BMJ Open

Validation of a delirium predictive model in patients admitted to surgical intensive care units: a multicentre prospective observational cohort study

Onuma Chaiwat [ID],[1,2] Kaweesak Chittawatanarat,[3,4] Sirirat Mueankwan,[5] Sunthiti Morakul,[6] Pitchaya Dilokpattanamongkol,[7] Chayanan Thanakiattiwibun,[2] Arunotai Siriussawakul[1,2]

For numbered affiliations see end of article.

Correspondence to
Dr Onuma Chaiwat;
onuma.cha@mahidol.ac.th

ABSTRACT

Objective To internally and externally validate a delirium predictive model for adult patients admitted to intensive care units (ICUs) following surgery.

Design A prospective, observational, multicentre study.

Setting Three university-affiliated teaching hospitals in Thailand.

Participants Adults aged over 18 years were enrolled if they were admitted to a surgical ICU (SICU) and had the surgery within 7 days before SICU admission.

Main outcome measures Postoperative delirium was assessed using the Thai version of the Confusion Assessment Method for the ICU. The assessments commenced on the first day after the patient's operation and continued for 7 days, or until either discharge from the ICU or the death of the patient. Validation was performed of the previously developed delirium predictive model: age+(5×SOFA)+(15×benzodiazepine use)+(20×DM)+(20×mechanical ventilation)+(20×modified IQCODE>3.42).

Results In all, 380 SICU patients were recruited. Internal validation on 150 patients with the mean age of 75±7.5 years resulted in an area under a receiver operating characteristic curve (AUROC) of 0.76 (0.683 to 0.837). External validation on 230 patients with the mean age of 57±17.3 years resulted in an AUROC of 0.85 (0.789 to 0.906). The AUROC of all validation cohorts was 0.83 (0.785 to 0.872). The optimum cut-off value to discriminate between a high and low probability of postoperative delirium in SICU patients was 115. This cut-off offered the highest value for Youden's index (0.50), the best AUROC, and the optimum values for sensitivity (78.9%) and specificity (70.9%).

Conclusions The model developed by the previous study was able to predict the occurrence of postoperative delirium in critically ill surgical patients admitted to SICUs.

Trial registration number Thai Clinical Trail Registry (TCTR20180105001).

BACKGROUND

Delirium, a disturbance of consciousness, is both acute and fluctuating. Delirium is an extremely common condition among hospitalised patients. Its incidence varies with

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The developed delirium predictive model consists of six risk factors was able to predict the occurrence of postoperative delirium in critically ill surgical patients.
⇒ The internal and external validation demonstrated moderate to good statistical performance, with the area under a receiver operating characteristic curve being comparable to that of the development cohort.
⇒ The optimum cut-off value to discriminate between a high and low probability of postoperative delirium in surgical intensive care unit patients was 115.

the study population, but higher rates are observed among geriatric, postsurgical, intensive care unit (ICU), cardiac surgery and hip-fracture patients.[1–4] Postoperative delirium (POD) among patients who have been treated with surgery and anaesthesia is typically found during the first 3 postoperative days.[5] Although the POD can be transient, it is linked to poor outcomes. These include long stays in postanesthesia care units, ICUs and hospitals; high medical-complication rates; and raised mortality levels.[6]

Several tools for assessing delirium have been validated. Among those is the Confusion Assessment Method for the ICU (CAM-ICU), which shows high sensitivity and specificity.[7] The CAM-ICU has been translated into Thai, and it, too, has demonstrated good sensitivity and specificity for critically ill patients.[8] In Thailand, there are limited data relating to POD as well as delirium among critically ill patients. Muangpaisan *et al*[9] reported the incidence of delirium was 22.5% in hip surgery. Their investigation also identified the following risk factors: age, premorbid function, dementia/cognitive impairment, the non-stop administration of non-steroidal anti-inflammatory drugs and postoperative

sedative use. Another study reported a 44.0% prevalence of delirium among critically ill, old patients at a medical ICU in northeastern Thailand. That work found that the independent factors related to delirium were the use of physical restraints, a history of stroke and multiple bed changes.[10]

Given that delirium can result in poor clinical outcomes, predictions of its occurrence among patients who are at risk of delirium are especially important. During the recent decade, some predictive scoring systems for delirium have been proposed for use with various populations. For instance, the PREdiction of DELIRIum in ICU patients delirium risk prediction tool was developed for intensive care patients.[11] This model uses 10 parameters. It had an area under a receiver operating characteristic curve (AUROC) of 0.87 (95% CI 0.85 to 0.89). Temporal validation and external validation resulted in an AUROC of 0.89 (0.86 to 0.92) and 0.84 (0.82 to 0.87), respectively.[11] Another tool, the Risk Model for Delirium, assesses a number of predisposing risk factors for delirium in hip fracture patients. This model showed good intraclass correlation coefficient (0.77), sensitivity (80.4%) and AUROC (0.73).[12] Furthermore, Kim et al developed the DELirium Prediction based on Hospital Information (Delphi) system for general surgery patients. Delphi demonstrated good AUROCs for both the developed (0.91) and validated models (0.98).[13] Nevertheless, each of the above models was developed for specific application with medical critically ill, general surgical or particular orthopaedic patients, and the scoring systems tend to be overly complicated.

