

Routine Use of the Confusion Assessment Method for the Intensive Care Unit

A Multicenter Study

Maarten M. van Eijk¹, Mark van den Boogaard², Rob J. van Marum³, Paul Benner⁴, Piet Eikelenboom^{5,6}, Marina L. Honing⁷, Ben van der Hoven⁸, Janneke Horn⁹, Gerbrand J. Izaks¹⁰, Annette Kalf¹¹, Attila Karakus¹², Ine A. Klijn¹³, Michael A. Kuiper¹⁴, Frank-Erik de Leeuw¹⁵, Tjarda de Man¹⁶, Roos C. van der Mast¹⁷, Robert-Jan Osse¹⁸, Sophia E. de Rooij¹⁹, Peter E. Spronk²⁰, Peter H. van der Voort²¹, Willem A. van Gool⁵, and Arjen J. Slooter¹

¹Department of Intensive Care Medicine and ¹³Department of Psychiatry, University Medical Center Utrecht, Utrecht, The Netherlands;

²Department of Intensive Care and ¹⁵Department of Neurology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands;

³Department of Geriatrics, Jeroen Bosch Ziekenhuis, Den Bosch, The Netherlands; ⁴Department of Intensive Care, Kennemergasthuis, Haarlem, The Netherlands;

⁵Department of Neurology; ⁹Department of Intensive Care Medicine, and ¹⁹Department of Geriatrics, Academic Medical Center, Amsterdam, The Netherlands;

⁶Department of Psychiatry, VU Medical Center, Amsterdam, The Netherlands; ⁷Department of Intensive Care and

¹⁶GGZ Noord-Holland-Noord, Medical Center Alkmaar, The Netherlands; ⁸Department of Intensive Care, and ¹⁸Department of Psychiatry, Erasmus MC-University Medical Center, Rotterdam, The Netherlands;

¹⁰University Center for Geriatric Medicine, University Medical Center, Groningen, The Netherlands; ¹¹Department of Geriatrics, Gelre Hospitals Apeldoorn/Zutphen, The Netherlands; ¹²Department of Intensive Care, Diaconessenhuis, Utrecht, The Netherlands;

¹⁴Department of Intensive Care, Medical Center Leeuwarden, The Netherlands; ¹⁷Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands;

²⁰Department of Intensive Care, Gelre Hospitals Apeldoorn, The Netherlands; and

²¹Department of Intensive Care, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands

Rationale: Delirium is often unrecognized in ICU patients and associated with poor outcome. Screening for ICU delirium is recommended by several medical organizations to improve early diagnosis and treatment. The Confusion Assessment Method for the ICU (CAM-ICU) has high sensitivity and specificity for delirium when administered by research nurses. However, test characteristics of the CAM-ICU as performed in routine practice are unclear.

Objectives: To investigate the diagnostic value of the CAM-ICU in daily practice.

Methods: Teams of three delirium experts including psychiatrists, geriatricians, and neurologists visited 10 ICUs twice. Based on cognitive examination, inspection of medical files, and *Diagnostic and Statistic Manual of Mental Disorders, 4th edition, Text Revision* criteria for delirium, the expert teams classified patients as awake and not delirious, delirious, or comatose. This served as a gold standard to which the CAM-ICU as performed by the bedside ICU-nurses was compared. Assessors were unaware of each other's conclusions.

Measurements and Main Results: Fifteen delirium experts assessed 282 patients of whom 101 (36%) were comatose and excluded. In the remaining 181 (64%) patients, the CAM-ICU had a sensitivity of 47% (95% confidence interval [CI], 35%–58%); specificity of 98% (95% CI, 93%–100%); positive predictive value of 95% (95% CI, 80%–99%); and negative predictive value of 72% (95% CI, 64%–79%). The positive likelihood ratio was 24.7 (95% CI, 6.1–100) and the negative likelihood ratio was 0.5 (95% CI, 0.4–0.8).

