.752-0.908) versus the PCS's AUC of 0.6719 (95% CI: 0.570-0.774); this difference remained statistically significant (DeLong test, P=0.0148). These results confirmed robust generalizability and stable performance during internal validation (**Figure 5B**). In the external validation cohort, the model consistently outperformed the PCS: its AUC of 0.7716 (95% CI: 0.690-0.853) was significantly higher than that of the PCS score (0.6303, 95% CI: 0.531-0.730) (DeLong test, P=0.0273) (**Figure 5C**). This suggests that the model's reliable discriminative ability renders it a valuable clinical tool for independent patient populations. Collectively, across all three datasets, the GEE-based dynamic prediction model outperformed the PCS scoring system in all aspects, demonstrating greater discriminative efficacy and applicability for predicting early postoperative pain.

Interaction effects between key predictors

Interaction analyses were performed using GEE models to examine whether the effects of major predictors varied across postoperative timepoints or differed between clinical subgroups. In the time × CPOT interaction model, the CPOT score remained a significant independent predictor of early postoperative pain (OR=1.373, 95% CI: 1.180-1.597). However, neither of the interaction terms reached statistical significance (30 min × CPOT: OR=1.14, P=0.124; 2 h × CPOT: OR=1.10, P=0.170). This indicates that the effect of nociceptive behavioral cues on pain risk was stable across 30 min, 1 h, and 2 h post-extubation (**Figure 6A**). In the BMI × intraoperative sedative use interaction model, both BMI (OR=1.321, 95% CI: 1.064-1.641) and sedative administration (OR=204.15, 95% CI: 0.47-88,250.97) independently contributed to pain risk. The interaction term was not statistically significant (OR=0.836, P=0.127), suggesting no synergistic or antagonistic modification of BMI's effect by sedative use (**Figure 6B**). The predicted probability curves in **Figure 6** are consistent with these statistical test results: the CPOT-related risk gradients remained parallel over time; while patients receiving sedatives consistently exhibited higher predicted probabilities at all BMI levels, and the slopes of the BMI-pain risk relationship were similar across sedative subgroups. Full numerical results of the interaction analyses are presented in Table S3.

Discussion

This study successfully constructed a postoperative pain prediction model using the GEE method, based on longitudinal multi-timepoint data from 373 ICU surgical patients. The model exhibited excellent discriminative ability, with AUC values of 0.820 in the training set and 0.785 in the testing set, as well as good calibration (Brier scores of 0.165 and 0.205, respectively). Validation in an independent cohort (The People's Hospital of Rugao, n=124) confirmed the model's validity, with an AUC of 0.772 and a Brier score of 0.192, along with good calibration. The consistent performance of the model in both internal and external datasets demonstrates its reliability across institutions. Kagerer et al. developed a gradient boosting model based on 70,000 patients, which also achieved an AUC of 0.82, confirming the value of advanced statistical methods in postoperative pain prediction. Through multivariable analysis, eight independent predictors were identified: CPOT score at extubation (OR=1.247), BMI (OR=1.134), smoking history (OR=2.612), age (OR=1.036), APACHE-II score (OR=1.075), and intraoperative sedative use (OR=4.113) were risk factors; temporal progression (OR=0.609) and post-extubation analgesic pump use (OR=0.358) were protective factors. DCA across a broad range of threshold probabilities (0.04-0.98) validated the model's clinical utility. DeLong tests showed that the model's discriminative ability was superior to the traditional PCS score in the training set (P<0.001) and testing set (P=0.019), as well as in the external validation set (P=0.027). Thus, the model is clinically applicable.

Mechanistic underpinnings and clinical significance of predictive factors

The CPOT score at extubation was identified as the most robust predictor. The CPOT is a reliable scale proven effective for assessing pain in critically ill patients with communication difficulties, based on objective evaluations of facial expressions, body movements, muscle tension, and ventilator compliance [17]. The inherent link between pain intensity at extubation

