Predictive Models/Analytic Tools for Delirium in Acute Hospital Setting Evidence Review

Clinical Question:

To answer the clinical question, "In adult patient's 18 years and above, what is the quality and consistency of the evidence with the use of predictive models/analytics compared to current practice in the accuracy of predicting early/identification of delirium within the hospital stay? An extensive search across multiple databases, including PubMed, Clinical Key, CINAHL, Cochrane Library, and Google Scholar, as well as Professional Nursing Organizations websites were with the search strategy methodology confirmed by a librarian. The search revealed that there was an extensive amount of literature, with **twenty-four final articles and one professional nursing news brief** that met the inclusion criteria for the review.

Using the Johns Hopkins Evidence-Based Model Hierarchical and Appraisal Tools (2022) an evaluation of the quality and consistency of the evidence was conducted. Out of twenty-five references, twenty-two were rated Level III for consistent results, good methodology, and appropriate sample sizes, while three were Level IV. All twenty-five references received a good quality B grade appraisal.

The synthesis of the evidence revealed three main predictive modeling themes:

1. Delirium Predictive Modeling Tools, which underscore advancements in delirium prediction through machine learning and deep learning techniques, and the use of electronic health records (EHR) to enhance accuracy and performance.

2. Machine Learning Models non-ICU and ICU settings, including Extreme Gradient Boosting (XBG) and Bidirectional Long Short-Term Memory (BilSTM). These models' performance was evaluated using metrics like the area under the receiver operating characteristics curve (AUROC) to determine their specificity and sensitivity for delirium predictive accuracy, clinical integration, and interpretability.

3. Other delirium prediction models, such as the PREdiction of DELIRium in ICU patients (PRE-DELIRIC), Early PREdiction model for DELIRium in ICU (E-PRE-DELERIC), Lanzhou Model, DELLiriuM, and The Mayo Delirium Prediction (MDP) tool, along with other assessment methods for early delirium detection, including CAM-ICU, Intensive Care Delirium Screening Checklist (ICDC), and Delirium Rate Scale Revised (DRS-R98) were described within the body of the science.

Additional risk predictors were identified for predicting delirium among specific populations such as patients undergoing hip and knee arthroplasty surgeries, utilizing sleep modalities, lumbar interbody fusions, and EEG data.

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

<table>
<tr><td colspan="2">There are essential nursing implications from literature, particularly utilizing nursing data or flowsheets—such as mobility, visual and hearing function, Glasgow Coma Scale (GCS), lines, tubes, and Braden Skin Scale score (BSS) to operationalize candidate predictive models. Thus, nursing is vital in accurate documentation that serves as risk predictors for predictive modeling. Nursing plays a critical role in managing delirium in ICU/non-ICU settings, and careful stratification of patients' delirium risk using highly sensitive delirium risk assessment tools like the Delirium Risk Assessment Score (DRAS), Delirium Risk Assessment Tool (DRAT), and Delirium Elderly At-Risk (DEAR) can guide surgical planning and postoperative management.</td></tr>
<tr><th>Delirium Predictive Modeling Tools</th><th>Key Summary</th></tr>
<tr><td>The Mayo Delirium Predictive (MPD) Tool Pegali et al. (2021) was developed from a large heterogenous patient population with good validation results and appear to be a reliable automated tool for delirium risk prediction with hospitalization. The area under the receiver operating characteristics curve (AUROC) for derivation cohort was 0.85 (95% confidence interval [CI], .846 to .855). Regression coefficients were obtained from the derivation cohort, predicted probability of delirium was calculated for each patient in the validation cohort. For the validation cohort, AUROC was 0.84 (95% CI, .834 to .847). Patients were classified into 1 of 3 risk groups based on their predictive probability of delirium: low (<5%), moderate (6% to 29%), and high (>30%). In derivation cohort, observed incidence of delirium was 1.7%, 12.8% and 44.8% (low, moderate, and high risk which is like the incidence rates in the validation cohort of 1.9%, 12.7%, and 46.3%. Limitation to the study is it was done at only one site, and electronic extraction is only as good as the assessments entered so retrospective data pulling. Further research is needed to use tools in other areas.

Singurska et al. (2023) developed a process using literature review and expertise “Iatrogenic Conditions Task Force” who judged whether the variables found from the literature is associated with the development of hospital-induced delirium for predictors. Nursing involvement has unique expertise from ongoing and purposeful assessments of patient data from the EHR. 40 clinical concepts were determined to be “expert” predictors of hospital-induced delirium based on the ICTF members. Results showed that the most common</td><td>Themes:
1. The Mayo Delirium Predictive (MPD) Tool Pegali et al. (2021) had good validation results and appears to be reliable. AUROC had high confidence intervals and predicted probability. Patients were classified into 1 to 3 risk groups (low, moderate, and high). Warrants further research in other areas.

2. The development of an expert multidisciplinary team “Iatrogenic Conditions Task Force” would be valuable in confirming hospital-induced delirium for predictors. Three variables that were full agreement and top results include: physical restraints, sedation status, and withdrawal. See Table 2 for variables noted by Iatrogenic Conditions Task Force.</td></tr>
</table>

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

predictors from at least half of the studies include age, sex, alcohol use. Three variables with complete agreement from ICTF include: physical restraints, sedation status, and withdrawal (Singurska et al., 2023). Other variables but were excluded due to its unstructured text found in EHR: duration of anesthesia, physical status (measured with the American Society of Anesthesiologist), severity of acute illness. Singurska et al. (2023) recommends using additional nursing flowsheets in the operationalization of candidate predictors.

Table 2 "Expert" Candidate Predictors of Hospital-induced Delirium (n = 40)

Administrative	Diagnosis	Laboratory Test	Medication	Nursing Assessment	Surgery
Age*	Alcohol use*	Prolonged bleeding (due to overanticoagulation and/or procedure)	Fentanyl	Cognitive status*	General anesthesia*
Length of hospital stay*	Alzheimer's disease*		Minerals and electrolytes*	NPO state	Neurosurgery
Length of intensive care unit stay*	Anxiety*		Opioids	Pain*	Surgery*
	Cerebral edema		Propofol*	Psychological status	Trauma surgery
	Dementia*		Psychotherapeutic agents*	Risk of falls	
	Depression*		Sedatives*	Sleep deprivation*	
	History of delirium*				
	Mental disorder*				
	Neurologic disease*				
	Post-surgical complications				
	Psychosis				
	Respiratory disease				
	Respiratory failure*				
	Respiratory infection				
	Sepsis				
	Shock				
	Sleep disorder				
	Substance use				
	Trauma*				
	Urinary tract infection*				

Note. The candidate predictors are organized by operational category based on how they were most accurately and reliably represented in our local EHR system, Epic. EHR = electronic health record; NPO = nil per os ("nothing by mouth").

* Empirical predictors of hospital-induced delirium (candidate predictors that were included in the final prognostic models of hospital-induced delirium in the model development-and-validation studies based on statistical methods (see Table S1 in the Supplementary Material for the list of studies)).