The Siriraj Integrated Perioperative Geriatric Excellent Research Center has studied the incidence, risk factors and predictive scores of POD in critically ill surgical patients. The independent risk factors for delirium identified by a multivariate analysis were age, diabetes mellitus, severity of disease (assessed by the sequential organ failure assessment (SOFA) score), perioperative use of benzodiazepine, mechanical ventilation and dementia defined by the Thai version of the Modified Informant Questionnaire on Cognitive Decline in the Elderly (modified IQCODE) scores >3.42. The following predictive model was created:

$$\text{Age} + (5 \times \text{SOFA}) + (15 \times \text{benzodiazepine use}) +$$
$$(20 \times \text{DM}) + (20 \times \text{mechanical ventilation}) +$$
$$(20 \times \text{modified IQCODE} > 3.42)$$

Its AUROC was 0.84 (95% CI 0.786 to 0.897). A cut-off value of 125 demonstrated a sensitivity of 72.1% and a specificity of 80.9%.[14] Thus, we were interested in validating the model. To this end, internal validation was performed at our hospital, while external validation was conducted at two other academic hospitals. There has been no previous investigation of a predictive model for POD in patients in surgical ICUs (SICUs). The aim of this study was to validate the use of the proposed POD predictive scoring tool in SICUs in order to identify patients who tend to develop delirium.

METHODS

Study esign

This was a multicentre prospective, observational, cohort study.

Study population

The study was conducted on 380 SICU patients at three hospitals: Siriraj, Ramathibodi and Maharaj Nakorn Chiang Mai.

The study population comprised patients who were at least 18 years of age and were admitted to an SICU within 7 days of surgery at Siriraj, Ramathibodi or Maharaj Nakorn Chiang Mai Hospital (table 1). In addition, patients for the internal validation cohort were 65 years or older and had been admitted to a Siriraj Hospital SICU[15 16] for a stay anticipated to exceed 24 hours. At all three hospitals, we excluded SICU patients who had (1) not undergone any operations; (2) communication problems (unable to communicate in Thai, or having a severe visual or auditory impairment interfering with communication); or (3) a Richmond Agitation Sedation Scale (RASS) score of −4 or −5 during the whole of their ICU stay. A flowchart illustrating the patient selection processes for the development and validation cohorts is presented in figure 1.

Patient and public involvement

No patient involved.

Table 1 Characteristics of participating hospitals

Participating hospital	Institution	ICU beds for adults	ICU population	CAM-ICU screenings
Siriraj Hospital	Faculty of Medicine Siriraj Hospital, Mahidol University	14 beds	Surgery	2/day; IRR>0.8
Ramathibodi Hospital	Faculty of Medicine Ramathibodi Hospital, Mahidol University	12 beds	Surgery	2/day; IRR>0.8
Maharaj Nakorn Chiang Mai Hospital	Faculty of Medicine, Chiang Mai University	7 beds	Surgery	2/day; IRR not measured

CAM, Confusion Assessment Method; ICU, intensive care unit; IRR, inter-rater reliability expressed as Cohen's κ.
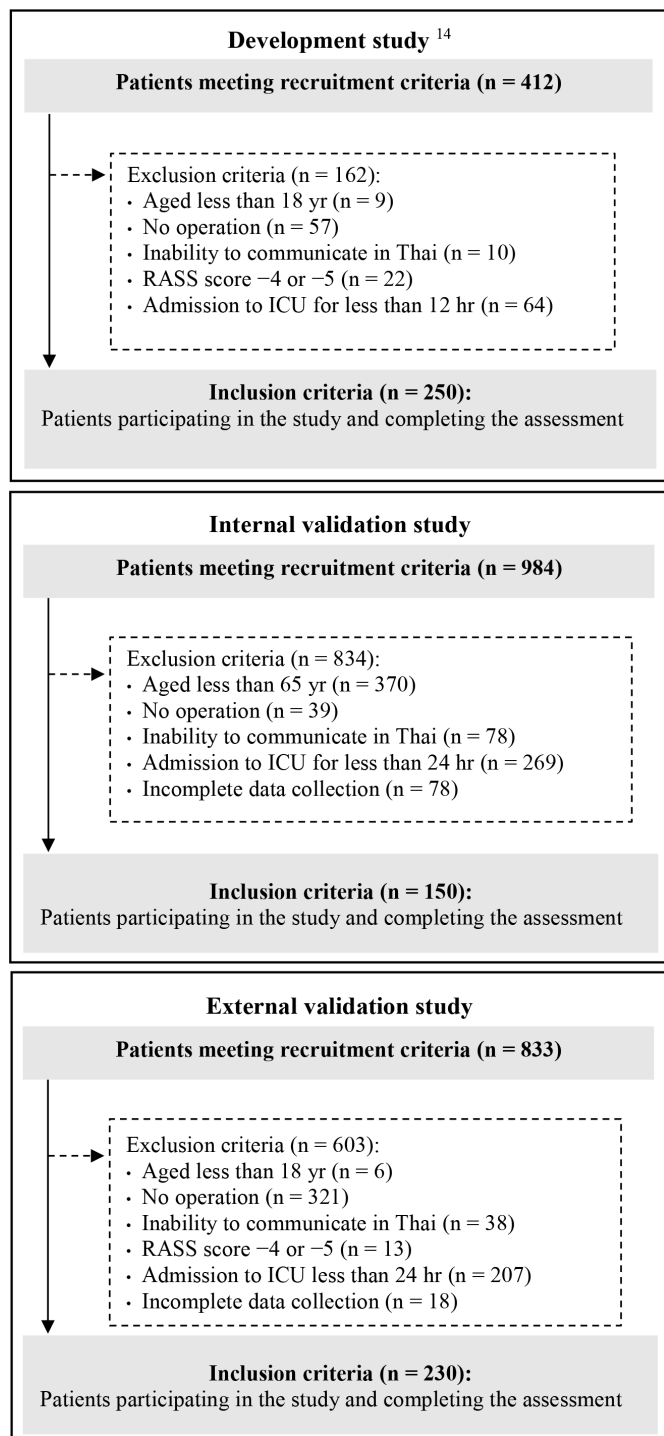
Figure 1 Flowchart of development and validation studies. ICU, intensive care unit; RASS, Richmond Agitation Sedation Scale.

