CLINICAL SCIENCE

The validity and reliability of the Portuguese versions of three tools used to diagnose delirium in critically ill patients

Dimitri Gusmao-Flores,1,7 Jorge Ibrain Figueira Salluh,II,III Felipe Dal-Pizzol,IV Cristiane Ritter,IV Cristiane Damiani Tomasi,IV Marco Antônio Sales Dantas de Lima,V Lauro Reis Santana,VI Rita Márcia Pacheco Lins,VI Patrícia Pimenta Lemos,VI Gisele Vasconcelos Serpa,VI Jenisson Oliveira,VI Ricardo Ávila Chalhub,VII Melissa Tassano Pitrowsky,II Acioy L.T. Lacerda,VIII Karestan C. Koenen,IX Lucas C. QuarantiniVI

1Universidade Federal da Bahia, University Hospital Prof. Edgar Santos, Intensive Care Unit, Salvador, BA/Brazil. 2D’Or Institute of Research and Education, Rio de Janeiro, RJ/Brazil. 3Instituto Nacional do Câncer (INCA), Intensive Care Unit, Rio de Janeiro, RJ/Brazil. 4Universidade do Extremo Sul Catarinense, Laboratório de Fisiopatologia Experimental e Instituto Nacional de Ciência e Tecnologia Translacional em Medicina (INCT-TM), Programa de Pós-Graduação em Ciências da Saúde, Unidade Acadêmica de Ciências da Saúde, Cricuíma/SC, Brazil. 5Instituto Nacional do Câncer (INCA), Neurosurgery Section, Rio de Janeiro, RJ/Brazil. 6University Hospital Prof. Edgar Santos, Universidade Federal da Bahia, Department of Psychiatry, Salvador, Bahia/Brazil. 7Universidade Federal da Bahia, Programa de Pós-graduação em Processos Interativos dos Órgãos e Sistemas, Instituto de Ciências da Saúde (ICS), Salvador, BA/Brazil. 8Universidade Federal de São Paulo, Laboratório Interdisciplinar de Neurociências Clínicas – LINC, Department of Psychiatry, São Paulo/SP, Brazil. 9Harvard School of Public Health, Departments of Society, Human Development, and Health & Epidemiology.

OBJECTIVES: The objectives of this study are to compare the sensitivity and specificity of three diagnostic tools for delirium (the Intensive Care Delirium Screening Checklist, the Confusion Assessment Method for Intensive Care Units and the Confusion Assessment Method for Intensive Care Units Flowsheet) in a mixed population of critically ill patients, and to validate the Brazilian Portuguese Confusion Assessment Method for Intensive Care Units.

METHODS: The study was conducted in four intensive care units in Brazil. Patients were screened for delirium by a psychiatrist or neurologist using the Diagnostic and Statistical Manual of Mental Disorders. Patients were subsequently screened by an intensivist using Portuguese translations of the three tools.

RESULTS: One hundred and nineteen patients were evaluated and 38.6% were diagnosed with delirium by the reference rater. The Confusion Assessment Method for Intensive Care Units had a sensitivity of 72.5% and a specificity of 96.2%; the Confusion Assessment Method for Intensive Care Units Flowsheet had a sensitivity of 72.5% and a specificity of 96.2%; the Intensive Care Delirium Screening Checklist had a sensitivity of 96.0% and a specificity of 72.4%. There was strong agreement between the Confusion Assessment Method for Intensive Care Units and the Confusion Assessment Method for Intensive Care Units Flowsheet (kappa coefficient = 0.96).

CONCLUSION: All three instruments are effective diagnostic tools in critically ill intensive care unit patients. In addition, the Brazilian Portuguese version of the Confusion Assessment Method for Intensive Care Units is a valid and reliable instrument for the assessment of delirium among critically ill patients.

KEYWORDS: CAM-ICU; ICSC; CAM-ICU Flowsheet; Critical care; Delirium.


