Critical Care Explorations

OPEN

ICU Delirium-Prediction Models: A Systematic Review

Matthew M. Ruppert, BS[1,2]; Jessica Lipori, BA[1,2]; Sandip Patel, MD[1,2]; Elizabeth Ingersent[1,2]; Julie Cupka, BS[1,2]; Tezcan Ozrazgat-Baslanti, PhD[1,2]; Tyler Loftus, MD[2,3]; Parisa Rashidi, PhD[2,4]; Azra Bihorac, MD, MS[1,2]

Objective: Summarize performance and development of ICU delirium-prediction models published within the past 5 years.

Data Sources: Systematic electronic searches were conducted in April 2019 using PubMed, Embase, Cochrane Central, Web of Science, and Cumulative Index to Nursing and Allied Health Literature to identify peer-reviewed studies.

Study Selection: Eligible studies were published in English during the past 5 years that specifically addressed the development, validation, or recalibration of delirium-prediction models in adult ICU populations.

Data Extraction: Screened citations were extracted independently by three investigators with a 42% overlap to verify consistency using the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies.

Data Synthesis: Eighteen studies featuring 23 distinct prediction models were included. Model performance varied greatly, as assessed by area under the receiver operating characteristic curve (0.62–0.94), specificity (0.50–0.97), and sensitivity (0.45–0.96). Most models used data collected from a single time point or window to predict the occurrence of delirium at any point during hospital or ICU admission, and lacked mechanisms for providing pragmatic, actionable predictions to clinicians.

Conclusions: Although most ICU delirium-prediction models have relatively good performance, they have limited applicability to clinical practice. Most models were static, making predictions based on data collected at a single time-point, failing to account for fluctuating conditions during ICU admission. Further research is needed to create clinically relevant dynamic delirium-prediction models that can adapt to changes in individual patient physiology over time and deliver actionable predictions to clinicians.

Key Words: delirium; intensive care unit; prediction model; risk prediction; systematic review

Delirium is a transient condition consisting of altered attention and consciousness common in hospital settings (1). Delirium has a particularly high prevalence in the ICU, ranging from 25% to 87% (2–4). Some factors associated with increased risk for ICU delirium include the following: older age, lower levels of education, history of hypertension, alcohol abuse, higher Acute Physiology, Age, Chronic Health Evaluation (APACHE) II scores, and use of sedative and analgesic medications (2, 5, 6). The use of benzodiazepines for mechanical ventilation carries a particularly high risk for delirium compared with other sedatives (7, 8). Environmental factors, including isolation, use of physical restraints, and prolonged exposure to light and sound have also been associated with delirium (9, 10).

ICU delirium is strongly associated with adverse outcomes, including increased hospital length of stay, greater morbidity and mortality, poor cognitive recovery, slower rates of overall recovery, and increased cost of care (3, 11, 12). Delirium assessments such as the confusion-assessment method for the ICU (CAM-ICU) and the Intensive Care Delirium Screening Checklist (ICDSC) have been shown to be effective in diagnosing delirium (13, 14) and their use is recommended under current clinical practice guidelines (15). However, these assessments are sometimes not trusted or understood by ICU staff and are therefore inconsistently applied (16–18).

The use of prediction models has shown promise in predicting several types of delirium, including postoperative and subsyndromal delirium as well as delirium in the ICU. These predictions can be used by clinicians as decision support for preventing and treating delirium (19, 20). However, clinical adoption of delirium-prediction models has been limited, perhaps because most models

[1]Department of Medicine, University of Florida, Gainesville, FL.

[2]Precision and Intelligent Systems in Medicine (Prisma[P]), Division of Nephrology, Hypertension, & Renal Transplantation, Department of Medicine, University of Florida, Gainesville, FL.

[3]Department of Surgery, University of Florida, Gainesville, FL.

[4]Department of Biomedical Engineering, University of Florida, Gainesville, FL.

Crit Care Expl 2020; 2:e0296

DOI: 10.1097/CCE.0000000000000296

are neither readily integrable into physician workflows or provide little clinical utility. Machine-learning techniques may abrogate these weaknesses, but contemporary descriptions of these techniques are sparse.

This systematic review is meant to build on the work of van Meenen et al (21), which summarized delirium-prediction model efficacy and characteristics up through 2014 and fills in the gaps of more recent delirium-prediction reviews that were limited to a specific patient population (22–24) or type of study (25).

MATERIALS AND METHODS

Search Strategy and Selection Criteria

PubMed, Embase, Cochrane Central, Web of Science, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) were systematically searched for articles relating to delirium-prediction models among adult ICU patients.

An ICU delirium-prediction model was defined as any model or algorithm applied to critical care patients that incorporated at least one clinical factor measured during a hospital admission to assign an estimated risk of developing delirium during a hospital stay. Studies that specifically addressed the development, validation, or recalibration of prediction models in adult ICU populations were included. Models that were designed to predict delirium in the context of substance abuse or withdrawal were excluded. Abstract only studies were excluded.

