BMJ Open

Validation of a delirium predictive model in patients admitted to surgical intensive care units: a multicenter study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057890
Article Type:	Original research
Date Submitted by the Author:	03-Oct-2021
Complete List of Authors:	Chaiwat, Onuma; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Siriraj Integrated Perioperative Geriatric Excellent Research Center Chittawatanarat, Kaweesak ; Chiang Mai University Faculty of Medicine, Department of Surgery Mueankwan, Sirirat ; Chiang Mai University Faculty of Medicine, Department of Surgery Morakul, Sunthiti ; Mahidol University Faculty of Medicine Ramathibodi Hospital, Department of Anesthesiology Dilokpattanamongkol, Pitchaya ; Mahidol University Faculty of Pharmacy, Department of Pharmacy Thanakiattiwibun, Chayanan ; Mahidol University Faculty of Medicine Siriraj Hospital, Siriraj Integrated Perioperative Geriatric Excellent Research Center Siriussawakul, Arunotai; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Siriraj Integrated Perioperative Geriatric Excellent Research Center
Keywords:	Adult intensive & critical care < ANAESTHETICS, INTENSIVE & CRITICAL CARE, Delirium & cognitive disorders < PSYCHIATRY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Validation of a delirium predictive model in patients admitted to surgical intensive care units: a multicenter study

Onuma Chaiwat[1, 5*],

Kaweesak Chittawatanarat[2],

Sirirat Mueankwan[2],

Sunthiti Morakul[3],

Pitchaya Dilokpattanamongkol[4],

Chayanan Thanakiattiwibun[5],

Arunotai Siriussawakul[1,5]

***Correspondence:** onuma.cha@mahidol.ac.th

Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700 Thailand

Full list of author information is available at the end of the article

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, multicenter 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) within a week of undergoing an operation.

Main outcome measures 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 resulted in an area under a receiver operating characteristic curve (AUROC) of 0.76 (0.683 to 0.837). External validation on 230 patients 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 cutoff value to discriminate between a high and low probability of postoperative delirium in SICU patients was 115. This cutoff 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. Consequently, high-risk patients are able to be identified, and both nonpharmacological and pharmacological prevention protocols can be subsequently implemented to improve the clinical outcomes.

Keywords: postoperative delirium**,** predictive model, surgical intensive care units, validation

Strengths and limitations of this study

- The developed delirium predictive model consists of 6 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 AUROC being comparable to that of the development cohort
- The optimum cutoff value to discriminate between a high and low probability of POD in SICU patients was 115

BACKGROUND

Delirium, a disturbance of consciousness, is both acute and fluctuating. It is characterized by the lessened ability of an individual to focus, maintain, or shift attention. It is also associated with cognitive changes and disruptions in perception that are secondary to a general medical condition. Delirium is an extremely common condition among hospitalized patients. Its incidence varies with 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 anesthesia 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 (PACUs), 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.,* 2015 [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 nonstop administration of nonsteroidal 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 PRE-DELIRC (PREdiction of DELIRlum in ICU patients) delirium risk prediction tool was developed for intensive care patients [11]. This model utilizes 10 parameters. It had an area under a receiver operating characteristic curve (AUROC) of 0.87 (95% confidence interval [CI], 0.85 to 0.89). Temporal validation and external validation resulted in an AUROC of 0.89 (0.86–0.92) and 0.84 (0.82–0.87), respectively [11]. Another tool, the Risk Model for Delirium, assesses a number of predisposing risk factors for delirium in hip fracture patients. They are delirium during previous hospitalization; the presence of dementia; substandard performance in a clock-drawing test; advanced age; hearing impairment; visual impairment; the need for domestic help, assistance with meal preparation, or help with physical care; the use of heroin, methadone, or morphine; and the consumption of alcoholic beverages. 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 orthopedic patients, and the scoring systems tend to be overly complicated.

The Siriraj Integrated Perioperative Geriatric (SIPG) 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, and mechanical ventilation. The following predictive score 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, 0.786–0.897). A cutoff 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 2 other academic hospitals. There has been no previous investigation of a predictive model for POD in patients in surgical ICUs (SICUs). By identifying critically ill patients at high risk of developing POD, a model would enable the implementation of non-pharmacological and pharmacological preventive measures to avoid severe complications. The aim of this study was to validate the use of the proposed POD predictive scoring tool in SICUs.

METHODS

Design

A prospective, observational, cohort study was conducted. 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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Study population

The study was conducted on 380 SICU patients at 3 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 a SICU within 7 days of surgery at Siriraj, Ramathibodi, or Maharaj Nakorn Chiang Mai Hospital. 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 3 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.

Measurement instruments

Delirium was assessed using the Thai version of the CAM-ICU. Delirium was identified by the following 4 features: 1) a change or fluctuation in baseline mental status; 2) inattention; and either 3) an altered level of consciousness; or 4) disorganized 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 utilizes 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 (hypo- and hyperactive) [18]. With regard to dementia, it was evaluated via the Thai version of the Modified Informant Questionnaire on Cognitive Decline in the Elderly (modified IQCODE). The questionnaire consists of 32 items, with assessments of patients being made by their caregivers. The optimal cutoff score for the modified IQCODE is 3.42 (sensitivity, 90%; specificity, 95%; and accuracy, 92%) [19]. Lastly, 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.

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 anesthetic records. They consisted of the surgical type (abdominal, vascular, orthopedic, urological, gynecological, 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 mmHg or the need to be treated with medications. Intraoperative hypoxemia 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.

Patient assessments

Patients provided informed consent in writing. Delirium was evaluated at least twice daily (once during the 12 hours from 6.00 AM, and once during the 12 hours after 6.00 PM), and whenever patients developed a mental change. Delirium was screened routinely utilizing a 2-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.

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 2 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 twice daily (Table 1). The validation process was conducted according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement[20], a guideline specifically designed for the reporting of studies developing or validating a multivariable prediction model, whether for diagnostic or prognostic purposes.

Statistical analysis

Demographic variables are presented as mean ± standard deviation and median (interquartile range [IQR]) for continuous data, and frequency and percentage for categorical data. Group

comparisons were performed using the independent-samples t-test, Mann–Whitney U test, chi-squared test, or Fisher's exact test, as appropriate.

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 cutoff 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. Maximizing this index allows an optimal cutoff value to be found from the ROC curve, independently from the prevalence [21, 22]. 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 [23, 24]. Statistics were analyzed using PASW Statistics for Windows (version 18; SPSS Inc., Chicago, IL, USA); and MedCalc statistical software (version 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, 1,437 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 2 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 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 6 factors (2 quantitative factors, and 4 binary factors). The formula of the predictive model was:

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

(20 × mechanical ventilation) + (20 × modified IQCODE > 3.42)

The AUROC was 0.84 (95% CI, 0.786−0.897; Figure. 2A). The cutoff 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. 2B), with a calibration slope of 0.972 and an intercept of 0.009 (Figure. 3A).

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), with a calibration slope of 0.929 and an intercept of 0.006 (Figure. 3B).

Optimal cutoff 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; Figure. 2). As no differences in prediction existed between the 3 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. 2D). The recalibration of all validation study showed a calibration slope of 0.945 and an intercept of 0.007. (Fig 3C) The optimum cutoff value to discriminate between a high and low probability of POD in SICU patients was 115. This cutoff 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 2 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 multicenter-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 cutoff 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, 2 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, 25, 26]. These 2 models are currently used in clinical practice and in research to predict the development of delirium in ICUs. The PRE-DELIRIC model consists of 10 predictors that are available during the first 24 hours after admission to an ICU [25]. The E-PRE-DELIRIC is composed of 9 parameters available at time of ICU admission. Wassenaar *et al*., 2019 [27] 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 2 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 recognized as strong predictors for delirium in hospitalized patients, [28, 29] 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. Firstly, it was developed specifically for surgical patients, and it demonstrated high accuracy. In addition, it employs only 5 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 5 times higher when evaluated by the modified IQCODE tool than when using information obtained solely from history taking [30]. Consequently, the proposed predictive model was validated in the same hospital and in 2 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 2 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 cutoff value to predict the occurrence of delirium in SICUs. This cutoff 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 [31]. A good model must have sufficient prevalence, high sensitivity, and high specificity, and it should allow diagnosis before a patient displays symptoms [31, 32].

Strengths and limitations

The significant strength of our study is that it was the first multicenter 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 mobilization, and enhanced sleep [33, 34] or pharmacological (like the prophylactic administration of dexmedetomidine [35] to high-risk patients).

Several limitations need to be addressed. Firstly, 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 [36]. 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 cutoff 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. Lastly, 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]Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai Thailand.

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

[4]Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand.

[5]Siriraj Integrated Perioperative Geriatric Excellent Research Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Contributions

OC and AS contributed to the design of the study. OC, KC and S. Morakul were involved in data management and oversaw the project. KC, S. Mueankwan, S. Morakul, PD, contributed to data collection. CT contributed to data analysis. OC and CT contributed to the interpretation of the results and drafting the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand (IO: R016132015). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics statements

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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Acknowledgements

The authors gratefully acknowledge the patients who generously agreed to participate in this

study, and Assist. Prof. Dr. Chulaluk Komoltri, M.P.H. Biostatistics, for the statistical analyses.

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. https://doi:10.1001/jama.1994.03510260066030.
2. Serafim RB, Dutra MF, Saddy F, Tura B, de Castro JE, Villarinho LC, et al. Delirium in postoperative nonventilated intensive care patients: risk factors and outcomes. *Ann Intensive Care* 2012;2:1–6. https://doi: 10.1186/2110-5820-2-51.
3. Pandharipande P, Cotton BA, Shintani A., et al. Prevalence and risk factors for development of delirium in trauma and surgical ICU patients. *J Trauma* 2008;65:34–41. https://doi: 10.1097/TA.0b013e31814b2c4d.
4. Schenning KJ, Deiner SG. Postoperative delirium: a review of risk factors and tools of prediction. *Curr Anesthesiol Rep* 2015; 5: 48–56. https://doi:10.1007/s40140-014-0086-1
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. https://doi: 10.1001/archinternmed.2007.4.
7. Gusmao-Flores D, Salluh JI, Chalhub RÁ, 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. https://doi: 10.1186/cc11407.
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. https://doi: 10.2147/CIA.S62660.
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. https://doi: 10.1080/13607863.2015.1035695.
11. Van den Boogaard M, Pickkers P, Slooter AJC, et al. Development and validation of PRE-DELIRC (PREdiction of DELIRlum in ICU patients) delirium prediction model for intensive care patients: observational multicente study. *BMJ* 2012;344:e420. https://doi:10.1136/bmj.e420.
12. Moerman S, Tuinebreijer WE, Boo MD, Pilot P, Nelissen RG, Vochteloo AJ. Validation of the risk model for delirium in hip fracture patients. *Gen Hosp Psychiatry* 2012;34:153–9. https://doi: 10.1016/j.genhosppsych.2011.11.011.
13. Kim MY, Park UJ, Kim HT, et al. DELirium Prediction Based on Hospital Information (Delphi) in General Surgery Patients. *Medicine* 2016;95:2–7. https://doi: 10.1097/MD.0000000000003072.
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. doi: 10.1186/s12871-019-0732-8.
15. Robinson TN, Raeburn CD, Tran ZV, et al. Postoperative delirium in the elderly. risk factors and outcomes. *Ann Surg* 2009;249:173–8. https://doi: 10.1097/SLA.0b013e31818e4776.
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. https://doi: 10.1097/01.sla.0000077920.38307.5f.

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. https://doi: 10.1097/00003246-200107000-00012.

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. https://doi: 10.1111/j.1532-5415.2005.00621.x.