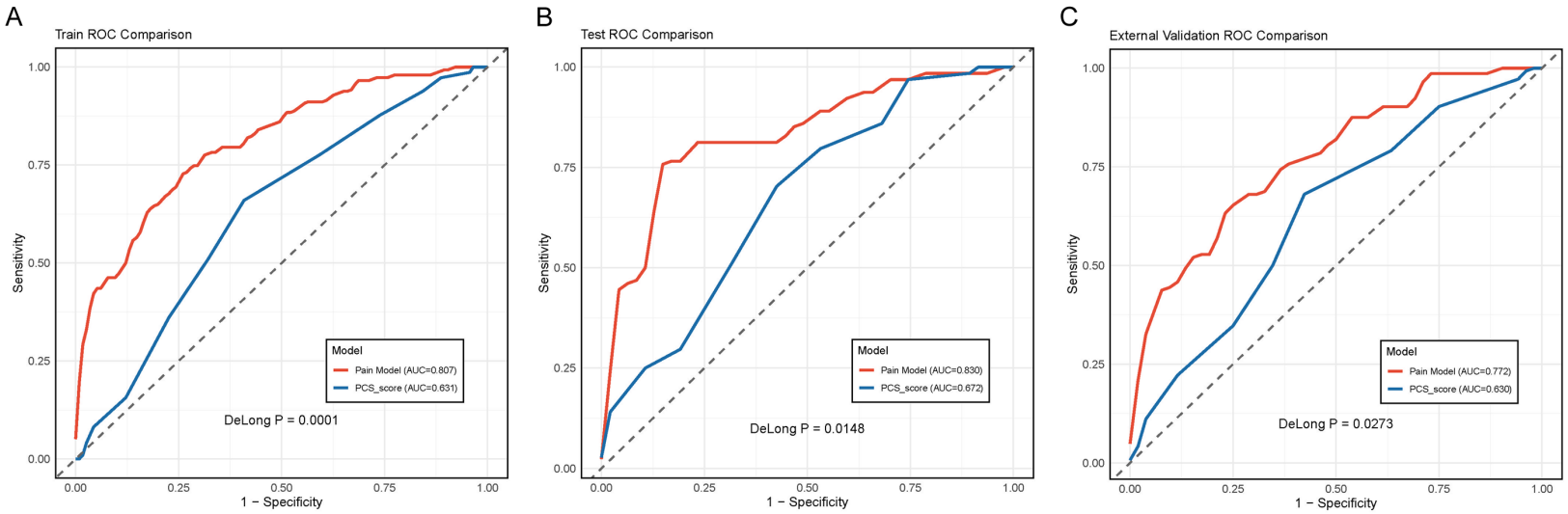

Figure 5. ROC curve comparison between model and PCS score based on Delong test. A. ROC curve comparison between the model and PCS score in the training set; B. ROC curve comparison between the model and PCS score in the testing set. C. ROC curve comparison between the model and PCS score in the external validation set. Note: ROC: Receiver Operating Characteristic, AUC: Area Under the Curve, PCS: Pain Catastrophizing Scale.