Search terms were tailored to use medical subject headings or subject headings embedded in each database. Each search query was the union of three search components: delirium, ICU, and prediction. The first component, delirium, contained delirium-associated terms and subject headings with words including but not limited to "delirium," "ICU syndrome," "acute confusion," and "CAM." The second component, ICU, contained ICU-associated terms and subject headings with words including but not limited to "ICU," "Intensive Care Unit," "Critical Care," and "Critically Ill." The third component, prediction, contained prediction-associated terms and subject headings with words including but not limited to "predict," "model," "risk," and "risk assessment." A full list of the search terms for each database is available in **Supplement A** (http://links.lww.com/CCX/A447).

In addition to the three query components above, search results were restricted to papers published in English within the past 5 years. Database searches were performed on April 25, 2019. There were 4,940 articles remaining after the search results were compiled and the duplicates removed (**Fig. 1**). The 4,940 articles were divided among four authors. Each article's title and abstract were reviewed by two authors independently to verify if the article described a delirium-prediction model that was applied to critically ill patients that did not focus on delirium as a result of substance abuse or on the terminally ill. All disagreements were settled by the lead author, reducing the number of articles to 20.

Data Extraction and Synthesis

The 20 articles were divided into three groups of 10, with a three-article overlap between each group, to verify consistency across authors. The data from each article group were then extracted independently by three authors using the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies Checklist for Systematic Reviews of Prediction Studies, excluding the treatments received element (26). Extracted data elements included study design, participant descriptions and recruitment methods, predicted outcomes, candidate predictors, final predictors, sample size, model development, model performance, model evaluation, study results, interpretation of those results, and treatment of missing data. During data extraction, one article was removed, because it was not a true prediction model and one article was removed as its main focus was not on ICU patients.

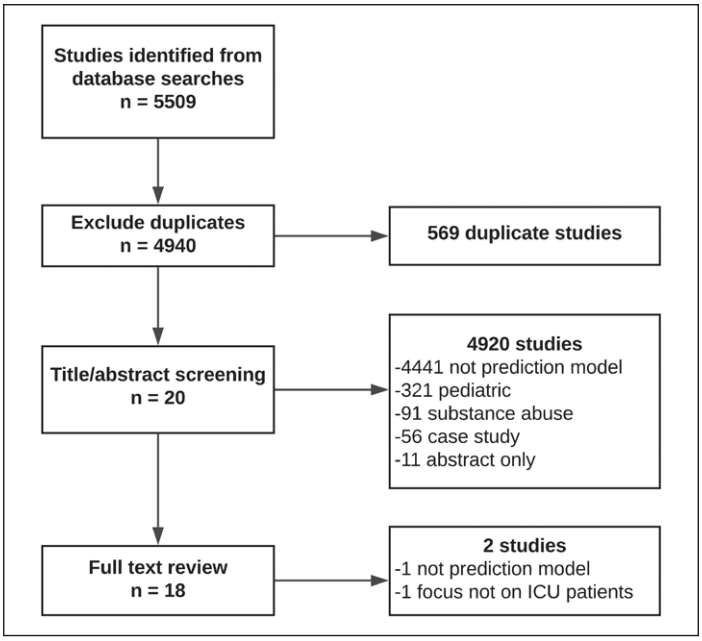

Figure 1. Consort diagram of studies included in review.

Assessment of Bias

Bias was assessed via the Prediction model Risk Of Bias ASsessment Tool (27) that was specifically developed by a panel of experts to evaluate bias in studies of prognostic and prediction models. The risk of bias is evaluated with respective to four categories (participants, predictors, outcome, and analysis) with two levels (high vs low risk). Risk of bias was evaluated independently by three authors during data extraction.

RESULTS

Study Characteristics

Of the 18 included studies, 12 were primarily concerned with the development of new prediction models (28–39) and six with the validation of existing models (20, 22, 40–43) (**Table 1**). These studies included 23 risk-prediction models, of which 12 were developed in the included studies. Of the 12 model development studies, nine were prospective cohorts and three were retrospective cohorts. Sample sizes ranged from 94 to 3,284 participants (33, 34). Those studies that validated existing models had a sample size ranging from 38 to 2,178 participants (42, 43). Delirium was most commonly assessed using CAM-ICU, though several studies used CAM (31),

TABLE 1. Overview of the Cohorts and Modeling Methodologies Used in Each of the Included Models Along With Their Respective Model Performances