(Received in original form January 13, 2011; accepted in final form April 24, 2011)

Author contributions: M.M.v.E., M.v.d.B., and A.J.S. designed and coordinated the study; M.M.v.E. and A.J.S. performed the statistical analyses and drafted the manuscript; M.M.v.E., M.v.d.B., R.J.v.M., P.E., G.J.I., A.K., I.A.K., M.A.K., F-E.d.L., T.d.M., R.C.v.d.M., R.J.O., S.E.d.R., W.A.v.G., and A.J.S. participated as delirium experts; M.M.v.E., M.v.d.B., P.B., M.L.H., B.v.d.H., J.H., A.K., M.A.K., P.E.S., and P.H.v.d.V. were local investigators. All authors interpreted the results, critically revised the manuscript, and approved the final version of the manuscript.

Correspondence and requests for reprints should be addressed to Arjen J. Slooter, M.D., Ph.D., Department of Intensive Care Medicine, University Medical Center Utrecht, Heidelberglaan 100, room F06.149, 3584 CX, Utrecht, The Netherlands. E-mail: a.slooter-3@umcutrecht.nl

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 184, pp 340–344, 2011

Originally Published in Press as DOI: 10.1164/rccm.201101-0065OC on May 11, 2011

Internet address: www.atsjournals.org

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The Confusion Assessment Method for the ICU is a frequently used method to screen for delirium.

What This Study Adds to the Field

Half of the patients with delirium are not detected by the Confusion Assessment Method for the ICU as performed in daily practice. This hampers the clinical use of this instrument.

Conclusions: Specificity of the CAM-ICU as performed in routine practice seems to be high but sensitivity is low. This hampers early detection of delirium by the CAM-ICU.

Keywords: CAM-ICU; delirium; intensive care unit; diagnostic characteristics

Delirium is characterized by an acute disturbance of consciousness and attention with cognitive or perceptual changes and often a fluctuating course (1). It is common in ICU patients, with an incidence during ICU stay ranging up to 89% (2–8). In patients in the ICU, delirium is associated with poor outcome, including increased mortality, increased ICU and hospital length of stay, more cognitive impairment after hospital discharge, and higher healthcare-related costs (4, 9–11).

Despite its frequency and impact, delirium in the ICU often goes unrecognized, which hampers early treatment (12). The clinical impression of ICU physicians and nurses whether or not an ICU patient was delirious had a sensitivity of 29% and 35%, respectively, compared with the conclusion of delirium experts (13, 14). To improve early recognition of delirium, several easy-to-use screening methods have been developed (15), such as the Confusion Assessment Method adopted for the ICU (CAM-ICU) (16) and the Intensive Care Delirium Screening Checklist (ICDSC) (17). Of these, the CAM-ICU

had higher sensitivity (64%; specificity 88%) than the ICDSC (sensitivity 43%; specificity 95%) within the same population of mixed ICU patients (13), and is therefore the most frequently used delirium detection tool (18).

Several medical organizations, including the Society of Critical Care Medicine and the American Psychiatric Association, advise standard screening for delirium in critically ill patients (12, 19). The CAM-ICU showed high sensitivity (range, 97%–100%) and specificity (range, 89%–100%) in several validation studies (16, 20–22). It should, however, be noted that all these investigations were performed in research settings (16, 20–22), which may differ from day-to-day critical care (23). Test characteristics of the CAM-ICU as performed by bedside ICU nurses are unknown. The aim of this study was to investigate the diagnostic value of the CAM-ICU when performed by bedside ICU nurses in routine daily practice. Some of the results of this study have been previously reported in the form of an abstract (24–26).

METHODS

Design and Setting

This prospective multicenter study was performed in 10 ICUs of university, teaching, and rural hospitals in The Netherlands, which were selected based on a previous nation-wide survey on the use of delirium monitoring (18). The study population consisted of mixed medical and surgical ICU patients who were admitted to one of the participating ICUs during visits of delirium experts, as described later. Patients who were unable to speak Dutch or English and those who could not be examined because of logistic reasons were excluded. This study was approved by the institutional review board of the University Medical Center Utrecht, Utrecht, The Netherlands, and a waiver for informed consent was obtained.

