

Development of a Preoperative Clinical Risk Assessment Tool for Postoperative Delirium in Elderly Patients Undergoing Hip Fracture Surgery

Jitpattra Nganlasom, M.D.¹, Ariyachart Kalawa, M.P.H.^{2*}, Piyarat Boonphila, B.N.S.²

1 Anesthesiologist, Department of Anesthesiology, Sakon Nakhon Hospital, Sakon Nakhon 47000, Thailand

2 Nurse Anesthetist, Division of Nurse Anesthesia, Sakon Nakhon Hospital, Sakon Nakhon 47000, Thailand

**Corresponding author: e-mail: Ariya.er@gmail.com*

Original article

OPEN ACCESS

Citation: Nganlasom, J., Kalawa, A., & Boonphila, P. - Development of a Preoperative Clinical Risk Assessment Tool for Postoperative Delirium in Elderly Patients Undergoing Hip Fracture Surgery: -. *Journal of Health Policy, Law and Administration*, 12(3), 185–197. retrieved from https://so05.tci-thaijo.org/index.php/journal_law/article/view/289356

Received: 13 May 2026

Revised: 7 June 2026

Accepted: 10 June 2026

Abstract

Postoperative delirium (POD) is a prevalent and severe complication in elderly patients undergoing hip fracture surgery. This retrospective cohort study aimed to develop a practical clinical risk prediction score for POD. We evaluated 115 patients (aged ≥ 60 years) who underwent hip fracture repair. Clinical data were retrospectively extracted and analyzed using multivariable logistic regression to identify independent predictors. The retained variables were assigned weighted points to construct a simple integer-based scoring system. The model's discriminative performance was assessed utilizing the area under the receiver operating characteristic curve (AUC), alongside calibration analysis.

The overall incidence of POD was 36.52%. The final prediction model incorporated seven preoperative predictors: advanced age (≥ 80 years), ASA physical status III–IV, dementia, urinary catheter use, insomnia, benzodiazepine use, and antipsychotic drug use. The total risk score ranged from 0 to 37 points. The model demonstrated excellent discriminative ability (AUC = 0.88) and adequate calibration. The optimal cut-off score of ≥ 11 achieved a sensitivity of 69.05%, specificity of 95.89%, and overall accuracy of 86.09%. Furthermore, a high-risk threshold of ≥ 13 provided an excellent specificity of 98.63% and a markedly elevated positive likelihood ratio of 46.92. In conclusion, this highly practical preoperative bedside tool effectively stratifies POD risk, facilitating early identification and the implementation of targeted prevention strategies in geriatric surgical care.

Keywords: Clinical prediction score; Elderly; Hip fracture; Postoperative delirium; Preoperative risk assessment

Introduction

Postoperative delirium (POD) is an acute neuropsychiatric syndrome characterized by fluctuating disturbances in attention, awareness, and cognition that develops in the perioperative period (American Psychiatric Association, 2022). This complication represents one of the most prevalent and clinically significant adverse events encountered in elderly surgical patients, particularly those undergoing hip fracture repair. Meta-analytic evidence indicates that the pooled prevalence of POD following hip fracture surgery ranges from 4.00% to 53.30%, with large-scale systematic reviews estimating an overall incidence of approximately 23.00% to 28.00% across diverse populations (Hu et al., 2026; Bai et al., 2020; Bruce et al., 2007). The clinical significance of POD extends beyond the acute episode, as it has been consistently associated with prolonged hospitalization, increased healthcare costs, functional decline, accelerated cognitive deterioration, institutionalization, and elevated short- and long-term mortality (Witlox et al., 2010; Inouye et al., 2014).

The pathophysiology of POD is multifactorial, involving complex interactions among neuroinflammation, neurotransmitter imbalance, disruption of the blood-brain barrier, oxidative stress, and neuroendocrine dysregulation (Maldonado, 2013; Wilson et al., 2020). The neuroinflammatory hypothesis posits that surgical trauma triggers a systemic inflammatory response that activates microglial cells and disrupts neuronal homeostasis, leading to the clinical manifestations of delirium (Maldonado, 2013; Wilson et al., 2020).

Simultaneously, the cholinergic deficiency hypothesis suggests that reduced central cholinergic activity, often exacerbated by anticholinergic medications, contributes to the attention and cognitive deficits characteristic of delirium (Hsieh et al., 2008). Elderly patients are particularly

vulnerable due to diminished cognitive reserve, age-related neurodegeneration, increased blood-brain barrier permeability, and the cumulative burden of medical comorbidities (Inouye et al., 2014).

Hip fracture patients represent a uniquely susceptible population owing to several converging risk factors. These patients are predominantly elderly with significant preexisting frailty, cognitive impairment, and polypharmacy. The acute traumatic event itself induces substantial physiological stress, pain, and immobility, all of which are recognized precipitants of delirium (Mosk et al., 2017). Furthermore, the urgency of surgical intervention often limits the opportunity for comprehensive preoperative optimization. Given that approximately 30.00% to 40.00% of delirium episodes are considered preventable through multicomponent interventions (Hsieh et al., 2015), the identification of modifiable and non-modifiable risk factors before surgery assumes critical importance.