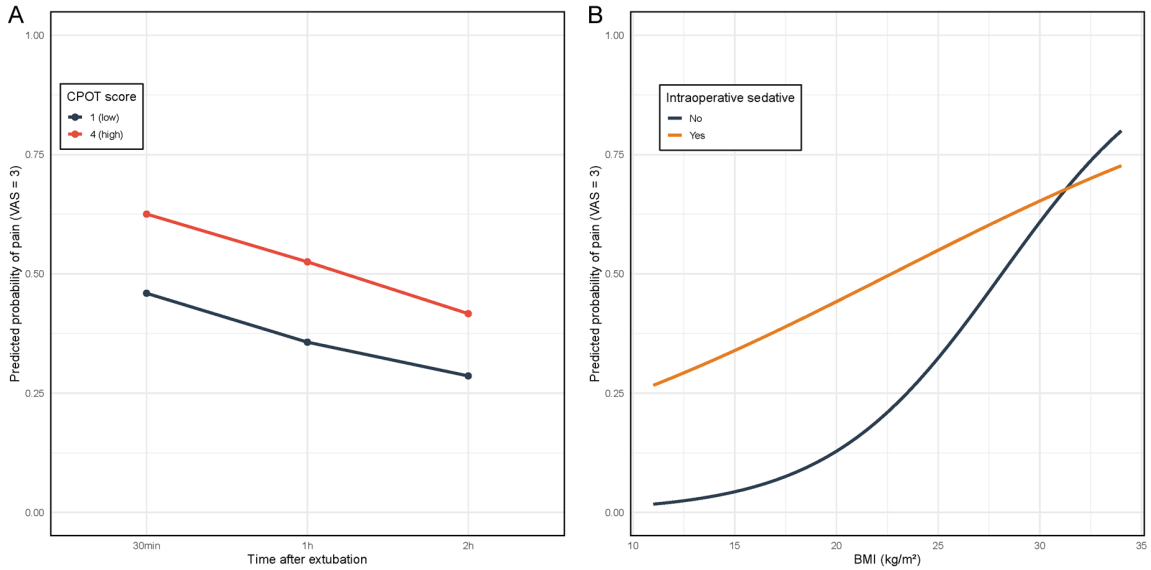

Figure 6. Interaction analysis of key predictors based on generalized estimating equations (GEE). A. Predicted probability of early postoperative pain across different time points (30 min, 1 h, and 2 h after extubation) stratified by CPOT score (low: 1; high: 4). B. Predicted probability of pain across the BMI range by intraoperative sedative use (Yes vs. No). Shaded curves and trajectory patterns indicate that CPOT score and BMI independently influence pain risk without significant interaction effects over time or between sedative use.

and subsequent pain persistence is reflected by its predictive value. According to the literature [4], the severity of acute pain 24 hours after surgery is an important predictor of chronic pain 3 months after hepatectomy; therefore, early pain control is crucial.

BMI, age, and smoking history affect postoperative pain susceptibility through multiple mechanisms. Patients with a higher BMI often have chronic inflammation due to metabolic syndrome, which enhances pain sensitivity via the release of pain mediators [18]. Lunde et al. [19] identified an association between chronic postsurgical pain and low BMI (P=0.032) and reported a non-linear relationship between BMI and pain. Metabolomic studies have provided evidence for BMI's role in pain: 19 lipid metabolites were found to be elevated in patients with chronic pain [20]. Sun et al. constructed an XGBoost model using data from 1,720 elderly patients and confirmed that age, smoking history, and APACHE-II score are major predictors of postoperative acute pain (validation set AUC=0.920). Consistent with this, the present study found smoking history to be a risk factor (OR=2.612). Smoking impairs pain regulation through nicotinic receptor modulation, chronic inflammation, and vascular dysfunction, and smokers often require higher opioid doses. For critically ill patients with severe disease, analgesic use may be limited or less effective due to the severity of their illness. The APACHE-II score reflects a hyperadrenergic state, which increases the release of inflammatory mediators [21, 22].

The strongest risk factor identified was intraoperative sedative use (OR=4.113), potentially due to sedation-analgesia dissociation. This phenomenon involves the formation of an "analgesic gap" caused by excessive sedation without adequate analgesia, which becomes apparent when the patient regains consciousness [23, 24]. The time factor indicates that pain tends to diminish over time, particularly with analgesic intervention. Zhang et al. [25] demonstrated that postoperative patient-controlled analgesia is an independent protective factor against moderate to severe post-anesthesia care unit pain in 22,600 patients (training set AUC=0.817). Literature also indicates that past pain scores and total opioid doses are important predictors of future pain [26]. Zhang et al. [27]. found that paravertebral nerve blockade significantly reduced the risk of chronic pain in patients undergoing thoracoscopic surgery, and severe acute postoperative pain was

an important predictor of chronic pain (training set AUC=0.878). This study further reinforces the value of early aggressive analgesia.

Comparisons with prior research and innovations

The GEE method used in this study enables the utilization of repeated-measures longitudinal data. By applying GEE, the temporal evolution of postoperative pain can be captured with minimal information loss. Furthermore, this methodology accounts for intra-individual correlations, resulting in more robust parameter estimates. Shi et al. [10] applied random forest modeling to orthopedic surgery patients, achieving an AUC of 0.810 and demonstrating the applicability of machine learning approaches in surgical settings. Recent advancements in machine learning have driven significant progress in postoperative pain prediction. Liu et al. [28] developed the Postoperative Pain Score, an attention-based set embedding model that leverages preoperative electronic health record (EHR) data from 234,274 adult non-cardiac surgical patients. Testing results showed state-of-the-art performance in predicting maximum pain scores across multiple postoperative days, and prospective validation revealed that it outperformed clinician predictions. Similarly, Zhang et al. [29] developed DoseFormer, a model based on Graph Attention Networks and Graph Transformer Networks, for video-assisted thoracoscopic surgery (VATS) patients (n=1552). They reported excellent performance with an AUC of 0.98 and an F1 score of 0.85, which was significantly superior to that of attending anesthesiologists (F1: 0.49), fellows (F1: 0.43), and residents (F1: 0.16). These findings suggest that algorithmic prediction may outperform subjective clinical evaluations, highlighting the potential value of automated risk assessment systems in perioperative pain management.