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. 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 Medicine.* 2015;13:1. https://doi.org/10.1186/s12916-014-0241-z

21. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5. https://doi: 10.1002/1097-0142(1950)3:1<32:aid-cncr2820030106>3.0.co;2-3.

22. Galley HF. Editorial II: Solid as a ROC. *Br J Anaesth* 2004;93:623–6. https://doi: 10.1093/bja/aeh247.

23. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, et al. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Mak Int J Soc Med Decis Mak* 2001;21:45–56.

24. Van Calster, B., McLernon, D.J., van Smeden, M. et al. Calibration: the Achilles heel of predictive analytics. *BMC Med* 2019;17. https://doi.org/10.1186/s12916-019-1466-7.

25. 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. https://doi: 10.1007/s00134-013-3202-7.

26. 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. https://doi: 10.1007/s00134-015-3777-2.

27. 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. https://doi: 10.1097/CCM.0000000000003911.
28. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22. https://doi: 10.1016/S0140-6736(13)60688-1.
29. 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. https://doi: 10.1097/CCM.0000000000003299.
30. To-adithep P, Wongviriyawong T, Thanakiattiwibun C, Chanidnuan M, Pancharoen W, Chaiwat O. Prevalence of Dementia among Elderly Surgical Patients using Thai Modified IQ CODE: A Cross-Sectional Study. *J Med Assoc Thai* 2018;101:83.
31. Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet* 2002;359:881–4. https://doi: 10.1016/S0140-6736(02)07948-5.
32. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ* 2002;324:477–80. htpps://doi: 10.1136/bmj.324.7335.477.
33. Hanison J, Conway D: A multifaceted approach to prevention of delirium on intensive care. *BMJ Qual Improv Rep* 2015;12:4–12. https://doi: 10.1136/bmjquality.u209656.w4000.
34. 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. https://doi:10.1016/j.jcrc.2015.09.031.

35. 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. *The Lancet* 2016;388:1893–1902. https://doi: 10.1016/S0140-6736(16)30580-3.

36. an 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. https://doi: 10.1164/rccm.201101-0065OC.

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

Abbreviations: CAM, Confusion Assessment Method; ICU, intensive care unit; IRR, inter-rater reliability expressed as Cohen's κ

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%)
Type of delirium				
Hypoactive	44 (72%)	16 (26.7%)	26 (11.3%)	42 (11.1%)
Hyperactive	9 (15%)	17 (28.3%)	6 (2.6%)	23 (6.1%)
Mixed	8 (13%)	27 (45%)	17 (7.4%)	44 (11.6%)
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%)
Orthopedic	29 (11.6%)	8 (5.3%)	10 (4.3%)	18 (4.7%)
Gynecological	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%)

Data are presented as mean ± SD, median (IQR), or n (%).

Abbreviations: 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

Table 3. Cutoff values of the Delirium Prediction Score

Cutoff value	Development (n = 250)		J	Internal validation (n = 150)		J	External validation (n = 230)		J	All validation (n = 380)		J
	Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)	
≥ 100	95.1% (86.3–99.0)	50.3% (42.9–57.6)	0.45	100.0% (94.0–100.0)	21.1% (13.2–31.0)	0.21	87.8% (75.2–95.4)	64.6% (57.2–71.6)	0.5	94.5% (88.4–98.0)	50.2% (44.1–56.3)	0.45
≥ 105	90.2% (79.8–96.3)	56.1% (48.7–63.3)	0.46	96.7% (88.5–99.6)	27.8% (18.9–38.2)	0.25	79.6% (65.7–89.8)	66.7% (59.5–73.7)	0.4	89.0% (81.6–94.2)	53.9% (47.7–59.9)	0.43
≥ 110	83.6% (71.9–91.9)	63.0% (55.7–69.9)	0.47	90.0% (79.5–96.2)	34.4% (24.7–45.2)	0.24	75.5% (61.1–86.7)	75.1% (68.2–81.3)	0.5	83.5% (75.2–89.9)	61.6% (55.6–67.4)	0.45
≥ 115	**78.7% (66.3–88.1)**	**70.4% (63.3–76.8)**	**0.49**	**86.7% (75.4–94.1)**	**50.0% (39.3–60.7)**	**0.37**	**69.4% (54.6–81.8)**	**81.2% (74.8–86.7)**	**0.5**	**78.90% (70.0–86.1)**	**70.9% (65.1–76.2)**	**0.50**
≥ 120	75.4% (62.7–85.5)	74.1% (67.2–80.2)	0.50	83.3% (71.5–91.7)	60.0% (49.1–70.2)	0.47	61.2% (46.2–74.8)	86.2% (80.3–90.9)	0.4	73.3% (64.1–81.4)	77.5% (72.1–82.3)	0.50
≥ 125	72.1% (59.2–82.9)	81.0% (74.6–86.3)	0.53	78.3% (65.8–87.9)	68.9% (58.3 to 78.2)	0.47	55.1% (40.2–69.3)	90.1% (84.7–94.0)	0.4	67.9% (58.3–76.5)	83.0% (78.0–87.3)	0.50
≥ 130	67.2% (54.0–78.7)	87.3% (81.7–91.7)	0.54	61.7% (48.2–73.9)	72.2% (61.8 to 81.2)	0.34	46.9% (32.5–61.7)	93.4% (88.7–96.5)	0.4	55.1% (45.2–64.6)	86.4% (81.7–90.2)	0.42

Abbreviation: J, Youden's index

Figure legends

Figure 1. Flowchart of development and validation studies

Figure 2. Receiver operating characteristic (ROC) curves and calculated areas under the curve (AUC)

(A) Development study of the delirium predictive score

(B) Internal validation study of the delirium predictive score

(C) External validation study of the delirium predictive score

(D) All validation study of the delirium predictive score

Figure 3. Calibration plot of pooled data

(A) Internal validation study of the delirium predictive score (n=150)

(B) External validation study of the delirium predictive score (n=230)

(C) All validation study of the delirium predictive score (n=380)

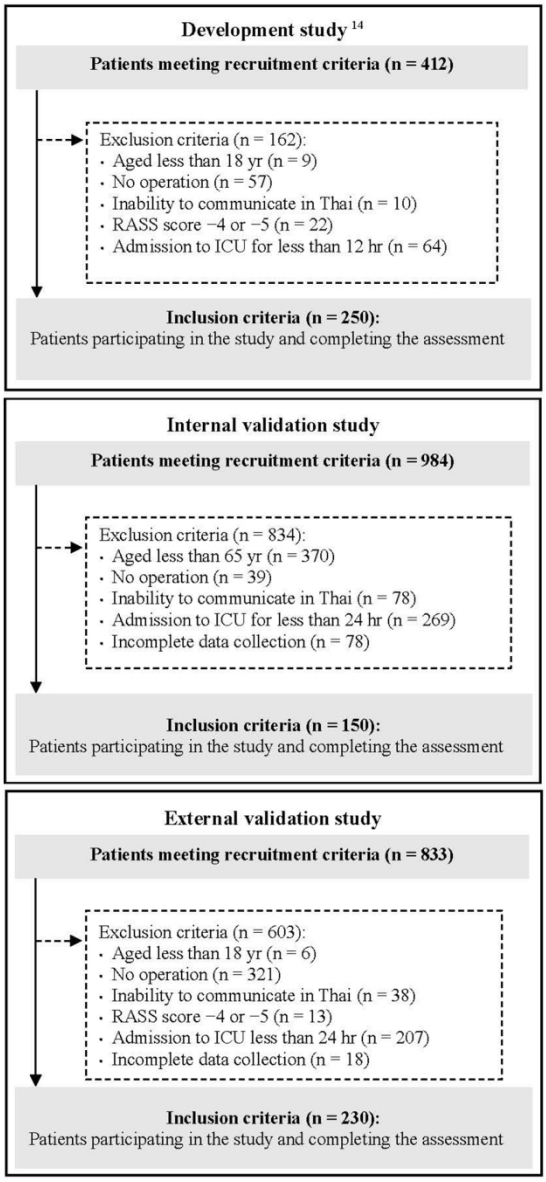


Figure 1

90x185mm (300 x 300 DPI)

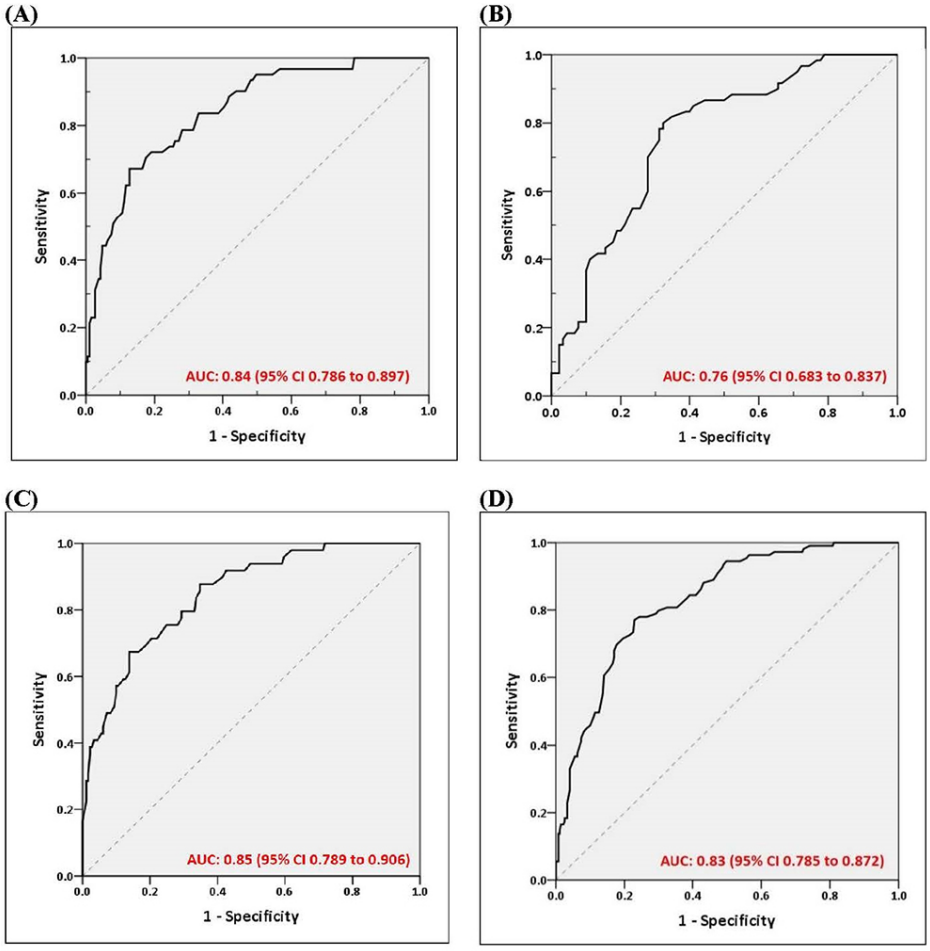


Figure 2

90x90mm (300 x 300 DPI)

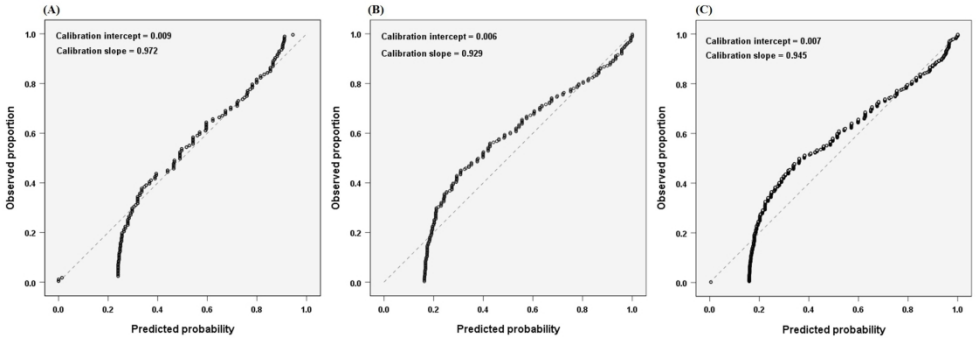


Figure 3

241x90mm (300 x 300 DPI)