Several prediction models for POD in hip fracture patients have been developed and reported in the literature. A systematic review by Hua et al. (2023) identified eight prediction models with AUC values ranging from 0.67 to 0.79, and noted that preoperative dementia, ASA classification, and advanced age were the most commonly incorporated predictors. Similarly, Zhou et al. (2024) critically appraised 16 prediction models and found considerable variability in the predictors included, sample sizes, and validation approaches. Despite these efforts, several limitations persist. Many existing models incorporate intraoperative or postoperative variables that are unavailable at the time of preoperative assessment, limiting their utility for proactive risk stratification. Others demonstrate inadequate discriminative performance, lack external validation, or include variables that require specialized laboratory testing not readily accessible in all clinical settings.

Therefore, the development of a practical, preoperative, bedside-applicable risk assessment tool that relies on routinely available clinical variables is warranted. Such a tool would enable perioperative care teams to systematically identify high-risk patients and implement targeted, evidence-based preventive strategies—including medication review, sleep promotion, early mobilization, and enhanced cognitive stimulation—prior to surgical intervention.

This study aims to develop a clinical risk prediction score for postoperative delirium in elderly patients undergoing hip fracture surgery at a provincial hospital in northeastern Thailand.

Materials and Methods

Study Design and Setting

This was a retrospective cohort study conducted at the Department of Anesthesiology, Sakon Nakhon Hospital, a tertiary-level provincial hospital in Sakon Nakhon Province, northeastern Thailand. The study was designed to develop a preoperative clinical risk prediction score for POD in elderly hip fracture patients. The study reporting adhered to the updated TRIPOD+AI guidelines for clinical prediction models (Collins et al., 2024).

Study Population

The study enrolled consecutive elderly patients (aged ≥ 60 years) who were admitted to Sakon Nakhon Hospital and underwent surgical repair for hip fractures between October 1, 2024, and September 30, 2025. Eligible patients included those diagnosed with proximal femoral fractures, specifically femoral neck, intertrochanteric, and subtrochanteric fractures (ICD-10 codes: S720, S721, S722). A key inclusion criterion was the presence of postoperative delirium assessments utilizing either the Confusion Assessment Method for the ICU (CAM-ICU) or the 4 'A's Test (4AT) as documented in the patient's medical records. Patients were excluded from the analysis if their

medical records were incomplete, particularly lacking progress note documentation regarding their level of consciousness, cognition, or acute confusion within the first three days following surgery

Sample Size Considerations

The sample size was calculated a priori using the sample size estimation formula for sensitivity (Ngamjarus, 2021). Based on previous literature, the expected sensitivity was set at 85.0% (Somnuke et al., 2022), and the expected prevalence of postoperative delirium was estimated at 21.0% (Guo et al., 2016). Using a 95% confidence level and an allowable margin of error of 0.15, the minimum required sample size was calculated to be 104 patients. To account for a potential 10% rate of incomplete or missing data from retrospective medical records, the final required sample size was determined to be 115 participants.

Data Collection and Variables

Following institutional ethical approval, data were retrospectively extracted from electronic medical records. Candidate variables were categorized into three main domains: (1) Demographic factors, including age, sex, American Society of Anesthesiologists (ASA) physical status, multiple comorbidities (>4 diseases), history of dementia, previous stroke, visual impairment, alcohol consumption, time to surgery >72 hours, chronic or acute illness, electrolyte imbalance, and dehydration; (2) clinical factors, including blood transfusion, preoperative urinary catheter use, and moderate-to-severe postoperative pain (score ≥ 4); and (3) medication-related factors, specifically the preoperative use of benzodiazepines, anticholinergic drugs, first-generation antihistamines, antipsychotics, and antidepressants.

Outcome Measurement

The primary outcome was the occurrence of postoperative delirium (POD) within the first 72 hours following surgery.

Given the retrospective design, POD was identified from medical records based on a documented physician's diagnosis or standardized clinical assessments utilizing either the Confusion Assessment Method for the ICU (CAM-ICU) or the 4 'A's Test (4AT). All documented delirium subtypes (hyperactive, hypoactive, and mixed) were included as outcome events.

Ethical Considerations

The study protocol was reviewed and approved by the Institutional Ethics Committee of Sakon Nakhon Hospital (Approval No. COA/1 No.003/2026, Date of Approval: 1 April 2026).

Statistical Analysis

Statistical analyses were performed using STATA version 18.0. Baseline characteristics were compared between patients with and without postoperative delirium (POD) using the Fisher's exact test for categorical variables, and the independent-samples t-test for continuous variables. A two-sided p-value < 0.05 was considered statistically significant.

The prediction model was developed using a two-stage approach. First, candidate variables were evaluated via univariate logistic regression. Variables demonstrating a significant association ($p < 0.05$) or strong clinical relevance were subsequently entered into a multivariable logistic regression model. To construct a practical risk scoring system, the regression coefficient (β) of each retained predictor was divided by the smallest coefficient in the model and rounded to the nearest integer (Sullivan et al., 2004). A patient's total risk score was calculated by summing these individual weighted points.

Model discriminative performance was evaluated using the area under the receiver operating characteristic curve (AUC). Calibration was assessed via a calibration bubble plot comparing predicted and observed probabilities across risk score quintiles (Hosmer, Lemeshow, & Sturdivant, 2013).

Diagnostic performance (sensitivity, specificity, accuracy, and likelihood ratios) was evaluated at various cut-off values, with the optimal threshold determined by maximizing the Youden index (Youden, 1950). Based on these cut-offs, patients were stratified into low-, moderate-, and high-risk categories, and positive likelihood ratios with 95% confidence intervals were calculated to confirm the clinical utility of this stratification.