Implementation of the CAM-ICU

Investigators in each participating center completed a questionnaire concerning the implementation of the CAM-ICU at their ICU. They registered whether lectures had been given to nurses preceding implementation in the daily routine of that specific ICU, and whether written information was provided explaining the use of the CAM-ICU. Furthermore, local investigators were asked if individual bedside training had been given to the nurses and whether the ICU physicians always or regularly used the CAM-ICU results at their daily rounds. Lastly, the local investigators registered compliance rate of the CAM-ICU in daily practice.

Delirium Assessment

During visits to the participating centers, a group of three experts made rounds along all admitted ICU patients at that time. One of these experts was either a research-physician (MMvE) or a nurse-scientist (MvdB), who guaranteed that all assessments and study-procedures were performed uniformly. The other two experts were, in different combinations guaranteeing a multidisciplinary team, psychiatrists ($n = 5$), geriatricians ($n = 4$), or neurologists ($n = 4$), who had on average 16 years (SD 5) clinical experience after their medical specialist registration and who saw an estimated 20 (mean SD 8) delirious patients monthly. To guarantee masking for CAM-ICU scores of preceding days, experts were not allowed to evaluate patients in their own center. The expert groups assessed the patients using the *Diagnostic and Statistical Manual of Mental Disorders, 4th edition, Text Revision* (1) criteria for delirium, based on clinical assessment for cognitive dysfunction and a review of the medical charts, but remained masked to reported CAM-ICU scores. The group of experts classified each patient as (1) awake and not delirious; (2) delirious; or (3) comatose (i.e., not assessable because of a low level of consciousness). If they diagnosed a patient as delirious they had to classify whether they thought the patient suffered from a hypoactive-, a hyperactive-, or

a mixed-type of delirium. The expert groups remained masked for the CAM-ICU as scored by the nurses throughout the visit.

The bedside nurses assessed all patients using the Dutch version of the CAM-ICU (21) within 3 hours of the expert assessment, without extra training for this study. We further registered the CAM-ICU scores on the day before and the day after the experts visits. The ICU nurses were masked for the assessment by the expert groups and received no notice before the study visits were made.

Other Data Collection

Local investigators supplied data on age; sex; Acute Physiology and Chronic Health Evaluation (APACHE) II score; admitting discipline; and the ability to verbally communicate (e.g., intubation or tracheotomy) at the moment of assessment. The timing of the administration of psychoactive medication (e.g., antipsychotics, opiates, or benzodiazepines) and the timing of assessment by the expert group and the bedside nurse were also noted by the local investigator.

Statistical Analyses

After exclusion of patients who were nonassessable by either the expert groups or the nurses, we calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for the CAM-ICU, based on 2×2 tables, with the classification of the expert groups as reference. Furthermore, because delirium may vary in time, we analyzed test characteristics using the Richmond Agitation and Sedation Score (RASS) and CAM-ICU results from the day before and the day after the experts' visits, based on the following classification: always RASS less than -3 during this 48-hour period, or never a positive CAM-ICU during this 48-hour period, or a positive CAM-ICU at one or more moments during this 48-hour period, and the reference described previously. Prespecified stratified analyses were performed on type of delirium (hypoactive-, hyperactive-, or mixed-type); study center; and ability of verbal communication. Agreement between the CAM-ICU results and the expert groups was computed with Cohen's kappa coefficient (κ).

RESULTS

Between April 2009 and April 2010 all 10 participating centers were attended twice. The expert groups visited 306 different patients of whom 282 (92%) were assessable as either awake and not delirious, or delirious or comatose. We excluded 14 patients (5%) who could not speak Dutch or English. Ten patients (3%) could not be assessed because they underwent an examination or a procedure when the expert group made their round. The average age in the included patients was 59 years (SD 18) and the average APACHE II score was 18.6 (SD 7.5) (Table 1).