Measurement instruments

Delirium was assessed using the Thai version of the CAM-ICU (online supplemental file S1). Delirium was identified by the following four features: (1) a change or fluctuation in baseline mental status; (2) inattention; and either (3) an altered level of consciousness; or (4) disorganised thinking.[17] The Thai version has demonstrated satisfactory validity and reliability (specificity, 94.7%;

sensitivity, 92.3%).[8] As to the level of consciousness, it was assessed by the RASS. It uses a 10-point scale ranging from −5 to +4. The delirium subtypes were recorded as hypoactive (RASS −1 to −3), hyperactive (RASS +1 to +4) and mixed type (hypoactive and hyperactive).[18] With regard to dementia, it was evaluated via the Thai version of modified IQCODE (online supplemental file S2). The questionnaire consists of 32 items, with assessments of patients being made by their caregivers. The optimal cut-off score for the modified IQCODE is 3.42 (sensitivity, 90%; specificity, 95% and accuracy, 92%).[19] Finally, the severity of illness at SICU admission was evaluated using the Acute Physiology and Chronic Health Evaluation II (APACHE II) scoring system, and SOFA scores.

Patients provided informed consent in writing. Delirium was evaluated at least two times per day (once during the 12 hours from 06.00, and once during the 12 hours after 18.00), and whenever patients developed a mental change. Delirium was screened routinely using a two-step process. Initially, the patients' level of consciousness was assessed using the RASS. If the score was between −3 and +4, the evaluators proceeded to step 2 (assessment of the patient with the Thai version of CAM-ICU). However, if step 1 produced a −4 RASS score (responsive only to physical stimulus) or a −5 RASS score (unresponsive to physical and verbal stimulus), step 2 was not performed. If a patient was found to be sedated in the first step, the dose of the sedative medication was adjusted. The patient was later assessed with the CAM-ICU once a RASS score of −3 or higher was achieved.

The second step involved the determination of the patient's delirium level using the Thai version of CAM-ICU, employing standard methodology. The assessments commenced on the first day after the patient's operation and continued for 7 days, or until either discharge from the ICU or the death of the patient. Patients with delirium were further assessed until the CAM-ICU was negative for 24 hours. Thereafter, the ICU attending physician was notified for further management.

Data collection

The predisposing and precipitating factors potentially linked to the onset of delirium were grouped as preoperative, intraoperative and postoperative variables. The preoperative risk factors were demographic variables obtained from a review of an individual patient's medical records and interviews with any proxies. Each patient's cognitive status was measured using the modified IQCODE.[19]

The intraoperative variables were obtained from anaesthetic records. They consisted of the surgical type (abdominal, vascular, orthopaedic, urological, gynaecological, and head and neck); admission type (emergency or elective); operation time; intraoperative blood loss; amount of blood transfused; and total fluid intake. Intraoperative hypotension was deemed to be either a systolic pressure below 90 mm Hg or the need to be treated with medications.[20 21] Intraoperative hypoxaemia was defined

as an oxygen saturation (derived from pulse oximetry) of below 90% for any duration.

The postoperative variables were primarily obtained from the SICU data records. They were the use of mechanical ventilation, physical restraints or a Foley's catheter; the presence of sleep deprivation or shock; exposure to psychoactive drugs (benzodiazepines, opioids and sedatives); and the presence of coma (indicated by a RASS score of −4 or −5).

Preparation of research team

The clinical researchers administering the Thai CAM-ICU were physicians and nurses who had been trained by the principal investigator. To ensure reliability among the assessors, inter-rater reliability scores were calculated. Once their kappa score reached 0.8, the trained physicians and nurses were permitted to perform the Thai CAM-ICU assessments.

Internal and external validation

After development of a predictive model from a prospective cohort study that took place between February 2016 and February 2017, we did a second prospective cohort study in the same hospital for internal validation of the model between April 2018 and December 2019. In the meantime, we externally validated the predictive model with data from intensive care surgical patients admitted to two other university hospitals in Thailand. They were Ramathibodi Hospital, Mahidol University, and Maharaj Nakorn Chiang Mai Hospital, Chiang Mai University. Trained intensive care nurses at those hospitals used the CAM-ICU at least two times per day. The validation process was conducted according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis Statement,[22] a guideline specifically designed for the reporting of studies developing or validating a multivariable prediction model, whether for diagnostic or prognostic purposes.

Statistical analysis

The sample size was estimated based on the reported 78% accuracy of development predictive score.[14] Based on the estimated accuracy of 80% (p=0.80) and a 4% error (d=0.04), an 5% alpha (α=0.05), the sample size of 380 cases was calculated. The sample size calculation was estimated using PASS V.14 (NCSS, Kaysville, Utah, USA).

Demographic variables are presented as mean±SD or median (IQR) for continuous data, and frequency and percentage for categorical data.