Model	Cohort/ Study Type	Sample Size Development/ Validation	Delirium Prevalence in Development/ Validation Cohorts, *n* (%)	Patient Description	Variable Selection Methodology	Model Methodology	Performance Area Under the Receiver Operating Characteristic Curve (95% CI)	Significant Predictors, *n*
PRE-DELIRIC (Azuma et al [20], 2019)	Retrospective external validation	NA/70	NA 14(20)	SC adult MICU	NA	Logistic regression	0.89 (not reported)	10
Chaiwat et al (28), 2019	Prospective development	250/NA	61(24) NA	SC adult SICU	Logistic regression to narrow variables from literature review	Logistic regression	0.84 (0.79–0.90)	5
Lanzhou model (Chen et al [29], 2017)	Prospective development	310/310	160(26)[a]	SC adult ICU	Logistic regression	Logistic regression	0.78 (not reported)	11
Fan et al (30), 2019	Prospective development	336/24	68(20) 46 (20)	SC adult ICU	Univariate analysis and backward stepwise logistic regression	Logistic regression	0.90 (0.86–0.94)	7
PRE-DELIRIC (Green et al [40], 2019)	Retrospective external validation	NA/455	NA 160(35)	SC adult MICU/ SICU	NA external validation	Logistic regression	0.79 (0.75–0.83)	10
R-PRE-DELIRIC (Green et al [40], 2019)	Retrospective external validation	NA/455	NA 160(35)	SC adult MICU/ SICU	NA external validation	Logistic regression	0.79 (0.75–0.83)	10
E-PRE-DELIRIC (Green et al [40], 2019)	Retrospective external validation	NA/455	NA 160(35)	SC Adult MICU/ SICU	NA external validation	Logistic regression	0.72 (0.67–0.77)	9
Lanzhou model (Green et al [40], 2019)	Retrospective external validation	NA/455	NA 160(35)	SC adult MICU/ SICU	NA external validation	Logistic regression	0.77 (0.72–0.81)	11
Kim et al (31), 2016	Prospective development	561/553	112(20) 99(18)	SC elderly SICU	Backwards stepwise logistic regression	Logistic regression	0.94 (0.91–0.97)	9
R-PRE-DELIRIC (Lee et al [22], 2017)	Prospective external validation	NA/600	NA 83(14)	SC adult CICU	NA external validation	Logistic regression	0.75 (0.72–0.79)	10
Katznelson (Lee et al [22], 2017)	Prospective external validation	NA/600	NA 83(14)	SC adult CICU	NA external validation	Logistic regression	0.62 (0.58–0.66)	6
PRE-DELIRIC (Linkaitė et al [43], 2018)	Prospective external validation	NA/38	NA 22(58)	SC adult ICU	NA external validation	Logistic regression	0.71 (0.54–0.89)	10
Marra et al (32), 2018	Prospective development	810/NA	606(75) NA	MC adult SICU/ MICU	Maximum likelihood estimation	Logistic regression	Not reported	14
Moon et al (33), 2018	Retrospective development	2299/985	485(21) 203(21)	SC adult SICU/ MICU	Information value and logistic regression	Logistic regression	0.9	11
Moon et al (33), 2018	Prospective internal validation	NA/263	NA 48(15)	SC adult SICU/ MICU		Logistic regression	0.94	11
Moon et al (33), 2018	Prospective internal validation	NA/431	NA 55(21)	SC adult SICU/ MICU		Logistic regression	0.88	11

(*Continued*)

TABLE 1. (*Continued*). Overview of the Cohorts and Modeling Methodologies Used in Each of the Included Models Along With Their Respective Model Performances

Model	Cohort/ Study Type	Sample Size Development/ Validation	Delirium Prevalence in Development/ Validation Cohorts, *n* (%)	Patient Description	Variable Selection Methodology	Model Methodology	Performance Area Under the Receiver Operating Characteristic Curve (95% CI)	Significant Predictors, *n*
Moon et al (33), 2018	Prospective external validation	NA/325	NA 114(26)	SC adult SICU/ MICU		Logistic regression	0.72	11
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	SVM with RBF kernels	Not reported	1*
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	Linear SVM	Not reported	1*
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	Linear discriminant analysis	Not reported	1*
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	Quadratic discriminant analysis	Not reported	1*
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	ELM with RBF kernels	Not reported	1*
Oh et al (34), 2018	Prospective development	94/NA	39(42) NA	SC adult SICU/ MICU	Normalized mutual information feature selection	Linear ELM	Not reported	1*
PRE-DELIRIC (Paton et al [41], 2016)	Prospective external validation	NA/44	NA 15(36)	SC adult ICU	NA external validation	Logistic regression	Not reported	10
Sakaguchi et al (35), 2018	Retrospective development	120/NA	38(32) NA	SC adult CICU	Forward stepwise logistic regression	Logistic regression	0.89 (not reported)	6
Stukenberg et al (36), 2016	Retrospective development	996/NA	161(16) NA	SC elderly CICU	Univariate and multivariate analyses	Logistic regression	Not Reported	3
VR-PRE-DELIRIC (van den Boogaard et al [37], 2014)	Prospective development	1824	363(20) NA	MC adult ICU	PRE-DELIRIC model variables	Logistic regression	0.77 (0.74–0.79)	10
Wang et al (38), 2018	Prospective development	1692/NA	Not Reported	SC adult SICU	Expert opinion	Logistic regression	Not reported	1
PRE-DELIRIC (Wassenaar et al [42], 2018)	Prospective external validation	NA/2178	NA 467(21)	MC adult ICU	NA external validation	Logistic regression	0.74 (0.71–0.76)	10
E-PRE-DELIRIC (Wassenaar et al [42], 2018)	Prospective external validation	NA/2178	NA 467(21)	MC adult ICU	NA external validation	Logistic regression	0.68 (0.66–0.71)	9
E-PRE-DELIRIC (Wassenaar et al [39], 2015)	Prospective development	1692/952	481(25) 208(22)	MC adult ICU	Backward selection with logistic regression	Logistic regression	0.75 (0.71–0.79)	9