Received for publication on May 19, 2011; First review completed on June 14, 2011; Accepted for publication on July 18, 2011

E-mail: dimitrigusmao@gmail.com
Tel.: 00 55 71 3245-2515

INTRODUCTION

Delirium is an acute and fluctuating disturbance of the consciousness that occurs in up to 80% of patients in intensive care units (ICU) and is associated with increased mortality, longer hospital stays, and long-term cognitive impairment.1-4 Despite its high prevalence and its negative impact on outcomes, the epidemiology and clinical management of delirium have long been compromised by the lack of uniform terminology and validated instruments for detecting and monitoring at-risk patients. Recently, an international effort culminated in a uniform definition and terminology for delirium.5 The need for a specialized professional to evaluate patients according to the Diagnostic and Statistical Manual of...
Mental Disorders (DSM-IV) and the dependence on clinical evaluations rather than validated diagnostic tools frequently leads to the underdiagnosis of delirium. In 2001, an adapted Confusion Assessment Method (CAM) was validated in a cohort of critically ill patients. Since then, the CAM-ICU and other tools, such as the Intensive Care Delirium Screening Checklist (ICDSC), have been tested in various ICU populations. Compared to the delirium checklist (ICDSC), the CAM-ICU demonstrated good agreement with delirium detection for critical care surgical subjects. In a large prospective evaluation, Pun et al. showed that the CAM-ICU demonstrated good compliance and excellent interrater reliability when implemented on a large scale by nursing staff. Recently, the CAM-ICU Flowsheet derived from the CAM-ICU was developed to reduce the time required for patient assessment.

Although a Brazilian national survey of ICU physicians showed that the Brazilian Portuguese version of the CAM-ICU is the most widely used diagnostic tool for delirium diagnosis in critically ill patients, no validation of this tool had been performed prior to the present study.

The main objectives of the present study were to compare the sensitivity and specificity of three diagnostic tools for delirium (the ICDSC, the CAM-ICU and the CAM-ICU Flowsheet) in critically ill patients and to validate the CAM-ICU in Brazilian Portuguese.

MATERIALS AND METHODS

The study was conducted in four medical-surgical intensive care units (two general ICUs in university hospitals, one medical-surgical ICU in a tertiary hospital and one medical-surgical ICU in a comprehensive cancer center) in three cities in diverse regions of Brazil (Salvador/Bahia in the Northeast, Rio de Janeiro/RJ in the Southeast, and Criciúma/Santa Catarina in the South). Each institution recruited a different number of patients. Two units in Salvador (one general ICU in a university hospital and one in a tertiary hospital) enrolled a total of 30 patients, one center in Rio de Janeiro (medical-surgical ICU) recruited 25 patients, and one center in Criciúma (general ICU) recruited 64 patients.

Data collection was conducted between July and November 2010. The local ethics committees approved the study.

Non-consecutive patients over 18 years of age and hospitalized in the ICU for more than 48 hours were included. This convenience sample was obtained with two evaluations every week according to the availability of participating neurologists and psychiatrists to perform evaluations using DSM-IV criteria. All patients had to be arousable (with a score of greater than or equal to 3 according to the Richmond Agitation Sedation Scale) for the evaluation. To prevent the effects of withdrawal, patients were excluded if they had a history of alcohol or narcotic abuse. Those who were unable to communicate (i.e., because of hearing and/or visual impairment) or who did not understand Portuguese were also excluded.

Only one intensivist in each unit was responsible for the application of delirium scales. All intensivists who applied the CAM-ICU and CAM-ICU Flowsheet were trained and had expertise in the use of the tools. With the exception of one center that used two psychiatrists to simultaneously rate the patients, the DSM-IV evaluation was conducted by only one neurologist or psychiatrist.

The ICDSC and the CAM-ICU were previously translated into Portuguese using the recommendations of the International Society of Pharmacoeconomics and Outcomes Research (ISPOR) by Salluh and Dal-Pizzol (www.icudelirium.com). Thirty minutes after the intensivist’s initial evaluation based on the CAM-ICU, a psychiatrist or neurologist applied the DSM-IV criteria as a reference standard. Subsequently, the patient was evaluated by the same intensivist using the CAM-ICU Flowsheet and the ICDSC. The items evaluated by the ICDSC included the following: changes in the level of consciousness, inattention, disorientation, hallucinations, delusions, psychosis, psychomotor agitation or retardation, speech or inappropriate mood, sleep/wake cycle disturbance, and symptom fluctuations. Patients were considered to have delirium if the ICDSC was equal to or greater than 4. Scores between 1 and 3 indicated subsyndromal delirium. The CAM-ICU Flowsheet was developed from the CAM-ICU and involves switching the original numbering of features 3 and 4, as most ICU patients with delirium are given positive scores in the order of the Flowsheet; switching the numbering allows the CAM-ICU Flowsheet to be completed with only three features and the fourth feature is only necessary in a minority of patients.

