2020

Diagnostic Performance of Delirium Assessment Tools in Critically Ill Patients: A Systematic Review and Meta-Analysis

Mu-Hsing Ho  
*University of Wollongong*, mhh838@uowmail.edu.au

Amy E. Montgomery  
*University of Wollongong*, asaunder@uow.edu.au

Victoria Traynor  
*University of Wollongong*, vtraynor@uow.edu.au

Chia-Chi Chang  
*University of Wollongong*

Ken N. Kuo  
*Taipei Medical University*

See next page for additional authors

Follow this and additional works at: [https://ro.uow.edu.au/smhpapers1](https://ro.uow.edu.au/smhpapers1)

**Publication Details Citation**


Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Diagnostic Performance of Delirium Assessment Tools in Critically Ill Patients: A Systematic Review and Meta-Analysis

Abstract

Background Critical care nurses are in the best position to detect and monitor delirium in critically ill patients. Therefore, an optimum delirium assessment tool with strong evidence should be identified with critical care nurses to perform in the daily assessment. Aim To evaluate and compare the diagnostic performance of delirium assessment tools in diagnosing delirium in critically ill patients. Methods We searched five electronic databases including the Cochrane Library, PubMed, Embase, CINAHL, and a Chinese database for eligible diagnostic studies published in English or Mandarin up to December 2018. This diagnostic test accuracy meta-analysis was limited to studies in intensive care unit (ICU) settings, using the Diagnostic and Statistical Manual of Mental Disorders (DSM) as a standard reference to test the accuracy of delirium assessment tools. Eligible studies were critically appraised by two investigators independently. The summary of evidence was conducted for pooling and comparing diagnostic accuracy by a bivariate random effects meta-analysis model. The pooled sensitivities and specificities, summary receiver operating characteristic curve (sROC), the area under the curve (AUC), and diagnostic odds ratio (DOR) were calculated and plotted. The possibility of publication bias was assessed by Deeks’ funnel plot. Data Synthesis We identified and evaluated 23 and 8 articles focused on CAM-ICU and ICDSC, respectively. The summary sensitivities of 0.85 and 0.87, and summary specificities of 0.95 and 0.91 were found for CAM-ICU and ICDSC, respectively. The AUC of the CAM-ICU was 0.96 (95% CI, 0.94–0.98), with DOR at 99 (95% CI, 55–177). The AUC of the ICDSC was 0.95 (95% CI, 0.92–0.96), and the DOR was 65 (95% CI, 27–153). Linking Evidence to Action CAM-ICU demonstrated higher diagnostic test accuracy and is recommended as the optimal delirium assessment tool. However, the results should be interpreted with caution due to the between-study heterogeneity of this diagnostic test accuracy meta-analysis.

Keywords

critically, tools, delirium, performance, diagnostic, assessment, review, meta-analysis, systematic, patients; ill

Publication Details


Authors

Mu-Hsing Ho, Amy E. Montgomery, Victoria Traynor, Chia-Chi Chang, Ken N. Kuo, H.C. Chang, and Kee-Hsin Chen

This journal article is available at Research Online: https://ro.uow.edu.au/smhpapers1/1554
Delirium assessment tool

Diagnostic performance of delirium assessment tools in critically ill patients: A systematic review and meta-analysis (Manuscript) please access the article from


First published: 12 August 2020

How to cite


ABSTRACT

Background: Critical care nurses are in the best position to detect and monitor delirium in critically ill patients. Therefore, an optimum delirium assessment tool with strong evidence should be identified with critical care nurses to perform in the daily assessment.

Aim: To evaluate and compare the diagnostic performance of delirium assessment tools in diagnosing delirium in critically ill patients.

Methods: We searched five electronic databases including the Cochrane Library, PubMed, Embase, CINAHL and a Chinese database for eligible diagnostic studies published in English or Mandarin up to December 2018. This diagnostic test accuracy meta-analysis was limited to studies in intensive care unit (ICU) settings, using diagnostic and statistical manual of mental disorders (DSM) as a standard reference to test the accuracy of delirium assessment tools. Eligible studies were critically appraised by two investigators independently. The summary of evidence was conducted for pooling and comparing diagnostic accuracy by a bivariate random effects meta-analysis model. The pooled sensitivities and specificities, summary receiver operating characteristic curve (sROC), the area under the curve (AUC), and diagnostic odds ratio (DOR) were calculated and plotted. The possibility of publication bias was assessed by
Deeks’ funnel plot.