The participating centers admitted on average 1,545 patients a year (SD 500) and had on average 25 beds (SD 12) all with capability for mechanical ventilation (Table 2). All ICUs worked according to closed format formula with on average nine (SD 3) intensivists. Concerning the implementation of the CAM-ICU, all participating ICUs reported to have provided lectures and written information to their nurses before its introduction. Most centers (60%) offered individual bedside teaching before or during the introduction of the CAM-ICU. In 3 out of 10 participating centers, CAM-ICU test results as performed by the ICU nurse were always part of the standard evaluation by the attending intensivist. In the other centers, CAM-ICU results were regularly used. The average time from implementation of the CAM-ICU to participation in this study was 2 years (SD 0.5).

The expert groups reached consensus in all 282 cases, who were classified as awake and not-delirious ($n = 106$; 38%), delirious ($n = 80$; 28%), or comatose ($n = 96$; 34%) (Table 3). Of these 282 patients, 159 (56%) patients were scored CAM-ICU

TABLE 1. CHARACTERISTICS OF THE STUDY POPULATION

Characteristic	Total (n = 282)	Delirium* (n = 80)	No Delirium* (n = 106)	Coma* (n = 96)
Age (years), mean (SD)	59 (18)	62 (15)	59 (16)	57 (21)
Gender, male, n (%)	172 (61)	54 (68)	64 (60)	54 (56)
APACHE-II score, mean (SD)	18.6 (7.5)	20.1 (7)	16.2 (6.9)	20.2 (7.8)
Admitting discipline, n (%)				
Internal medicine	96 (34)	29 (36)	37 (36)	30 (31)
General surgery	90 (32)	23 (29)	29 (26)	38 (40)
Cardiology and cardiothoracic surgery	62 (22)	21 (25)	24 (22)	17 (17)
Neurology and neurosurgery	34 (12)	7 (10)	16 (16)	11 (12)
Able to communicate verbally, n (%)	107 (38)	36 (45)	71 (67)	0 (0)

Definition of abbreviation: APACHE = Acute Physiology and Chronic Health Evaluation.

*As defined by *Diagnostic and Statistical Manual of Mental Disorders, 4th ed., Text Revision* criteria and assessed by experts.

negative by the bedside nurses; 44 subjects (16%) as CAM-ICU positive; and 79 patients (28%) as RASS less than -3 (not assessable). In total, 101 patients were identified as comatose, either by the expert groups or by the bedside nurses, and excluded to calculate sensitivity, specificity, and predictive values of the CAM-ICU. The kappa score for agreement between CAM-ICU results and expert conclusions was $\kappa = 0.63$.

As shown in Table 4, delirium was detected in 75 out of 181 remaining patients by the experts. The CAM-ICU as administered by the bedside nurses was positive in 35 of these 75 subjects. This yielded an overall sensitivity of 47% (95% confidence interval [CI], 35%–58%) and a specificity of 98% (95% CI, 93%–100%). The overall PPV and NPV were 95% (95% CI, 80%–99%) and 72% (95% CI, 64%–79%), respectively. The positive likelihood ratio was 24.7 (95% CI, 6.1–100) and the negative likelihood ratio was 0.5 (95% CI, 0.4–0.8) (see Table E1 in the online supplement). When this analysis was based on the RASS and CAM-ICU scores the day before, the day of, and the day after the expert assessment, we found the sensitivity to be 53% (95% CI, 41%–65%); the specificity 86% (95% CI, 77%–92%); the PPV 73% (95% CI, 60%–83%); and the NPV 72% (95% CI, 64%–79%).

The median duration between evaluation by the expert group and assessment with the CAM-ICU was 86 minutes (interquartile range, 41–168 min). Based on this interval, data were divided into quartiles and analyses were repeated. No substantial differences were found between the lowest quartile (interval <41 min: sensitivity 36% [95% CI, 14%–64%]; specificity 100% [95% CI, 86%–100%]; PPV 100% [95% CI, 46%–100%]; and NPV 78% [95% CI, 61%–89%]) and the highest quartile (interval >168 min: sensitivity 43% [95% CI, 24%–65%]; specificity 95% [95% CI, 74%–99%]; PPV 90% [95% CI, 57%–99%]; and NPV 61% [95% CI, 42%–77%]). In 31% of patients

(n = 87) the evaluation of the experts preceded the assessment of the nurses; in 69% of patients (n = 195) the assessment of the nurses preceded the visit of the experts. When we compared test characteristics between these two groups, no substantial differences in sensitivity, specificity, PPV, and NPV were found (data not shown).