CICU = cardiac ICU, ELM = extreme learning machine, E-PRE-DELIRIC = early prediction model for delirium, PRE-DELIRIC = PREdiction of DELIRium in ICu patients, MC = multicenter, MICU = medical ICU, NA = not applicable, RBF = radial basis function, R-PRE- DELIRIC = Recalibrated PREdiction of DELIRium in ICu patients, SC = single-center, SICU = surgical ICU, SVM = support vector machine.

Nursing Delirium Screening Scale (NuDESC) (31, 39), diagnostic and statistical manual of mental disorders (DSM) (20, 36), and ICDSC (20, 35, 42).

Risk Factors

Six studies evaluated existing delirium-prediction models, including the prediction model for delirium (PRE-DELIRIC) and the early prediction model for delirium (E-PRE-DELIRIC). The PRE-DELIRIC model consists of nine risk factors: age, APACHE II score, coma, sedative use, morphine use, serum urea, metabolic acidosis, urgent admission, and admission category (Table 1). The E-PRE-DELIRIC model considered 18 candidate predictors chosen by literature review and input from an expert panel of physicians, and nine of these predictors were included in the final model (39). The PRE-DELIRIC and E-PRE-DELIRIC share three predictors in common: age, admission category, and urgent admission. Additionally, both use some marker of renal function, namely, blood urea nitrogen (BUN) and urea concentration. Among other differences, E-PRE-DELIRIC includes a greater number of predictors relating to patient's predisposing factors, such as history of cognitive impairment and history of alcohol use.

Although rationale for candidate predictors was not consistently available, many studies that sought to develop new models for delirium prediction employed a combination of literature review and expert opinion. Some of the more frequently selected candidate predictors were also common to the aforementioned PRE-DELIRIC and E-PRE-DELIRIC models (**Fig. 2**). For example, age was included in six studies (28, 29, 31, 32, 37, 39), APACHE II score was included in four studies (29, 30, 32, 37), and some marker of renal function (blood urea concentration, serum creatinine, or a BUN/Cr ratio) was used in three studies (31, 35, 37). Other commonly considered predictors included mechanical ventilation, urgent admission, and use of antipsychotics, sedatives, and benzodiazepines.

Two studies evaluated greater than 40 candidate predictors. One employed machine learning to develop delirium-prediction models (34) and the other used an informative value calculation and multiple regression analyses "based on literature review" to narrow down the number of predictors in a stepwise manner to use in modeling (31). A few studies went beyond the existing approach of selecting candidate predictors within 24 hours of admission and instead included various predictors to be collected daily. This was performed in an effort to predict delirium dynamically and evaluate recurrent or ongoing delirium, which are not assessed by PRE-DELIRIC and E-PRE-DELIRIC.

Predictive Model Development and Performance

Our review included 23 prediction models, of which 12 were developed in the included studies (Table 1). These models used various numbers of predictors ranging from one to 4,211 (33, 38). Predictors were most commonly chosen for inclusion by logistic regression, although one model was developed using machine-learning techniques (34) and one recalibrated an existing model (37). Of the 11 models developed using logistic regression, three employed additional bootstrapping to allow for better calibration and adjust for overfitting. Multiple methods were used for determining, which variables should be included in multiple logistic regression or the final model. These included preselection based on literature review, univariate regression, machine-learning techniques, and preselection of factors from a previous model. In the final models, most used regression coefficients to establish either a sum score or a score chart with scores stratified into different risk subgroups. Eighteen of the models measured discrimination with area under the receiver operating characteristic curve (AUROC), reporting values between 0.62 and 0.94 (22, 31). Studies that statistically assessed the calibration of their models used the Hosmer-Lemeshow goodness of fit test and calibration plots. Although most studies were concerned with a binary outcome (delirious vs nondelirious) occurring at any time, one (32) developed a model aimed at predicting daily transitions between multiple states (normal, delirious, comatose, discharge, or death).