In both validation studies, we multiplied regression coefficients for each risk factor in the predictive model by the observed patients' values. The outcome was a calculated predicted probability, on which we built a new AUROC. Finally, an ROC curve was plotted to determine the best cut-off in terms of Youden's index, sensitivity, specificity and 95% CI. The Youden's index was the difference between the true and the false positive rates. Maximising this index allows an optimal cut-off value to

be found from the ROC curve, independently from the prevalence.[23 24] Finally, to examine how well the model was calibrated, we calculated linear predictor values for each patient of every cohort by using the coefficients from the model. We used these linear predictors in a logistic regression model to test whether the prediction rule was well calibrated, resulting in a calibration slope and an intercept. A calibration slope of 1 and an intercept of 0 show a perfect calibration.[25 26] Statistics were analysed using PASW Statistics for Windows (V.18; SPSS); and MedCalc statistical software (V.17.6; MedCalc Software BVBA, Ostend, Belgium).

RESULTS
Patients

The patients were enrolled between February 2016 and February 2017[14] for the development cohort, and between April 2018 and December 2019 for the internal and external validation studies. In all, 1437 SICU patients were excluded for the reasons given in figure 1 and 380 were recruited. The mean age of the patients in the internal validation cohort was 75.1±7.5 years, while the mean for the patients in the external validation cohort was 56.9±17.3 years. The mean age of all of the patients in the two validation cohorts was 64.1±16.8 years. More than half of the patients in the validation cohort were males. Details relating to the demographic and intraoperative data, ICU admission and the medications used are given in table 2. There was a higher proportion of patients with hypertension, diabetes mellitus (DM) and cardiac disease in the internal validation cohort than the external validation cohort. The incidence of delirium was 40.0%, 21.3% and 28.7% in the internal, external, and all validation cohorts, respectively, compared with 24.4% in the development cohort. The majority of patients in all cohorts underwent intra-abdominal surgery. The median SOFA score was 4 (IQR 1–6) for all validation cohorts, which was higher than the median of 3 (IQR 2–6) for the development cohort. The percentage of benzodiazepine use in all external validation cohort was less than half of the development cohort (10% vs 25.2%; table 2).

Development study

Of the 412 recruited patients, a total of 162 were excluded for the reasons detailed in figure 1. As a result, 250 patients were enrolled, 61 of whom (24.4%) developed delirium (table 2). The predictive model was derived from a multiple logistic regression that used significant risk factors. The final formula required six factors (two quantitative factors, and four binary factors). The formula of the model was:

$$\text{Age} + (5 \times \text{SOFA}) + (15 \times \text{benzodiazepine use}) + (20 \times \text{DM}) +$$

$$(20 \times \text{mechanical ventilation}) + (20 \times \text{modified IQCODE} > 3.42)$$

Table 2 Characteristics of patients in development and validation groups

Variable	Development (n=250)	Internal validation (n=150)	External validation (n=230)	All validation (n=380)
Demographic data				
Age (years)	64.2±16.4	75.1±7.5	56.9±17.3	64.1±16.8
Sex; male	121 (48.4%)	84 (56.0%)	128 (55.7%)	212 (55.8%)
Comorbidities				
Hypertension	155 (62.0%)	101 (67.3%)	109 (47.4%)	210 (55.3%)
DM	63 (25.2%)	41 (27.3%)	49 (21.3%)	90 (23.7%)
Cardiac disease	64 (25.6%)	37 (24.7%)	33 (14.3%)	70 (18.4%)
ESRD or CKD stage 4–5	34 (13.6%)	30 (20.0%)	75 (32.6%)	105 (27.6%)
Modified IQCODE score ≥3.42	16 (6.4%)	20 (13.3%)	27 (11.7%)	47 (12.4%)
Current alcohol consumption	17 (6.8%)	12 (8.0%)	41 (17.8%)	53 (13.9%)
Incidence of delirium	61 (24.4%)	60 (40.0%)	49 (21.3%)	109 (28.7%)
Onset of delirium (days)	1 (1–2)	2 (1–4)	1 (1–1)	1 (1–3)
Type of delirium				
Hypoactive	44 (72.1%)	16 (26.7%)	26 (53.1%)	42 (38.5%)
Hyperactive	9 (14.8%)	17 (28.3%)	6 (12.2%)	23 (21.1%)
Mixed	8 (13.1%)	27 (45%)	17 (34.7%)	44 (40.4%)
Intraoperative data				
Emergency surgery	108 (43.2%)	64 (42.7%)	73 (31.7%)	137 (36.1%)
Type of surgery				
Vascular	52 (20.8%)	43 (28.7%)	26 (11.3%)	69 (18.2%)
Intra-abdominal	88 (35.2%)	79 (52.7%)	81 (35.2%)	160 (42.1%)
Orthopaedic	29 (11.6%)	8 (5.3%)	10 (4.3%)	18 (4.7%)
Gynaecological	26 (10.4%)	1 (0.7%)	4 (1.7%)	5 (1.3%)
Other	55 (22.0%)	19 (12.7%)	109 (47.4%)	128 (33.7%)
Hypoxia	10 (4.0%)	8 (5.3%)	2 (0.9%)	10 (2.6%)
Intraoperative hypotension	196 (78.4%)	127 (84.7%)	93 (40.4%)	220 (57.9%)
ICU admission				
Sepsis	61 (24.4%)	39 (26.0%)	30 (13.0%)	69 (18.2%)
APACHE II score	9 (6–11)	14 (11–19)	12 (8–17)	12 (9–17)
SOFA score	3 (2–6)	4 (3–6)	4 (1–6)	4 (2–6)
Mechanical ventilation	185 (74.0%)	126 (84.0%)	153 (66.5%)	279 (73.4%)
Medication				
Benzodiazepine	63 (25.2%)	19 (12.7%)	19 (8.3%)	38 (10.0%)
Opioid	244 (97.6%)	140 (93.3%)	203 (88.3%)	343 (90.3%)
Outcomes				
Duration of mechanical ventilation (days)	1 (0–4)	3 (1–7)	1 (0–3)	2 (1–4)
Nosocomial infection in ICU	29 (11.6%)	41 (27.3%)	4 (1.7%)	45 (11.8%)
ICU length of stay (days)	3 (2–5)	5 (3–10)	3 (3–6)	4 (3–7)
ICU mortality	9 (3.6%)	5 (3.3%)	5 (2.2%)	10 (2.6%)
Hospital length of stay (days)	16 (10–29)	20 (13–34)	16 (13–27)	18 (13–29)
Hospital mortality	26 (10.4%)	25 (16.7%)	13 (5.7%)	38 (10.0%)