TRIPOD Checklist: Prediction Model Validation

Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	6
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	7
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	7
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	7
	5b	Describe eligibility criteria for participants.	7
	5c	Give details of treatments received, if relevant.	
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	8-10
	6b	Report any actions to blind assessment of the outcome to be predicted.	
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	5, 8
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	
Sample size	8	Explain how the study size was arrived at.	
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	10
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	10-11
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	
Risk groups	11	Provide details on how risk groups were created, if done.	
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	10-11
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	12, Fig 1
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	12
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	13
Model performance	16	Report performance measures (with CIs) for the prediction model.	13
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	17
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	15-16
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	17
Implications	20	Discuss the potential clinical use of the model and implications for future research.	17-18
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	
Funding	22	Give the source of funding and the role of the funders for the present study.	19

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057890.R1
Article Type:	Original research
Date Submitted by the Author:	01-Apr-2022
Complete List of Authors:	Chaiwat, Onuma; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center Chittawatanarat, Kaweesak ; Chiang Mai University Faculty of Medicine, Department of Surgery; Chiang Mai University Faculty of Medicine, Clinical surgical research center Mueankwan, Sirirat ; Chiang Mai University Faculty of Medicine, Surgical critical care unit, Department of Surgery Morakul, Sunthiti ; Mahidol University Faculty of Medicine Ramathibodi Hospital, Department of Anesthesiology Dilokpattanamongkol, Pitchaya ; Mahidol University Faculty of Pharmacy, Department of Pharmacy Thanakiattiwibun, Chayanan ; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center Siriussawakul, Arunotai; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, INTENSIVE & CRITICAL CARE, Delirium & cognitive disorders < PSYCHIATRY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

Onuma Chaiwat[1, 7*],

Kaweesak Chittawatanarat[2, 3],

Sirirat Mueankwan[4],

Sunthiti Morakul[5],

Pitchaya Dilokpattanamongkol[6],

Chayanan Thanakiattiwibun[7],

Arunotai Siriussawakul[1,7]

***Correspondence:** onuma.cha@mahidol.ac.th

Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700 Thailand

Full list of author information is available at the end of the article

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, multicenter 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 CAM-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 year 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 year 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 cutoff value to discriminate between a high and low probability of postoperative delirium in SICU patients was 115. This cutoff 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.

Registration: Thai Clinical Trail Registry (TCTR ID: TCTR20180105001), 05 January 2018

Keywords: postoperative delirium**,** predictive model, surgical intensive care units, validation

Strengths and limitations of this study

- The developed delirium predictive model consists of 6 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 AUROC being comparable to that of the development cohort
- The optimum cutoff value to discriminate between a high and low probability of POD in SICU patients was 115

BACKGROUND

Delirium, a disturbance of consciousness, is both acute and fluctuating. Delirium is an extremely common condition among hospitalized patients. Its incidence varies with 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 anesthesia 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 (PACUs), 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.,* 2015 [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 nonstop administration of nonsteroidal 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 PRE-DELIRC (PREdiction of DELIRlum in ICU patients) delirium risk prediction tool was developed for intensive care patients [11]. This model utilizes

10 parameters. It had an area under a receiver operating characteristic curve (AUROC) of 0.87 (95% confidence interval [CI], 0.85 to 0.89). Temporal validation and external validation resulted in an AUROC of 0.89 (0.86–0.92) and 0.84 (0.82–0.87), 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 orthopedic patients, and the scoring systems tend to be overly complicated.

The Siriraj Integrated Perioperative Geriatric (SIPG) 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, 0.786–0.897). A cutoff 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 2 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

Design

A prospective, observational, cohort study was conducted. 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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Study population

The study was conducted on 380 SICU patients at 3 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 a 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 3 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.

Measurement instruments

Delirium was assessed using the Thai version of the CAM-ICU (S1). Delirium was identified by the following 4 features: 1) a change or fluctuation in baseline mental status; 2) inattention; and either 3) an altered level of consciousness; or 4) disorganized 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 utilizes 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 (hypo- and hyperactive) [18]. With regard to dementia, it was evaluated via the Thai version of modified IQCODE (S2). The questionnaire consists of 32 items, with assessments of patients being made by their caregivers. The optimal cutoff score for the modified IQCODE is 3.42 (sensitivity, 90%; specificity, 95%; and accuracy, 92%) [19]. Lastly, 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 twice daily (once during the 12 hours from 6.00 AM, and once during the 12 hours after 6.00 PM), and whenever patients developed a mental change. Delirium was screened routinely utilizing a 2-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 anesthetic records. They consisted of the surgical type (abdominal, vascular, orthopedic, urological, gynecological, 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 mmHg or the need to be treated with medications.[20, 21] Intraoperative hypoxemia 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 2 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 twice daily. The validation process was conducted according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) 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.

Demographic variables are presented as mean ± standard deviation or median (interquartile range [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 cutoff 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. Maximizing this index allows an optimal cutoff 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 analyzed using PASW Statistics for Windows (version 18; SPSS Inc., Chicago, IL, USA); and MedCalc statistical software (version 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, 1,437 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 2 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 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 6 factors (2 quantitative factors, and 4 binary factors). The formula of the model was:

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

(20 × mechanical ventilation) + (20 × modified IQCODE > 3.42)

The AUROC was 0.84 (95% CI, 0.786−0.897). The cutoff 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 cutoff 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 differences in prediction existed between the 3 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 cutoff value to discriminate between a high and low probability of POD in SICU patients was 115. This cutoff 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 2 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 multicenter-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 cutoff 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, 2 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 2 models are currently used in clinical practice and in research to predict the development of delirium in 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 9 parameters available at time of ICU admission. Wassenaar *et al*., 2019 [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 2 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 recognized as strong predictors for delirium in hospitalized 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. Firstly, it was developed specifically for surgical patients, and it demonstrated high accuracy. In addition, it employs only 5 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 5 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 2 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 2 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 cutoff value to predict the occurrence of delirium in SICUs. This cutoff 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 multicenter 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 mobilization, and enhanced sleep) [35, 36] or pharmacological (like the prophylactic administration of dexmedetomidine [37] to high-risk patients).

Several limitations need to be addressed. Firstly, 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 cutoff 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. Lastly, 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]Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai Thailand.

[3]Clinical surgical research center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Surgical critical care unit, Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

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

[6]Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand.

[7]Integrated Perioperative Geriatric Excellent Research Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Contributions

OC and AS contributed to the design of the study. OC, KC and S. Morakul were involved in data management and oversaw the project. KC, S. Mueankwan, S. Morakul, PD, contributed to data collection. CT contributed to data analysis. OC and CT contributed to the interpretation of the results and drafting the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand (IO: R016132015). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics statements

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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Acknowledgements

The authors gratefully acknowledge the patients who generously agreed to participate in this study, and Assist. Prof. Dr. Chulaluk Komoltri, M.P.H. Biostatistics, for the statistical analyses.

Supplemental file

S1. Confusion Assessment method for the ICU (CAM-ICU) tool

S2. Modified Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) tool

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. https://doi:10.1001/jama.1994.03510260066030.
2. Serafim RB, Dutra MF, Saddy F, Tura B, de Castro JE, Villarinho LC, et al. Delirium in postoperative nonventilated intensive care patients: risk factors and outcomes. *Ann Intensive Care* 2012;2:1–6. https://doi: 10.1186/2110-5820-2-51.
3. Pandharipande P, Cotton BA, Shintani A., et al. Prevalence and risk factors for development of delirium in trauma and surgical ICU patients. *J Trauma* 2008;65:34–41. https://doi: 10.1097/TA.0b013e31814b2c4d.
4. Schenning KJ, Deiner SG. Postoperative delirium: a review of risk factors and tools of prediction. *Curr Anesthesiol Rep* 2015; 5: 48–56. https://doi:10.1007/s40140-014-0086-1
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. https://doi: 10.1001/archinternmed.2007.4.
7. Gusmao-Flores D, Salluh JI, Chalhub RÁ, 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. https://doi: 10.1186/cc11407.
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. https://doi: 10.2147/CIA.S62660.
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. https://doi: 10.1080/13607863.2015.1035695.
11. Van den Boogaard M, Pickkers P, Slooter AJC, et al. Development and validation of PRE-DELIRC (PREdiction of DELIRlum in ICU patients) delirium prediction model for intensive care patients: observational multicente study. *BMJ* 2012;344:e420. https://doi:10.1136/bmj.e420.
12. Moerman S, Tuinebreijer WE, Boo MD, Pilot P, Nelissen RG, Vochteloo AJ. Validation of the risk model for delirium in hip fracture patients. *Gen Hosp Psychiatry* 2012;34:153–9. https://doi: 10.1016/j.genhosppsych.2011.11.011.
13. Kim MY, Park UJ, Kim HT, et al. DELirium Prediction Based on Hospital Information (Delphi) in General Surgery Patients. *Medicine* 2016;95:2–7. https://doi: 10.1097/MD.0000000000003072.
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. doi: 10.1186/s12871-019-0732-8.
15. Robinson TN, Raeburn CD, Tran ZV, et al. Postoperative delirium in the elderly. risk factors and outcomes. *Ann Surg* 2009;249:173–8. https://doi: 10.1097/SLA.0b013e31818e4776.
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. https://doi: 10.1097/01.sla.0000077920.38307.5f.

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. https://doi: 10.1097/00003246-200107000-00012.

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. https://doi: 10.1111/j.1532-5415.2005.00621.x.

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. doi: 10.1097/ALN.0b013e3182a10cce.

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. doi:10.1097/ALN.0b013e3182a10e26

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 Medicine.* 2015;13:1. https://doi.org/10.1186/s12916-014-0241-z

23. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5. https://doi: 10.1002/1097-0142(1950)3:1<32:aid-cncr2820030106>3.0.co;2-3.

24. Galley HF. Editorial II: Solid as a ROC. *Br J Anaesth* 2004;93:623–6. https://doi: 10.1093/bja/aeh247.

25. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, et al. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Mak Int J Soc Med Decis Mak* 2001;21:45–56.

26. Van Calster, B., McLernon, D.J., van Smeden, M. et al. Calibration: the Achilles heel of

predictive analytics. *BMC Med* 2019;17. https://doi.org/10.1186/s12916-019-1466-7.

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. https://doi: 10.1007/s00134-013-3202-7.
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. https://doi: 10.1007/s00134-015-3777-2.
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. https://doi: 10.1097/CCM.0000000000003911.
30. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22. https://doi: 10.1016/S0140-6736(13)60688-1.
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. https://doi: 10.1097/CCM.0000000000003299.
32. To-adithep P, Wongviriyawong T, Thanakiattiwibun C, Chanidnuan M, Pancharoen W, Chaiwat O. 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. https://doi: 10.1016/S0140-6736(02)07948-5.
34. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ* 2002;324:477–80. htpps://doi: 10.1136/bmj.324.7335.477.
35. Hanison J, Conway D: A multifaceted approach to prevention of delirium on intensive

care. *BMJ Qual Improv Rep* 2015;12:4–12. https://doi: 10.1136/bmjquality.u209656.w4000.