A narrative review of 61 machine learning studies on postoperative pain and opioid use identified substantial variability in sample sizes, predictors, and outcome definitions, with the authors advocating for standardization in such studies [30]. Importantly, patient-reported predictors were emphasized as highly valuable for algorithm performance. Our GEE-based approach (AUC: 0.785-0.820) is competitive and offers additional advantages: it is applicable to longitudinal repeated-measures data and provides clinically interpretable insights through traditional statistical methods. Although DoseFormer [29] generally exhibits better discriminative performance (AUROC: 0.98), GEE is more clinically interpretable, modular, and less computationally complex than deep learning. It also allows for straightforward understanding of how predictors contribute to the model, making it better suited for implementation in most resource-limited clinical settings. Our findings regarding the range of risk factors are highly consistent with the literature [9].

Notably, the model's successful performance in the external validation cohort demonstrates its portability and reliability across multiple institutions - an important advantage that many postoperative pain prediction studies lack, as they are typically only validated at the derivation center. Consistent findings from ROC curve analysis, calibration assessments, and DCA across both hospitals confirm the robustness of this model architecture, thereby enhancing confidence in its clinical application and generalizability.

Adequate management of acute pain not only impacts postoperative recovery but also influences long-term pain outcomes. Hong et al. found that acute pain scores recorded 24 hours after surgery are key predictors of chronic pain. For patients undergoing knee arthroplasty, baseline pain status, kinesiophobia, and knee function are important predictors of chronic pain (validation set AUC=0.897, Brier score 0.119). This evidence indicates that the pain prediction model developed in this study not only demonstrates clinical value but also helps identify populations at high risk of developing chronic pain. In this study, preoperative VAS scores showed no significant correlation with early postoperative pain, which the authors attribute to the low preoperative pain scores in the study population (median 1.0). The characteristics of the ICU patient population may also contribute to this observation. Surgical duration and intraoperative blood loss were not significant predictors in multivariable modeling, indicating the complex nature of pain-influencing factors in the ICU setting.

Clinical application and methodological advantages

Multimodal analgesic regimens should be implemented in high-risk patients. These regi-

mens include optimized preoperative preventive analgesia, with the appropriate adjunctive use of non-opioid analgesics and regional nerve blocks. A systematic review by Holm et al. [31] (involving 946 VATS patients) confirmed that compared with systemic opioids, epidural analgesia reduced resting pain and movement pain by 0.8 and 1.1 points, respectively, during the first three postoperative days; when combined with baseline multimodal analgesia, this effect was even more pronounced. Patients with higher CPOT scores, higher BMI, or higher APACHE-II scores should be prioritized for analgesic pump therapy. Evidence indicates that fully automated prediction systems based solely on preoperative EHR data can be deployed for all patients undergoing or considering surgery, enabling preoperative risk stratification and the planning of preventive interventions. Model implementation can facilitate automated risk assessment by integrating into EHR systems or developing bedside rapid assessment tools to assist nursing staff with real-time evaluations.

In terms of methodology, the longitudinal multitimepoint design accurately reflects the dynamic evolution of postoperative pain. Furthermore, the GEE methodology is well suited to handling correlated responses among repeated measures. For the prediction of lymph node metastasis in gastric cancer using clustered data, Yang et al. [11] applied the GEE-logistic model, which yielded excellent performance with AUC values of 0.944 and 0.836, respectively. Literature has documented that GEE logistic regression effectively analyzes clustered medical data and exhibits robustness in clinical prediction modeling. The inclusion of an additional external validation cohort further strengthens evidence for the model's stability and reproducibility. The stratified training-testing split, excellent baseline comparability (all SMD <0.2), and comprehensive model evaluation (including discrimination, calibration, and DCA collectively underscore the methodological rigor of this study.