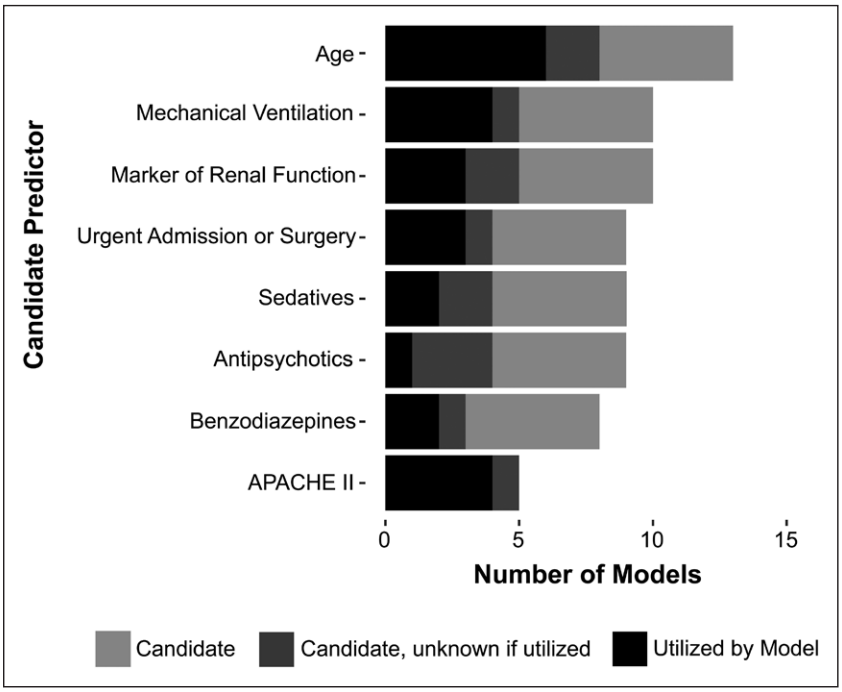

Figure 2. Prevalence of predictors considered in at least five models. APACHE II = Acute Physiology, Age, Chronic Health Evaluation II.

Methodological Evaluation of Models and Risk of Bias

This review identified 23 predictive models for delirium in an ICU setting. Of the 23 models, five were externally validated by a separate study (20, 22, 40–43) and another seven were internally validated by a separate or split cohort (28–33, 39). Five studies were retrospective and assessed delirium incidence through electronic health record data (20, 33, 35, 36, 40). Studies had various methods of assessing delirium including CAM-ICU, CAM, NuDESC, DSM, and ICDSC.

Models predicted delirium using various numbers of risk factors ranging from 1 to 14 (32, 38). Thirteen of

the 23 models, which reported the number of predictors in the final model, included six to 11 predictors (22, 29–31, 33, 35, 37, 39). Across the models reviewed, there was a significant amount of overlap between candidate and final predictors, including age, APACHE II, renal function, sedative use, benzodiazepine use, mechanical ventilation, and urgent admission (Fig. 2). Many of these common factors are consistent with Zaal et al (44) who found 11 strongly supported factors in literature (age, dementia, hypertension, trauma, emergency surgery, APACHE II, coma [sedative-associated], previous delirium within 24 hr, mechanical ventilation, metabolic acidosis, and multiple organ failure/Sequential Organ Failure Assessment score). Despite the overlap between risk factors, there were differences in how predictors were defined and measured, making it difficult to compare directly the relative importance or weight of the predictor in each model. For example, mental status was assessed by mini–mental state examination score (30), history of dementia (29), use of Alzheimer medication (32), and history of cognitive impairment (39).

Most studies used logistic regression in the development of their models, one study used machine-learning techniques (34), and one recalibrated an existing model (37). One of these models took the unique approach of applying machine-learning techniques to analyze electrocardiograms to predict delirium (34). Additionally, one developed a model aimed at predicting daily transitions between multiple states (normal, delirious, comatose, discharge, or death) rather than predicting the development of delirium at any time (32). Eighteen models and six validation studies reported AUROC for the development and validation of their model, with values ranging from 0.62 to 0.94 (22, 31). Models tended to exhibit poorer discrimination in validation than in development, and calibration was inconsistently reported.

The major risks of bias in included papers were assessment of outcome, selection of candidate predictors, sample size, and treatment of missing data (26, 27). Two models were at risk for bias due to having multiple delirium-assessment measures (31, 39). Retrospective studies had an increased risk of bias due to both the problem of assessing for delirium retrospectively and that the outcome was often assessed by the same researchers that selected candidate predictors. Selection of predictors was a further source of bias. Although some selected candidate predictors from literature review (31–33, 39), many gave a little or no justification for candidate predictor selection (28, 30, 35), and some selected final predictors without prior analysis (29). Most studies did not report missing data or management of missing data. Five of 12 models that were developed in the included studies were at high risk for overfitting due to a low ratio of delirium incidence compared with the number of candidate predictors (28, 30, 31, 33, 35).