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. https://doi:10.1016/j.jcrc.2015.09.031.
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. *The Lancet* 2016;388:1893–1902. https://doi: 10.1016/S0140-6736(16)30580-3.
38. an 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. https://doi: 10.1164/rccm.201101-0065OC.

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

Abbreviations: CAM, Confusion Assessment Method; ICU, intensive care unit; IRR, inter-rater reliability expressed as Cohen's κ

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%)
Type of delirium				
Hypoactive	44 (72%)	16 (26.7%)	26 (53.1%)	42 (38.5%)
Hyperactive	9 (15%)	17 (28.3%)	6 (12.2%)	23 (21.1%)
Mixed	8 (13%)	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%)
Orthopedic	29 (11.6%)	8 (5.3%)	10 (4.3%)	18 (4.7%)
Gynecological	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 (%).

Abbreviations: 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

Table 3. Cutoff values of the Delirium Prediction Score

Cutoff value	Development (n = 250)		J	Internal validation (n = 150)		J	External validation (n = 230)		J	All validation (n = 380)		J
	Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)	
≥ 100	95.1% (86.3–99.0)	50.3% (42.9–57.6)	0.45	100.0% (94.0–100.0)	21.1% (13.2–31.0)	0.21	87.8% (75.2–95.4)	64.6% (57.2–71.6)	0.5	94.5% (88.4–98.0)	50.2% (44.1–56.3)	0.45
≥ 105	90.2% (79.8–96.3)	56.1% (48.7–63.3)	0.46	96.7% (88.5–99.6)	27.8% (18.9–38.2)	0.25	79.6% (65.7–89.8)	66.7% (59.5–73.7)	0.46	89.0% (81.6–94.2)	53.9% (47.7–59.9)	0.43
≥ 110	83.6% (71.9–91.9)	63.0% (55.7–69.9)	0.47	90.0% (79.5–96.2)	34.4% (24.7–45.2)	0.24	75.5% (61.1–86.7)	75.1% (68.2–81.3)	0.5	83.5% (75.2–89.9)	61.6% (55.6–67.4)	0.45
≥ 115	**78.7% (66.3–88.1)**	**70.4% (63.3–76.8)**	**0.49**	**86.7% (75.4–94.1)**	**50.0% (39.3–60.7)**	**0.37**	**69.4% (54.6–81.8)**	**81.2% (74.8–86.7)**	**0.5**	**78.90% (70.0–86.1)**	**70.9% (65.1–76.2)**	**0.50**
≥ 120	75.4% (62.7–85.5)	74.1% (67.2–80.2)	0.50	83.3% (71.5–91.7)	60.0% (49.1–70.2)	0.47	61.2% (46.2–74.8)	86.2% (80.3–90.9)	0.47	73.3% (64.1–81.4)	77.5% (72.1–82.3)	0.50
≥ 125	72.1% (59.2–82.9)	81.0% (74.6–86.3)	0.53	78.3% (65.8–87.9)	68.9% (58.3 to 78.2)	0.47	55.1% (40.2–69.3)	90.1% (84.7–94.0)	0.46	67.9% (58.3–76.5)	83.0% (78.0–87.3)	0.50
≥ 130	67.2% (54.0–78.7)	87.3% (81.7–91.7)	0.54	61.7% (48.2–73.9)	72.2% (61.8 to 81.2)	0.34	46.9% (32.5–61.7)	93.4% (88.7–96.5)	0.42	55.1% (45.2–64.6)	86.4% (81.7–90.2)	0.42

Abbreviation: J, Youden's index

Figure legends

Figure 1. Flowchart of development and validation studies

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

Figure 1. Flowchart of development and validation studies

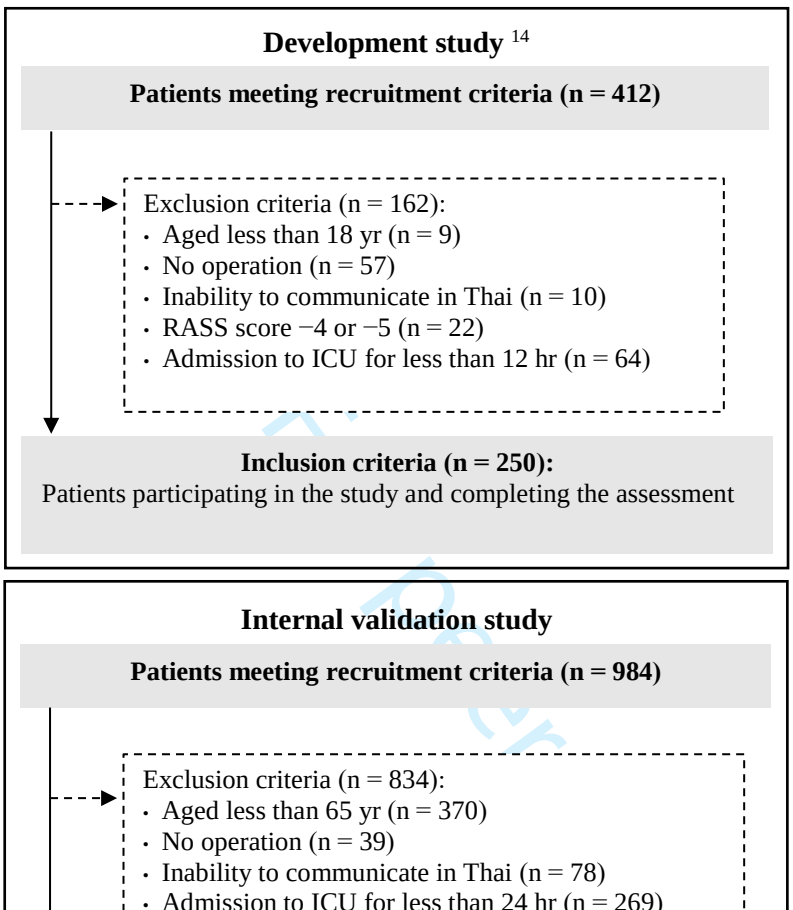
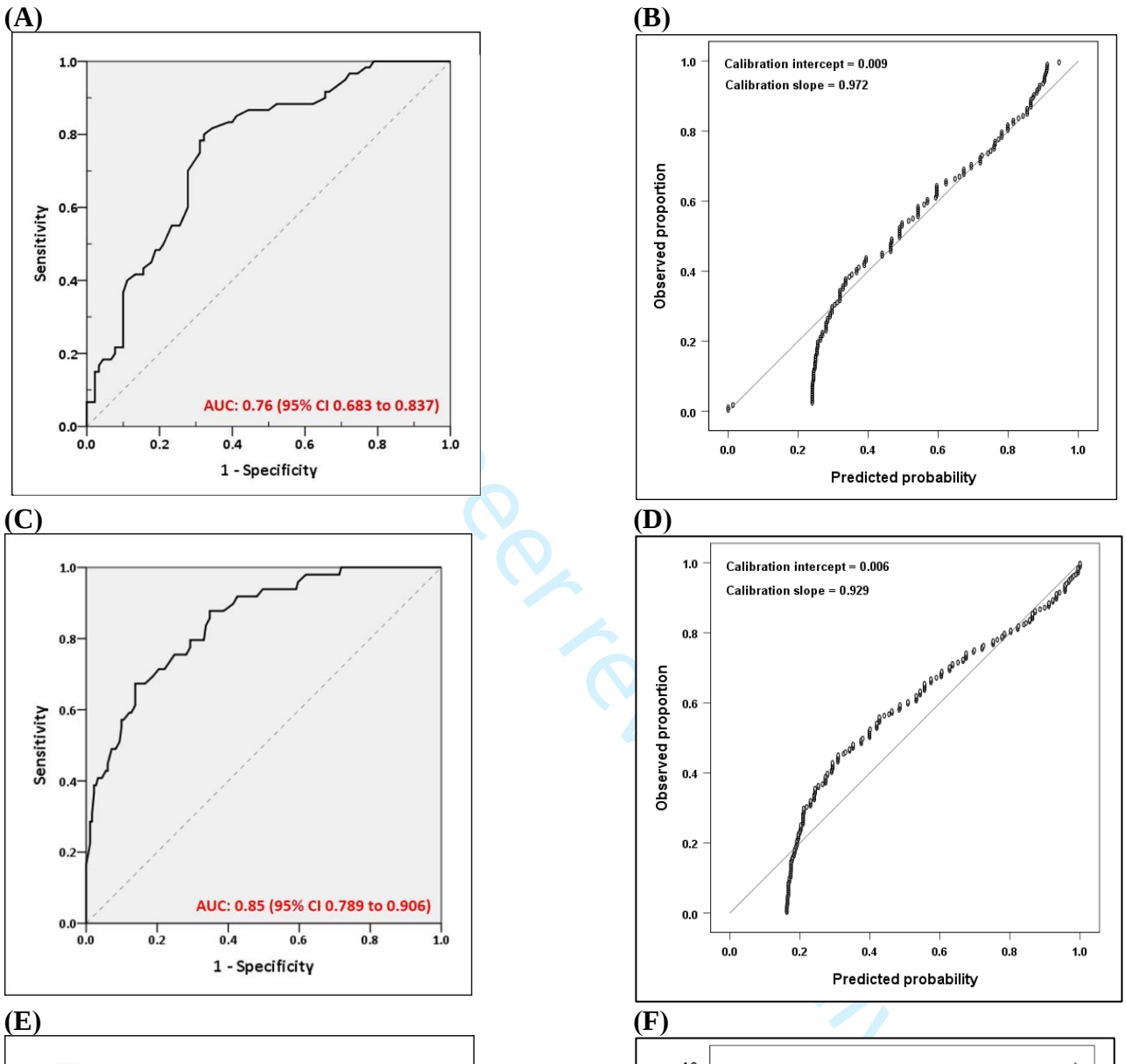


Figure 2. Receiver operating characteristic (ROC) curves and calculated areas under the curve (AUC), and Calibration plot of pooled data;

