Evaluation of delirium in critically ill patients: Validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)

E. Wesley Ely, MD, MPH; Richard Margolin, MD; Joseph Francis, MD, MPH; Lisa May, RN, BSN; Brenda Truman, RN, MSN; Robert Dittus, MD, MPH; Theodore Speroff, PhD, MPH; Shiva Gautam, PhD; Gordon R. Bernard, MD; Sharon K. Inouye MD, MPH

Objective: **To develop and validate an instrument for use in the intensive care unit to accurately diagnose delirium in critically ill patients who are often nonverbal because of mechanical ventilation.**

Design: **Prospective cohort study.**

Setting: **The adult medical and coronary intensive care units of a tertiary care, university-based medical center.**

Patients: **Thirty-eight patients admitted to the intensive care units.**

Measurements and Main Results: **We designed and tested a modified version of the Confusion Assessment Method for use in intensive care unit patients and called it the CAM-ICU. Daily ratings from intensive care unit admission to hospital discharge by two study nurses and an intensivist who used the CAM-ICU were compared against the reference standard, a delirium expert who used delirium criteria from the *Diagnostic and Statistical Manual of Mental Disorders* (fourth edition). A total of 293 daily, paired evaluations were completed, with reference standard diagnoses of delirium in 42% and coma in 27% of all observations. To include only interactive patient evaluations and avoid repeat-observer bias for patients studied on multiple days, we used only the first-alert or lethargic comparison evaluation in each patient. Thirty-three of 38 patients (87%) developed delirium during their intensive care unit stay, mean duration of 4.2 ± 1.7 days. Excluding evaluations of comatose patients because of lack of characteristic delirium features, the two critical care study nurses and intensivist demonstrated high interrater reliability for their CAM-ICU ratings with kappa statistics of 0.84, 0.79, and 0.95, respectively ($p < .001$). The two nurses' and intensivist's sensitivities when using the CAM-ICU compared with the reference standard were 95%, 96%, and 100%, respectively, whereas their specificities were 93%, 93%, and 89%, respectively.**

Conclusions: **The CAM-ICU demonstrated excellent reliability and validity when used by nurses and physicians to identify delirium in intensive care unit patients. The CAM-ICU may be a useful instrument for both clinical and research purposes to monitor delirium in this challenging patient population. (Crit Care Med 2001; 29:1370–1379)**

KEY WORDS: delirium; dementia; aging; geriatrics; cognitive impairment; mechanical ventilation; sedatives; analgesics; protocols; respiratory diseases; critical care; Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)

Delirium occurs in up to 60% of older hospitalized patients and is the most frequent hospital complication in these patients (1, 2). Delirium has been associated with poor hospital outcomes, including increased morbidity and mortality, prolonged length of stay, institutionalization, and functional decline (3–9). Patients in the intensive care unit (ICU) are at very high risk for the development of delirium because of factors such as multiple-system illnesses and comorbidities, the use of psychoactive medications, and age. Because of the variable definitions for delirium in the critical care literature, including the commonly used misnomers "intensive care syndrome" and "ICU psychosis" (10, 11), the true incidence of delirium in the ICU is unknown. In a recent investigation, the development of delirium was selected as one of the top three most important target areas for quality of care improvement in vulnerable older adults (12). However, publications focusing on outcomes of mechanical ventilation in the elderly have been limited by the lack of meaningful cognitive assessments in these patients (13–15). Almost all delirium investigations have excluded mechanically ventilated patients because of the communication challenges they pose for most clinicians (1, 2, 4, 16–18). There are currently no instruments available to measure delirium in nonverbal, intubated patients receiving mechanical ventilation.

The Confusion Assessment Method (CAM), developed by Inouye et al. (18), has helped to improve the assessment of delirium by nonpsychiatrists. Inouye and colleagues based the CAM on expert opinion and definitions of the American Psychiatric Association (published in the *Diagnostic and Statistical Manual of Mental Disorders* [DSM], third revised edition) to assist clinicians without formal psychiatric training (18). The CAM has been used over the past decade in numerous studies investigating delirium rates, risk factors, outcomes, and interventions (2, 16, 17). Other instruments have been validated (19–23), but combinations of these instruments do not yield

From the Departments of Internal Medicine, Divisions of General Internal Medicine and Center for Health Services Research (EWE, JF, LM, BT, RD, TS, SG, GRB), and Geriatric Psychiatry and Neuroscience (RM), Vanderbilt University School of Medicine, Nashville, TN; and the Department of Internal Medicine (SKI), Yale University School of Medicine, New Haven, CT.