The intensivists who performed the CAM-ICU and ICDSC were kept unaware of the clinical diagnoses made by the psychiatrist or neurologist.

Patients diagnosed with delirium on any scale were classified into one of three groups: hypoactive, hyperactive, or mixed. Delirium subtypes were classified into motoric subtype groupings according to the Richmond Agitation Sedation Scale (RASS). Patients were considered to have hypoactive delirium if their RASS ratings were between −3 and 0, and were considered to have hyperactive delirium if their RASS ratings were between 1 and 4; mixed-type delirium was defined as alternating between these two ranges.

Demographic and clinical characteristics were collected in addition to the APACHE II score. Patients were followed for up to 28 days, and the following outcomes were recorded: the ICU length of stay, the duration of mechanical ventilation, and 28-day mortality.

Statistical Analysis

Standard descriptive statistics were applied. Using 2x2 tables, the diagnostic values of the CAM-ICU, the CAM-ICU Flowsheet and the ICDSC were described with regard to sensitivity (true positive/[true-positive + false-negative]), specificity (true-negative/[false-positive + true-negative]), positive predictive value (true-positive/[true-positive + false-positive]), and negative predictive value (true-negative/[false-negative + true-negative]). The kappa test was used to verify the reproducibility between instruments, and the chi-square test was used to detect differences in the diagnoses based on the instruments. A receiver operating characteristic (ROC) curve was used to evaluate the performance of the ICDSC in classifying delirium.

Statistical analyses were performed with the statistical software package STATA (version 10.0) using a significance level of 5%.

RESULTS

The characteristics of the 119 patients who met the inclusion criteria are presented in Table 1.
As demonstrated by their severity-of-illness scores (APACHE II scores of 15 ± 6), the sample of patients was determined to be critically ill. The patients represented a mixed ICU population. Using standard evaluation technique (DSM-IV), delirium was observed in 38.6% (46/119) of the patients. The types of delirium observed in these patients included hypoactive (69.5%), hyperactive (19.5%), and mixed (11%) delirium. Most patients (76.4%) were easily arousable at the time of their evaluation (RASS-0 or RASS-1). The CAM-ICU identified 26.8% of the delirious patients unable to speak after being intubated and ventilated. As demonstrated by their severity-of-illness scores (APACHE II scores of 15 ± 6), the sample of patients was determined to be critically ill. The patients represented a mixed ICU population. Using standard evaluation technique (DSM-IV), delirium was observed in 38.6% (46/119) of the patients. The types of delirium observed in these patients included hypoactive (69.5%), hyperactive (19.5%), and mixed (11%) delirium. Most patients (76.4%) were easily arousable at the time of their evaluation (RASS-0 or RASS-1). The CAM-ICU identified 26.8% of the delirious patients unable to speak after being intubated and ventilated. The first validation study of the CAM-ICU included only 38 patients. Two nurses and two intensivists compared the CAM-ICU Flowsheet and the ICDSC. The McNemar test was used to detect the correlation between the diagnostic tools. We observed a concordance of 98.32% with a kappa of 0.96 between the CAM-ICU and the ICDSC (Table 3). The McNemar test (p = 1.00) suggested that there were no significant differences between the two instruments. To assess the correlation between the ICDSC and the CAM-ICU, it was necessary to exclude patients with a diagnosis of subsyndromal delirium that was diagnosed based only on the ICDSC (27 patients). We found a kappa of 0.59 (Table 4). As expected, similar findings were observed when comparing the CAM-ICU Flowsheet and the ICDSC.

The ICDSC classified delirium adequately when compared with the DSM-IV (area under the ROC curve = 0.91). The ROC curve is displayed in Figure 1. A diagnostic cutoff value of 5 or more for the ICDSC total score provided 67.5% sensitivity and 96.2% specificity for diagnosing delirium. With this cutoff, delirium was correctly classified in 86.5% of cases.