**Data Synthesis:** We identified and evaluated 23 and 8 articles focused on CAM-ICU and ICDSC, respectively. The summary sensitivities of 0.85, 0.87, and summary specificities of 0.95, 0.91 for CAM-ICU, ICDSC respectively. The AUC of the CAM-ICU was 0.96 (95% CI, 0.94-0.98), with DOR at 99 (95% CI, 55-177). The AUC of the ICDSC was 0.95 (95% CI, 0.92-0.96), and the DOR was 65 (95% CI, 27-153).

**Linking Evidence to Action:** CAM-ICU demonstrated higher diagnostic test accuracy and is recommended as the optimal delirium assessment tool. However, the results should be interpreted with caution due to the between-study heterogeneity of this diagnostic test accuracy meta-analysis.

**Key Words:** Advanced practice/Advanced nursing practice, Critical care/Intensive care, Delirium, Evidence-based practice, Meta-analysis/Data pooling, Neurology, Nursing Practice
INTRODUCTION

Delirium is a common neuropsychiatric complication in intensive care unit (ICU) with an incidence rate of 26% and the prevalence of delirium in mechanically ventilated patients is as high as 80% (Sanchez-Hurtado et al., 2018; van den Boogaard et al., 2012; Setters & Solberg, 2017). Delirium defined as resulting disturbances in attention, awareness, orientation, cognition (Neufeld & Thomas, 2013) and often leads to adverse outcomes such as prolonged ICU admission, persisting cognitive dysfunction and increased mortality rate (Marcantonio, 2017; McCoy, 2018; Van Rompaey, Sabbe, Dilles, & van den Boogaard, 2018). Delirium is underdiagnosed in ICUs, particularly in cases of mechanically ventilated patients. Critical care nurses are often in the frontline care of assessing and detecting delirium. Therefore, it is crucial for critical care nurses to use diagnostic tools with efficient sensitivity and specificity to detect delirium in critically ill patients (Pun & Devlin, 2013; Ritter et al., 2018). The Diagnostic and Statistical Manual of Mental Disorders (DSM), fifth edition is the gold standard criteria for diagnosing delirium. Several assessment tools such as the Confusion Assessment Method for the ICU (CAM-ICU) and the Intensive Care Delirium Screening Checklist (ICDSC) are based on the DSM criteria to assist clinicians detect delirium in ICUs (Bergeron, Dubois, Dumont, Dial, & Skrobik, 2001; Ely, Margolin, et al., 2001).

To the best of our knowledge, a previous meta-analysis reported low pooled sensitivity on CAM-ICU and low pooled specificity on ICDSC (Neto et al., 2012). The possible reason might be that previous studies were carried out with a small population and sample size. This study aimed to evaluate and compare the diagnostic performance of the CAM-ICU and ICDSC for delirium detection in critically ill patients. It is the purpose of the study to more strongly define inform evidence-based practice for critical
METHODS

The protocol of our study was registered in International Prospective Register of Systematic Reviews (PROSPERO).

Search Strategy

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines. The review question addressed in PICO was which delirium assessment tool (I, intervention) performs the highest diagnostic accuracy in detecting delirium (C, compare to standard reference in diagnostic studies; O, outcome) among critically ill patients (P, Population)? Critically ill patients were defined as patients who were admitted to ICU. Five core electronic databases were searched (Cochrane Library, PubMed, Embase, CINAHL, Chinese Electronic Periodical Services) using the following key words: ("Critical Care"[Mesh] OR "Critical Care Nursing"[Mesh] OR "Intensive Care Units"[Mesh]) AND "Delirium"[Mesh] AND ("Nursing Assessment"[Mesh] OR "assessment tool" [Text Word] OR screening [Text Word] OR instrument [Text Word] OR scale [Text Word] OR diagnosis [Mesh]). English and Mandarin language filters were applied in searching databases. Additional papers were obtained through citation chasing and scrutinizing the reference lists.

Study Selection and Inclusion Criteria

The inclusion criteria were papers written in English or Mandarin, published from inception to October 2018, published in a peer-reviewed journal, focused on the use of a delirium assessment tool in ICU, described appropriate reference criteria (DSM) by an expert in delirium and patients included were 18 years and older.