Supported, in part, by an AFAR Pharmacology in Aging Grant and the Geriatric Research and Education Clinical Center (to EWE); and a Midcareer Award (K24AG00949) from the National Institute on Aging and a Donaghue Investigator Award (DF98-105) from the Patrick and Catherine Weldon Donaghue Medical Research Foundation (to SKI).

Address requests for reprints to: E. Wesley Ely, MD, MPH, Center for Health Services Research, 6th Floor Medical Center East, Vanderbilt University Medical Center, Nashville, TN 37232-8300; E-mail: wes.ely@mcmail.vanderbilt.edu

substantial improvements in performance (21, 24). In fact, the CAM has been compared with other instruments by external reviewers and found to have the best combination of ease, speed of use, data acquisition, reliability, and validity (24).

Use of delirium assessment tools has only recently been brought into the arena of the ICU. Marcantonio et al. (16) used the conventional CAM in an elective surgical ICU population to derive and validate a delirium clinical prediction rule, but their patients were not being mechanically ventilated during the assessments. Hart et al. (25, 26) studied the use of a visual memory and attention instrument in mechanically ventilated patients called the Cognitive Test for Delirium. Even in these investigations, however, heavily sedated, stuporous, or comatose patients were excluded. To better understand the demographics, causes, and prevention of delirium in the ICU, an objective and valid tool for verbal and nonverbal patients is necessary. The goal of this investigation was to develop and validate a brief, accurate, and reliable instrument for use by nurses and physicians to identify delirium in ICU patients.

METHODS

Patients. The study population included both ventilated and nonventilated adult medical ICU patients admitted to the Vanderbilt University Medical Center. The institutional review board approved this study, and informed consent was obtained from the patient or the surrogate (family member). To determine the sample size, we chose a target for the test sensitivity of 90% to ensure a statistically sound screening instrument. As an assumption for these calculations, the prevalence of delirium in the 2×2 table was set at 80% (a number consistent with pilot work in our ICU). It would require 11 patients to ensure that the lower range of the confidence interval (CI) for the CAM-ICU was 75% and 25 patients to ensure a lower range of 80% for the CI. Our enrollment exceeded both of these sample size estimates.

During the study period, 86 patients were admitted to the ICU, of whom 48 (56%) were enrolled. Enrollment criteria included any adult admitted to the medical ICU during the study interval from June to August 1999. Exclusion criteria defined *a priori* included a history of severe dementia, psychosis, or neurologic disease (as defined in the chart or by family interview) that would confound the diagnosis of delirium (n = 12); patient or family refusal to participate (n = 8); and admission to the ICU after the predefined cap of ten study patients per day had been reached because of research staffing limitations (n = 18). If the census approached ten on any given day, the patients were prioritized according to age, with the oldest patients enrolled first to test the population with greatest potential challenges to the CAM-ICU, such as hearing and/or visual deficits. In addition, five of the 48 patients enrolled remained comatose throughout the investigation and five others were never evaluated by the reference standard geriatric psychiatric specialist because of early death after enrollment and therefore were excluded from further analysis. Thus, the final sample size for the present study was 38 patients. As described in Statistical Analysis, we used the first alert or lethargic comparison evaluation for each patient. Comparison of the demographic variables of the enrolled (n = 38) vs. excluded (n = 48) patients showed no significant differences in age, gender, or race (all $p > .2$), but the excluded patients were less often receiving mechanical ventilation (28% vs. 58%, $p = .01$).

Validation Study Procedures. Two critical care study nurses enrolled patients each morning and managed the daily census for the investigation. Both nurses (LM and BT) and an intensivists (EWE and GRB) performed daily, independent CAM-ICU ratings during the patients' entire hospital stay. The reference standard DSM-IV evaluations (27) were performed independently by the delirium experts (JF or RM) as described subsequently. All cognitive assessments by the nurses, intensivists, and delirium experts were conducted independently in a blinded fashion in the afternoon between 3 and 7 pm. The rationale for allowing ≤4 hrs between assessments was to balance the intent of having the evaluations as close together as possible with the reality of clinicians' busy schedules and difficulties introduced by patient care (e.g., procedures). The mean time between the assessments conducted by the nurses and those of the intensivists and the reference standard delirium raters was 1.4 ± 0.8 hrs. None of the raters had access to the other's CAM-ICU or DSM ratings. From these ratings, reliability and validity measures were calculated as described in the Statistical Analysis section.

Development and Implementation of the CAM-ICU. The CAM (18) has four features, which are determined by the patient, nurse, and family interview. These features are as follows: 1) an acute onset of mental status changes or a fluctuating course; 2) inattention; 3) disorganized thinking; and 4) an altered level of consciousness (i.e., other than alert). The patient is diagnosed as delirious (i.e., CAM positive) if he or she has both features 1 and 2 and either feature 3 or 4 (Appendix 1). To use the CAM in nonverbal mechanically ventilated patients, we modified the approach by incorporating nonverbal, objective assessment instruments as described subsequently, and we called this delirium assessment the CAM-ICU.