**DISCUSSION**

The aim of this study was to validate the Brazilian Portuguese version of the CAM-ICU according to DSM-IV criteria. The scale showed an overall sensitivity of 72.5% (95% CI: 55.9% – 84.9%) and specificity of 96.2% (95% CI: 88.5% – 99%). Moreover, the scale demonstrated high positive (90.6%; 95% CI: 73.8% – 97.5%) and negative (87.4%; 95% CI: 78.1% – 93.2%) predictive values, which suggests that very few cases of delirium remain unidentified when the scale is used systematically. Thus, our data demonstrate that the CAM-ICU is valid in Brazilian Portuguese. In addition, there is high accuracy for delirium diagnosis among the three tools (CAM-ICU, CAM-ICU Flowsheet, and ICDSC), and the CAM-ICU and CAM-ICU Flowsheet can be used interchangeably.

The development and validation of diagnostic tools is important to a thorough understanding of clinical disorders. Unfortunately, studies have demonstrated that a clinical impression is insufficient for delirium diagnosis. Recently, a Dutch group observed a low sensitivity for delirium diagnosis with only clinical observation (45%); making it necessary to develop and validate diagnostic tools. In 1990, using DSM-III-R criteria, Inouye created and validated the Confusion Assessment Method (CAM), an algorithmic technique that uses only four of the DSM-III-R criteria for delirium. In the intensive care environment, the CAM has been adapted as the CAM-ICU because many patients are unable to speak after being intubated and ventilated.

The first validation study of the CAM-ICU included only 38 patients. Two nurses and two intensivists compared the CAM-ICU method with the standard DSM-IV. In addition to a high specificity and sensitivity, an excellent interrater correlation was observed. The same investigators published a second study that included 111 patients who were on mechanical ventilation; in addition to confirming a high interrater correlation (kappa coefficient: 0.99, 95% CI: 0.92 – 0.99), they reported a sensitivity and specificity of approximately 100%.

Our study differs in some respects from the studies by Ely et al. described above. First, we did not observe as high a sensitivity for the CAM-ICU, which varied in Ely et al. from 93% to 100%. In our study, the sensitivity of the CAM-ICU was 72.5%. Although there is not a clear explanation, the difference is not likely related to the implementation of the CAM-ICU in Portuguese, as similar
One possible explanation is a change in the sensitivity of the CAM-ICU when it is used in a cohort of mechanically ventilated and sedated patients. We observed that most patients had a RASS score of zero (60.5%), which may not only represent the lower degree of severity in our cohort but may also represent a current trend toward less sedation in ICU patients. When comparing diagnostic instruments for delirium, Luetz et al. demonstrated that the sensitivity of the CAM-ICU is higher in patients with a RASS score of higher than 0. However, a common feature of every published study is the high specificity of the CAM-ICU.

The CAM-ICU has been translated and validated in many languages and has become the most frequently employed tool for diagnosing delirium in ICU patients. A distinct advantage of this tool is that it does not require the patient to speak, which can be useful in patients who are on mechanical ventilation. In our study, we observed no difference in the accuracy of the CAM-ICU between the patients who were ventilated and those who were not. We also observed different scales in patients undergoing noninvasive ventilation (NIV). Eleven patients were observed with NIV (three of whom presented with delirium), and there was a 100% correlation among the CAM-ICU Flowsheet, the CAM-ICU, and the ICDSC.

The CAM-ICU Flowsheet was developed to decrease the time required for the evaluation of patients with suspected delirium. However, to the best of our knowledge, only a single study has evaluated its performance. In our study, we observed an excellent correlation between the CAM-ICU and the CAM-ICU Flowsheet, with a kappa of 0.96. Guenther et al. evaluated the CAM-ICU Flowsheet in German (with a duration of application that did not exceed 2 minutes) and noted a sensitivity of 88% to 92% and a specificity of 100% with close interobserver correlation. In our study, the sensitivity was 72.5%, and the specificity was 96.2%. In some cases, less than one minute was necessary for completion of the instrument.