Haight & Marsh (2020) created a prediction model based on risk factors admitted to ICU or after ischemic stroke or intracranial hemorrhage with symptom onset within 72 hours, presence of delirium as defined by positive CAM-ICU at any point during admission. Multiple logistic regression was used to create predictive model with delirium as dependent variable and evaluated by using receiver operating characteristics (ROC) analysis to calculate area under the curve (AUC). The patient's probability of developing delirium with an AUC of 0.90 and to a unique cohort of patients with an AUC of 0.82. Risk factors identified in the study include age greater than 64, presence of intraventricular hemorrhage (IVH), intubation, presence of acute kidney injury (AKI), and strokes with either cognitive deficit, neglect, or aphasia remained significant and were most strongly associated with delirium. Strokes presenting with cognitive dysfunction was highly predictive of delirium with AUC of 0.76 in ROC analysis.

3. **Haight & March (2020)** use of risk factors such age, presence of intraventricular hemorrhage (IVH), intubation, presence of acute kidney injury (AKI), and stroke with either cognitive deficit, neglect, or aphasia remained significant and were most strongly associated with delirium. Patients who developed delirium during hospitalization were older, have more medical comorbidities, and take more medications.

These themes highlight the importance of utilizing risk factors as part of predictive delirium models. The use of receiver operating characteristics (ROC) analysis to calculate area under the curve (AUC) is an appropriate calculation to predict the probability of developing delirium. An AUC greater than 0.70 demonstrates good predictive value.

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

<table>
<tr><th>Machine Learning Models in non-ICU setting</th><th>Key Summary</th></tr>
<tr><td> Friedman et al. (2025) developed and validated a machine learning model that uses electronic medical records to predict the risk of delirium in <u>non-ICU patients aged 60</u> and older as quality improvement project. Data collection included demographics, medications, labs, and unstructured clinical notes. The Confusion Assessment Method (CAM) served as a diagnostic reference for identifying delirium. This data was used to train and test a machine learning model. The Elixhauser comorbidity index was calculated using ICD-10-CM secondary diagnoses. Results showed a significant improvement in delirium detection in non-ICU patients aged 60 and older. Post implementation detection rates rose from 4.42% to 17.17% and patients received lower doses of benzodiazepines and olanzapine indicating better clinical management. AUC of 0.94 with combined EMR and nurses notes. Nurses can use the model's EMR-integrated risk scores to focus assessments more effectively, leading to earlier intervention.
Holler et al. (2025) developed and externally validated a machine learning model that can accurately predict <u>postoperative delirium 50 years and older</u> and identified preoperative EHR-based predictors of POD. Sociodemographic variables used to build predictive models include: 1) Age patient reported, 2) patient-reported race (categorized as Black, White, Asian, and other or unknown for analytic purposes), 3) insurance type (Medicare or Medicaid), self-pay or other/unknown, 4) smoking status at time of surgery (current, former, never smoker), 5) BMI from the visit nearest to the index, 6) The initial American Society of Anesthesiologists (ASA) class of 5 of E were also included 7) Surgical specialty was assigned based on National Surgical Quality Improvement Program inclusion and exclusion criteria, 8) diagnosis variables were generated using ICD-9/ICD-10-CM codes (see p. 3), 9) medication using medication order data i.e., anticholinergic medications using anticholinergic cognitive burden (ACB) scale, representing score: ACB 1, 2, 3. Utilizing Extreme gradient boosting (XGB) and a </td>
<td> Themes:

Machine Learning Models for Delirium Prediction: Several studies have developed and validated machine learning models to predict delirium in various patient populations. These models often show better performance compared to traditional statistical models. For example, Racine et al. (2021) found that machine learning models performed better than traditional stepwise logistic regression. Similarly, Bishara et al. (2022) validated machine learning models that outperformed clinician-guided models.
Deep Learning and Advanced Techniques: Advanced techniques like deep learning and the combination of long short-term memory (LSTM) with machine learning have been used to improve delirium prediction. Liu et al. (2022) developed a model that combined LSTM and machine learning, achieving high accuracy.
</td></tr>
</table>

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

multilayer neural network for complex nonlinear relationship between variables, the model was evaluated between the three hospitals using AUROC evaluation which had an AUROC=.79. The model demonstrated good predictive accuracy with XGBa, XGBb, XGBc. This model's most important feature was the ASA class which supports existing literature linking higher ASA has greater risk of POD. **Nursing implications**: Patients included were those who developed POD and had at least 1 positive CAM, perioperative nursing assessments which include Braden scale score and captured functional status are highly correlated to delirium which explains why their model had a higher AUROC. **Shaw et al. (2025)** developed a model that can provide automated delirium screening based on data readily available from EHR. Retrospective data and randomly selected 20% for the test database n=4511. The remaining 80% of patients were used for model selection, model training, and hyperparameter tuning. Inclusion criteria: At least 1 positive CAM within 24 hours (5am-5am), age, gender, vital signs, lab values, medications, and prior CAM assessments i.e. 1=positive 0=negative. This would produce a single prediction each morning of whether the patient will have delirium later that day using all the information available. XGBoost Library was used to fit boosted tree models and datasets. Step by step models and code are available from the article. The area underneath the receiver operating characteristics curve (AUROC) was used to evaluate the performance of each model. To capture the effects of population prevalence on performance, the area under the precision-recall curves (AUPRC) was used. Results of the n=4583 patients, 19.9% had positive CAM screening. Boosted tree model with highest AUROC (0.92 CI= 0.913-9.22). Random forest model with AUROC (0.85, 95% CI=0.841-0.852). Although the model had high AUROC results, the models significantly varied in their ability to maintain a high positive predictive value as sensitivity was increased. In other words, those identified with delirium had an incidence of 13%, those with no current delirium was 6%,	3. **Predictive Tools and Models**: Various predictive tools and models have been developed to assess the risk of delirium in hospitalized patients. The Mayo Delirium Prediction (MDP) tool developed by Pagali et al. (2021) is one such example, showing high accuracy in predicting delirium at hospital admission. Castro et al. (2021) developed a model to predict delirium in hospitalized COVID-19 patients. 4. **Use of Electronic Health Records (EHR)**: Many studies have utilized data from electronic health records (EHR) to build predictive models. These models often include variables such as demographics, vital signs, medications, and lab values. For instance, Harsha et al. (2025) analyzed risk factors for postoperative delirium using an optimized XGBoost model. 5. **Accuracy and Performance Metrics**: The performance of these models is often evaluated using metrics like the area under the