Validation of Models

Our review included six studies primarily focused on the validation of existing models (20, 22, 40–43) (Table 1). These studies used logistic regression and calibration curves for their analysis with the exception of one (41) that did not include any statistical analysis. For model development studies, many included internal validation of their models. Sixteen of the 23 models were validated either internally through a split cohort or externally using data from a separate institution (20, 22, 29–31, 33, 39–43). The remaining seven models either lacked any validation or merely used bootstrapping (28, 32, 34–36, 38). Of the seven internally validated models, cohorts were split either temporally or randomly. Five studies externally validated the PRE-DELIRIC model (20, 40–43), two validated the E-PRE-DELIRIC (40, 42), two the recalibrated PRE-DELIRIC (22, 40), one the model proposed by Green et al (40), and one the Katznelson model (22).

The reported AUROC of the externally validated models ranged from 0.62 to 0.89 (20, 22), with the majority of models exhibiting poorer performance during validation. Of the eight models with an AUROC of 0.75 or greater in development, six were validated in a split or separate cohort (29–31, 33, 37, 39) and only four maintained an average AUROC of 0.75 or higher during validation (29–31, 33).

Risk-Prediction Performance

Twenty models stratified patients into two to five risk groups using cutoff values calculated with Youden index (28, 39). Studies most commonly reported the sensitivity, specificity, positive predictive value, and negative predictive value associated with the determined threshold, though positive and negative likelihood ratios were also reported by some studies (28, 45). Although models varied as to whether they had a higher sensitivity (0.45–0.96) (33, 34) or specificity (0.50–0.97) (37, 43), the negative predictive value was generally much higher than the positive predictive value, indicating that a higher proportion of patients were erroneously assigned to the high-risk of delirium group than the proportion of delirious patients to the low-risk group.

DISCUSSION

Application to Practice

Delirium has classically been described as a transient, waxing and waning condition. Interest in creating delirium-prediction models emerges from clinicians' difficulties in recognizing the signs and risk factors of this multifactorial and dynamic condition that can evolve on an hourly basis. Given time constraints, uncertainty, and changing conditions, delirium is often unrecognized in the ICU or is recognized in a delayed fashion. This difficulty is compounded by the lack of a gold standard for delirium assessment and the low reliability of delirium assessments without significant and continued training (46). The studies reviewed generally seemed to voice an appreciation for the challenge of predicting delirium amidst the high demands of the critical care setting. Thus, considerations such as the ease-of-use of the developed models were often noted. Wassenaar et al (42), for example, noted that ICU physicians rated the user convenience of E-PRE-DELIRIC superior to PRE-DELIRIC despite the latter having superior performance in predicting delirium. However, the trade-off between the ease of implementing delirium-prediction models in clinical practice and the actual predictive power or clinical utility of these models need not persist, given recent advances in automation and machine learning. Moon et al (33) created and implemented a delirium-prediction algorithm in an electronic medical record system, which updated every day at midnight, making the system readily accessible to healthcare providers. Unfortunately, the algorithm's low positive predictive value (0.52) may lead to alarm fatigue. This model, along with Marra et al's (32)

model, avoids the temporally static prediction paradigm that is a limiting factor of most ICU delirium-prediction models. The most frequently studied and cited models rely on a set of factors collected at a single time-point to predict whether delirium will occur at any point during the remainder of the ICU admission. Such models are unable to account for the dynamic condition of patients and delirium itself, each of which can change on an hourly basis.

Strengths and Limitations

We conducted a thorough review of PubMed, Embase, Cochrane Central, Web of Science, and CINAHL databases to ascertain a comprehensive picture of what is available in the literature on delirium predictive models in the ICU. This review builds on the work of van Meenen et al (21) by including articles published from 2014 to April 25, 2019. This is important, because machine-learning approaches have emerged and evolved during these recent years. Other systematic reviews have exclusively studied older adults (23, 24), excluded validation studies (25), or were restricted to cardiac surgery patients (22). The studies included in our review were summarized in detail in terms of candidate predictors, final model predictors, and risk of bias (26). Limitations of our review include the following: limiting the inclusion criteria to studies published within the past 5 years, including only ICU delirium-prediction models and excluding non-English studies. Studies that developed and validated delirium-prediction models not limited to the context of the ICU, especially in surgical patients, have used the attending surgeon (47) or innovative methods such as electroencephalography (48, 49) and near-field infrared spectroscopy (48) to predict delirium, which may provide valuable information for postoperative ICU patients.

CONCLUSIONS

Many ICU delirium-prediction models have been developed and validated within the last 5 years. Most of these models were developed with similar statistical methods and use common predictive factors, though inconsistencies in how these factors were assessed and used obviate a consensus, as does the risk of bias. External validation efforts have primarily focused on a few select models, especially PRE-DELIRIC and E-PRE-DELIRIC, making external validation of competing models an area where further research is needed. Most delirium-prediction models use a single snapshot in time, usually within 24 hours of admission and do not account for fluctuations in patients' conditions during ICU admission. This is inconsistent with critical illness and delirium pathophysiology. Further research is needed to create clinically relevant dynamic delirium-prediction models, which can not only adapt over time but deliver pragmatic and actionable predictions to clinicians.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccxjournal).