Articles that evaluated the outcomes of sensitivity, specificity, receiver operating
characteristics (ROC) curve, positive and negative likelihood ratio of the results of delirium assessment tools were included. Articles that adopted prospective, retrospective, observational (case-control, cross-sectional, cohort and longitudinal) research designs which met the inclusion criteria were considered eligible for inclusion.

The titles and abstracts of potentially eligible studies were independently screened by two reviewers. A third author served as an arbitrator for adjudication if consensus was not reached. Studies that were not published in full-text papers (i.e., abstract in conference proceedings) were excluded.

**Data Extraction**

A specific data collection sheet for data extraction was constructed. Two authors extracted data independently and were blinded to each other’s data. The following data were extracted from included studies: the setting of the studies, sample sizes, participants, and outcome. In case of missing information from the published paper, we will contact the original study authors.

**Quality Assessment**

The methodological quality assessment of the included studies was independently evaluated by two review authors using the Joanna Briggs Institute (JBI) critical appraisal checklist for diagnostic test accuracy (Campbell et al., 2015). This checklist addresses ten questions including patient selection and index test. Each question is defined as Yes, No, Unclear or Not Applicable. Differences between the reviewers were referred to a third reviewer. A detailed quality assessment is provided in Table S1.

**Statistical Analysis**

The extracted data from included studies were summarized in two-by-two tables and entered in a bivariate random effects model, which estimates pairs of logit-transformed sensitivity and specificity from studies and considers the correlation.
between the sensitivity and specificity observed among studies. The pooled sensitivity, specificity, likelihood ratios (LRs), and diagnostic odds ratios (DORs) were also estimated to compare the diagnostic performance in different delirium assessment tools. The 95% confidence intervals (CIs) were reported long with all statistical tests. As part of the assessment for heterogeneity between studies, the relationship between sensitivity and specificity was explored using a graphical approach (a plot of sensitivity and specificity in a receiver operating characteristics curve). We plotted summary ROC curve which summarized all possible combinations of sensitivity and specificity on the curve and estimated the areas under the ROC curves (AUCs) to examine the accuracy with a graphic approach. Deeks' funnel plot asymmetry test was used to examine publication bias by regressing log DOR, with \( p \) value less than .10 for the slope coefficient which indicating significant asymmetry. We also conducted a sensitivity analysis to identify the potential influential study by using the Cook distance and generated a scatter plot for detecting outliers by using standardized predicted random effects (standardized level 2 residuals). We excluded highly influential and outlier studies detected by sensitivity analysis from the model and re-calculated the summary of sensitivity and specificity in order to examine the robustness of the results. All Data analyses were displayed by using Stata Version 15 (midas commands) and Review Manager 5.3.

**RESULTS**

The flow diagram in Figure 1 shows the systematic literature search process. The initial search identified 2,728 published articles (484 articles from PubMed, 1 article from Cochrane Library, 2,000 articles from Embase, 228 articles from CINAHL and 15 articles from the Chinese Electronic Periodical Services database). After removing duplicate articles and excluding on the basis of the title and abstract screening, there
were 42 full-text articles reviewed. Thirteen articles were excluded after full-text assessment, the reasons of exclusion were presented in Table S2. The remaining 29 articles met the inclusion criteria for systematic review and meta-analysis.