(A and B) Internal validation study of the delirium predictive score

(C and D) External validation study of the delirium predictive score

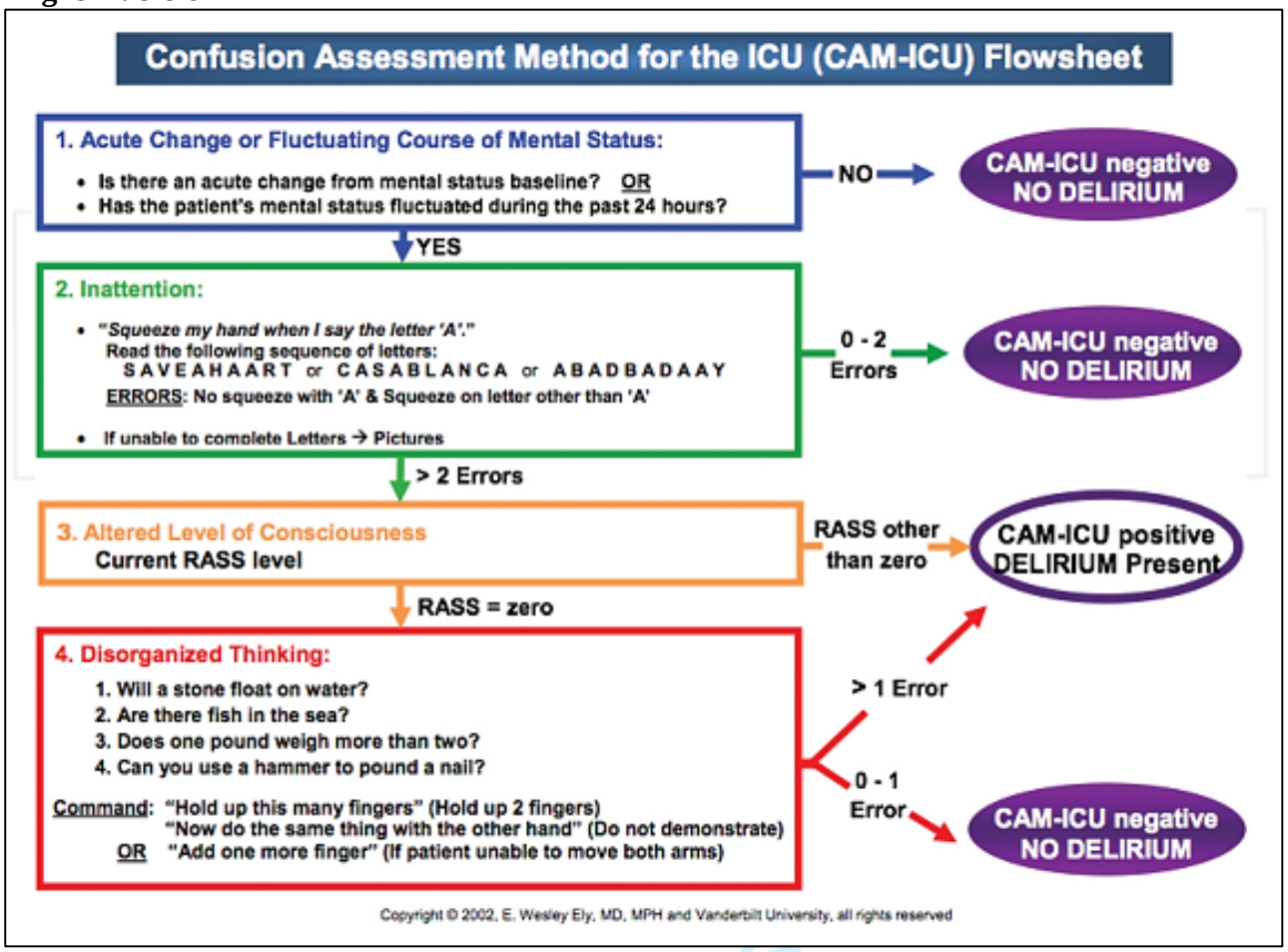
(E and F) All validation study of the delirium predictive score



Supplemental file

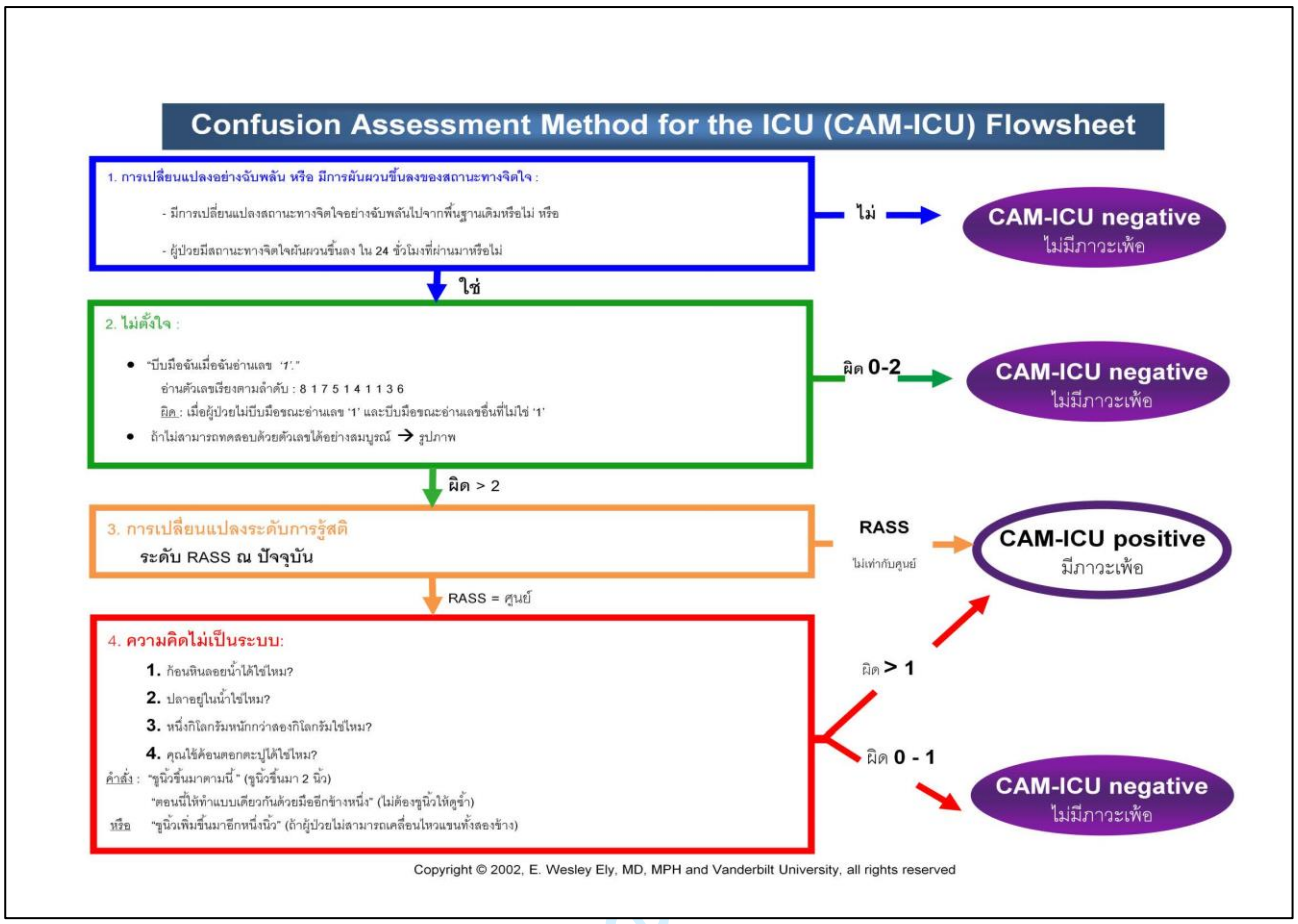
S1. Confusion Assessment method for the ICU (CAM-ICU) tool

English Version



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. https://doi: 10.1097/00003246-200107000-00012.

Thai Version



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. https://doi: 10.2147/CIA.S62660.

Supplemental file

S2. Modified Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) tool

English version

Items	Comparing the elder's change with the previous 10 years				
	1 Much improve	2 A bit improve	3 Not much change	4 A bit worse	5 Much worse
1. Recognizing the faces of family and friends					
2. Remembering the names of family and friends					
3. Remembering things about family and friends eg, occupations, birthdays, addresses					
4. Remembering things that have happened recently					
5. Recall conversations a few day later					
6. Forgetting what he/she want to say in the middle of a conversation					
7. Remembering his/her address and telephone number					
8. Remembering what day and month it is					
9. Remembering where things are usually kept					
10. Remembering where to find things which have been put in a different place from usual					
11. Remembering things that happened to him/her when he/she was young					
12. Remembering things he/she learned when he/she was young					
13. Knowing about important historical events of the past					
14. Adjusting to any change in his/her day-to-day					
15. Knowing how to work familiar machines around the house					
16. Learning to use a new gadget or machines around the house					
17. Learning the new things that in general					
18. Understanding the meaning of unusual words					
19. Understanding magazine or newspaper articles					
20. Following a story in a book or on TV					
21. Contacting with friends or for business purposes					
22. Making decisions on everyday matters					
23. Handing money for shopping					
24. Handing financial matters					
25. Handing other everyday arithmetic problems, eg knowing how much food to buy, knowing a period of time for doing activity					
26. Using his/her intelligence to understand what's going on and to reason things through					
27. Able to sing or pray the used one					
28. Selecting appropriate instrument					
29. Keep speak repeating					
30. Carrying out daily activities					
31. Traveling to familiar place					
32. Working ability					

Thai Version

	การเปลี่ยนแปลงระหว่าง 10 ปีที่แล้วกับปัจจุบัน				
	1 ดีขึ้นมาก	2 ดีขึ้นเล็กน้อย	3 ไม่เปลี่ยนแปลง	4 แย่ลงเล็กน้อย	5 แย่ลงมาก
1. ความจำเกี่ยวกับหน้าตาคนในครอบครัวหรือญาติ					
2. ความจำเกี่ยวกับชื่อคนในครอบครัวหรือญาติ					
3. ความจำในรายละเอียดของคนในครอบครัวหรือญาติเกี่ยวกับอาชีพที่อยู่					
4. ความจำในเหตุการณ์ที่เกิดขึ้นเมื่อ 2-3 วันที่ผ่านมา					
5. ความจำในเรื่องที่สนทนาไปเมื่อ 2-3 วันที่ผ่านมา					
6. พูดคุยอย่างต่อเนื่องโดยไม่ลืมสิ่งที่จะพูด					
7. จำได้ว่าตอนนี้พักอาศัยอยู่ที่ไหน					
8. จำได้ว่าวันนี้เป็น วัน เดือน อะไร					
9. ความจำเกี่ยวกับที่ประจำที่ใช้เก็บของในบ้าน					
10. จำได้ว่าวางของไว้ที่ไหน					
11. จำเหตุการณ์เกี่ยวกับตนในวัยเด็ก					
12. จำสิ่งที่ตนได้เรียนรู้ในวัยเด็ก					
13. ทราบเหตุการณ์ที่สำคัญในอดีต					
14. ความสามารถในการปรับตัวเข้ากับการเปลี่ยนแปลงในชีวิตประจำวัน					
15. สามารถใช้เครื่องมือที่คุ้นเคยภายในบ้าน					
16. สามารถเรียนรู้การใช้เครื่องมือ เครื่องใช้ใหม่ๆ ในบ้าน					
17. สามารถเรียนรู้สิ่งใหม่ ๆ ทั่ว ๆ ไป					
18. สามารถเข้าใจความหมายของคำแปลก ๆ					
19. สามารถเข้าใจบทความในหนังสือพิมพ์หรือนิตยสาร					
20. สามารถติดตามเรื่องราวต่างๆ ในวิทยุ หรือโทรทัศน์					
21. สามารถติดต่อลูกหลาน ญาติหรือกิจธุระทั่วๆ ไป					
22. ความสามารถในการตัดสินใจเรื่องต่าง ๆ ในชีวิตประจำวัน					
23. ความสามารถในการใช้จ่าย					
24. ความสามารถในการจัดสรรเรื่องเงิน					
25. สามารถประมาณได้ว่าจะใช้สิ่งของประมาณเท่าไร เช่น จะซื้ออาหารเท่าไร หรือกะเวลาที่ใช้ในการทำกิจกรรมต่างๆ เช่น ใช้เวลาในการเดินทางเท่าไร					
26. สามารถที่จะเข้าใจในสิ่งที่เกิดขึ้น พร้อมกับให้เหตุผลในสิ่งนั้นได้					
27. สามารถร้องเพลงที่เคยร้อง หรือ สวดมนต์ที่เคยสวด					
28. สามารถเลือกใช้เครื่องมือเครื่องใช้ต่าง ๆได้อย่างเหมาะสมกับงาน					
29. การพูดจาหรือถามซ้ำๆ					
30. สามารถปฏิบัติกิจวัตรประจำวันของตนเอง					
31. สามารถเดินทางไป-กลับสถานที่ที่คุ้นเคยได้โดยลำพัง					
32. สามารถทำงานที่เคยทำ					

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.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057890.R2
Article Type:	Original research
Date Submitted by the Author:	29-May-2022
Complete List of Authors:	Chaiwat, Onuma; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center Chittawatanarat, Kaweesak ; Chiang Mai University Faculty of Medicine, Department of Surgery; Chiang Mai University Faculty of Medicine, Clinical surgical research center Mueankwan, Sirirat ; Chiang Mai University Faculty of Medicine, Surgical critical care unit, Department of Surgery Morakul, Sunthiti ; Mahidol University Faculty of Medicine Ramathibodi Hospital, Department of Anesthesiology Dilokpattanamongkol, Pitchaya ; Mahidol University Faculty of Pharmacy, Department of Pharmacy Thanakiattiwibun, Chayanan ; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center Siriussawakul, Arunotai; Mahidol University Faculty of Medicine Siriraj Hospital, Department of Anesthesiology; Mahidol University Faculty of Medicine Siriraj Hospital, Integrated Perioperative Geriatric Excellent Research Center
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, INTENSIVE & CRITICAL CARE, Delirium & cognitive disorders < PSYCHIATRY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

Onuma Chaiwat[1, 7*],

Kaweesak Chittawatanarat[2, 3],

Sirirat Mueankwan[4],

Sunthiti Morakul[5],

Pitchaya Dilokpattanamongkol[6],

Chayanan Thanakiattiwibun[7],

Arunotai Siriussawakul[1,7]

***Correspondence:** onuma.cha@mahidol.ac.th

Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700 Thailand

Full list of author information is available at the end of the article

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, multicenter 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 CAM-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 year 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 year 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 cutoff value to discriminate between a high and low probability of postoperative delirium in SICU patients was 115. This cutoff 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.