Supported, in part, by National Institutes of Health (NIH)/National Institute of General Medical Sciences (NIGMS) RO1 GM-110240 (to Mr. Ruppert, Dr. Ozrazgat-Baslanti, Dr. Rashidi, and Dr. Bihorac), Davis Foundation – University of Florida (to Mr. Ruppert), NIH/NIGMS P50 GM111152 (to Drs. Ozrazgat-Baslanti and Bihorac), NIH/NIGMS T32 GM-008721 (to Dr. Loftus), National Science Foundation Faculty Early Career Development Program 1750192 (to Dr. Rashidi), and NIH/National Institute of Biomedical Imaging and Bioengineering 1R21EB027344-01 (to Drs. Rashidi and Bihorac).

The authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: abihorac@ufl.edu

REFERENCES

1. Setters B, Solberg LM: Delirium. *Prim Care* 2017; 44:541–559
2. Kanova M, Sklienka P, Roman K, et al: Incidence and risk factors for delirium development in ICU patients - a prospective observational study. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2017; 161:187–196
3. Maldonado JR: Acute brain failure: Pathophysiology, diagnosis, management, and sequelae of delirium. *Crit Care Clin* 2017; 33:461–519
4. Foroughan M, Delbari A, Said SE, et al: Risk factors and clinical aspects of delirium in elderly hospitalized patients in Iran. *Aging Clin Exp Res* 2016; 28:313–319
5. Mori S, Takeda JR, Carrara FS, et al: Incidence and factors related to delirium in an intensive care unit. *Rev Esc Enferm USP* 2016; 50:587–593
6. Vasilevskis EE, Han JH, Hughes CG, et al: Epidemiology and risk factors for delirium across hospital settings. *Best Pract Res Clin Anaesthesiol* 2012; 26:277–287
7. Pandharipande PP, Pun BT, Herr DL, et al: Effect of sedation with dexmedetomidine vs lorazepam on acute brain dysfunction in mechanically ventilated patients: The MENDS randomized controlled trial. *JAMA* 2007; 298:2644–2653
8. Ouimet S, Kavanagh BP, Gottfried SB, et al: Incidence, risk factors and consequences of ICU delirium. *Intensive Care Med* 2007; 33:66–73
9. Van Rompaey B, Elseviers MM, Schuurmans MJ, et al: Risk factors for delirium in intensive care patients: A prospective cohort study. *Crit Care* 2009; 13:R77
10. van de Pol I, van Iterson M, Maaskant J. Effect of nocturnal sound reduction on the incidence of delirium in intensive care unit patients: An interrupted time series analysis. *Intensive Crit Care Nurs*. 2017; 41:18–25.
11. Slooter AJ, Van De Leur RR, Zaal IJ: Delirium in critically ill patients. *Handb Clin Neurol* 2017; 141:449–466
12. Vasilevskis EE, Chandrasekhar R, Holtze CH, et al: The cost of ICU delirium and coma in the intensive care unit patient. *Med Care* 2018; 56:890–897
13. 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; 29:1370–1379
14. 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
15. Barr J, Fraser GL, Puntillo K, et al; American College of Critical Care Medicine: Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med* 2013; 41:263–306
16. Elliott SR: ICU delirium: A survey into nursing and medical staff knowledge of current practices and perceived barriers towards ICU delirium in the intensive care unit. *Intensive Crit Care Nurs* 2014; 30: 333–338
17. Zamoscik K, Godbold R, Freeman P: Intensive care nurses' experiences and perceptions of delirium and delirium care. *Intensive Crit Care Nurs* 2017; 40:94–100
18. Trogrlić Z, Ista E, Ponssen HH, et al: Attitudes, knowledge and practices concerning delirium: A survey among intensive care unit professionals. *Nurs Crit Care* 2017; 22:133–140
19. Liang CK, Chu CL, Chou MY, et al: Developing a prediction model for post-operative delirium and long-term outcomes among older patients receiving elective orthopedic surgery: A prospective cohort study in Taiwan. *Rejuvenation Res* 2015; 18:347–355
20. Azuma K, Mishima S, Shimoyama K, et al: Validation of the prediction of delirium for intensive care model to predict subsyndromal delirium. *Acute Med Surg* 2019; 6:54–59