In addition, the model is intended for use in adult postoperative ICU patients who require extubation following mechanical ventilation, regardless of surgical type or specific perioperative characteristics. Instead of developing separate subgroup models, we emphasize that the identified independent predictors - including advanced age, smoking history, elevated BMI, higher CPOT scores at extubation, higher APACHE-II scores, and intraoperative sedative use - already identify patient populations at increased risk of early postoperative pain. These populations may benefit from prioritized postoperative analgesia assessment and targeted preventive strategies. Moreover, the current model focuses on predicting very early postoperative pain (within 2 hours after extubation) in adult ICU patients, as opposed to long-term prognostication. From a clinical perspective, this window presents an opportunity to titrate analgesia, prevent pain peaks, and thereby optimize cardiovascular stability. Given the reported associations between high acute pain intensity and subsequent chronic postsurgical pain in other surgical populations, our findings suggest that patients identified as high-risk by this model (e.g., older individuals, smokers, patients with higher BMI, higher APACHE-II and CPOT scores, and those who received intraoperative sedatives) may also require closer long-term follow-up. However, this hypothesis was not directly tested in the current study and should be formally examined in future prospective cohorts.

Limitations and future directions

This study has several limitations that should be acknowledged. First, although subgroup analyses by surgical type, ASA classification, and mechanical ventilation duration were recommended, the number of events in each subgroup was insufficient to yield stable estimates, consistent with EPV principles and TRIPOD guidelines. We therefore clarified the intended application population and highlighted high-risk groups identified by the independent predictors. Second, two key predictors - CPOT score at extubation and post-extubation analgesic pump use - are only obtainable at or immediately after extubation. The CPOT represents the earliest reliable behavioral indicator of nociception in ICU patients, who often cannot self-report pain immediately before extubation; however, this means the model does not function as a purely pre-extubation prediction tool. Although dynamic nomograms or score-to-risk conversion tables could enhance bedside usability, we intentionally refrained from constructing a finalized scoring system at this

stage, as premature release without broader validation may compromise clinical authority. A standardized score-probability conversion table and recommended intervention thresholds will be developed following formal publication and multicenter verification.

Third, although the model demonstrated good discrimination and calibration in both internal and external validation cohorts, both datasets were derived from tertiary hospitals. Thus, even though the study includes two independent centers, the model's adaptability across different healthcare levels (e.g., secondary hospitals), diverse surgical populations, and varying analgesic practices remain uncertain. Future work should explore recalibration strategies (e.g., intercept and slope adjustment) and evaluate the model's transportability across heterogeneous clinical environments. Fourth, pain-related inflammatory or molecular biomarkers (e.g., TNF-α, IL-6, NF-κB) were not routinely measured and thus could not be incorporated into the model, limiting mechanistic interpretation. Fifth, despite external validation, the retrospective study design may limit generalizability, and only very early postoperative pain (within 2 hours) was evaluated. Psychosocial variables and long-term outcomes such as chronic postoperative pain, readmission, and psychological distress were not collected. Langford et al. [29] emphasized the critical importance of including patient-reported predictors where possible, as these have been shown to be particularly informative in machine learning algorithms for postoperative pain prediction. Future studies should conduct multicenter prospective validation, compare pre-extubation and dynamic post-extubation prediction models, integrate biological and physiological markers to elucidate mechanistic pathways, establish clinically actionable risk-threshold strategies, and evaluate whether model-guided analgesic interventions improve patient outcomes. Embedding the model within EHR systems may further facilitate real-time bedside implementation.

Conclusion

Based on data from 373 ICU surgical patients, this study developed a dynamic prediction model for early postoperative pain. The model demonstrated strong discriminative ability, with AUC values of 0.820 in the training set, 0.785 in the testing set, and 0.772 in the external validation cohort; it also exhibited good calibration and significantly outperformed the PCS score in discriminative power. The eight identified independent predictors - including CPOT score at extubation, BMI, smoking history, age, APACHE-II score, intraoperative sedative use, temporal progression, and post-extubation analgesic pump use - may facilitate clinical risk stratification and the formulation of individualized analgesic strategies.