Registration: Thai Clinical Trail Registry (TCTR ID: TCTR20180105001), 05 January 2018

Keywords: postoperative delirium**,** predictive model, surgical intensive care units, validation

Strengths and limitations of this study

- The developed delirium predictive model consists of 6 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 AUROC being comparable to that of the development cohort
- The optimum cutoff value to discriminate between a high and low probability of POD in SICU patients was 115

BACKGROUND

Delirium, a disturbance of consciousness, is both acute and fluctuating. Delirium is an extremely common condition among hospitalized patients. Its incidence varies with 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 anesthesia 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 (PACUs), 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.,* 2015 [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 nonstop administration of nonsteroidal 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 PRE-DELIRC (PREdiction of DELIRlum in ICU patients) delirium risk prediction tool was developed for intensive care patients [11]. This model utilizes

10 parameters. It had an area under a receiver operating characteristic curve (AUROC) of 0.87 (95% confidence interval [CI], 0.85 to 0.89). Temporal validation and external validation resulted in an AUROC of 0.89 (0.86–0.92) and 0.84 (0.82–0.87), 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 orthopedic patients, and the scoring systems tend to be overly complicated.

The Siriraj Integrated Perioperative Geriatric (SIPG) 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, 0.786–0.897). A cutoff 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 2 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

Design

A prospective, observational, cohort study was conducted. 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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Study population

The study was conducted on 380 SICU patients at 3 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 a 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 3 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.

Measurement instruments

Delirium was assessed using the Thai version of the CAM-ICU (S1). Delirium was identified by the following 4 features: 1) a change or fluctuation in baseline mental status; 2) inattention; and either 3) an altered level of consciousness; or 4) disorganized 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 utilizes 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 (hypo- and hyperactive) [18]. With regard to dementia, it was evaluated via the Thai version of modified IQCODE (S2). The questionnaire consists of 32 items, with assessments of patients being made by their caregivers. The optimal cutoff score for the modified IQCODE is 3.42 (sensitivity, 90%; specificity, 95%; and accuracy, 92%) [19]. Lastly, 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 twice daily (once during the 12 hours from 6.00 AM, and once during the 12 hours after 6.00 PM), and whenever patients developed a mental change. Delirium was screened routinely utilizing a 2-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 anesthetic records. They consisted of the surgical type (abdominal, vascular, orthopedic, urological, gynecological, 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 mmHg or the need to be treated with medications.[20, 21] Intraoperative hypoxemia 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 2 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 twice daily. The validation process was conducted according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) 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 ± standard deviation or median (interquartile range [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 cutoff 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. Maximizing this index allows an optimal cutoff 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 analyzed using PASW Statistics for Windows (version 18; SPSS Inc., Chicago, IL, USA); and MedCalc statistical software (version 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, 1,437 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 2 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 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 6 factors (2 quantitative factors, and 4 binary factors). The formula of the model was:

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

(20 × mechanical ventilation) + (20 × modified IQCODE > 3.42)

The AUROC was 0.84 (95% CI, 0.786−0.897). The cutoff 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 cutoff 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 3 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 cutoff value to discriminate between a high and low probability of POD in SICU patients was 115. This cutoff 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 2 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 multicenter-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 cutoff 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, 2 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 2 models are currently used in clinical practice and in research to predict the development of delirium in 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 9 parameters available at time of ICU admission. Wassenaar *et al*., 2019 [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 2 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 recognized as strong predictors for delirium in hospitalized 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. Firstly, it was developed specifically for surgical patients, and it demonstrated high accuracy. In addition, it employs only 5 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 5 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 2 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 2 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 cutoff value to predict the occurrence of delirium in SICUs. This cutoff 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 multicenter 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 mobilization, and enhanced sleep) [35, 36] or pharmacological (like the prophylactic administration of dexmedetomidine [37] to high-risk patients).

Several limitations need to be addressed. Firstly, 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 cutoff 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. Lastly, 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]Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai Thailand.

[3]Clinical surgical research center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Surgical critical care unit, Department of Surgery, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

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

[6]Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand.

[7]Integrated Perioperative Geriatric Excellent Research Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Contributions

OC and AS contributed to the design of the study. OC, KC and S. Morakul were involved in data management and oversaw the project. KC, S. Mueankwan, S. Morakul, PD, contributed to data collection. CT contributed to data analysis. OC and CT contributed to the interpretation of the results and drafting the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand (IO: R016132015). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics statements

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 Prof. Chairat Shayakul, M.D.) 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 Asst. Prof. Chusak Okascharoen, M.D.) 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 Prof. Panja Kulapongs, M.D.) on 28 November 2017. Written informed consent was obtained from the participants before their entry into the study. The trial was registered with the Thai Clinical Trials Registry (TCTR20180105001).

Acknowledgements

The authors gratefully acknowledge the patients who generously agreed to participate in this study, and Assist. Prof. Dr. Chulaluk Komoltri, M.P.H. Biostatistics, for the statistical analyses.

Supplemental file

S1. Confusion Assessment method for the ICU (CAM-ICU) tool

S2. Modified Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) tool

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. https://doi:10.1001/jama.1994.03510260066030.
2. Serafim RB, Dutra MF, Saddy F, Tura B, de Castro JE, Villarinho LC, et al. Delirium in postoperative nonventilated intensive care patients: risk factors and outcomes. *Ann Intensive Care* 2012;2:1–6. https://doi: 10.1186/2110-5820-2-51.
3. Pandharipande P, Cotton BA, Shintani A., et al. Prevalence and risk factors for development of delirium in trauma and surgical ICU patients. *J Trauma* 2008;65:34–41. https://doi: 10.1097/TA.0b013e31814b2c4d.
4. Schenning KJ, Deiner SG. Postoperative delirium: a review of risk factors and tools of prediction. *Curr Anesthesiol Rep* 2015; 5: 48–56. https://doi:10.1007/s40140-014-0086-1
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. https://doi: 10.1001/archinternmed.2007.4.
7. Gusmao-Flores D, Salluh JI, Chalhub RÁ, 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. https://doi: 10.1186/cc11407.
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. https://doi: 10.2147/CIA.S62660.
9. Muangpaisan W, Wongprikron A, Srinonprasert V, et al. Incidence and risk factors of

457 acute delirium in older patients with hip fracture in Siriraj Hospital. *J Med Assoc Thai*
458 2015;98:423–30.
459 10. Limpawattana P, Panitchote A, Tangvoraphonkchai K, et al. Delirium in critical care: a
460 study of incidence, prevalence, and associated factors in the tertiary care hospital of older
461 Thai adults. *Aging Ment Health* 2016;20:74–80. https://doi: 10.1080/13607863.
462 2015.1035695.
463 11. Van den Boogaard M, Pickkers P, Slooter AJC, et al. Development and validation of PRE-
464 DELIRC (PREdiction of DELIRlum in ICU patients) delirium prediction model for
465 intensive care patients: observational multicente study. *BMJ* 2012;344:e420. https://
466 doi:10.1136/bmj.e420.
467 12. Moerman S, Tuinebreijer WE, Boo MD, Pilot P, Nelissen RG, Vochteloo AJ. Validation
468 of the risk model for delirium in hip fracture patients. *Gen Hosp Psychiatry* 2012;34:153–
469 9. https://doi: 10.1016/j.genhosppsych.2011.11.011.
470 13. Kim MY, Park UJ, Kim HT, et al. DELirium Prediction Based on Hospital Information
471 (Delphi) in General Surgery Patients. *Medicine* 2016;95:2–7. https://doi: 10.1097/MD.
472 0000000000003072.
473 14. Chaiwat O, Chanidnuan M, Pancharoen W, et al. Postoperative delirium in critically ill
474 surgical patients: incidence, risk factors, and predictive scores. *BMC Anesthesiol*
475 2019;19:39. doi: 10.1186/s12871-019-0732-8.
476 15. Robinson TN, Raeburn CD, Tran ZV, et al. Postoperative delirium in the elderly. risk
477 factors and outcomes. *Ann Surg* 2009;249:173–8. https://doi: 10.1097/SLA.
478 0b013e31818e4776.
479 16. Böhner H, Hummel TC, Habel U, et al. Predicting Delirium After Vascular Surgery. A
480 Model Based on Pre- and Intraoperative Data. *Ann Surg* 2003;238:149–56. https://doi:
481 10.1097/01.sla.0000077920.38307.5f.

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. https://doi: 10.1097/00003246-200107000-00012.
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. https://doi: 10.1111/j.1532-5415.2005.00621.x.
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. doi: 10.1097/ALN.0b013e3182a10cce.
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. doi:10.1097/ALN.0b013e3182a10e26
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 Medicine.* 2015;13:1. https://doi.org/10.1186/s12916-014-0241-z
23. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5. https://doi: 10.1002/1097-0142(1950)3:1<32:aid-cncr2820030106>3.0.co;2-3.
24. Galley HF. Editorial II: Solid as a ROC. *Br J Anaesth* 2004;93:623–6. https://doi: 10.1093/bja/aeh247.
25. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, et al. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Mak Int J Soc Med Decis Mak* 2001;21:45–56.
26. Van Calster, B., McLernon, D.J., van Smeden, M. et al. Calibration: the Achilles heel of

predictive analytics. *BMC Med* 2019;17. https://doi.org/10.1186/s12916-019-1466-7.

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. https://doi: 10.1007/s00134-013-3202-7.
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. https://doi: 10.1007/s00134-015-3777-2.
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. https://doi: 10.1097/CCM.0000000000003911.
30. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22. https://doi: 10.1016/S0140-6736(13)60688-1.
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. https://doi: 10.1097/CCM.0000000000003299.
32. To-adithep P, Wongviriyawong T, Thanakiattiwibun C, Chanidnuan M, Pancharoen W, Chaiwat O. 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. https://doi: 10.1016/S0140-6736(02)07948-5.
34. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ* 2002;324:477–80. htpps://doi: 10.1136/bmj.324.7335.477.
35. Hanison J, Conway D: A multifaceted approach to prevention of delirium on intensive

care. *BMJ Qual Improv Rep* 2015;12:4–12. https://doi: 10.1136/bmjquality.u209656.w4000.