Results

Baseline Characteristics

A total of 115 elderly patients were included in the analysis, of whom 42 (36.52%) developed postoperative delirium (POD) and 73 (63.48%) did not. The baseline characteristics of the participants, stratified by POD status, are detailed in Table 1.

Demographically, patients who developed POD were significantly older, with a higher proportion aged ≥ 80 years compared to the non-POD group (47.62% vs. 20.55%), although the overall age distribution did not reach statistical significance ($p = 0.07$). Gender distribution and ASA physical status did not differ significantly between the two groups, although a high proportion of all patients were classified as ASA III–IV (95.24% vs. 89.04%, $p = 0.322$).

In the univariate analysis, several clinical and medication-related factors were significantly associated with the development of POD. Specifically, the POD group exhibited higher prevalence of dementia (9.52% vs. 1.37%, $p = 0.05$), insomnia (40.48% vs. 2.74%, $p < 0.001$), urinary catheter use (28.57% vs. 8.22%, $p = 0.007$), and blood transfusion (21.43% vs. 6.85%, $p = 0.035$). Furthermore, the preoperative use of specific medications was markedly more frequent among patients with POD, including benzodiazepines (50.00% vs. 4.11%, $p < 0.001$), antipsychotics (11.90% vs. 1.37%, $p = 0.024$), and antidepressants (16.67% vs. 2.74%, $p = 0.011$).

Table 1. Baseline characteristics of participants stratified by postoperative delirium status

Variables	POD (n = 42)	Non-POD (n = 73)	p-value
Demographic Characteristics			
Age, n (%)			0.07
60-70 years	6 (14.29)	23 (31.51)	
71-80 years	16 (38.10)	35 (47.95)	
≥ 80 years	20 (47.62)	15 (20.55)	
Gender, male, n (%)	11 (26.19)	23 (31.51)	0.672
ASA physical status III–IV, n (%)	40 (95.24)	65 (89.04)	0.322
Clinical Factors			
Multiple comorbidities (>4 diseases), n (%)	14 (33.33)	23 (31.51)	0.839
Dementia, n (%)	4 (9.52)	1 (1.37)	0.05*
Previous stroke, n (%)	9 (21.43)	9 (12.33)	0.286
Visual impairment, n (%)	6 (14.29)	4 (5.48)	0.167
Alcohol consumption, n (%)	2 (4.76)	4 (5.48)	1.00
Time to surgery >72 hours, n (%)	20 (47.62)	32 (43.84)	0.702
Insomnia, n (%)	17 (40.48)	2 (2.74)	<0.001*
Electrolyte imbalance, n (%)	4 (9.52)	3 (4.11)	0.256
Dehydration, n (%)	8 (19.05)	8 (10.96)	0.268
Blood transfusion, n (%)	9 (21.43)	5 (6.85)	0.035*
Urinary catheter use, n (%)	12 (28.57)	6 (8.22)	0.007*
Postoperative pain score ≥4, n (%)	3 (7.14)	4 (5.48)	0.705
Medication-Related Factors			
Benzodiazepine use, n (%)	21 (50.00)	3 (4.11)	<0.001*
Anticholinergic drug use, n (%)	2 (4.76)	1 (1.37)	0.553
First-generation antihistamine use, n (%)	1 (2.38)	1 (1.37)	1.00
Antipsychotic drug use, n (%)	5 (11.90)	1 (1.37)	0.024*
Antidepressant use, n (%)	7 (16.67)	2 (2.74)	0.011*

Note: Data are presented as number (percentage). Abbreviations: POD, Postoperative delirium; ASA, American Society of Anesthesiologists. *Statistically significant at $p < 0.05$.

Multivariable Logistic Regression and Risk Score Development

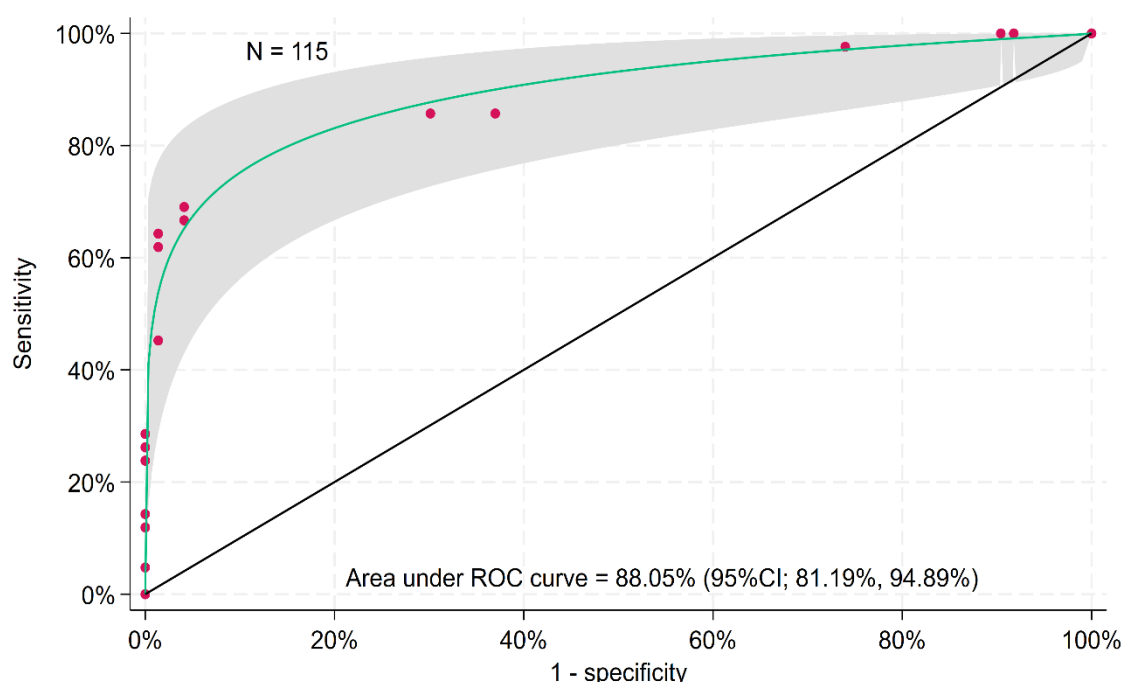
Based on the univariate analysis and clinical relevance, seven predictors were included in the multivariable logistic regression model (Table 2). Antipsychotic drug use emerged as the strongest predictor of POD ($\beta = 4.12$; 95% CI: 0.25–7.99, $p =$