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

and those who never had delirium 4%, thus the decrement in the AUPRC is related to the decreased incidence in these subgroups (Shaw et al., 2025). The AUROC results from this study are favorable in that it is greater than most .75 (which is considered satisfactory performance of a clinical test) and like a commonly used diagnostic test such as D-dimer for suspected of PE. Recommendations to include other features from flowsheet data: LOS, clinical notes, sleep disruption. **Wong et al. (2018)** developed and validated a machine learning model to predict hospital-acquired incident delirium in patients without baseline cognitive impairment. The Gradient Boosting Machine (GBM) model performed best with AUC of 0.855. Setting specificity at 90%, the model had a 59.7% (95% CI, 52.4%-66.7%) sensitivity, 23.1% (95% CI, 20.5%-25.9%) positive predictive value, 97.8% (95% CI, 97.4%-98.1%) negative predictive value, and a number needed to screen of 4.8. Variables identified by an expert panel of healthcare professionals relevant to delirium prediction and available in the EHR within 24 hours of admission (admitting diagnoses, medications, lab values, vital signs, demographic and nursing data i.e., mobility, visual and hearing function, Glasgow Coma Scale, lines, and tubes); microbiology, radiobiology, pathology, and procedures were not included. See Table 2 shows variable importance by Gradient Boosting Machine. **Racine et al. (2021)** identified optimal machine learning approach to predict delirium after surgery compared to traditional statistical prediction model. Results under the receiver operating characteristic curve (AUC) were higher in the large feature set conditions (range of AUC, 0.62-0.71 across algorithms) versus selected feature set conditions (AUC range, 0.53-0.57). The machine learning approach was better than chance comparable with traditional stepwise logistic regression. **Castro et al. (2021)** developed a statistical model that can be applied to HER data to predict the probability of a	receiver operating characteristics curve (AUROC). High AUROC values indicate good predictive accuracy. For example, Mulkey et al. (2022) showed high accuracy in predicting delirium using gamma band power ratios from EEG data. *These themes highlight the advancements in delirium prediction using machine learning and deep learning techniques, the importance of predictive tools, and the use of EHR data to improve accuracy and performance.*

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

hospitalized patient with COVID-19 developing delirium. The application of the predictive model in the external validation cohort of 755 patients, the c-index was 0.75 (0.71-0.79) and the lift in top quintile was 2.1. At sensitivity of 80%, the specificity was 56%, negative predictive value 92%, and positive predictive value 92% and positive value 30%. Variables that were included in the predictive model include: Charlson comorbidity index (calculating diagnostic codes to represent differential risk associated with specific co-occurring medical conditions, prior diagnostic codes were collapsed to the second level of healthcare utilization project, vital signs, body mass index, smoking status, age, race, ethnicity). **Pagali et al. (2021)** developed a delirium risk prediction tool that is applicable across different clinical patient populations and can predict the risk of delirium at admission to hospital. Table 1 demographics and risk factors abstracted from electronic medical records were used as variables. The area under the receiver operating characteristics curve (AUROC) for **Mayo Delirium Prediction (MDP) tool** using derivation cohort was 0.85 (95% confidence interval [CI], .846 to .855). For the validation cohort AUROC was .84 (95% CI, .834 to .847). MDP tool results are reassuring for health care providers to risk stratify delirium patients in hospitalized adults at admission effectively. MDP tool derivation and validation cohorts are large and include both medical and surgical patient populations. **Bishara et al. (2022)** developed and internally validated a machine learning post operative delirium risk prediction model. The AUCORC for **Neutral Net** was 0.841 [95% CI .816-.863] and **XGBoost** was 0.851 [95% CI .827-.874] and significantly better than the clinician-guided AUCROC of 0.763. Bishara et al. (2022) validated two machine learning-derived risk prediction models from preoperative data prior to start of surgery to predict incident postoperative delirium in a broad surgical patient population. Machine learning models offer better performance than traditional clinician-based regression models.	

Liu et al. (2022) developed and evaluated an accurate deep learning model for predicting new onset delirium in hospitalized adult patients. Developed model combining long short-term memory (LSTM) and machine learning to predict new onset delirium and compared its performance with machine-learning-only models. The **LightGB** model achieved the best performance (AUC 0.927 [0.924, 0.929]. The LSTM model, the final model's performance improved with AUC 0.952. The precision value of the combined model improved from 0.497 to 0.751 with a fixed recall of 0.8. Using SHAP values, identification of 20 features included: age, heart rate, Richmond Agitation-Sedation scale score, Morse Fall risk score, pulse, respiratory rate, and level of care. Leveraging LSTM to capture temporal trends and combining it with the LightGBM model can significantly improve the prediction of new onset delirium, providing an algorithmic basis for the subsequent development of clinical decision support tools for proactive delirium interventions. **Harsha et al. (2025)** main objective is to analyze the risk factors for POD in light of the characteristics through these key variables: sex, age, inpatient or outpatient status, transfer status, type of anesthesia administered, discharge destination, patient height and weight, surgical specialty, whether the surgery was elective, smoking status, presence of dyspnea, cancer diagnosis, diabetes status, levels of preoperative sodium and albumin, hematocrit levels, whether the procedure was an emergency, surgery duration, presence of renal insufficiency, use of steroids, wound classification, presence of wound infection, and the occurrence of sepsis. The optimized XGBoost model demonstrated superior performance, achieving an AUC of 0.85. The model exhibited an accuracy of 0.926, precision of 0.934, recall of 0.945 and an F1 score of 0.939 indicating a high level of predictive accuracy and balance between sensitivity and specificity.	
Machine Learning Models in ICU setting	**Key Summary**
Machine learning-based risk prediction model for delirium in early COPD patients with respiratory	**Themes:**

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

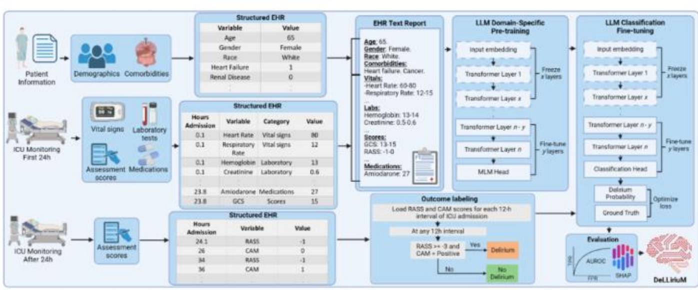
<u>failure in ICU</u> We et al. (2025) from publicly available MIMIC-IV v.2.0 database for clinical records of patients admitted to Beth Israel Deaconess Medical Center at Harvard Medical School. SQL Server 2012 installed to extract demographics i.e., age, gender, marital status, insurance, race, weight), laboratory clinical characteristics i.e. 1st day in ICU: SOFA, APS III, SIRS, GCS, Vital signs i.e., heart rate SBP, DBP, MAP, RR, SpO2, urine output. Results of the study obtained 65 predictor variables for selected model construction. Two groups were validated on: Group A: Delirium and Group B: non-delirium. Significant differences between groups in terms of LOS and ICU stay. SOFA, APS III scores, GCS verbal, LOS, mean SPO2, mean DBP, and MDRD equation were top five contributors to the model. SHAP values for hospital days indicated longer hospital stays and elevated model's prediction. A low-risk (2.48 score) patient predicted risk score was influenced by mean DBP and mean finger pulse oximetry. A high-risk (1.69 score) predicted patient risk was influenced by hospital LOS and gender. Decision curve analysis (DCA) of the test dataset demonstrated that interventions based on predictive modeling yielded favorable results. Predictive models can be easily integrated into existing ICU workflows. Limitations suggest despite advantages in machine learning, most models function as black boxes, making it difficult to interpret traditional linear methods raising concerns among clinicians. SHAP offers a means of model interpretation, thus allowing clinical staff to better understand predictive models. Wu et al. (2025) consistent with existing research with Li and colleagues who also identified age, BMI, hypertension, APACHE II score, CPOT, sedation, and PaO2 as key factors for delirium in this population. Wu et al. (2025) need further research validation. XGBoost algorithm exhibited the best performance with an AUC of 0.879 (CI:0.819-0.940) like Wu et al. (2025 findings) with an AUC of 0.921 which surpasses these benchmarks particularly for ICU patients with COPD and RF. Conflicting evidence, Song et al in Wu et al (2025) found that logistic regression outperformed complex machine learning algorithms in predicting postoperative delirium.	1. **Machine Learning Algorithms and Techniques**: Various machine learning algorithms, such as XGBoost, Bidirectional Long Short-Term Memory (BilSTM), and large language models (LLMs), are employed to predict delirium in ICU patients. These models utilize different techniques to enhance predictive accuracy and interpretability. 2. **Predictive Variables and Data Sources**: The models incorporate a wide range of predictor variables, including demographics, clinical characteristics, vital signs, and laboratory results. Data is often extracted from large databases like MIMIC-IV, eICU, and MIMIC-III, providing a comprehensive dataset for model training and validation. 3. **Model Performance and Evaluation**: The performance of these models is evaluated using metrics such as the area under the receiver