The ICDSC checklist is an eight-item screening tool (one point for each item) that is based on DSM criteria and applied to data that can be collected through medical records or on information obtained from a multidisciplinary team. Bergeron et al. observed a high sensitivity (99%) when a cutoff of 4 was used, but a moderate specificity was observed (64%). Similarly, our study found a high sensitivity (96%) and a moderate specificity (72.4%). Other studies have reported a low sensitivity for the ICDSC (43%) compared with the standard method of diagnosis (DSM-IV). In the mixed population of patients in this earlier study, the CAM-ICU showed a higher sensitivity (64%) but a lower specificity. More recently, the German version of the ICDSC was compared with the CAM-ICU in a population of surgical patients with a close correlation (kappa coefficient: 0.8; 95% CI: 0.78 – 0.84; p<0.001). In our study, we observed a low correlation between the CAM-ICU and the ICDSC (kappa coefficient: 0.59). A change in the cutoff would likely change the correlation between these diagnostic tools. A cutoff of 5 correctly identified 86.5% of cases, whereas a cutoff of 4 correctly identified 80.6% of cases. For this analysis, it was necessary to exclude cases that were considered to be subsyndromal delirium (a cutoff of 3).

Evaluating all 119 cases, we observed a high degree of accuracy with this tool and the DSM-IV, with an area under the ROC curve of 0.91. Because the CAM-ICU and CAM-ICU Flowsheet responses are dichotomous (yes or no), it was not possible to draw an ROC curve.

We found the CAM-ICU and the CAM-ICU Flowsheet to be similar and to be highly accurate for delirium diagnosis, which suggests that these are appropriate tools for developing a diagnostic profile. However, because of its high specificity and only moderate sensitivity, the ICDSC may be more useful in stratifying patients with delirium. Recently, Tomasi et al. suggested that the CAM-ICU is a better predictor of outcomes than the ICDSC, which is probably because of the high rate of false positives with the ICDSC.

Our study has some notable limitations. First, as the study was performed in different regions, the tools evaluated (CAM-ICU, CAM-ICU Flowsheet, and ICDSC) were applied by different intensivists in different ICUs. Thus, we could not perform an interrater correlation with the tools that were applied. However, we believe that the evidence is strong enough to demonstrate a close correlation between the raters because the tool is simple and easily applied. We measured the performance of the scales against the DSM-IV, which is considered to be the standard technique for clinical assessment. Therefore, the application of the CAM-ICU and the CAM-ICU Flowsheet by the same investigator does not imply an evaluation bias.

Our study also has several strengths. Not only was our study the first to validate the CAM-ICU and the CAM-ICU Flowsheet for Brazilian Portuguese, but it was performed as a multicenter evaluation in three different and representative regions of Brazil. The study evaluated a mixed population of critically ill patients, including ventilated and non-ventilated patients. These methodological characteristics increase the external validity of the results.

The present data demonstrate that the CAM-ICU and the ICDSC are valid tools that can be used in Brazilian Portuguese with a high degree of accuracy. The CAM-ICU Flowsheet has an excellent agreement with the CAM-ICU (kappa coefficient = 0.96) and can be employed as a fast,
practical and reliable tool. Finally, the ICDSC has a high sensitivity for diagnosing delirium but a moderate specificity and a poor correlation with the CAM-ICU and CAM-ICU Flowsheet.

ACKNOWLEDGMENTS

This project was partially supported by the National Council of Technological and Scientific Development (CNPq): [474869/2010-5] - Edital Universal MCT/CNPq 14/2010.

AUTHOR CONTRIBUTIONS

Gusmao-Flores D and Salluh JIF designed the study and wrote the protocol, applied the scales and performed the data collection, undertook the statistical analysis, and contributed to the writing of the first draft of the manuscript. Pitrowsky MT applied the scales, performed the data collection, and contributed to the writing of the first draft of the manuscript. Quarantini LC designed the study and wrote the protocol, applied the scales and performed the data collection, undertook the statistical analysis, and contributed to the writing of the first draft of the manuscript. Lima MASD, Santana LR, Lins RMP, Lemos PP, Serpa G and Oliveira J applied the scales, performed the data collection and contributed to the writing of the first draft of the manuscript. dal-Pizzol F, Ritter C, Tomasi CD designed the study and wrote the protocol, applied the scales and performed the data collection. Lima DP, Ochiai ME, Lima AB, Curiati JA, Farfel JM, Filho WJ. Delirium in hospitalized elderly patients and post-discharge mortality. Clinics. 2010;65:251-5, doi: 10.1590/S1807-59322010000300003.

REFERENCES