0.009) and was assigned the highest weight of 9 points. Benzodiazepine use contributed 6 points, while advanced age (≥ 80 years) and ASA physical status III–IV contributed 5 points each. The remaining predictors—dementia, insomnia, and urinary catheter use—were assigned 4 points each. The total possible risk score ranged from 0 to 37 points.

Table 2. Regression Coefficients and Assigned Risk Scores for Predicting Postoperative Delirium

Variables	β	95% CI	p-value	Assigned Score
Age				
60-70 years	-	ref.	-	0
71-80 years	0.45	-1.19 – 2.10	0.588	1
≥ 80 years	2.06	0.38 – 3.73	0.016	5
ASA physical status				
I-II	-	ref.	-	0
III-IV	2.46	-1.21 – 6.12	0.190	5
Dementia	2.00	-0.64 – 4.60	0.140	4
Urinary catheter use	1.80	0.29 – 3.14	0.019	4
Insomnia	1.90	-0.19 – 4.00	0.076	4
Benzodiazepine use	2.54	0.68 – 4.39	0.037	6
Antipsychotic drug use	4.12	0.25 – 7.99	0.009	9

Note: The assigned scores were derived from a multivariable logistic regression model. Each score was calculated by dividing the regression coefficient (β) of the respective predictor by the smallest coefficient in the model and rounding to the nearest integer. Abbreviations: β , regression coefficient; CI, confidence interval; ASA, American Society of Anesthesiologists; ref., reference category.

**Figure 1.** Receiver Operating Characteristic (ROC) Curve of the Clinical Prediction Score for Predicting Postoperative Delirium

Model Discrimination

The ROC curve analysis demonstrated that the clinical prediction score had excellent discriminative

performance, with an AUC of 0.881 (95% CI: 0.812–0.949). This value indicates that the model correctly ranked a randomly selected patient with POD higher than a randomly selected patient without POD 88.1% of the time (Figure 1).

Internal Validation

Internal validation of the final clinical prediction score was performed using bootstrap resampling with 1,000 repetitions. The apparent AUC was 0.8805, with a mean optimism of -0.00013 . After optimism correction, the AUC remained essentially unchanged at 0.8806, demonstrating stable discriminative performance with minimal evidence of overfitting.

Diagnostic Performance at Various Cut-off Values

The diagnostic performance of the clinical prediction score across multiple cut-off values is detailed in Table 3. Based on the maximum Youden index, a score of ≥ 11 was identified as the optimal threshold for clinical application, balancing a sensitivity of 69.05% with a high specificity of 95.89%, and achieving an overall accuracy of 86.09% (+LR = 16.80). A lower cut-off of ≥ 9 maximized sensitivity (85.71%) but yielded moderate specificity (63.01%). Conversely, the highest evaluated cut-off of ≥ 13 provided excellent specificity (98.63%) and a markedly elevated positive likelihood ratio (46.92), albeit with reduced sensitivity (64.29%).

Table 3. Diagnostic Performance of the Clinical Prediction Score at Different Cut-off Values for Predicting Postoperative Delirium

Cut-off Value	Sensitivity (%)	Specificity (%)	+LR	-LR	Accuracy (%)
≥ 9	85.71	63.01	2.32	0.23	71.30
≥ 10	85.71	69.86	2.84	0.20	75.65
≥ 11	69.05	95.89	16.80	0.32	86.09
≥ 12	66.67	95.89	16.22	0.34	85.22
≥ 13	64.29	98.63	46.92	0.36	86.09
≥ 14	61.90	98.63	45.19	0.38	85.22

Note: The optimal cut-off value of ≥ 11 was determined by maximizing the Youden index. Abbreviations: +LR, positive likelihood ratio; -LR, negative likelihood ratio.