**Study Characteristics**

Table S3 presents the characteristics of the 29 studies. Twenty-three and 8 studies reported accuracy estimates for CAM-ICU (Adamis et al., 2012; Aljuaid et al., 2018; Barman et al., 2018; Boettger et al., 2018; Chanques et al., 2018; Chuang et al., 2007; Ely, Inouye, et al., 2001; Ely, Margolin, et al., 2001; Guenther et al., 2010; Gusmao-Flores et al., 2011; Heo et al., 2011; Karlicic et al., 2016; Koga et al., 2015; Lin et al., 2004; Luetz et al., 2010; Mitasova et al., 2012; Nishimura et al., 2016; Pipanmekaporn et al., 2014; Selim et al., 2018; van Eijk et al., 2011; van Eijk et al., 2009; Vreeswijk et al., 2009; Wang et al., 2013) and ICDSC (Barman et al., 2018; Bergeron et al., 2001; Boettger et al., 2018; Chanques et al., 2018; Gusmao-Flores et al., 2011; Kose, Bolu, Ozdemir, Acikel, & Hatipolu, 2016; Nishimura et al., 2016; van Eijk et al., 2009), respectively; two studies applied delirium detection score (DDS) (Luetz et al., 2010; Otter et al., 2005), one applied cognitive test for delirium (CTD) alone (Hart et al., 1996), one applied Stanford proxy test for delirium (S-PTD) alone (Aljuaim et al., 2018), one applied Neelon and Champagne confusion scale (NEECHAM) alone (Immers, Schuurmans, & Van De Bijl, 2005) and one applied nursing delirium screening scale (Nu-DESC) for delirium assessment in critically ill patients (Luetz et al., 2010). Due to a small number of studies examining the diagnostic accuracy of DDS, CTD, S-PTD, NEECHAM and Nu-DESC, diagnostic accuracy estimates were extracted but were not plotted in the summary ROC curve. Among the 29 analyzed studies, 14 studies employed the psychometric methodological approach. Eight studies recruited participants from surgical ICUs only, seven recruited participants from
Delirium assessment tool

medical ICUs only, and the remaining studies (n= 14) recruited participants from both medical and surgical ICUs. The studies were published from 1996 to 2018. The sample size ranged from 22 to 1073 with a mean age ranged from 47.9 to 74.0 years.

Publication Bias

Figure S1 shows Deeks’ funnel plots with superimposed regression lines for each included assessment tool. The statistically non-significant $p$ values (.61, and .43 for CAM-ICU, and ICDSC, respectively) for the slope coefficient suggested symmetry in data and no significant publication bias.

Pooled Results

Table 1 summarizes the pooled sensitivity, specificity, LR+, LR−, and DORs for each included study. Figure 2 illustrates the forest plots of sensitivities and specificities with 95% CIs for CAM-ICU and ICDSC in this meta-analysis. Among the studies (n= 23) on CAM-ICU, the summary of sensitivity and specificity were 0.85 (95% CI 0.77–0.91) and 0.95 (95% CI 0.90–0.97), respectively. The DOR was 99.0 (95% CI 55.0–177.0). In addition, the pooled sensitivity, specificity, and DOR among studies (n= 8) on ICDSC were 0.87 (95% CI 0.70–0.95), 0.91 (95% CI 0.85–0.95), and 65.0 (95% CI 27.0–153.0), respectively. Figure 3 demonstrates summary ROC curves and AUCs for the included assessment tools. The AUCs were 0.96, and 0.95 for CAM-ICU, and ICDSC, respectively. The bivariate random effects model revealed substantial between-study heterogeneity for CAM-ICU and ICDSC (both $I^2 > 50%$).

Sensitivity Analysis

According to the Cook distance, studies conducted by Boettger et al. (2018), Chuang et al. (2007) and Ely, Inouye et al. (2001) were the most influential (Figure S2.) for CAM-ICU. However, only Chuang et al. (2007) and Nishimura et al. (2016) were detected as outliers, with the highest extreme value for sensitivity and specificity.
Delirium assessment tool (Figure S2.). After excluding these two studies, there was no change in sensitivity and specificity (0.85 and 0.95, respectively); however, the AUC increased from 0.96 to 0.97.

DISCUSSION

The results suggested that the ICDSC had a higher sensitivity than CAM-ICU for delirium detection for patients in ICUs. The high specificity of the CAM-ICU makes it clinically useful for excluding delirium. However, the heterogeneity of the results was significantly high. It could be explained by differences in studies characteristics (medical ICU vs. both medical and surgical ICU) as well as the setting (daily and routine practice vs. research setting). Our results support that both CAM-ICU and ICDSC have more accurate diagnostic performance in sensitivity, specificity, and DORs than any other delirium assessment tools. The findings suggest that CAM-ICU has better diagnostic performance and is therefore recommended for the optimal delirium assessment tool.