All analyses were repeated after exclusion of 46 patients (25%) who had received psychoactive medication between both assessments, which did not differ from the previously described overall results: sensitivity 43% (95% CI, 30%–58%); specificity 99% (95% CI, 93%–100%); PPV 96% (95% CI, 76%–100%); and NPV 74% (95% CI, 64%–82%).

After stratification according to type of delirium, sensitivity of the CAM-ICU was lowest in the hypoactive subgroup (31%; 95% CI, 17%–48%); highest in the hyperactive delirious patients (100%; 95% CI, 56%–100%); and intermediate in the mixed-type patients (53%; 95% CI, 35%–74%). As further shown in Table 4, the CAM-ICU showed particularly poor test characteristics in neurocritical care patients (sensitivity 17%; 95% CI, 1%–64%). Centers always using the CAM-ICU result to adapt clinical practice on a daily base showed better test characteristics than centers in which the CAM-ICU was regularly used, especially with regard to sensitivity (range, 50%–71%; respectively, 29%–80%).

DISCUSSION

In this multicenter evaluation of daily practice, the CAM-ICU was found to have a sensitivity and specificity of 47% and 98%, respectively, and PPV and NPV of 95% and 72%, respectively. Sensitivity was particularly poor in neurocritical care patients, in patients with hypoactive delirium, and in centers where the test results were not always part of the standard

TABLE 2. CHARACTERISTICS OF PARTICIPATING ICUS

Characteristics	Center A	Center B	Center C	Center D	Center E	Center F	Center G	Center H	Center I	Center J
Beds per center, n	33	10	32	12	10	24	50	32	30	10
Intensivists (full-time equivalents) per center, n	20	4	11	5	4	7	19	11	10	4
Annual admissions per center, n	2,250	600	2,500	640	800	1,456	1,952	2,000	2,000	713
Time from implementation CAM-ICU to participation in study, mo	36	48	36	24	12	36	12	12	24	12
Lectures given	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Written information available	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Individual bedside training	Y	Y	Y	N	N	Y	N	N	Y	Y
Duration of individual training, min	20	30	10	N/A	N/A	20	N/A	N/A	20	10
Use of dedicated training nurses	Y	Y	N	Y	N	N	Y	N	N	N
Standard use of CAM-ICU in daily decision making	A	A	R	A	R	R	R	R	R	R
Trained nurses, %	95	95	80	N/A	N/A	90	N/A	N/A	80	90
Compliance to the daily CAM-ICU, %	93	95	90	70	90	80	95	80	85	95
Frequency of CAM-ICU per day	3	3	3	3	2	2	2	2	3	2

Definition of abbreviations: A = always; CAM-ICU = Confusion Assessment Method for the Intensive Care Unit; N = no; N/A = not applicable; R = regularly; Y = yes.

TABLE 3. OVERALL CLASSIFICATION OF THE STUDY POPULATION

	Delirium*	No Delirium*	Coma*	Total
CAM-ICU positive	35	2	7	44
CAM-ICU negative	40	104	15	159
RASS less than -3	5	0	74	79
Total	80	106	96	282

Definition of abbreviations: CAM-ICU = Confusion Assessment Method for the Intensive Care Unit administered by the bedside nurse; RASS = Richmond Agitation and Sedation score.

*According to the delirium experts and *Diagnostic and Statistical Manual of Mental Disorders, 4th ed., Text Revision* criteria.

evaluation by the attending intensivist. The sensitivity remained low when CAM-ICU results from a 48-hour period were considered.

The CAM-ICU in daily practice showed thus not quite as good test characteristics as presented in the original validation studies (16, 20–22), where a limited number of specially trained research nurses performed the test and some categories of ICU patients were excluded (e.g., patients with a neurologic disorder). The discrepancy in findings may also be caused by inadequate training or by incomplete implementation of the CAM-ICU in daily routine. Because training was comprehensive and did not essentially differ between centers, a possible explanation for our results may be that bedside nurses lack motivation to perform the CAM-ICU correctly if it is not standard to always evaluate the result in the treatment of their patients. However, even in centers where the test results were always part of the standard evaluation by the intensivist, sensitivity was still substantially lower (50%–71%) than in the original validation studies (97%–100%) (16, 20–22).