Data are presented as mean±SD, median (IQR) or n (%).
APACHE II score, Acute Physiology and Chronic Health Evaluation II score; CKD, chronic kidney disease; DM, diabetes mellitus; ESRD, end state renal disease; Modified IQCODE, Modified Informant Questionnaire on Cognitive Decline in the Elderly; SOFA score, Sequential Organ Failure Assessment score.
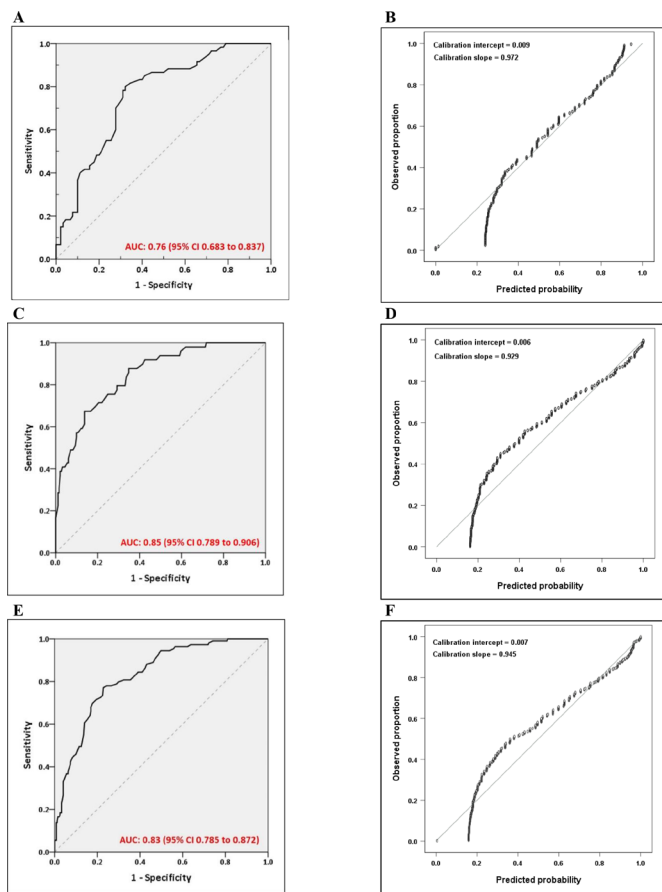
Figure 2 Receiver operating characteristic (ROC) curves and calculated areas under the curve (AUC), and Calibration plot of pooled data. (A and B) Internal validation study of the delirium predictive score. (C and D) External validation study of the delirium predictive score. (E and F) All validation study of the delirium predictive score.

The AUROC was 0.84 (95% CI 0.786 to 0.897). The cut-off value of ≥125 demonstrated a sensitivity of 72.1% and a specificity of 80.9%.[14]

Validation study

Internal validation of predictive model

For the prospective validation study, we recruited 984 consecutive patients who were aged over 65 years; however, 834 were subsequently excluded (figure 1). Of the remaining 150 patients, 60 (40%) developed delirium (table 2). The internal validation resulted in an AUROC of 0.76 (0.683 to 0.837; figure 2A), and this AUROC was not significantly different from the AUROC of the developed predictive model (p=0.092), with a calibration slope of 0.972 and an intercept of 0.009 (figure 2B).

External validation of predictive model

We performed the external validation study on critically ill surgical patients admitted to SICUs at Ramathibodi and Maharaj Nakorn Chiang Mai Hospitals. Of the 833 recruited patients, 603 were excluded (figure 1). As a result, 230 patients were enrolled: 62 (27%) at Ramathibodi Hospital, and 168 (73%) at Maharaj Nakorn Chiang

Mai Hospital. The incidence of delirium in the external validation cohort was 21% (table 2). The external validation resulted in an AUROC of 0.85 (0.789 to 0.906; figure 2C), and it was not significantly different from the AUROC of the developed predictive model (p=0.865), with a calibration slope of 0.929 and an intercept of 0.006 (figure 2D).

Optimal cut-off value of predictive model

The AUROCs of the development, internal and external validation cohorts were comparable (0.84 for the development cohort, 0.76 for the internal validation cohort and 0.85 for the external validation cohort). As no significant differences in prediction existed between the three validation studies, we pooled the data of all validation cohorts (n=380). That revealed that 109 patients (29%) developed delirium (table 2). Consequently, the AUROC of all of the validation cohorts was 0.83 (0.785 to 0.872; figure 2E). The recalibration of all validation study showed a calibration slope of 0.945 and an intercept of 0.007 (figure 2F). The optimum cut-off value to discriminate between a high and low probability of POD in SICU patients was 115. This cut-off presented the highest value of Youden's index (0.50), the best AUROC, and the optimum values for sensitivity (78.9%) and specificity (70.9%; table 3). The last two values were similar to the sensitivity (78.8%) and specificity (70.4%) of the development cohort.