Nursing implications: Wu et al. (2025) utilized the patient's first-day GCS verbal and motor scores on the first day in ICU admission. In another study Teng et al. in Wu et al (2025) utilized the MIMIC IV data with GCS score, mechanical ventilation, and sedation were top 3 most important influencing features. ***GCS and individual subscale behavior i.e., GCS verbal score is important early recognition of delirium** **Cheng et al. (2023)** explored the relationship between Braden Skin Score (BSS) and delirium and whether BSS could be used as an assessment tool to predict the risk of delirium in 65 and older years patients in ICU setting. BSS captures assessment in cognitive function, nutritional status and mobility, providing comprehensive patient information (Sardo et al., 2015 in Cheng et al., 2023). In older patients with low BSS, these problems further affect the neurological function of older patients, placing their brains at risk for delirium (Hayhurst et al., 2016 in Cheng et al., 2023). Utilized Navicat Premium software version 16.0.11 sensitivity analyses for matching covariates. Multifactorial logistic regression model was constructed to analyze the matched population. Variables extracted: sex, age, race, weight, length of ICU stay, GCS, BBS, Sequential Organ Failure Assessment (SOFA) score, Acute Physiology Score III (APS-III), vital signs (temp, HR, RR, mean blood pressure), labs (sodium, K, WBC, RBC, platelets, hemoglobin and glucose), comorbidities (MI, CHF, dementia, chronic lung disease, DM, renal disease, liver disease, malignant CA, Hypertension and cerebrovascular disease), sedative drugs, vasoactive agents, renal replacement therapy, mechanical ventilation, delirium, ICU death, in-hospital death. Results showed the lower BSS score had a higher prevalence of delirium and increased ICU mortality. All the five subscales for BSS were significantly associated with delirium prevalence (p<.001) except mobility. There was a strong association between BSS and discharge location of older patients. Survival rate was significantly lower in the delirium group than in the no delirium group (p<.001). Three predictive tools in delirium management were discussed by (Inouye et al., 1999; Marcantionio et	operating characteristics curve (AUROC). High AUROC values indicate good predictive accuracy, and models are often optimized to maximize performance during specific time frames, such as peak delirium onset times. 4. **Clinical Application and Integration**: The models are designed to be integrated into existing ICU workflows, providing real-time predictions that can inform clinical decision-making. This integration aims to improve patient outcomes by enabling early intervention and personalized care. 5. **Interpretability and Usability**: Techniques like SHAP values are used to interpret models, making it easier for clinicians to understand and trust predictions. This addresses the challenge of machine learning models being perceived as "black boxes." 6. **Nursing Implications**: The studies highlight the importance of nursing

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

al., 2001 in Cheng et al., 2023). 1) Prediction Model for Delirium (PRE-DELRIC), 2) Early Prediction Model for Delirium (E-PRE-DELIRIC), and 3) Lanzhou model. All have some clinical applications, but practical applicability is limited, hindering real-time action by clinicians.

Contreara et al. (2020) developed and validated a DeLLiriuM, a novel large language model (LLM) delirium prediction tool which employs structured EHR data in text form. The primary outcome predicted by DelliriuM model is the risk developing delirium at any point during a patient's ICU admission after 24 hours. The presence of delirium is defined as a positive CAM score along with a Richmond Agitation Sedation Scale (RASS) score of -3 or higher at any 12-hour interval after 24 hours of ICU admission. Figure 2 show data from three ICUs used for training and validating the DeLLiriuM model. Validated this LLM on ICU admissions. The area under the receiver operating characteristics curve (AUROC) showed that DeLLiriuM outperformed all baselines in two external validation sets with 0.77 [95% CI 0.83-0.85] across 77,543 patients spanning 194 hospitals in the U.S. ICU settings. First LLM-based delirium prediction tool for the ICU based structured HER data.



Bhattacharyya et al. (2022) developed a delirium prediction model that can be used as a screening tool. Bidirectional Long Short-Term Memory (BilSTM) models where precision and recall changed from 37.52% [95% CI, 36.00%-39.05%]. Use of eICU and collaborative research databases and MIMIC-III databases to extract variables at real time. Employed a sliding window for prediction and incorporated the trajectory of each variable over time. Bhattacharyya et al. (2022) looked at peak delirium onset

assessments and interventions in predicting and managing delirium. Variables such as the Braden Skin Score (BSS) and Glasgow Coma Scale (GCS) scores are emphasized for their role in early recognition and management of delirium.

These themes reflect the advancements in using machine learning models for delirium prediction in ICU settings, emphasizing the importance of predictive accuracy, clinical integration, and interpretability.

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

<table>
<tr><td>time in our population and optimized the model to maximize predictive accuracy in that time frame.

American Association of Critical Care Nurses: News brief from AACN May 2023 shared about a dynamic model aimed to predict the onset of delirium up to 12 hours in advance using data throughout an ICU stay. This model identified 16.1% to 20.9% of cases, producing a mean AUC of 0.845 with a maximum of 0.859 at the shortest lead time. The author Stevens believes that predictive models are the basis for a paradigm shift toward preventive and personalized medicine. With validation clinicians will be able to leverage model outputs to mitigate risk for delirium through targeted interventions.