Strengths of this study include the sample size. This study is the largest study on this topic, with 181 included noncomatose

patients from 10 different ICUs. Earlier CAM-ICU validation studies (16, 20–22) included fewer patients (range, 30–129) and all had a single-center design, potentially hampering external validity. Most importantly, in all previous studies, assessments were performed by a limited number of research nurses, whereas our investigation is an evaluation of daily life. The gold standard classification was made by multidisciplinary expert groups comprising physicians from other centers with significant expertise and experience in assessing patients with delirium. Furthermore, it was ensured that the assessments and study procedures by the expert groups were always performed similarly. Because the expert groups were unaware of the CAM-ICU as registered by the bedside nurses, and bedside nurses were blinded to the examinations and conclusions of the expert groups, our findings are not subject to bias.

This study has also some limitations. The classification of the type of delirium by the experts may have lacked accuracy because it was based on an assessment at a given moment in time, whereas delirium symptoms tend to fluctuate over the day.

Second, expert assessment and the CAM-ICU could not always be performed immediately after each other. Because delirium tends to fluctuate during the course of the day, discrepancies between the two assessments might result from differences in clinical presentation over time. However, our results were not related to the time interval in between assessments, did not change when we excluded patients who had received psychoactive medication in between evaluations, and were essentially similar when we stratified on the order of the assessments. Therefore our findings seem not to be subject to bias. Moreover, time between assessments was comparable with the original validation studies (16, 20).

Third, we stratified our results according to study center and related these findings to differences in training and implementation. These observations should, however, be interpreted with

TABLE 4. TEST CHARACTERISTICS OF THE CONFUSION ASSESSMENT METHOD FOR THE INTENSIVE CARE UNIT

Subpopulation (n)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total population (n = 181)	47% (35%–58%)	98% (93%–100%)	95% (80%–99%)	72% (64%–79%)
Psychoactive medication between assessments				
Yes (n = 46)	54% (33%–74%)	95% (75%–99%)	93% (64%–99%)	65% (47%–81%)
No (n = 135)	43% (30%–58%)	99% (93%–100%)	96% (76%–100%)	74% (64%–82%)
Delirium subtypes*				
Hypoactive (delirious n = 36; not delirious n = 106)	31% (17%–48%)	98% (92%–99%)	85% (54%–97%)	81% (72%–87%)
Hyperactive (delirious n = 7; not delirious n = 106)	100% (56%–100%)	98% (93%–100%)	78% (40%–96%)	100% (95%–100%)
Mixed-type (delirious n = 32; not delirious n = 106)	53% (35%–74%)	98% (93%–100%)	90% (65%–98%)	87% (80%–93%)
Admitting discipline				
Internal medicine (n = 52)	54% (33%–73%)	96% (78%–100%)	93% (64%–100%)	69% (52%–83%)
General surgery (n = 64)	38% (21%–59%)	97% (85%–100%)	91% (57%–100%)	70% (55%–81%)
Cardiology and cardiothoracic surgery (n = 43)	58% (34%–79%)	100% (83%–100%)	100% (68%–100%)	75% (56%–88%)
Neurology and neurosurgery (n = 22)	17% (1%–64%)	100% (76%–100%)	100% (1%–100%)	76% (52%–91%)
Communication ability				
Verbal communication possible (n = 109)	42% (26%–61%)	99% (91%–100%)	93% (66%–99%)	79% (68%–86%)
Verbal communication not possible (n = 72)	50% (34%–66%)	97% (82%–100%)	95% (74%–100%)	61% (46%–74%)
Center (number of included patients)				
A (n = 26)	71% (42%–90%)	92% (60%–100%)	91% (57%–100%)	73% (45%–91%)
B (n = 9)	50% (4%–91%)	100% (46%–100%)	100% (20%–100%)	71% (30%–95%)
C (n = 34)	29% (10%–58%)	100% (80%–100%)	100% (40%–100%)	67% (47%–82%)
D (n = 20)	60% (17%–93%)	93% (66%–100%)	75% (22%–99%)	88% (60%–98%)
E (n = 10)	33% (2%–87%)	100% (56%–100%)	100% (5%–100%)	78% (40%–96%)
F (n = 15)	N/A	100% (31%–100%)	N/A	20% (5%–49%)
G (n = 26)	44% (21%–69%)	100% (66%–100%)	100% (56%–100%)	53% (29%–75%)
H (n = 14)	80% (30%–99%)	100% (63%–100%)	100% (40%–100%)	90% (54%–99%)
I (n = 20)	44% (15%–77%)	100% (68%–100%)	100% (39%–100%)	69% (41%–88%)
J (n = 7)	N/A	100% (46%–100%)	N/A	71% (30%–95%)