DISCUSSION

Given the high costs of managing delirium and its consequential complications, it is essential to identify individuals at high risk of developing the condition and to deliver evidenced-based preventive measures. This multicentre study demonstrated the performance of the internal and external validation of a proposed model[14] that had been developed to predict POD in patients admitted postoperatively to an SICU. It is essential to confirm the predictive performance of the model before its use outside the development setting. The external validation showed moderate to good statistical performance, with the AUROC of the external cohort being comparable to that of the development cohort. In addition, the new cut-off value also demonstrated optimum sensitivity and specificity values that were equivalent to those achieved for the development cohort. However, the performance of the internal validation cohort was not as high as the development and external validation cohort (AUROC, 0.76). This was because the internal validation cohort only included patients aged 65 years or older, resulting in a higher incidence of delirium.

Recently, two ICU delirium predictive models-the early predictive model for ICU delirium (E-PRE-DELIRIC), and the recalibrated predictive model for ICU delirium (PRE-DELIRIC) have been developed and validated.[11 27 28] These two models are currently used in clinical practice and in research to predict the development of delirium in

Table 3 Cut-offvalues of the delirium prediction score

Cut-off value	Development (n=250)			Internal validation (n=150)			External validation (n=230)			All validation (n=380)		
	Sensitivity (95% CI)	Specificity (95% CI)	J	Sensitivity (95% CI)	Specificity (95% CI)	J	Sensitivity (95% CI)	Specificity (95% CI)	J	Sensitivity (95% CI)	Specificity (95% CI)	J
≥100	95.1% (86.3 to 99.0)	50.3% (42.9 to 57.6)	0.45	100.0% (94.0 to 100.0)	21.1% (13.2 to 31.0)	0.21	87.8% (75.2 to 95.4)	64.6% (57.2 to 71.6)	0.52	94.5% (88.4 to 98.0)	50.2% (44.1 to 56.3)	0.45
≥105	90.2% (79.8 to 96.3)	56.1% (48.7 to 63.3)	0.46	96.7% (88.5 to 99.6)	27.8% (18.9 to 38.2)	0.25	79.6% (65.7 to 89.8)	66.7% (59.5 to 73.7)	0.46	89.0% (81.6 to 94.2)	53.9% (47.7 to 59.9)	0.43
≥110	83.6% (71.9 to 91.9)	63.0% (55.7 to 69.9)	0.47	90.0% (79.5 to 96.2)	34.4% (24.7 to 45.2)	0.24	75.5% (61.1 to 86.7)	75.1% (68.2 to 81.3)	0.51	83.5% (75.2 to 89.9)	61.6% (55.6 to 67.4)	0.45
≥115	78.7% (66.3 to 88.1)	70.4% (63.3 to 76.8)	0.49	86.7% (75.4 to 94.1)	50.0% (39.3 to 60.7)	0.37	69.4% (54.6 to 81.8)	81.2% (74.8 to 86.7)	0.51	78.90% (70.0 to 86.1)	70.9% (65.1 to 76.2)	0.50
≥120	75.4% (62.7 to 85.5)	74.1% (67.2 to 80.2)	0.50	83.3% (71.5 to 91.7)	60.0% (49.1 to 70.2)	0.47	61.2% (46.2 to 74.8)	86.2% (80.3 to 90.9)	0.47	73.3% (64.1 to 81.4)	77.5% (72.1 to 82.3)	0.50
≥125	72.1% (59.2 to 82.9)	81.0% (74.6 to 86.3)	0.53	78.3% (65.8 to 87.9)	68.9% (58.3 to 78.2)	0.47	55.1% (40.2 to 69.3)	90.1% (84.7 to 94.0)	0.45	67.9% (58.3 to 76.5)	83.0% (78.0 to 87.3)	0.50
≥130	67.2% (54.0 to 78.7)	87.3% (81.7 to 91.7)	0.54	61.7% (48.2 to 73.9)	72.2% (61.8 to 81.2)	0.34	46.9% (32.5 to 61.7)	93.4% (88.7 to 96.5)	0.40	55.1% (45.2 to 64.6)	86.4% (81.7 to 90.2)	0.42

J, Youden's index.

ICUs. The PRE-DELIRIC model consists of 10 predictors that are available during the first 24 hours after admission to an ICU.[27] The E-PRE-DELIRIC is composed of nine parameters available at the time of ICU admission. Wassenaar et al[29] recently conducted an external validation of both assessment tools, using either the CAM-ICU or the Intensive Care Delirium Screening Checklist for delirium assessment. The researchers reported moderate-to-good statistical performances. Nevertheless, the formulas for those two models were quite complicated, using several parameters, and they were developed in a mixed-ICU setting (medical and surgical populations). Given that cognitive impairment (including dementia) and severity of illness have been recognised as strong predictors for delirium in hospitalised patients,[30 31] the E-PRE-DELIRIC system included only a history of cognitive impairment but no severity scores. In contrast, the PRE-DELIRIC model included only APACHE II scores, but no information on cognitive impairment.