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. https://doi:10.1016/j.jcrc.2015.09.031.
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. *The Lancet* 2016;388:1893–1902. https://doi: 10.1016/S0140-6736(16)30580-3.
38. an 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. https://doi: 10.1164/rccm.201101-0065OC.

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

Abbreviations: CAM, Confusion Assessment Method; ICU, intensive care unit; IRR, inter-rater reliability expressed as Cohen's κ

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%)	16 (26.7%)	26 (53.1%)	42 (38.5%)
Hyperactive	9 (15%)	17 (28.3%)	6 (12.2%)	23 (21.1%)
Mixed	8 (13%)	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%)
Orthopedic	29 (11.6%)	8 (5.3%)	10 (4.3%)	18 (4.7%)
Gynecological	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 (%).

Abbreviations: 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

Table 3. Cutoff values of the Delirium Prediction Score

Cutoff value	Development (n = 250)		J	Internal validation (n = 150)		J	External validation (n = 230)		J	All validation (n = 380)		J
	Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)		Sensitivity (95% CI)	Specificity (95% CI)	
≥ 100	95.1% (86.3–99.0)	50.3% (42.9–57.6)	0.45	100.0% (94.0–100.0)	21.1% (13.2–31.0)	0.21	87.8% (75.2–95.4)	64.6% (57.2–71.6)	0.5	94.5% (88.4–98.0)	50.2% (44.1–56.3)	0.45
≥ 105	90.2% (79.8–96.3)	56.1% (48.7–63.3)	0.46	96.7% (88.5–99.6)	27.8% (18.9–38.2)	0.25	79.6% (65.7–89.8)	66.7% (59.5–73.7)	0.46	89.0% (81.6–94.2)	53.9% (47.7–59.9)	0.43
≥ 110	83.6% (71.9–91.9)	63.0% (55.7–69.9)	0.47	90.0% (79.5–96.2)	34.4% (24.7–45.2)	0.24	75.5% (61.1–86.7)	75.1% (68.2–81.3)	0.5	83.5% (75.2–89.9)	61.6% (55.6–67.4)	0.45
≥ 115	**78.7% (66.3–88.1)**	**70.4% (63.3–76.8)**	**0.49**	**86.7% (75.4–94.1)**	**50.0% (39.3–60.7)**	**0.37**	**69.4% (54.6–81.8)**	**81.2% (74.8–86.7)**	**0.5**	**78.90% (70.0–86.1)**	**70.9% (65.1–76.2)**	**0.50**
≥ 120	75.4% (62.7–85.5)	74.1% (67.2–80.2)	0.50	83.3% (71.5–91.7)	60.0% (49.1–70.2)	0.47	61.2% (46.2–74.8)	86.2% (80.3–90.9)	0.47	73.3% (64.1–81.4)	77.5% (72.1–82.3)	0.50
≥ 125	72.1% (59.2–82.9)	81.0% (74.6–86.3)	0.53	78.3% (65.8–87.9)	68.9% (58.3 to 78.2)	0.47	55.1% (40.2–69.3)	90.1% (84.7–94.0)	0.46	67.9% (58.3–76.5)	83.0% (78.0–87.3)	0.50
≥ 130	67.2% (54.0–78.7)	87.3% (81.7–91.7)	0.54	61.7% (48.2–73.9)	72.2% (61.8 to 81.2)	0.34	46.9% (32.5–61.7)	93.4% (88.7–96.5)	0.42	55.1% (45.2–64.6)	86.4% (81.7–90.2)	0.42

Abbreviation: J, Youden's index

Figure legends

Figure 1. Flowchart of development and validation studies

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

Figure 1. Flowchart of development and validation studies

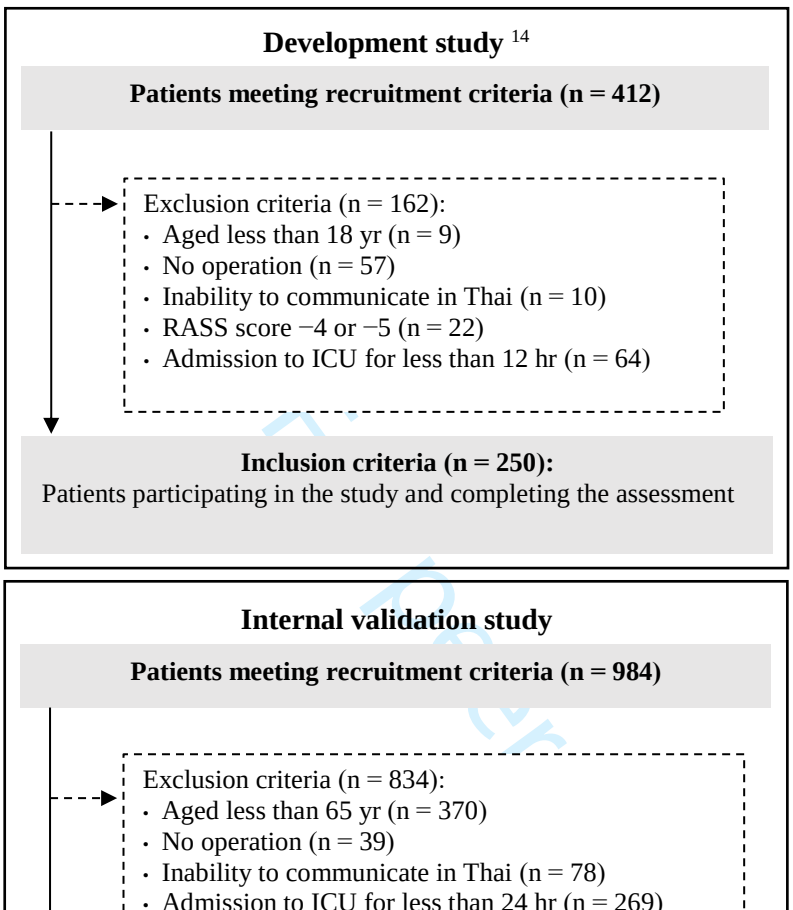
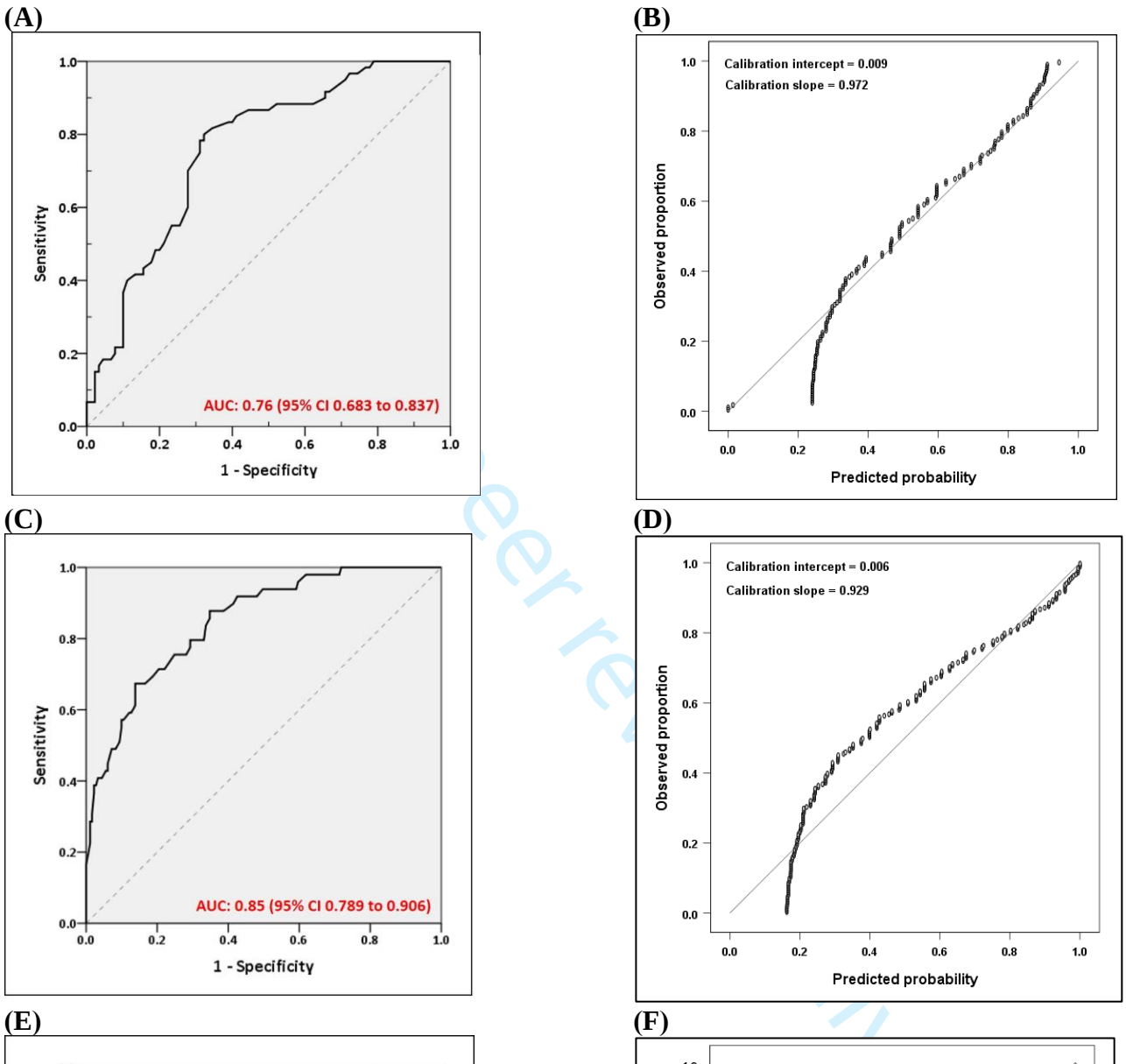


Figure 2. Receiver operating characteristic (ROC) curves and calculated areas under the curve (AUC), and Calibration plot of pooled data;