Disclosure of conflict of interest

None.

Address correspondence to: Zhihong Tang, Department of Critical Care Medicine, West China Hospital, Sichuan University, No. 37, Wainan Guoxue Lane, Chengdu 610041, Sichuan, China. E-mail: tzhihong@126.com

References

[1] Wheller KM, Stewart RD, Breedlove KL, Feesler F, Padrutt GC and Bigham MT. Standardising discharge post-surgical pain control: a quality improvement project. Cardiol Young 2025; 1-6.

[2] Olsen SW, Lehmkuhl L, Hamborg LW, Torkov AA, Fog-Nielsen R and Lauridsen J. Postoperative pain relief after ambulatory laparoscopic surgery a nonmatched case-control study. J Perianesth Nurs 2024; 39: 254-262.

[3] Juma SL, Ma J, Rasmussen JL, Shorts Z, Magardichian M, Gentis P, Patrick B and Pettegrew J. Nonopioid pain protocols vs. opioid analgesia for postoperative pain control following arthroscopic surgeries: a literature review. Cureus 2025; 17: e83357.

[4] Hong Y, Li Y, Ye M, Yan S, Yang W and Jiang C. Identifying an optimal machine learning model generated circulating biomarker to predict chronic postoperative pain in patients undergoing hepatectomy. Front Surg 2023; 9: 1068321.

[5] Yan Z, Liu M, Wang X, Wang J, Wang Z, Liu J, Wu S and Luan X. Construction and validation of machine learning algorithms to predict chronic post-surgical pain among patients undergoing total knee arthroplasty. Pain Manag Nurs 2023; 24: 627-633.

[6] Yang MMH, Riva-Cambrin J, Cunningham J and Casha S. Validation of the calgary postoperative pain after spine surgery score for poor postoperative pain control after spine surgery. Can J Neurol Sci 2023; 50: 687-693.

[7] Chiu LYL, Sun T, Ree R, Dunsmuir D, Dotto A, Ansermino JM and Yarnold C. The evaluation of smartphone versions of the visual analogue scale and numeric rating scale as postoperative pain assessment tools: a prospective randomized trial. Can J Anaesth 2019; 66: 706-715.

[8] Kagerer C, Jauk S, Kramer D, Avian A, Neumann T, Schlögl M and Sandner-Kiesling A. A machine learning-based risk assessment model for poor postoperative pain outcome. Stud Health Technol Inform 2025; 324: 98-104.

[9] Sun Y, Yu K, Du L, Hu X, Sheng W, Wang D and Miao H. Application of XGBoost in the prediction of acute postoperative pain after major noncardiac surgery in older patients. Mol Pain 2025; 21: 17448069251376199.

[10] Shi G, Liu G, Gao Q, Zhang S, Wang Q, Wu L, He P and Yu Q. A random forest algorithm-based prediction model for moderate to severe acute postoperative pain after orthopedic surgery under general anesthesia. BMC Anesthesiol 2023; 23: 361.

[11] Yang L, Ding Y, Zhang D, Yang G, Dong X, Zhang Z, Zhang C, Zhang W, Dai Y and Li Z. Predictive value of enhanced CT and pathological indicators in lymph node metastasis in patients with gastric cancer based on GEE model. BMC Med Imaging 2025; 25: 36.

[12] Hwang WY, Kim K, Cho HY, Yang EJ, Suh DH, No JH, Lee JR, Hwang JW, Do SH and Kim YB. The voiding VAS score is a simple and useful method for predicting POUR after laparoscopy for benign gynaecologic diseases: a pilot study. J Obstet Gynaecol 2022; 42: 2469-2473.

[13] Kaur H, Chandran VP, Rashid M, Kunhikatta V, Poojari PG, Bakkannavar SM, Balakrishnan JM and Thunga G. The significance of APACHE II as a predictor of mortality in paraquat poisoning: a systematic review and meta-analysis. J Forensic Leg Med 2023; 97: 102548.

[14] Balakrishnan B, VanDongen-Trimmer H, Kim I, Hanson SJ, Zhang L, Simpson PM and Farias-Moeller R. GCS-pupil score has a stronger association with mortality and poor functional outcome than GCS alone in pediatric severe traumatic brain injury. Pediatr Neurosurg 2021; 56: 432-439.