Risk Stratification

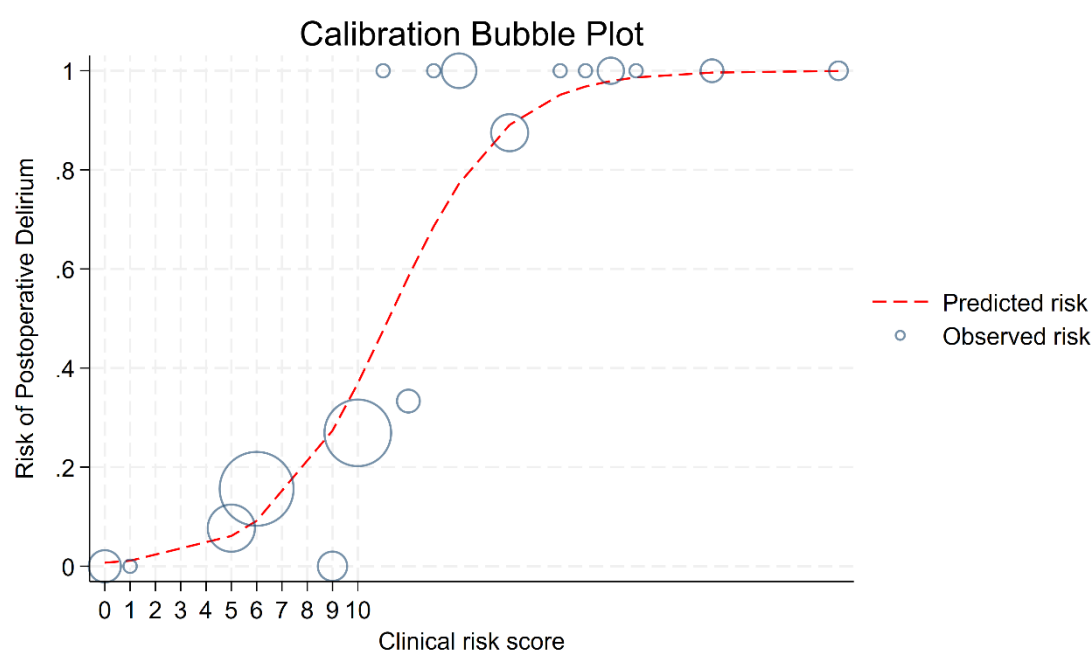
Based on the clinical prediction score, patients were stratified into low- (≤ 9), moderate- (10–12), and high-risk (≥ 13) categories (Table 4). The mean total score was significantly higher in the POD group than in the non-POD group (14.73 ± 6.08 vs. 6.82 ± 3.17 , $p < 0.001$). The incidence of POD increased progressively across the risk

tiers: 10.53% (6/57) in the low-risk, 30.00% (9/30) in the moderate-risk, and 96.43% (27/28) in the high-risk group. Notably, classification in the high-risk category yielded a markedly elevated positive likelihood ratio of 46.92 (95% CI: 6.61–332.99), demonstrating strong clinical utility for identifying patients at the highest risk of developing POD.

Table 4. Risk Stratification According to the Clinical Prediction Score for Postoperative Delirium

Risk categories	Score Range	POD (n=42)	Non-POD (n=73)	+LR	95%CI	p-value
Low risk	≤9	6 (10.53)	51 (89.47)	0.20	0.09 – 0.43	<0.001
Moderate risk	10-12	9 (30.00)	21 (70.00)	0.74	0.37 - 1.47	<0.001
High risk	≥13	27 (96.43)	1 (3.57)	46.92	6.61 – 332.99	<0.001
Mean ±SD	-	14.73±6.08	6.82±3.17	-	-	<0.001

Note: Data are presented as number (percentage) for categorical variables, and as mean ± standard deviation (SD) for continuous variables. Abbreviations: POD, Postoperative delirium; +LR, positive likelihood ratio; CI, confidence interval; SD, standard deviation.

**Figure 2.** Calibration Bubble Plot of the Clinical Prediction Score for Predicting Postoperative Delirium

Calibration

The calibration bubble plot (Figure 2) demonstrated good agreement between the predicted and observed probabilities of POD. The progressive increase in observed POD rates aligned proportionately with higher predicted probabilities, confirming adequate model calibration and reliability across all risk categories. **Discussion**

This study developed a clinical risk prediction score for postoperative delirium in elderly patients undergoing hip fracture surgery. Based on seven readily assessable predictors,

the scoring system demonstrated excellent discriminative performance (AUC = 0.881). The risk stratification scheme effectively differentiated patients into low-, moderate-, and high-risk categories, with the high-risk group showing a markedly elevated probability of developing POD. These findings support the tool's potential clinical utility for risk assessment in the geriatric hip fracture population.

The observed POD incidence of 36.52% is generally consistent with international literature, albeit slightly higher than the pooled prevalence of 23.00% and 28.00% reported in recent meta-analyses (Hu et

al., 2026; Bai et al., 2020). This relatively higher incidence likely reflects the specific characteristics of our study population, wherein 91.30% of patients were classified as ASA III–IV, indicating a substantial burden of comorbidities. Furthermore, methodological differences across studies, including the timing, frequency, and specific diagnostic criteria used for delirium assessment, may contribute to this variation.

Comparison of Model Performance with Existing Tools

The present prediction model achieved an apparent AUC of 0.88 and an optimism-corrected AUC of 0.88 following bootstrap internal validation, demonstrating stable discriminative performance and minimal evidence of overfitting. The model also showed higher discriminative ability than most previously published tools for POD in hip fracture patients. Recent systematic reviews by Hua et al. (2023) and Zhou et al. (2024) indicated that previous models generally yielded lower AUC values (ranging from 0.67 to 0.79) despite sharing common traditional predictors such as advanced age, dementia, and ASA classification. The superior discriminative ability of our model is likely attributable to the incorporation of highly impactful but historically underrepresented variables: specific medication use (benzodiazepines and antipsychotics) and insomnia, an increasingly recognized modifiable risk factor.