Raghu et al. (2023) developed an automated machine learning model for delirium severity measurement and predict the level of delirium severity upon ICU admission. The cumulative doses of outpatient benzodiazepines and given upon ICU admission were top features identified to predict severe delirium.</td><td></td></tr>
<tr><th>Other Delirium Predictive Models and commonly used assessment methods</th><th>Key Summary</th></tr>
<tr><td>Abdelbaky et al. (2024) literature review provided three models to predict delirium and aid in preventing and treating delirium:
1) PREdiction of DELIRium in ICU patients (PRE-DELIRIC) (10 predictors) Predict delirium in the patients within 24 hours of admission to ICU (Linkaite et al., 2018 in Abedelbaky & Eldelphany, 2024), Liang et al. reported PRE-DELRIC has high predictive value in ICU patients
2) Early PREdiction model for DELIRim in ICU (E-PRE-DELERIC) (9 predictors)
3) Lanzhou Model (11 predictors) by Chen et al. validated and results supporting the use of this model in clinical practice to identify ICU delirium
<u>Other commonly used assessment methods:</u>
1) CAM-ICU- acute onset of mental status change, cognitive disturbance, inattention, and altered levels of consciousness. Validated scale measuring severity of</td><td>Themes: from the literature review by Abdelbaky et al. (2024):
1. Delirium Prediction Models: The review highlights three models designed to predict delirium in ICU patients:
○ PRE-DELIRIC: This model uses 10 predictors to forecast delirium within 24 hours of ICU admission. It has been reported to have high predictive value.
○ E-PRE-DELERIC: This model employs 9 predictors</td></tr>
</table>

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

delirium. The scale ranges from 0 to 7. Scores from 0 to 2 are linked with no delirium, mild and mod delirium (3-5), severe delirium (6 or 7). 78 sensitivity and 95% specificity. 2) ICDC- checklist of 8 items (score 1, resulting in a total score that ranges from 0 to 8). 0 score: no delirium, 1-3 score: subsyndromal delirium, and 4 and higher score, delirium. 98% sensitivity and 55% specificity for screening delirium. 3) Delirium Rate Scale (DRS-R98)- assessment of delirium and its severity. 13-item scale includes (hallucinations, mood lability, wake, and sleep cycle disturbances, thinking abnormalities, motor retardation, agitation, attention impairments) Challenge of this scale is the need for trained professionals to administer it leading to occasional variations in results. **Nursing Implications:** Importance of early detection and management of delirium in the ICU to improve pt outcomes and reduce mortality rates.	for early prediction of delirium in ICU patients. ○ **Lanzhou Model**: Validated by Chen et al., this model uses 11 predictors and is supported for clinical use to identify ICU delirium. 2. **Assessment Methods for Delirium**: The review discusses commonly used methods to assess delirium: ○ **CAM-ICU**: Measures acute onset of mental status change, cognitive disturbance, inattention, and altered levels of consciousness. It has a validated scale ranging from 0 to 7, with sensitivity of 78% and specificity of 95%. ○ **ICDC**: A checklist of 8 items used to screen for delirium, with sensitivity of 98% and specificity of 55%. ○ **Delirium Rate Scale (DRS-R98)**: A 13-item scale assessing various aspects of delirium and its severity. It requires trained professionals for administration, which can lead to variations in results. 3. **Nursing Implications**: The review emphasizes the importance of early detection and management of delirium in the ICU to improve

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

<table>
<tr><td></td><td>patient outcomes and reduce mortality rates.

These themes reflect the advancements in delirium prediction models, the effectiveness of various assessment methods, and the critical role of nursing in managing delirium in ICU settings.</td></tr>
<tr><th>Identifying Risk Predictors for specific populations</th><th>Key Summary</th></tr>
<tr><td>Post Hip Fracture Delirium (PHFD) Kim et al. (2020) developed a risk predictor for post operative delirium in patients with a hip fracture. Nine predictors were found in the final PHFD: 1) Preoperative delirium, 2) Preoperative dementia, 3) Age, 4) Medical co-management, 5) American Society of Anesthesiologists (ASA) physical status III-V, 6) Functional dependence, 7) Smoking, 8) Systematic inflammatory response syndrome/sepsis/septic shock, 9) preoperative use of mobility aid. Limitations of PHFD did not state if a validated assessment tool was used i.e., CAM or another tool and future studies should include other severity of delirium.

Delirium risk index (DRI) Zhong et al. (2021) conducted a retrospective cohort study utilizing Premier Healthcare Database of knee arthroplasty surgeries for external validation to predict postoperative delirium. DRI derived fourteen comorbidities and assigned comorbidities weights ranging from 1 to 6. DRI is recognized as a valuable tool for assessing the risk of postoperative delirium. It helps in adjusting for comorbidities and predicting the likelihood of delirium after surgery. Further research evaluating its performance on varying definitions of delirium and diverse patient populations. Limitations of DRI was it used ICD-9 codes i.e., billing information for drugs and may not have captured types of delirium such as hypoactive delirium.

<u>Other Attributes for Predicting risk for delirium:</u>

Sleep Attributes/Metrics predict at risk for delirium:</td><td>Themes:
1. Risk Predictors for Postoperative Delirium:

○ Post Hip Fracture Delirium (PHFD): Kim et al. (2020) identified nine predictors for postoperative delirium in hip fracture patients, including preoperative delirium, preoperative dementia, age, medical co-management, ASA physical status III-V, functional dependence, smoking, systemic inflammatory response syndrome/sepsis/septic shock, and preoperative use of mobility aid.
○ Delirium Risk Index (DRI): Zhong et al. (2021) validated the DRI for predicting postoperative delirium in knee arthroplasty surgeries. The DRI uses 14 comorbidities with</td></tr>
</table>

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

Jaiswal et al. (2021) tested specific attributes from their previous clinical trial. Actiware Software to analyze sleep and sleep activity metrics to predict older patients at high risk for delirium. Study results found total sleep was similar between patients with n=17 and without delirium n=53 had 135.8 min, p=.081. Mean sleep bout times were shorter in delirious versus never delirious patients. Patients with delirium had shorter bout times (<10 min) and fewer longer sleep bouts (>30 min) compared to those without delirium. Results of the study showed that delirium was associated with increased sleep fragmentation detected by actigraphy may indicate a useful biomarker for delirium prediction in future. **Predictive delirium risk assessment tools (postoperative patients with posterior lumbar interbody fusions):** **Singh et al. (2025)** evaluated the effectiveness of three delirium risk assessment tools: 1) Delirium Risk Assessment Score (DRAS), 2) Delirium Risk Assessment Tool (DRAT), and 3) Delirium Elderly At-Risk (DEAR) in identifying patients at risk of postoperative delirium (POD) following posterior lumbar interbody fusions. ROC curve analyses revealed that DRAS score of 5 (Sensitivity=62.9%, Specificity=63.9%), DRAT score of 3 (Sensitivity=31.4%, Specificity=81.0%), and DEAR score of 2 (Sensitivity=40.0%, Specificity=82.9%). Patients above these thresholds were 6.0, 2.0, and 3.2 times more likely to develop POD after posterior lumbar fusion. Using these simple risk assessment tools may be the most effective in guiding the implementation of target interventions. Nursing plays an important role with careful stratification of patients' risk of delirium using highly sensitive and specific tools like DRAS may guide surgical planning and postoperative management plans. **Use of EEG data to predict delirium:** **Van Sleuwen et al. (2022)** compared the delirium severity as a tool that can help close the diagnostic gap: the **electroencephalographic confusion assessment method**	assigned weights to assess delirium risk. 2. **Sleep Attributes and Delirium Prediction**: o **Sleep Metrics**: Jaiswal et al. (2021) analyzed sleep activity metrics to predict delirium in older patients. The study found that delirium was associated with increased sleep fragmentation, suggesting that sleep metrics could be useful biomarkers for delirium prediction. 3. **Delirium Risk Assessment Tools**: o **Assessment Tools for Posterior Lumbar Interbody Fusions**: Singh et al. (2025) evaluated the effectiveness of three delirium risk assessment tools (DRAS, DRAT, and DEAR) for predicting postoperative delirium in patients undergoing posterior lumbar interbody fusions. These tools were found to be effective in guiding targeted interventions. 4. **Delirium Severity Assessment**:

severity score (E-CAM-S and CAM-S were similar in their strength of association with hospital LOS (correlation =0.31 vs 0.41 p=0.082 and in hospital mortality (area under the curve= 0.77 vs 0.81; p=0.310. Patients with delirium were older, had lower RASS scores, and higher CAM-S and CCI. **Mulkey et al. (2022)** examined the use of machine learning methods evaluating the use of gamma band to predict delirium from EEG data derived from a limited lead rapid response handheld device. Data was obtained from a sample n=13 aged 50 or older requiring mechanical ventilation for more than 12 hours. Results showed that frequency band power ratios were higher in participants with delirium and remained consistent for all power ratios that included the gamma band. The accuracy of the step wise logistic discriminant analysis were between 87 to 97% when analyzing the gamma/delta and gamma/theta power ratios. This shows a strong correlation given the small size with a 92% accuracy. When using support vector machines, the gamma/delta power ratio was able to predict delirium with 82.1% accuracy. The use of a non-invasive 10-lead handheld rrEEG is feasible in clinical application of predicting delirium.	○ **E-CAM-S and CAM-S**: Van Sleuwen et al. (2022) compared the electroencephalographic confusion assessment method severity score (E-CAM-S) with the CAM-S. Both tools were similar in their association with hospital length of stay and in-hospital mortality, highlighting their utility in assessing delirium severity. ○ EEG data showed that frequency band power ratios were higher in participants with delirium (Mulkey et al., 2022). Use of a non-invasive test is feasible in clinical application of predicting delirium. *These themes emphasize the importance of identifying risk predictors, utilizing sleep metrics, employing effective risk assessment tools, and assessing delirium severity to improve patient outcomes in hip fracture and knee arthroplasty surgeries.*

Evidence Search Strategies: The evidence review was conducted in May 2025 to find the latest evidence on early predictive models for delirium. The aim of the search was to examine the best available evidence for early predictive models and/or early identification of predictors for delirium patients in the adult population (18 years and above) within the acute care setting. Key search terms were broad and included: Early predictors delirium, adult patients' delirium, predictive learning models delirium, screening tools delirium, delirium imposed on dementia and Parkinson's, geriatric delirium, elderly/older delirium, adults' delirium, alcohol withdrawal delirium, emergence delirium. Filters included: last 5 years, English, Humans, Full text, United States, and inpatients. The electronic databases included PubMed, Clinical Key, CINAHL, Cochrane Libraries, and Google Scholar. Search was individualized for each database and reviewed search methodology by the clinical library. Additionally, references from Professional Organization websites were also included in the search. After evaluation for inclusion and exclusion criteria, and relevance to the question, there were 23 final articles that answered the clinical PICOT question.

Electronic Database Search Methodology

Dates: Within the last 5 years May 2020-May 2025, United States (U.S.), English language, Full-text, peer-reviewed only articles

Literature search topic/clinical question: In adult patient's 18 years and above, what is the quality and consistency of the evidence with the use of predictive models/analytics compared to current practice in the accuracy of predicting early/identification of delirium within the hospital stay?

Inclusion Criteria: adults > 18 years, predictive models for early delirium, analytic tools for early delirium, hospital stay, all in-patient departments including ED

Exclusion Criteria: <18 years, not predictive analytic tools for delirium, outpatient setting i.e., clinic, home health patients, not maternal child health departments, assessment of delirium, risk factors for delirium, systematic, metanalysis, scoping reviews if outside U.S.

Key search terms were broad and included: Early predictors delirium, adult patients' delirium, predictive learning models delirium, screening tools delirium, delirium imposed on dementia and Parkinson's, geriatric delirium, elderly/older delirium, adults delirium, alcohol withdrawal delirium, emergence delirium.

Database	Key Word(s) and/or Controlled Vocabulary Terms #	Total References Identified (hits)	No. of Relevant References	No. of Total Duplicate Articles	No. of Articles Selected for Review	No. of Articles Excluded	Final Total Relevant References
Name: PubMed **Years:** 2020-2025	Delirium & Predictive Tools& tools	N= 120	17	0 *1 duplicate (Pagali et al. 2022)	10 (1st round) **4 (2nd round** 6 were excluded)	103	4

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

Name: PubMed **Years:** 2020-2025	("Delirium"[MeSH Terms] OR "Delirium Prediction") AND ("Predictive Analytics" OR "Predictive Models" OR "Machine Learning" OR "Artificial Intelligence" OR "Risk Assessment" OR "Clinical Decision Support Systems") AND ("Hospitalized Patients" OR "Inpatients" OR "Hospital Admission" OR "Hospital Stay")	N=42	10	1 *(Castro et al., 2021)	10 (1st round) **7 (2nd round)** *3 were excluded on 2nd round of reading full text articles	35	6
Name: CINAHL **Years:** 2020-2025	Delirium AND Predictive Tools	N=8	0	0	0	8	0
Name: Cochrane Library **Years:** 2020-2025	"delirium predictive tools" AND "adult" "delirium predict" AND "adult" Publish data 2020-2025	N=129	27 (8 if all non-US removed)	5	6 (1st round) **3 (2nd round)** *3 were removed after full text)	23	3
Name: Clinical Key **Years:** 2020-2025	Delirium predictive model	N=292	20	0	20 (1st round) **2 (2nd round)** *18 excluded systematic, metanalysis, did not meet setting, RCT	289	2

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

					international settings		
Name: Google Scholar **Years:** 2020-2025	Delirium prediction model	17,500	30	5	10 (1st round) **6 (2nd round)** *2 were excluded during the 2nd round of reading full text articles	24	6
Name: AACN **Years:** 2020-2025	"delirium predictive" "delirium" Published data 2020-2025	97 84	7	3	2	95	*2 (included in reference/contextual links)
Total							**21**

#Controlled vocabulary (subject terms, MESH terms, tagged terms specific to database)
*Use the first database as the main comparison for subsequent database searches and identifying duplicate articles