21. van Meenen LC, van Meenen DM, de Rooij SE, et al: Risk prediction models for postoperative delirium: A systematic review and meta-analysis. *J Am Geriatr Soc* 2014; 62:2383–2390
22. Lee A, Mu JL, Joynt GM, et al: Risk prediction models for delirium in the intensive care unit after cardiac surgery: A systematic review and independent external validation. *Br J Anaesth* 2017; 118:391–399
23. Kalimisetty S, Askar W, Fay B, et al: Models for predicting incident delirium in hospitalized older adults: A systematic review. *J Patient Cent Res Rev* 2017; 4:69–77
24. Lindroth H, Bratzke L, Purvis S, et al: Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open* 2018; 8:e019223
25. Kassie GM, Nguyen TA, Kalisch Ellett LM, et al: Do risk prediction models for postoperative delirium consider patients' preoperative medication use? *Drugs Aging* 2018; 35:213–222
26. Moons KG, de Groot JA, Bouwmeester W, et al: Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Med* 2014; 11:e1001744
27. Wolff RF, Moons KGM, Riley RD, et al; PROBAST Group†: PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019; 170:51–58
28. 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
29. Chen Y, Du H, Wei BH, et al: Development and validation of risk-stratification delirium prediction model for critically ill patients: A prospective, observational, single-center study. *Medicine (Baltimore)* 2017; 96:e7543
30. Fan H, Ji M, Huang J, et al: Development and validation of a dynamic delirium prediction rule in patients admitted to the intensive care units (DYNAMIC-ICU): A prospective cohort study. *Int J Nurs Stud* 2019; 93:64–73
31. Kim MY, Park UJ, Kim HT, et al: DELirium prediction based on hospital information (Delphi) in general surgery patients. *Medicine (Baltimore)* 2016; 95:e3072
32. Marra A, Pandharipande PP, Shotwell MS, et al: Acute brain dysfunction: Development and validation of a daily prediction model. *Chest* 2018; 154:293–301
33. Moon KJ, Jin Y, Jin T, et al: Development and validation of an automated delirium risk assessment system (Auto-DelRAS) implemented in the electronic health record system. *Int J Nurs Stud* 2018; 77:46–53
34. Oh J, Cho D, Park J, et al: Prediction and early detection of delirium in the intensive care unit by using heart rate variability and machine learning. *Physiol Meas* 2018; 39:035004
35. Sakaguchi T, Watanabe M, Kawasaki C, et al: A novel scoring system to predict delirium and its relationship with the clinical course in patients with acute decompensated heart failure. *J Cardiol* 2018; 71:564–569
36. Stukenberg S, Franck M, Spies CD, et al: How can postoperative delirium be predicted in advance? A secondary analysis comparing three methods of early assessment in elderly patients. *Minerva Anestesiol* 2016; 82:751–759
37. 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; 40:361–369
38. Wang S, Sigua NL, Manchanda S, et al: Preoperative STOP-BANG scores and postoperative delirium and coma in thoracic surgery patients. *Ann Thorac Surg* 2018; 106:966–972
39. 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; 41:1048–1056
40. Green C, Bonavia W, Toh C, et al: Prediction of ICU delirium: Validation of current delirium predictive models in routine clinical practice. *Crit Care Med* 2019; 47:428–435
41. Paton L, Elliott S, Chohan S: Utility of the PRE-DELIRIC delirium prediction model in a Scottish ICU cohort. *J Intensive Care Soc* 2016; 17:202–206
42. Wassenaar A, Schoonhoven L, Devlin JW, et al: Delirium prediction in the intensive care unit: Comparison of two delirium prediction models. *Crit Care* 2018; 22:114
43. Linkaitė G, Riauka M, Bunevičiūtė I, et al: Evaluation of PRE-DELIRIC (PREdiction of DELIRium in ICu patients) delirium prediction model for the patients in the intensive care unit. *Acta Med Litu* 2018; 25:14–22
44. Zaal IJ, Devlin JW, Peelen LM, et al: A systematic review of risk factors for delirium in the ICU. *Crit Care Med* 2015; 43:40–47
45. van den Boogaard M, Pickkers P, Slooter AJ, et al: Development and validation of PRE-DELIRIC (PREdiction of DELIRium in ICu patients) delirium prediction model for intensive care patients: Observational multicentre study. *BMJ* 2012; 344:e420
46. Maybrier HR, Mickle AM, Escallier KE, et al; PODCAST Research Group: Reliability and accuracy of delirium assessments among investigators at multiple international centres. *BMJ Open* 2018; 8:e023137
47. Davoudi A, Ebadi A, Rashidi P, et al: Delirium prediction using machine learning models on preoperative electronic health records data. *Proc IEEE Int Symp Bioinformatics Bioeng* 2017; 2017:568–573
48. Momeni M, Meyer S, Docquier MA, et al: Predicting postoperative delirium and postoperative cognitive decline with combined intraoperative electroencephalogram monitoring and cerebral near-infrared spectroscopy in patients undergoing cardiac interventions. *J Clin Monit Comput* 2019; 33:999–1009
49. Fritz BA, Kalarickal PL, Maybrier HR, et al: Intraoperative electroencephalogram suppression predicts postoperative delirium. *Anesth Analg* 2016; 122:234–242