Definition of abbreviations: CAM-ICU = Confusion Assessment Method for the Intensive Care Unit; CI = confidence interval; N/A = not applicable (no delirious patient identified); NPV = negative predictive value; PPV = positive predictive value.

*According to the expert groups.

caution because the number of patients per center was small and the exact process of training and implementation was difficult to objectify.

Fourth, a possible concern is the generalizability of our findings. The participating centers were selected from all Dutch ICUs based on a previous survey on routine delirium monitoring (18). Because these centers represent the ICUs where delirium monitoring was implemented earliest, these centers most likely are the most active sites with regard to delirium care. It seems unlikely that our selection of study centers has negatively influenced test characteristics of the CAM-ICU.

High sensitivity is an essential feature for a screening instrument, because screening is about identifying all patients with the disease. In our study, sensitivity of the CAM-ICU was overall 47% and 31% for hypoactive delirium, the delirium subtype most difficult to recognize for ICU physicians (13). This low sensitivity of the CAM-ICU hampers its use as a screening instrument for delirium in critically ill patients. The specificity and the PPV were, however, high. The higher sensitivity of the CAM-ICU found in centers always using the CAM-ICU results in daily care suggests that this may be a necessary condition for achieving adequate implementation in daily practice. Furthermore, sensitivity may be increased by combining CAM-ICU results with observations described in nursing files (27). Results from clinical efficacy trials often contradict results from “real-world” analyses (23). In this study we have shown that this may also apply for screening instruments.

In conclusion, in this multicenter study, specificity of the CAM-ICU as performed in daily critical care seems to be high but sensitivity is low. The low sensitivity of the CAM-ICU in routine practice hampers early detection of delirium.

Author Disclosure: None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Acknowledgment: The authors thank all intensive care unit nurses in the participating centers.