The CAM-ICU was developed from on the Confusion Assessment Method (CAM). The CAM was validated for the detection of delirium in a critical care setting, however, participants were excluded from this study if there were unable to communicate (Martins et al., 2015). The CAM-ICU can be applied in mechanically ventilated patients by alternative approach to assess features of delirium including attentiveness and disorganized thinking with visual and auditory assessment such as picture recognition components and the Vigilance A random letter test. (Ely, Inouye, et al., 2001; Ely, Margolin, et al., 2001). Although there are several clinical features of delirium, the CAM-ICU diagnostic criteria is characterized by four elements: 1) acute onset of mental status changes of fluctuating course; 2) inattention; 3) altered level of consciousness; and 4) disorganized thinking. A patient is diagnosed in delirium when both features 1 and 2, and either feature 3 or 4 are present. Before using the CAM-ICU
to assess delirium, it is necessary to evaluate the level of consciousness with the Richmond Agitation Sedation Scale (RASS). Only critically ill patients with a RASS score $\geq -3$ are considered to be assessed by CAM-ICU. This could explain the variations in sensitivity of the CAM-ICU due to the difference of characteristics, particularly in assessing patients with hypoactive delirium and mechanically ventilated patients (Luetz et al., 2010; Wang et al., 2013).

The ICDSC is based on observations during routine patient care and scoring in eight items. There is no need to communicate with patients. Comparing to the CAM-ICU which provides the state of a specific time point, items of ICDSC are observed and scored over 24 hours (Bergeron et al., 2001). Therefore, a previous study suggested the validation of the accuracy of ICDSC with different cut-off points is required in accordance with the utilization in different settings to detect delirium (Boettger et al., 2018).

The American College of Critical Care Medicine guidelines for the management of pain, agitation, and delirium (PAD) recommended regularly using delirium assessment tool such as CAM-ICU or ICDSC in daily practice (Garrett, 2016; Mansouri et al., 2013). As noted, the satisfactory diagnostic performance of CAM-ICU and ICDSC generated valuable diagnostic information for delirium assessments. However, in clinical settings, the CAM-ICU has strengths of a) easier for clinicians to use; b) requires less education or training; c) quicker to apply into daily practice compares to the ICDSC (Boettger et al., 2018; Neto et al., 2012; Nishimura et al., 2016).

Our study had several limitations. First, potential of underestimated of some analyses due to the limited number of studies with sufficient data. Second, the meta-analysis compared different assessment tools by using studies which compared several assessment tools in the same settings. However, different studies characteristics (e.g.,
Delirium assessment tool

sample size and study design) among studies might have possible causes for confounding the results, therefore, future diagnostic studies and meta-analyses which controlling the between-study heterogeneity are warranted.

LINKING EVIDENCE TO ACTION

- CAM-ICU demonstrated good sensitivity and specificity in detecting delirium in critically ill patients. It is the optimal tool for assessing delirium according to its accurate diagnostic performance.
- In clinical settings, CAM-ICU has strengths of easier to use, requires less education or training, and quicker to apply into daily practice for critical care nurses.
- Critical care nurses are suggested to use CAM-ICU regularly in assessing delirium of critically ill patients.

CONCLUSIONS

This meta-analysis aimed to identify the optimum assessment tool for detecting delirium in patients during ICU admission. Our findings revealed that CAM-ICU and ICDSC demonstrated good sensitivity and specificity in detecting delirium in critically ill patients. As many symptoms of delirium are not always detected by critical care nurses and delirium often remains underdiagnosed. The information provided by both methods would enhance critical care nurses’ keen observation of critically ill patients’ symptoms in ICUs. Our findings indicate that CAM-ICU is the optimal diagnostic tool for detecting delirium according to its accurate diagnostic performance.
References


Delirium assessment tool

Jenewein, J. (2017b). Delirium in the intensive care setting: A reevaluation of the validity of the CAM-ICU and ICDSC versus the DSM-IV-TR in determining a diagnosis of delirium as part of the daily clinical routine. *Palliative and Supportive Care, 15*(6), 675-683. doi:10.1017/s1478951516001176


Delirium assessment tool


Delirium assessment tool

*Psychosomatics, 37*(6), 533-546.


Delirium assessment tool

*Internal Medicine, 31*(3), 297-303. doi:10.1007/s11606-015-3514-0


Mitášová, A., Košťálová, M., Bednařík, J., Michalčáková, R., Kašpárek, T., Balabánová, P., . . .


Delirium assessment tool

*Interventions in Aging, 9*, 879-885. doi:10.2147/cia.s62660


van den Boogaard, M., Schoonhoven, L., van der Hoeven, J. G., van Achterberg, T., & Pickkers,