(A and B) Internal validation study of the delirium predictive score

(C and D) External validation study of the delirium predictive score

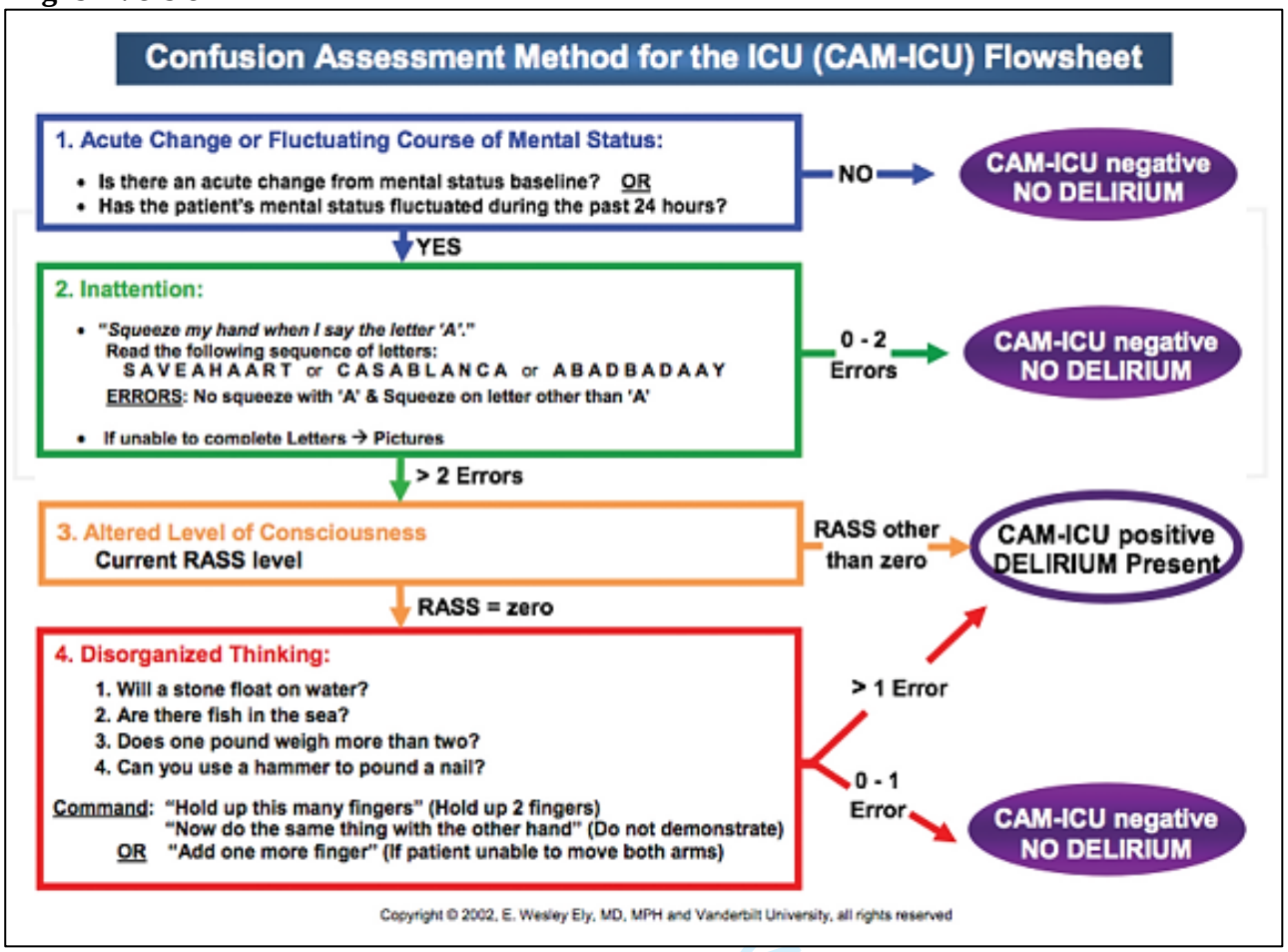
(E and F) All validation study of the delirium predictive score



Supplemental file

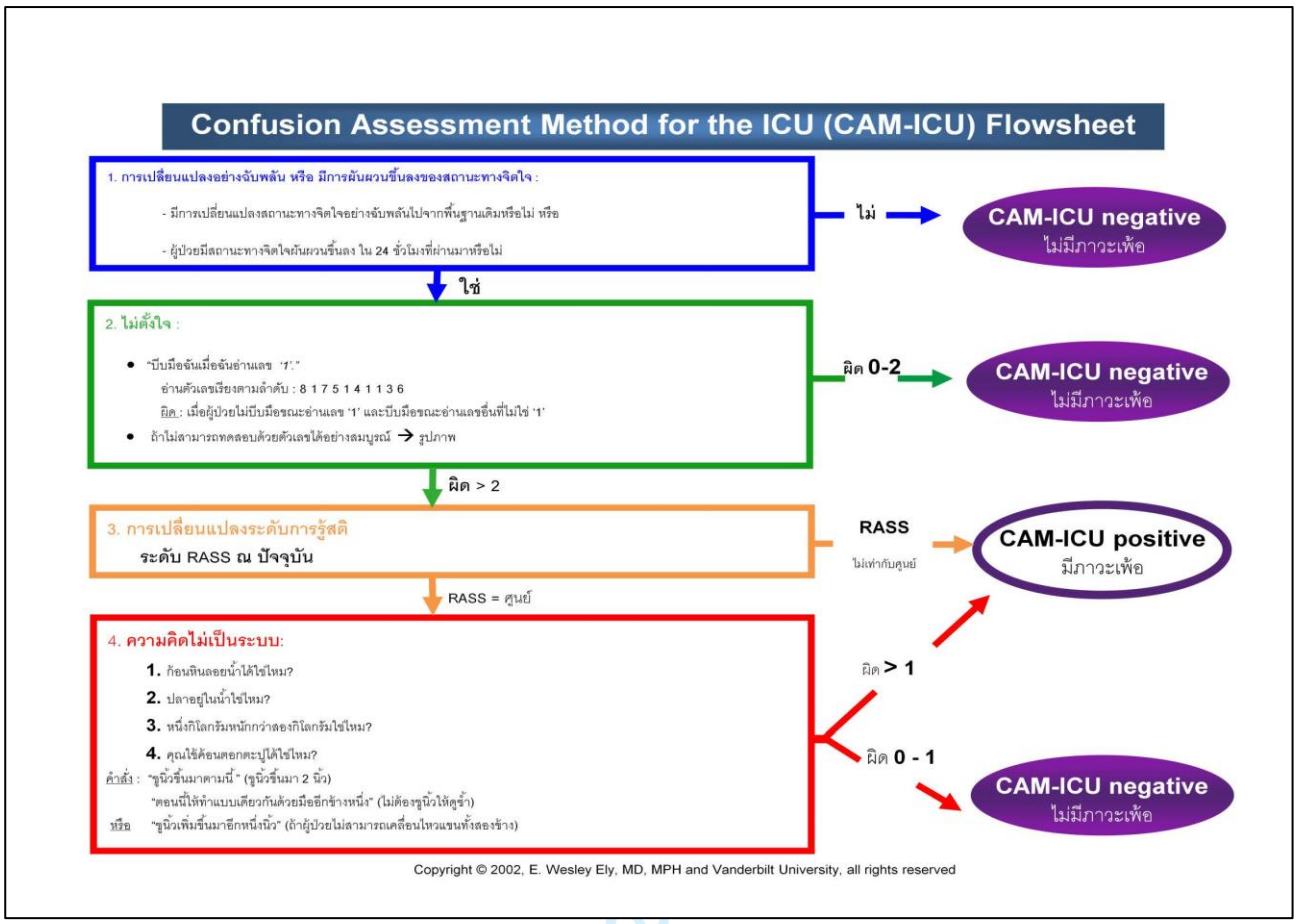
S1. Confusion Assessment method for the ICU (CAM-ICU) tool

English Version



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. https://doi: 10.1097/00003246-200107000-00012.

Thai Version



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. https://doi: 10.2147/CIA.S62660.

Supplemental file

S2. Modified Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) tool

English version

Items	Comparing the elder's change with the previous 10 years				
	1 Much improve	2 A bit improve	3 Not much change	4 A bit worse	5 Much worse
1. Recognizing the faces of family and friends					
2. Remembering the names of family and friends					
3. Remembering things about family and friends eg, occupations, birthdays, addresses					
4. Remembering things that have happened recently					
5. Recall conversations a few day later					
6. Forgetting what he/she want to say in the middle of a conversation					
7. Remembering his/her address and telephone number					
8. Remembering what day and month it is					
9. Remembering where things are usually kept					
10. Remembering where to find things which have been put in a different place from usual					
11. Remembering things that happened to him/her when he/she was young					
12. Remembering things he/she learned when he/she was young					
13. Knowing about important historical events of the past					
14. Adjusting to any change in his/her day-to-day					
15. Knowing how to work familiar machines around the house					
16. Learning to use a new gadget or machines around the house					
17. Learning the new things that in general					
18. Understanding the meaning of unusual words					
19. Understanding magazine or newspaper articles					
20. Following a story in a book or on TV					
21. Contacting with friends or for business purposes					
22. Making decisions on everyday matters					
23. Handing money for shopping					
24. Handing financial matters					
25. Handing other everyday arithmetic problems, eg knowing how much food to buy, knowing a period of time for doing activity					
26. Using his/her intelligence to understand what's going on and to reason things through					
27. Able to sing or pray the used one					
28. Selecting appropriate instrument					
29. Keep speak repeating					
30. Carrying out daily activities					
31. Traveling to familiar place					
32. Working ability					

Thai Version

	การเปลี่ยนแปลงระหว่าง 10 ปีที่แล้วกับปัจจุบัน				
	1 ดีขึ้นมาก	2 ดีขึ้นเล็กน้อย	3 ไม่เปลี่ยนแปลง	4 แย่ลงเล็กน้อย	5 แย่ลงมาก
1. ความจำเกี่ยวกับหน้าตาคนในครอบครัวหรือญาติ					
2. ความจำเกี่ยวกับชื่อคนในครอบครัวหรือญาติ					
3. ความจำในรายละเอียดของคนในครอบครัวหรือญาติเกี่ยวกับอาชีพที่อยู่					
4. ความจำในเหตุการณ์ที่เกิดขึ้นเมื่อ 2-3 วันที่ผ่านมา					
5. ความจำในเรื่องที่สนทนาไปเมื่อ 2-3 วันที่ผ่านมา					
6. พูดคุยอย่างต่อเนื่องโดยไม่ลืมสิ่งที่จะพูด					
7. จำได้ว่าตอนนี้พักอาศัยอยู่ที่ไหน					
8. จำได้ว่าวันนี้เป็น วัน เดือน อะไร					
9. ความจำเกี่ยวกับที่ประจำที่ใช้เก็บของในบ้าน					
10. จำได้ว่าวางของไว้ที่ไหน					
11. จำเหตุการณ์เกี่ยวกับตนในวัยเด็ก					
12. จำสิ่งที่ตนได้เรียนรู้ในวัยเด็ก					
13. ทราบเหตุการณ์ที่สำคัญในอดีต					
14. ความสามารถในการปรับตัวเข้ากับการเปลี่ยนแปลงในชีวิตประจำวัน					
15. สามารถใช้เครื่องมือที่คุ้นเคยภายในบ้าน					
16. สามารถเรียนรู้การใช้เครื่องมือ เครื่องใช้ใหม่ๆ ในบ้าน					
17. สามารถเรียนรู้สิ่งใหม่ ๆ ทั่ว ๆ ไป					
18. สามารถเข้าใจความหมายของคำแปลก ๆ					
19. สามารถเข้าใจบทความในหนังสือพิมพ์หรือนิตยสาร					
20. สามารถติดตามเรื่องราวต่างๆ ในวิทยุ หรือโทรทัศน์					
21. สามารถติดต่อลูกหลาน ญาติหรือกิจธุระทั่วๆ ไป					
22. ความสามารถในการตัดสินใจเรื่องต่าง ๆ ในชีวิตประจำวัน					
23. ความสามารถในการใช้จ่าย					
24. ความสามารถในการจัดสรรเรื่องเงิน					
25. สามารถประมาณได้ว่าจะใช้สิ่งของประมาณเท่าไร เช่น จะซื้ออาหารเท่าไร หรือกะเวลาที่ใช้ในการทำกิจกรรมต่างๆ เช่น ใช้เวลาในการเดินทางเท่าไร					
26. สามารถที่จะเข้าใจในสิ่งที่เกิดขึ้น พร้อมกับให้เหตุผลในสิ่งนั้นได้					
27. สามารถร้องเพลงที่เคยร้อง หรือ สวดมนต์ที่เคยสวด					
28. สามารถเลือกใช้เครื่องมือเครื่องใช้ต่าง ๆได้อย่างเหมาะสมกับงาน					
29. การพูดจาหรือถามซ้ำๆ					
30. สามารถปฏิบัติกิจวัตรประจำวันของตนเอง					
31. สามารถเดินทางไป-กลับสถานที่ที่คุ้นเคยได้โดยลำพัง					
32. สามารถทำงานที่เคยทำ					

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.