References

- Diagnostic and statistical manual of mental disorders, 4th ed., Text revision. Washington, DC: American Psychiatric Association; 2000.
- Peterson JF, Pun BT, Dittus RS, Thomason JW, Jackson JC, Shintani AK, Ely EW. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc* 2006;54:479–484.
- Dubois MJ, Bergeron N, Dumont M, Dial S, Skrobik Y. Delirium in an intensive care unit: a study of risk factors. *Intensive Care Med* 2001;27:1297–1304.
- Lin SM, Liu CY, Wang CH, Lin HC, Huang CD, Huang PY, Fang YF, Shieh MH, Kuo HP. The impact of delirium on the survival of mechanically ventilated patients. *Crit Care Med* 2004;32:2254–2259.
- Thomason JW, Shintani A, Peterson JF, Pun BT, Jackson JC, Ely EW. Intensive care unit delirium is an independent predictor of longer hospital stay: a prospective analysis of 261 non-ventilated patients. *Crit Care* 2005;9:R375–R381.
- McNicoll L, Pisani MA, Zhang Y, Ely EW, Siegel MD, Inouye SK. Delirium in the intensive care unit: occurrence and clinical course in older patients. *J Am Geriatr Soc* 2003;51:591–598.
- Ely EW, Shintani A, Truman B, Speroff T, Gordon SM, Harrell FE Jr., Inouye SK, Bernard GR, Dittus RS. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA* 2004;291:1753–1762.
- Ouimet S, Kavanagh BP, Gottfried SB, Skrobik Y. Incidence, risk factors and consequences of ICU delirium. *Intensive Care Med* 2007;33:66–73.
- Ely EW, Gautam S, Margolin R, Francis J, May L, Speroff T, Truman B, Dittus R, Bernard R, Inouye SK. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med* 2001;27:1892–1900.
- Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of post-discharge mortality, institutionalization, and dementia: a meta-analysis. *JAMA* 2010;304:443–451.
- Milbrandt EB, Deppen S, Harrison PL, Shintani AK, Speroff T, Stiles RA, Truman B, Bernard GR, Dittus RS, Ely EW. Costs associated with delirium in mechanically ventilated patients. *Crit Care Med* 2004;32:955–962.
- American Psychiatric Association. Practice guideline for the treatment of patients with delirium. *Am J Psychiatry* 1999;156:1–20.
- van Eijk MM, van Marum RJ, Klijn IA, de Wit N, Kesecioglu J, Slooter AJ. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med* 2009;37:1881–1885.
- Spronk PE, Riekerk B, Hofhuis J, Rommes JH. Occurrence of delirium is severely underestimated in the ICU during daily care. *Intensive Care Med* 2009;35:1276–1280.
- Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. *JAMA* 2010;304:779–786.
- Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, Speroff T, Gautam S, Bernard GR, Inouye SK. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med* 2001;29:1370–1379.
- Bergeron N, Dubois MJ, Dumont M, Dial S, Skrobik Y. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med* 2001;27:859–864.
- van Eijk MM, Kesecioglu J, Slooter AJ. Intensive care delirium monitoring and standardised treatment: a complete survey of Dutch intensive care units. *Intensive Crit Care Nurs* 2008;24:218–221.
- Jacobi J, Fraser GL, Coursin DB, Riker RR, Fontaine D, Wittbrodt ET, Chalfin DB, Masica MF, Bjerke HS, Coplin WM, et al. Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult. *Crit Care Med* 2002;30:119–141.
- Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, Truman B, Speroff T, Gautam S, Margolin R, et al. Delirium in mechanically ventilated patients: validity and reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *JAMA* 2001;286:2703–2710.
- Vreeswijk R, Honing M, Bakker K, Spronk P, De Man T, de Jonghe JF, Kalisvaart CJ. Translation, retranslation and validation of the Dutch Confusion Assessment Method for the Intensive Care Unit. *Crit Care* 2008;12:515.
- Toro AC, Escobar LM, Franco JG, az-Gomez JL, Munoz JF, Molina F, Bejarano J, Yepes D, Navarro E, Garcia A, et al. Spanish version of the method for evaluation of confusion in intensive cares, pilot study of validation. *Med Intensiva* 2010;34:14–21.
- Grapow MT, von WR, Guller U, Beyersdorf F, Zerkowski HR. Randomized controlled trials do not reflect reality: real-world analyses are critical for treatment guidelines! *J Thorac Cardiovasc Surg* 2006;132:5–7.
- van Eijk M, Slooter A. Validation of the CAM-ICU as performed by bed-side ICU nurses: a multicenter study. Presented at the European Delirium Association 5th Scientific Congress, November 11–12, 2010, Amsterdam, The Netherlands.
- van Eijk M, Slooter A. Validation of the CAM-ICU as performed by bed-side ICU nurses: a multicenter study. Presented at the Society of Critical Care Medicine 40th Critical Care Congress, January 15–19, 2011, San Diego, CA.
- van Eijk M, Slooter A. Validation of the CAM-ICU as performed by bed-side ICU nurses: a multicenter study. Presented at the 23rd European Society of Intensive Care Medicine Annual Congress, October 9–13, 2010, Barcelona, Spain.
- Pisani MA, Araujo KL, Van Ness PH, Zhang Y, Ely EW, Inouye SK. A research algorithm to improve detection of delirium in the intensive care unit. *Crit Care* 2006;10:R121.