[15] Hoshino H, Ikeda M, Matsuishi Y, Enomoto Y, Shimojo N, Kotani M, Kobayashi S, Kido T, Hayashi S, Furuya Y and Inoue Y. Validity and reliability of the critical-care pain observation tool (CPOT) for critically ill pediatric patients. PLoS One 2025; 20: e0320373.

[16] Lang-Illievich K, Wejbora M, Szilagyi IS and Bornemann-Cimenti H. A novel metric for reporting acute postoperative pain correlates more closely with patient satisfaction. Br J Anaesth 2020; 124: e228-e229.

[17] Li MMJ, Ocay DD, Larche CL, Vickers K, Saran N, Ouellet JA, Gélinas C and Ferland CE. Validation of the Critical-Care Pain Observation Tool (CPOT) in pediatric patients undergoing orthopedic surgery. Can J Pain 2023; 7: 2156332.

[18] Zengin M, Ulger G, Baldemir R, Sazak H and Alagoz A. Is there a relationship between body mass index and postoperative pain scores in thoracotomy patients with thoracic epidural analgesia? Medicine (Baltimore) 2021; 100: e28010.

[19] Lunde S, Petersen KK, Søgaard-Andersen E and Arendt-Nielsen L. Preoperative quantitative sensory testing and robot-assisted laparoscopic hysterectomy for endometrial cancer: can chronic postoperative pain be predicted? Scand J Pain 2020; 20: 693-705.

[20] Lunde S, Nguyen HT, Petersen KK, Arendt-Nielsen L, Krarup HB and Søgaard-Andersen E. Chronic postoperative pain after hysterectomy for endometrial cancer: a metabolic profiling study. Mol Pain 2020; 16: 1744806920923885.

[21] Wang K, Zhu J, Gao W, Guo W and Guo Y. The effect of hybrid blood purification combined with ulinastatin for the treatment of severe sepsis on APACHE II score and levels of miR-146a and miR-155. Int J Gen Med 2024; 17: 5897-5905.

[22] Kotfis K, Zegan-Barańska M, Strzelbicka M, Safranow K, Żukowski M and Ely EW; POL-CPOT Study Group. Validation of the Polish version of the Critical Care Pain Observation Tool (CPOT) to assess pain intensity in adult, intubated intensive care unit patients: the POL-CPOT study. Arch Med Sci 2018; 14: 880-889.

[23] Nazarnia S and Subramaniam K. Nonopioid analgesics in postoperative pain management after cardiac surgery. Semin Cardiothorac Vasc Anesth 2021; 25: 280-288.

[24] Jiang L, Wang C, Tong J, Han X, Miao C and Liang C. Comparison between thoracic epidural analgesia VS patient controlled analgesia on chronic postoperative pain after video-assisted thoracoscopic surgery: a prospective randomized controlled study. J Clin Anesth 2025; 100: 111685.

[25] Zhang Y, Dai Q, Xu K, Fu H, Zhang A and Du W. Predictors and influence of postoperative moderate-to-severe pain of PACU in the patients with malignancy. BMC Anesthesiol 2024; 24: 81.

[26] Joe YE, Ha N, Lee W and Byon HJ. Prediction of postoperative pain and side effects of patient-controlled analgesia in pediatric orthopedic patients using machine learning: a retrospective study. J Clin Med 2025; 14: 1459.

[27] Zhang XF, Peng CG, Guo HJ and Zhang ZM. Development and validation of a prediction

model for chronic post-surgical pain risk: a single-center prospective study of video-assisted thoracoscopic lung cancer surgery. J Cardiothorac Surg 2025; 20: 85.

[28] Liu R, Gutiérrez R, Mather RV, Stone TAD, Santa Cruz Mercado LA, Bharadwaj K, Johnson J, Das P, Balanza G, Uwanaka E, Sydloski J, Chen A, Hagood M, Bittner EA and Purdon PL. Development and prospective validation of postoperative pain prediction from preoperative EHR data using attention-based set embeddings. NPJ Digit Med 2023; 6: 209.

[29] Zhang C, He J, Liang X, Shi Q, Peng L, Wang S, He J and Xu J. Deep learning models for the prediction of acute postoperative pain in PACU for video-assisted thoracoscopic surgery. BMC Med Res Methodol 2024; 24: 232.