***Reference/Contextual Links**	***Reference/Contextual Links**
Citation: NICHE: 0	Citation: Society of Critical Care Medicine: 3
Citation: American Association Critical Nurse: 1	Citation: American Delirium Society: 0
Citation:	Citation:

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

Senior Surgical Care Program (SSCP):0	Mayo Clinic: 1 (Pagali et al. 2022- duplicate)

Total Articles Included in Literature Review: Database (20) + Contextual Links (3)= 25

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

Respectfully submitted,

Quincyann Tsai, MSN, RN
Nursing Research and EBP Consultant Regional Nursing Research & EBP Program
Kaiser Permanente Southern California & Hawai'i Patient Care Services

Veronica Timple, PhD, RN, CCRN-K
Nurse Scientist, Regional Nursing Research & EBP Program
Kaiser Permanente Southern California & Hawai'i Patient Care Services

Emma Aquino-Maneja DNP, MEd, RN, CCRN
Clinical Consultant V, Regional Professional Practice
Kaiser Permanente Southern California & Hawai'i Patient Care Services

Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC
Clinical Nurse Specialist
Kaiser Permanente-Orange County

Belen Thornfield MLIS, AHIP
Librarian Specialist II
Health Science/Medical Library

Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC
Director Clinical Practice & Operations
Kaiser Permanente Southern California & Hawai'i Patient Care Services

Hannah Jang Kim, PhD, RN, CNL, PHN
National Senior Operations Leader, Analyst
Nurse Research Scientist
Kaiser Permanente National Patient Care Services

Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS
Critical Care Clinical Nurse Specialist
Kaiser Permanente Los Angeles

Tennessee Meris, MSN, GCNS/GNP-C
Geriatric Clinical Nurse Specialist
Kaiser Permanente Woodland Hills

Lina Najib Kawar, PhD, R.N.R.M., RN, CNS, FAAN
Director Regional Nursing Research and EBP Program
Nurse Scientist
Kaiser Permanente Southern California & Hawai'i Patient Care Services

References

Abdelbaky, A.M., Eldelpshany, M.S. (2024). Patient outcomes and management strategies for Intensive Care Unit (ICU)- Associated Delirium: A literature review. Cureus, 6(6): e61527. DOI 10.7759/cureus.61527

American Association of Critical Care Nurses. (2023). *Clinical voices May 2023*. https://m.aacn.org/newsroom/clinical-voices-may-2023

Bhattacharyya, A., Sheikhalishahi, S., Torbic, H., Yeung, W., Wang, T., Birst, J., Duggal, A., Celi, L. A., & Osmani, V. (2022). Delirium prediction in the ICU: designing a screening tool for preventive interventions. *JAMIA open, 5*(2), ooac048. https://doi.org/10.1093/jamiaopen/ooac048

Bishara, A., Chiu, C., Whitlock, E. L., Douglas, V. C., Lee, S., Butte, A. J., Leung, J. M., & Donovan, A. L. (2022). Postoperative delirium prediction using machine learning models and preoperative electronic health record data. *BMC anesthesiology*, *22*(1), 8. https://doi.org/10.1186/s12871-021-01543-y

Castro, V. M., Sacks, C. A., Perlis, R. H., & McCoy, T. H. (2021). Development and External Validation of a Delirium Prediction Model for Hospitalized Patients With Coronavirus Disease 2019. *Journal of the Academy of Consultation-Liaison Psychiatry*, *62*(3), 298–308. https://doi.org/10.1016/j.jaclp.2020.12.005

Cheng et al. (2023). Can admission Braden skin score predict delirium in older adults in the intensive care unit? Results from a multicenter study. Journal of Clinical Nursing, 33. DOI: 10.1111/jocn.16962

Contreara, M., Kapoor, S., Zhang, J., Davidson, A., Ren, y., Guan, Y., Baslanli, T., Nerella, S, Bihorac, A., & Rashidi, P. 2020. deLLIirium: A large language model for delirium prediction in the ISU using structured EHR. Cornell University: Retrieved at 2410.17363v1

Haight, T.N. & Marsh, E.B. (2020). Identifying delirium early after Stroke: A new prediction tool for the intensive care unit. Journal of Stroke and Cerebrovascular Diseases, 29 (11)105219.

Harsha Boppana, Sri1; Tyagi, Divyansh2; Komati, Sachin3; Lasya Boppana, Sri4; Mintz, David5. 712: AI-Delirium: Predictive modeling of postoperative delirium in elderly surgical patients. Critical Care Medicine 53(1):, January 2025. | DOI: 10.1097/01.ccm.0001101512.19120.96

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

Holler et al. (2025). Development and Validation of a Routine Electronic Health Record-Based Delirium Prediction Model for Surgical Patients Without Dementia: Retrospective Case-Control Study. JMIR Perioperative Medicine,8, e59422. http://dx.doi.org/10.2196/59422

Friedman, JosephI., Parchure, P., Cheng, F.-Y., Fu, W., Cheertirala, S., Timsina, P., Raut, G., Reina, K., Joseph-Jimerson, J., Mazumdar, M., Freeman, R., Reich, DavidL., & Kia, A. (2025). Machine Learning Multimodal Model for Delirium Risk Stratification. *JAMA Network Open*, *8*(5), e258874. https://doi.org/10.1001/jamanetworkopen.2025.8874

Jaiswal, S. J., Kang, D. Y., Wineinger, N. E., & Owens, R. L. (2021). Objectively measured sleep fragmentation is associated with incident delirium in older hospitalized patients: Analysis of data collected from a randomized controlled trial. Journal of sleep research, 30(3), e13205. https://doi.org/10.1111/jsr.13205

Kim, E. M., Li, G., & Kim, M. (2020). Development of a Risk Score to Predict Postoperative Delirium in Patients with Hip Fracture. Anesthesia and analgesia, 130(1), 79–86. https://doi.org/10.1213/ANE.0000000000004386

Liu, S., Schlesinger, J. J., McCoy, A. B., Reese, T. J., Steitz, B., Russo, E., Koh, B., & Wright, A. (2022). New onset delirium prediction using machine learning and long short-term memory (LSTM) in electronic health record. *Journal of the American Medical Informatics Association : JAMIA, 30*(1), 120–131. https://doi.org/10.1093/jamia/ocac210

Mulkey, M., Albanese, T., Kim, S., Huang, H., & Yang, B. (2022). Delirium detection using GAMMA wave and machine learning: A pilot study. *Research in nursing & health, 45*(6), 652–663. https://doi.org/10.1002/nur.22268

Pagali, S. R., Fischer, K. M., Kashiwagi, D. T., Schroeder, D. R., Philbrick, K. L., Lapid, M. I., Pignolo, R. J., & Burton, M. C. (2022). Validation and Recalibration of Modified Mayo Delirium Prediction Tool in a Hospitalized Cohort. *Journal of the Academy of Consultation-Liaison Psychiatry, 63*(6), 521–528. https://doi.org/10.1016/j.jaclp.2022.05.006