Biological and Clinical Plausibility of Identified Predictors

The seven predictors retained in our model are supported by established pathophysiological mechanisms and clinical evidence. Advanced age (≥ 80 years) is a consistently identified risk factor for POD, driven by diminished cognitive reserve, cholinergic decline, and increased blood-brain barrier permeability (Inouye et al., 2014). Concordantly, a meta-analysis by Qi et al. (2022) demonstrated a 2.26-fold increased risk of POD among hip fracture patients older than 80 years. Similarly, higher ASA physical status (III–IV) reflects a substantial burden of

systemic disease and diminished physiological reserve, rendering patients highly susceptible to the surgery-induced neuroinflammatory cascade. The predictive value of this variable is strongly supported by a recent meta-analysis (Hua et al., 2023), which reported a pooled odds ratio of 2.34 (95% CI: 1.15–4.79) for POD development.

Preexisting dementia is a critical predisposing factor for POD. The underlying neurodegenerative processes—specifically cholinergic deficits, chronic neuroinflammation, and diminished cognitive reserve—render patients highly vulnerable to the acute physiological stress of surgery and anesthesia (Oh et al., 2015). Consistent with this biological plausibility, meta-analyses consistently identify dementia as a leading risk factor for POD, reporting odds ratios between 2.10 and 4.63 (Hua et al., 2023).

Urinary catheter use emerged as a significant predictor of POD, aligning with previous evidence in hospitalized older adults (Bo et al., 2019; OR = 2.45, 95% CI: 1.73–3.47). Mechanistically, catheterization acts as a deliriogenic precipitant via physical discomfort, restricted mobility (the "tethering effect"), and an increased susceptibility to urinary tract infections and subsequent systemic inflammation (Wilson et al., 2020). Given its highly modifiable nature, the judicious application and early removal of urinary catheters remain critical components of perioperative delirium prevention bundles.

Preoperative insomnia was strongly associated with POD, exhibiting a markedly higher prevalence in the POD group compared to the non-POD group (40.48% vs. 2.74%). This aligns with recent evidence confirming sleep disturbance as an independent predictor of POD in elderly surgical patients (Guo et al., 2024; incidence 17.00% vs. 6.20%). Mechanistically, insomnia disrupts circadian rhythms and impairs the glymphatic clearance of neurotoxic metabolites, thereby diminishing the brain's resilience to perioperative stress (Todd et al., 2017; Gao et al., 2023). Consequently, preoperative sleep optimization

represents a highly actionable and modifiable target for delirium prevention.

Preoperative benzodiazepine use emerged as the second strongest predictor in our model. This aligns with existing literature identifying benzodiazepines as potent deliriogenic agents in older adults, driven by age-related pharmacokinetic alterations and enhanced sensitivity to GABAergic agents (Clegg & Young, 2011). Concordantly, Fenta et al. (2025) reported a substantially increased risk of POD associated with benzodiazepine premedication (AOR = 4.61, 95% CI: 2.48–8.57). Given this robust association, the judicious minimization or avoidance of benzodiazepine exposure during the perioperative period remains a critical preventive strategy.

Antipsychotic drug use was identified as the strongest individual predictor, contributing the highest weight (9 points) to the risk score. Clinically, preoperative antipsychotic use likely serves as a proxy for severe underlying neuropsychiatric vulnerability, such as preexisting dementia with behavioral disturbances or a history of delirium. Pharmacologically, antipsychotics directly precipitate delirium through anticholinergic properties and the disruption of dopaminergic and serotonergic pathways, which exacerbate perioperative neurotransmitter imbalances (Maldonado, 2013). This profound association emphasizes the necessity of comprehensive medication reconciliation and meticulous neuropsychiatric evaluation during preoperative assessment.

Clinical Implications

The developed risk prediction score offers substantial clinical utility. First, it relies entirely on routinely available preoperative variables without requiring specialized laboratory or imaging tests, facilitating rapid bedside assessment by multidisciplinary care teams. Second, the three-tier stratification system provides actionable thresholds to guide clinical decision-making. High-risk patients

(score ≥ 13) warrant targeted, multicomponent prevention bundles—such as high-risk medication deprescribing, sleep optimization, early mobilization, and proactive geriatric consultation. Conversely, moderate- (score 10–12) and low-risk (score ≤ 9) patients can be effectively managed with standard preventive measures and routine vigilance, respectively.

Notably, the markedly elevated positive likelihood ratio (46.92) for the high-risk category provides clinicians with a definitive signal to initiate intensive preventive protocols. In parallel, the low-risk category (+LR = 0.20) reliably identifies patients with a minimal probability of POD, enabling the efficient and cost-effective allocation of hospital resources. Unlike static risk models that rely solely on unmodifiable factors like age, our tool identifies actionable targets. For a patient scoring in the high-risk tier (≥ 13) driven by preoperative insomnia and benzodiazepine use, the clinical mandate is clear: immediate implementation of non-pharmacological sleep protocols, coupled with a proactive medication review to safely taper or substitute deliriogenic agents prior to surgery.

Strengths and Limitations

This study possesses several notable strengths. The prediction model relies entirely on routinely available preoperative variables, making it highly practical for proactive risk assessment. It demonstrated excellent discriminative performance (AUC = 0.88), outperforming most previously published tools for this population. Furthermore, the inclusion of historically underrepresented medication-related variables enhances the model's comprehensiveness, while its simple integer-based scoring system facilitates rapid adoption in routine clinical practice.