The currently proposed predictive model for POD in critically ill surgical patients has several strengths. First, it was developed specifically for surgical patients, and it demonstrated high accuracy. In addition, it employs only five parameters, which makes it relatively easy to calculate. Furthermore, dementia is assessed by both the patient's history and the modified IQCODE assessment tool. A previous study found that the prevalence of dementia among elderly delirious patients was five times higher when evaluated by the modified IQCODE tool than when using information obtained solely from history taking.[32] Consequently, the proposed predictive model was validated in the same hospital and in two other academic hospitals. Although we recruited only elderly patients for the internal validation cohort, the AUROC showed an acceptable value. For the external validation cohort in the SICUs of the two other hospitals, we performed quality control by determining the inter-rater reliability of CAM-ICU assessment before commencing the study. There were differences in the patient case-mix of the external and development validation samples. In particular, relative to the development group, the external validation cohort had a lower age, a lower percentage of patients with mechanical ventilation, a higher percentage of dementia, and a lower percentage of benzodiazepine use. Despite that, the models' discriminative performance showed the same value (AUROC 0.84 for the development cohort, and 0.85 for the external validation cohort). In short, for the all-validation cohort, the AUROC was approximately the same as that for the development and the external validation cohorts. A score of ≥115 was the best cut-off value to predict the occurrence of delirium in SICUs. This cut-off presented the highest value for Youden's index (0.50), the best AUROC, and the optimum values for sensitivity (78.9%) and specificity (70.9%). Additionally, the predictive value depends on a disease's prevalence in the population group that is being diagnosed.[33] A good model must have sufficient prevalence, high sensitivity and high specificity, and it should allow diagnosis before a patient displays symptoms.[33 34]

STRENGTHS AND LIMITATIONS

The significant strength of our study is that it was the first multicentre study in Thailand to evaluate the performance of a proposed predictive model for delirium in SICUs. The early prediction of the development of delirium in ICU patients facilitates the implementation of prevention protocols. These interventions can be non-pharmacological (such as cognitive stimulation, early mobilisation, and enhanced sleep)[35 36] or pharmacological (like the prophylactic administration of dexmedetomidine[37] to high-risk patients).

Several limitations need to be addressed. First, only the CAM-ICU was used to assess delirium. In the current work, the researchers (physicians and nurses) who evaluated delirium using this tool were well-trained, and their ratings are therefore regarded as accurate. However, other research showed that the accuracies of delirium assessments performed by bedside nurses in daily practice demonstrated lower sensitivity and specificity than our clinical researchers achieved.[38] The skill level of staff undertaking assessments in a clinical setting may therefore influence the results of the predictive model. In addition, the internal validation cohort only included critically ill elderly patients. The optimum cut-off value that resulted in the best sensitivity and specificity might be different from the all-validation and development cohorts. Moreover, differences in risk factors might affect the predictive model. We did not perform a logistic regression for the validation cohort in order to identify independent risk factors for delirium. This is because the prognostic ability demonstrated by the AUROC of the internal and external validation groups showed moderate-to-good performance. Finally, the predictive model only used parameters available at the time of SICU admission. Any changes in patients' conditions during their stay can affect the probability of their developing delirium. Our model did not account for such changes.

CONCLUSIONS

The model reported in this study can predict which critically ill surgical patients will develop POD in SICUs. Consequently, high-risk patients can be identified, and both non-pharmacological and pharmacological prevention protocols can be implemented to improve the clinical outcomes. The use of this selective strategy is appropriate in a resource-limited country, in which the administration of a prevention protocol for all critically ill patients is not viable.

Author affiliations
[1]Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[2]Integrated Perioperative Geriatric Excellent Research Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

[3]Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Clinical Surgical Research Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Surgical Critical Care Unit, Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[6]Department of Anesthesiology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

[7]Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was conducted according to the ethical standards established by the 1964 Declaration of Helsinki. The study was approved by the Siriraj Institutional Review Board of the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (Si 623/2017, Chairperson Professor Chairat Shayakul, MD) on 20 October 2017; the Committee on Human Rights Related to Research Involving Human Subjects, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand (MURA 2017/574, Chairperson Assistant Professor Chusak Okascharoen, MD) on 27 November 2017; and Research Ethics Committee 4 at the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand (SUR-2560-05016, Chairperson Emeritus Professor Panja Kulapongs, MD) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD
Onuma Chaiwat http://orcid.org/0000-0001-5407-306X