Racine, A. M., Tommet, D., D'Aquila, M. L., Fong, T. G., Gou, Y., Tabloski, P. A., Metzger, E. D., Hshieh, T. T., Schmitt, E. M., Vasunilashorn, S. M., Kunze, L., Vlassakov, K., Abdeen, A., Lange, J., Earp, B., Dickerson, B. C., Marcantonio, E. R., Steingrimsson, J., Travison, T. G., Inouye, S. K., … , the RISE Study Group (2021). Machine Learning to Develop and Internally Validate a Predictive Model for Post-operative Delirium in a Prospective, Observational Clinical Cohort Study of Older Surgical Patients. *Journal of general internal medicine*, *36*(2), 265–273. https://doi.org/10.1007/s11606-020-06238-7

Raghu, R., Mohiuddin, A., Huang, Y., Herasevich, V., & Lindroth, H. (2023). Predicting severity of delirium on ICU admission: Development of an automated machine learning model. Research Snapshot Theater: Future of Critical Care; Critical Care Medicine**.**

Shaw, K.M. et al. (2025). Daily automated prediction of delirium risk in hospitalized patients: Model development and validation. JMIR, 13, e60442. http://dx.doi.org/10.2196/60442

Singh, M., Sahhar, M., Nassar, J. E., Farias, M. J., Rasquinha, R., Kim, J., Diebo, B. G., & Daniels, A. H. (2025). Analysis of Delirium Risk Assessment Tools for Prediction of Postoperative Delirium Following Lumbar Spinal Fusion. *Spine*. Advance online publication. https://doi.org/10.1097/BRS.0000000000005271

Singurska et al. (2023). Application of a practice-based approach in variable selection for a prediction model development study of hospital-induced delirium. BMC Medical Informatics and Decision Making, 23https://doi.org/10.1186/s12911-023-02278-1

Van Sleuwen, M., Sun, H., Eckhardt, C., Neelagiri, A., Tesh, R. A., Westmeijer, M., Paixao, L., Rajan, S., Velpula Krishnamurthy, P., Sikka, P., Leone, M. J., Panneerselvam, E., Quadri, S. A., Akeju, O., Kimchi, E. Y., & Westover, M. B. (2022). Physiological Assessment of Delirium Severity: The Electroencephalographic Confusion Assessment Method Severity Score (E-CAM-S). Critical care medicine, 50(1), e11–e19. https://doi.org/10.1097/CCM.0000000000005224

Wong, A., Young, A. T., Liang, A. S., Gonzales, R., Douglas, V. C., & Hadley, D. (2018). Development and validation of an electronic health record–based machine learning model to estimate delirium risk in newly hospitalized patients without known cognitive impairment. *JAMA network open*, *1*(4), e181018-e181018.

Wu, Z.B, Jian, Y.L., Li, S.S., Ki, A. (2025). Enhanced machine learning predictive modeling for delirium in elderly ICU patients with COPD and respiratory failure: A retrospective study based on MIMIC-IV. PLOS ONE. https://doi.org/10.1371/journal.pone.0319297

Zhong, X., Lin, J. Y., Li, L., Barrett, A. M., Poeran, J., & Mazumdar, M. (2021). Derivation and validation of a novel comorbidity-based delirium risk index to predict postoperative delirium using national administrative healthcare database. Health services research, 56(1), 154–165. https://doi.org/10.1111/1475-6773.13565

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

LEGAL NOTICE

This summary document (referred to generally as an "Evidence Review") was created and is presented by Kaiser Foundation Hospitals Nursing Research, on behalf of the Kaiser Permanente Medical Care Program. ***The following notices and provisions apply to all use of this Evidence Review for any purpose.***

Purpose/intended Audience

Because we want everyone in our communities to have the healthiest lives possible, we are making our Evidence Reviews available to the communities we serve to help Californians and others lead healthier lives. Evidence Reviews (also called "integrative reviews" and "evidence summaries") may include any and all of the following methodologies: integrative, scoping, systematic, rapid and literature reviews.

Evidence Reviews are provided as a community service for reference purposes only and are presented for use solely as specified in this disclaimer. The information presented is intended and designed for review by trained clinicians with experience in assessing and managing healthcare conditions. The information contained in the evidence reviews is not intended to constitute the practice of medicine or nursing, including telemedicine or advice nursing.

Limitations On Use

These documents have been developed to assist clinicians by providing an analytical framework for the effective evaluation and treatment of selected common problems encountered in patients. These documents are not intended to establish a protocol for all patients with a particular condition. While Evidence Reviews provide one approach to evaluating a problem, clinical conditions may vary significantly from individual to individual. Therefore, clinicians must exercise independent professional judgment and make decisions based upon the situation presented.

Kaiser Permanente's documents were created using an evidence-based process; however, the strength of the evidence supporting these documents differs. Because there may be differing yet reasonable interpretations of the same evidence, it is likely that more than one viewpoint on any given healthcare condition exists. Many reviews will include a range of recommendations consistent with the existing state of the evidence.

All of the Evidence Reviews were developed from published research and non-research evidence and do not necessarily represent the views of all clinicians who practice on behalf of the Kaiser Permanente Medical Care Program. These Evidence Reviews also may include recommendations that could differ from certain federal or state health care regulations or recommendation.

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C

No Endorsement or Promotional Use

Any reference in these documents to a specific commercial product, process, or service by trade name, trademark, or manufacturer, does not constitute or imply an endorsement or recommendation by Kaiser Foundation Hospitals or any other entity of the Kaiser Permanente Medical Care Program. The views and opinions expressed in these documents may not be used for any advertising, promotional, or product endorsement purposes.

Disclaimer of All Warranties and Liabilities

Finally, specific recommendations presented in Evidence Reviews derive from combining the best available evidence. Although Kaiser Foundation Hospitals has sought to ensure that its Evidence Reviews accurately and fully reflect its view of the appropriate combination of evidence at the time of initial publication, Kaiser Foundation Hospitals cannot anticipate changes and take no responsibility or assume any legal liability for the continued currency of the information or for the manner in which any person who references them may apply them to any particular patient. Neither Kaiser Foundation Hospitals nor any entity of the Kaiser Permanente Medical Care Program, assumes any legal liability or responsibility for the completeness, clinical efficacy or value of any apparatus, product, or process described or referenced in the documents. The entities of the Kaiser Permanente Medical Care Program make no warranties regarding errors or omissions and assume no responsibility or liability for loss or damage resulting from the use of these documents.

Quincyann Tsai, MSN, RN, Lina Najib Kawar, PhD, RN, CNS, Veronica Timple, PhD, RN, CCRN-K, Emma Aquino-Maneja PhD, DNP, MEd, RN, CCRN, Joria Rainbolt-Clemente, DNP, MSN, RN, Gero-BC, Belen Thornfield MLIS, AHIP, Becky Campbell MPH, MSHCM, BSN, CCRN, NE-BC, Hannah Jang Kim, PhD, RN, CNL, PHN, Olena Svetlov, DNP, RN, AGNP, CNS, ATCN, CCRN, PHN, NE, FCNS, Tennessee Meris, MSN, GCNS/GNP-C