Several limitations warrant consideration. First, although our EPV ratio of 6.0 falls below the historical threshold of 10, modern simulation studies demonstrate that a lower EPV does not inherently invalidate a model when the selection of candidate variables

is strongly constrained by biological plausibility and clinical evidence. To mitigate the risk of variable-selection bias from low EPV, we deliberately restricted our multivariable pool to established clinical pillars. Nevertheless, the wide confidence intervals observed for low-prevalence predictors like preexisting dementia highlight the need for larger, multi-center cohorts to refine coefficient precision. Second, the single-center retrospective design may limit the generalizability of the findings and introduce residual confounding from unmeasured intraoperative variables, such as anesthesia type or surgical duration.

Third, because this model was derived from a single-center cohort in northeastern Thailand, its transportability to alternative geographic cohorts remains unestablished. A prospective, multi-center external validation study involving both community and university hospitals is currently being planned to evaluate the calibration and clinical utility of this scoring system across a broader range of healthcare settings. Finally, internal validation of the final clinical prediction score was performed using bootstrap resampling with 1,000 repetitions. The apparent AUC was 0.8805, with a mean optimism of -0.00013 . After optimism correction, the AUC remained essentially unchanged at 0.8806, indicating stable discriminative performance with minimal evidence of overfitting. Nevertheless, external validation in independent patient populations remains necessary to confirm the generalizability and transportability of the scoring system.

References

- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425787>
- Bai, J., Liang, Y., Zhang, P., Liang, X., He, J., Wang, J., & Wang, Y. (2020). Association between postoperative delirium and mortality in elderly patients undergoing hip fractures surgery: A meta-analysis. *Osteoporosis International*, 31(2), 317–326. <https://doi.org/10.1007/s00198-019-05172-7>

Conclusion

This study developed a clinical risk prediction score for postoperative delirium in elderly patients undergoing hip fracture surgery. Incorporating seven readily accessible predictors—advanced age, ASA physical status III–IV, dementia, urinary catheter use, insomnia, benzodiazepine use, and antipsychotic drug use—the model demonstrated excellent discriminative performance (AUC = 0.88). Furthermore, the three-tier risk stratification system effectively differentiated patients across clinically meaningful risk categories. Internal bootstrap validation confirmed stable model performance with minimal optimism, supporting the robustness of the proposed scoring system. Ultimately, this practical bedside tool facilitates targeted preoperative risk assessment and guides the implementation of individualized delirium prevention strategies.

Acknowledgements

The authors sincerely thank the multidisciplinary clinical staff at the Department of Anesthesiology, Sakon Nakhon Hospital, for their dedication to patient care and meticulous clinical documentation, which provided the high-quality data essential for this retrospective study.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest relevant to this study.

- Bo, M., Porrino, P., Di Santo, S. G., Mazzone, A., Cherubini, A., Mossello, E., Bianchetti, A., Musicco, M., Ferrari, A., Ferrara, N., Filippini, C., Trabucchi, M., Morandi, A., Bellelli, G., & Italian Study Group on Delirium (ISGoD). (2019). The association of indwelling urinary catheter with delirium in hospitalized patients and nursing home residents: An explorative analysis from the "Delirium Day 2015". *Aging Clinical and Experimental Research*, 31(3), 411–420. <https://doi.org/10.1007/s40520-018-0974-1>
- Bruce, A. J., Ritchie, C. W., Blizard, R., Lai, R., & Raven, P. (2007). The incidence of delirium associated with orthopedic surgery: A meta-analytic review. *International Psychogeriatrics*, 19(2), 197–214. <https://doi.org/10.1017/S104161020600425X>
- Clegg, A., & Young, J. B. (2011). Which medications to avoid in people at risk of delirium: A systematic review. *Age and Ageing*, 40(1), 23–29. <https://doi.org/10.1093/ageing/afq140>
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A. L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., ... Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, Article e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Fenta, E., Teshome, D., Kibret, S., Hunie, M., Tiruneh, A., Belete, A., Molla, A., Dessie, B., & Geta, K. (2025). Incidence and risk factors of postoperative delirium in elderly surgical patients. *Scientific Reports*, 15(1), Article 1400. <https://doi.org/10.1038/s41598-024-84554-2>
- Gao, L., Li, P., Gaykova, N., Zheng, X., Gao, C., Lane, J. M., Saxena, R., Scheer, F. A. J. L., Rutter, M. K., Akeju, O., & Hu, K. (2023). Circadian rest-activity rhythms, delirium risk, and progression to dementia. *Annals of Neurology*, 93(6), 1145–1157. <https://doi.org/10.1002/ana.26617>
- Guo, H., Li, L. H., Lv, X. H., Su, F. Z., Chen, J., Xiao, F., Shi, M., & Xie, Y. B. (2024). Association between preoperative sleep disturbance and postoperative delirium in elderly: A retrospective cohort study. *Nature and Science of Sleep*, 16, 389–400. <https://doi.org/10.2147/NSS.S452517>
- Guo, Y., Jia, P., Zhang, J., Wang, X., Jiang, H., & Jiang, W. (2016). Prevalence and risk factors of postoperative delirium in elderly hip fracture patients. *The Journal of International Medical Research*, 44(2), 317–327. <https://doi.org/10.1177/0300060515624936>
- Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118548387>
- Hshieh, T. T., Fong, T. G., Marcantonio, E. R., & Inouye, S. K. (2008). Cholinergic deficiency hypothesis in delirium: A synthesis of current evidence. *The Journals of Gerontology: Series A, Biological Sciences and Medical Sciences*, 63(7), 764–772. <https://doi.org/10.1093/gerona/63.7.764>
- Hshieh, T. T., Yue, J., Oh, E., Puella, M., Dowal, S., Trivison, T., & Inouye, S. K. (2015). Effectiveness of multicomponent nonpharmacological delirium interventions: A meta-analysis. *JAMA Internal Medicine*, 175(4), 512–520. <https://doi.org/10.1001/jamainternmed.2014.7779>
- Hu, Y., Tang, W., & Jiang, X. (2026). Prevalence and risk factors for postoperative delirium after hip fracture in the elderly: A systematic review and meta-analysis. *Medicine*, 105(4), Article e47296. <https://doi.org/10.1097/MD.00000000000047296>
- Hua, Y., Yuan, Y., Wang, X., Liu, L., Zhu, J., Li, D., & Tu, P. (2023). Risk prediction models for postoperative delirium in elderly patients with hip fracture: A systematic review. *Frontiers in Medicine*, 10, Article 1226473. <https://doi.org/10.3389/fmed.2023.1226473>
- Inouye, S. K., Westendorp, R. G., & Saczynski, J. S. (2014). Delirium in elderly people. *The Lancet*, 383(9920), 911–922. [https://doi.org/10.1016/S0140-6736\(13\)60688-1](https://doi.org/10.1016/S0140-6736(13)60688-1)