REFERENCES

1 Marcantonio ER, Goldman L, Mangione CM, *et al*. A clinical prediction rule for delirium after elective noncardiac surgery. *JAMA* 1994;271:134–9.
2 Serafim RB, Dutra MF, Saddy F, *et al*. Delirium in postoperative nonventilated intensive care patients: risk factors and outcomes. *Ann Intensive Care* 2012;2:51–6.
3 Pandharipande P, Cotton BA, Shintani A, *et al*. Prevalence and risk factors for development of delirium in surgical and trauma intensive care unit patients. *J Trauma* 2008;65:34–41.
4 Schenning KJ, Deiner SG. Postoperative delirium: a review of risk factors and tools of prediction. *Curr Anesthesiol Rep* 2015;5:48–56.
5 Whitlock EL, Vannucci A, Avidan MS. Postoperative delirium. *Minerva Anestesiol* 2011;77:448–56.
6 Leslie DL, Marcantonio ER, Zhang Y, *et al*. One-Year health care costs associated with delirium in the elderly population. *Arch Intern Med* 2008;168:27–32.
7 Gusmao-Flores D, Salluh JIF, Chalhub Ricardo Ávila, *et al*. The confusion assessment method for the intensive care unit (CAM-ICU) and intensive care delirium screening checklist (ICDSC) for the diagnosis of delirium: a systematic review and meta-analysis of clinical studies. *Crit Care* 2012;16:R115.
8 Pipanmekaporn T, Wongpakaran N, Mueankwan S, *et al*. Validity and reliability of the Thai version of the confusion assessment method for the intensive care unit (CAM-ICU). *Clin Interv Aging* 2014;9:879–85.
9 Muangpaisan W, Wongprikron A, Srinonprasert V, *et al*. Incidence and risk factors of acute delirium in older patients with hip fracture in Siriraj Hospital. *J Med Assoc Thai* 2015;98:423–30.
10 Limpawattana P, Panitchote A, Tangvoraphonkchai K, *et al*. Delirium in critical care: a study of incidence, prevalence, and associated factors in the tertiary care hospital of older Thai adults. *Aging Ment Health* 2016;20:74–80.
11 van den Boogaard M, Pickkers P, Slooter AJC, *et al*. Development and validation of PRE-DELIRIC (prediction of delirium in ICU patients) delirium prediction model for intensive care patients: observational multicentre study. *BMJ* 2012;344:e420.
12 Moerman S, Tuinebreijer WE, de Boo M, *et al*. Validation of the risk model for delirium in hip fracture patients. *Gen Hosp Psychiatry* 2012;34:153–9.
13 Kim MY, Park UJ, Kim HT, *et al*. Delirium prediction based on hospital information (Delphi) in general surgery patients. *Medicine* 2016;95:e3072–7.
14 Chaiwat O, Chanidnuan M, Pancharoen W, *et al*. Postoperative delirium in critically ill surgical patients: incidence, risk factors, and predictive scores. *BMC Anesthesiol* 2019;19:39.
15 Robinson TN, Raeburn CD, Tran ZV, *et al*. Postoperative delirium in the elderly: risk factors and outcomes. *Ann Surg* 2009;249:173–8.
16 Böhner H, Hummel TC, Habel U, *et al*. Predicting delirium after vascular surgery: a model based on pre- and intraoperative data. *Ann Surg* 2003;238:149–56.
17 Ely EW, Margolin R, Francis J, *et al*. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). *Crit Care Med* 2001;29:1370–9.
18 Peterson JF, Pun BT, Dittus RS, *et al*. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc* 2006;54:479–84.
19 Siri S, Okanurak K, Chansirikanjana S, *et al*. Modified informant questionnaire on cognitive decline in the elderly (IQCODE) as a screening test for dementia for Thai elderly. *Southeast Asian J Trop Med Public Health* 2006;37:587–94.
20 Brady K, Hogue CW. Intraoperative hypotension and patient outcome: does "one size fit all?". *Anesthesiology* 2013;119:495–7.
21 Walsh M, Devereaux PJ, Garg AX, *et al*. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. *Anesthesiology* 2013;119:507–15.
22 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med* 2015;13:1.
23 Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5.
24 Galley HF. Editorial II: solid as a ROC. *Br J Anaesth* 2004;93:623–6.
25 Steyerberg EW, Eijkemans MJ, Harrell FE, *et al*. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;21:45–56.
26 Van Calster B, McLernon DJ, van Smeden M, *et al*. Calibration: the Achilles heel of predictive analytics. *BMC Med* 2019;17:230.
27 van den Boogaard M, Schoonhoven L, Maseda E, *et al*. Recalibration of the delirium prediction model for ICU patients (PRE-DELIRIC): a multinational observational study. *Intensive Care Med* 2014;40:361–9.
28 Wassenaar A, van den Boogaard M, van Achterberg T, *et al*. Multinational development and validation of an early prediction model for delirium in ICU patients. *Intensive Care Med* 2015;41:1048–56.
29 Wassenaar A, Schoonhoven L, Devlin JW, *et al*. External validation of two models to predict delirium in critically ill adults using either the confusion assessment Method-ICU or the intensive care delirium screening checklist for delirium assessment. *Crit Care Med* 2019;47:e827–35.

30 Inouye SK, Westendorp RGJ, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22.

31 Devlin JW, Skrobik Y, Gélinas C, *et al*. Clinical practice guidelines for the prevention and management of pain, Agitation/Sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med* 2018;46:e825–73.

32 To-adithep P, Wongviriyawong T, Thanakiattiwibun C. Prevalence of dementia among elderly surgical patients using Thai modified IQ code: a cross-sectional study. *J Med Assoc Thai* 2018;101:83.

33 Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet* 2002;359:881–4.

34 Knottnerus JA, van Weel C, Muris JWM. Evaluation of diagnostic procedures. *BMJ* 2002;324:477–80.

35 Hanison J, Conway D. A multifaceted approach to prevention of delirium on intensive care. *BMJ Qual Improv Rep* 2015;4:4–12.

36 Rivosecchi RM, Kane-Gill SL, Svec S, *et al*. The implementation of a nonpharmacologic protocol to prevent intensive care delirium. *J Crit Care* 2016;31:206–11.

37 Su X, Meng Z-T, Wu X-H, *et al*. Dexmedetomidine for prevention of delirium in elderly patients after non-cardiac surgery: a randomised, double-blind, placebo-controlled trial. *Lancet* 2016;388:1893–902.

38 van Eijk MM, van den Boogaard M, van Marum RJ, *et al*. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *Am J Respir Crit Care Med* 2011;184:340–4.