[30] Langford DJ, Reichel JF, Zhong H, Basseri BH, Koch MP, Kolady R, Liu J, Sideris A, Dworkin RH, Poeran J and Wu CL. Machine learning research methods to predict postoperative pain and opioid use: a narrative review. Reg Anesth Pain Med 2025; 50: 102-109.

[31] Holm JH, Bak M and Brøchner AC. Epidural analgesia versus systemic opioids for postoperative pain management after VATS: a systematic review and meta-analysis. Acta Anaesthesiol Scand 2025; 69: e70122.

Table S1. Comparison of QIC values across working correlation structures tested in the GEE model

Working correlation	QIC	QICu	CIC	QICC
Independence	831.6	809.77	25.91	833.56
Exchangeable	831.52	810.06	25.73	833.74
AR(1)	835.02	810.77	27.12	837.24

Note: QIC: Quasi-likelihood under the Independence model Criterion, QICu: QIC based on quasi-likelihood, CIC: Correlation Information Criterion, QICC: corrected QIC. QIC values are reported from the final fitted GEE model. The AR(1) structure yielded non-convergent and non-interpretable negative QIC values and was therefore excluded. The exchangeable structure, which produced the lowest plausible QIC, was selected as the final working correlation structure.

Table S2. Comparison of risk factor variables among the three sample groups

Variable	Total	Training Group (n=262)	Testing Group (n=111)	External validation Group (n=124)	$t/Z/\chi^2$	*P*-value
Smoking history					0.431	0.806
Yes	341 (68.61%)	178 (67.94%)	75 (67.57%)	88 (70.97%)		
No	156 (31.39%)	84 (32.06%)	36 (32.43%)	36 (29.03%)		
Preoperative analgesic use					1.062	0.588
Yes	27 (5.43%)	12 (4.58%)	8 (7.21%)	7 (5.65%)		
No	470 (94.57%)	250 (95.42%)	103 (92.79%)	117 (94.35%)		
Post-extubation analgesic pump use					0.046	0.977
Yes	431 (86.72%)	228 (87.02%)	96 (86.49%)	107 (86.29%)		
No	66 (13.28%)	34 (12.98%)	15 (13.51%)	17 (13.71%)		
Age	55.00 [41.00, 65.00]	55.00 [41.00, 66.00]	55.00 [44.00, 62.50]	54.50 [40.75, 64.25]	0.123	0.940
BMI	23.31±3.68	23.19±3.86	23.53±3.45	23.37±3.48	0.361	0.697
APACHE-II score at ICU admission	23.00 [17.00, 25.00]	23.00 [17.00, 24.00]	23.00 [17.50, 24.00]	23.00 [17.00, 25.00]	0.188	0.910
CPOT score at extubation	2.00 [2.00, 4.00]	2.00 [2.00, 4.00]	2.00 [1.00, 4.00]	2.00 [2.00, 4.00]	0.103	0.950

Note: BMI: Body mass index, ICU: Intensive Care Unit, APACHE-II: Acute Physiology and Chronic Health Evaluation II, CPOT: Critical-Care Pain Observation Tool.

Table S3. Interaction effects in GEE models

Model	Interaction Term	Coefficient (β)	Robust SE	OR	95% CI	*P* value
Time × CPOT score	time_factor1 h × CPOT	0.1313	0.0854	1.14	0.965-1.348	0.1242
	time_factor2 h × CPOT	0.0928	0.0676	1.097	0.961-1.253	0.1698
BMI × intraoperative sedative use	BMI × sedative (Yes)	-0.1791	0.1173	0.836	0.664-1.052	0.1266
Age × smoking history	Age × Smoking (Yes)	-0.0318	0.0184	0.969	0.934-1.004	0.0837

Note: All standard errors are robust (sandwich) standard errors. OR: odds ratio; CI: confidence interval; CPOT: Critical-Care Pain Observation Tool; BMI: body mass index. Reference categories: time_factor =30 min, sedative use = No, smoking history = No. None of the interaction terms reached statistical significance (P<0.05), indicating that the effects of CPOT score on pain risk were stable across time points, and no synergistic or antagonistic interactions were observed between BMI and sedative use or between age and smoking history.