- Maldonado, J. R. (2013). Neuropathogenesis of delirium: Review of current etiologic theories and common pathways. *The American Journal of Geriatric Psychiatry*, 21(12), 1190–1222. <https://doi.org/10.1016/j.jagp.2013.09.005>
- Mosk, C. A., Mus, M., Vroemen, J. P., van der Ploeg, T., Vos, D. I., Elmans, L. H., & van der Laan, L. (2017). Dementia and delirium, the outcomes in elderly hip fracture patients. *Clinical Interventions in Aging*, 12, 421–430. <https://doi.org/10.2147/CIA.S115945>
- Ngamjarus, C. (2021). *Sample size calculation for health science research*. Teaching and Learning Innovation Center, Khon Kaen University. (in Thai).
- Oh, E. S., Li, M., Fafowora, T. M., Inouye, S. K., Chen, C. H., Rosman, L. M., Lyketsos, C. G., Sieber, F. E., & Puhon, M. A. (2015). Preoperative risk factors for postoperative delirium following hip fracture repair: A systematic review. *International Journal of Geriatric Psychiatry*, 30(9), 900–910. <https://doi.org/10.1002/gps.4233>
- Qi, Y. M., Li, Y. J., Zou, J. H., Qiu, X. D., Sun, J., & Rui, Y. F. (2022). Risk factors for postoperative delirium in geriatric patients with hip fracture: A systematic review and meta-analysis. *Frontiers in Aging Neuroscience*, 14, Article 960364. <https://doi.org/10.3389/fnagi.2022.960364>
- Somnuk, P., Limprapassorn, P., Srinonprasert, V., Wongviriyawong, T., Suraarunsumrit, P., Morkphrom, E., Sura-Amonrattana, U., Phannarus, H., Choorerk, D., Radtke, F. M., & Chaiwat, O. (2022). The Thai version of the nursing delirium screening scale-Thai: Adaptation and validation study in postoperative patients. *Frontiers in Medicine*, 9, Article 956435. <https://doi.org/10.3389/fmed.2022.956435>
- Sullivan, L. M., Massaro, J. M., & D'Agostino, R. B., Sr. (2004). Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in Medicine*, 23(10), 1631–1660. <https://doi.org/10.1002/sim.1742>
- Todd, O. M., Gelrich, L., MacLulich, A. M., Driessen, M., Thomas, C., & Kreisel, S. H. (2017). Sleep disruption at home as an independent risk factor for postoperative delirium. *Journal of the American Geriatrics Society*, 65(5), 949–957. <https://doi.org/10.1111/jgs.14685>
- van Smeden, M., de Groot, J. A., Moons, K. G., Collins, G. S., Altman, D. G., Eijkemans, M. J., & Reitsma, J. B. (2016). No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology*, 16(1), Article 163. <https://doi.org/10.1186/s12874-016-0267-3>
- Wilson, J. E., Mart, M. F., Cunningham, C., Shehabi, Y., Girard, T. D., MacLulich, A. M. J., Slooter, A. J. C., & Ely, E. W. (2020). Delirium. *Nature Reviews Disease Primers*, 6(1), Article 90. <https://doi.org/10.1038/s41572-020-00223-4>
- Witlox, J., Eurelings, L. S., de Jonghe, J. F., Kalisvaart, K. J., Eikelenboom, P., & van Gool, W. A. (2010). Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: A meta-analysis. *JAMA*, 304(4), 443–451. <https://doi.org/10.1001/jama.2010.1013>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::aid-cnrcr2820030106>3.0.co;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cnrcr2820030106>3.0.co;2-3)
- Zhou, B., Wang, A., & Cao, H. (2024). Risk prediction models for postoperative delirium in elderly patients with fragility hip fracture: A systematic review and critical appraisal. *International Journal of Orthopaedic and Trauma Nursing*, 52, Article 101077. <https://doi.org/10.1016/j.ijotn.2023.101077>