> *The patient is diagnosed as delirious (i.e., CAM positive) if he or she has both features 1 and 2 and either feature 3 or 4 (Appendix 1).*

In this investigation, the completion of the CAM-ICU incorporated careful observations of the abilities of the patient and knowledge of the patient's former level of functioning. To operationalize the CAM-ICU, we defined a structured approach to completing the instrument that involved a bedside evaluation and screening for cognitive and attention deficits. Each CAM-ICU feature was approached as described in Appendix 1, and the evaluation was tailored to the individual patient to accommodate any reported or perceived auditory or visual deficits. The primary instrument used to assess inattention (feature 2) was the Attention Screening Examination (ASE). All patients were evaluated with a picture recognition tool as described by Hart et al. (25, 26) in the Cognitive Test for Delirium, and those with a history of visual impairment also received the Vigilance A random letter test (28). Both of these tools are described in Appendix 2. In completing the picture recognition component of the ASE, the patient was shown five simple pictures at 3-sec intervals and asked to remember them. The patient was then immediately shown ten subsequent pictures and asked to nod "yes" or "no" according to whether they had or had not just seen each of the pictures. Because five pictures had been shown to them already (correct nod = yes) and five others were new (correct nod = no), patients scored perfectly if they achieved ten correct yes or no nods. Rather than arbitrarily choose an absolute threshold below which inattention (feature 2) would be labeled as present, this feature was answered qualitatively based on the overall impression by the CAM-ICU rater after the entire patient evaluation was completed.

Reference Standard Evaluations. Independently, one of the delirium experts served as the reference standard for each patient and rated the patient as either normal, delirious, stuporous, or comatose by using the DSM-IV (27) criteria for delirium or standardized definitions for stupor and coma. Our two experts, one a geriatrician and delirium expert (JF) (4, 29–31) and the other a geriatric consult-liaison psychiatrist with >20 yrs of experience (RM), performed comparison patient evaluations independently—including interviews with family members and the patient's nurse

and chart review for lab data and nursing notes. To standardize the approach to DSM-IV ratings before the enrollment of patients, the delirium experts met with an attending neurologist and intensivist via roundtable discussions regarding their approach to ICU delirium assessment. They then met on four occasions to make ICU rounds and to discuss salient issues regarding their evaluations and to compare their ratings. To provide a comprehensive, reference standard assessment, the experts were permitted to administer any patient testing as well as use all sources of information (i.e., the patient's nurses and family members), but they were blinded to any of the CAM-ICU ratings of the nurses or intensivists. The reference standard also used DSM-IV criteria to render a diagnosis of suspected dementia whenever the criteria were found to be present during their comprehensive examinations (realizing that despite efforts at enrollment to exclude severely demented patients, mildly demented patients were admitted to the ICU).

Study Variables. At the time of enrollment, the following items were collected: demographics, severity of illness data by using the Acute Physiology and Chronic Health Evaluation II score (32), activities of daily living (33), and risk factors for delirium derived from three previous investigations (4, 16, 34). Although patients with a history of severe dementia were excluded, as mentioned previously, the Blessed Dementia Rating Scale (35) was used at enrollment to screen for previously unrecognized dementia via surrogate interviews. Patients were defined as having "suspected dementia" for the purposes of subgroup analyses if they met any of the following three criteria: the delirium expert rated them as demented, a Blessed Dementia Rating Scale score of ≥3, or a rating by the surrogate of ≥3 out of 5 as "possibly having dementia." Rather than the usual cutoff of 4, we defined "suspected dementia" as a score of ≥3 to increase sensitivity for the detection of mild dementia. The additional surrogate rating at the time of enrollment was an arbitrary one based on a 5-point scale as to whether the surrogate believed the patient was demented. At the time of hospital discharge, the patients completed a Folstein Mini-Mental State Examination (MMSE) (36), CAM rating, Geriatric Depression Scale (37), Short Form-12 (38), and Maugeri Respiratory Foundation (39) pulmonary disease specific quality of life instruments.

Statistical Analysis. To compare continuous variables between groups, the Student's *t*-test was used. To compare proportions and rates, chi-square tests were used when sample sizes were large and Fisher's exact tests when appropriate (40). The operating characteristics for the CAM-ICU were calculated by using standard definitions. Interrater reliability was determined by comparing the CAM-ICU ratings of nurse 1 vs. nurse 2 vs. the intensivist by using the kappa coefficient. Criterion validity was determined by comparing the three CAM-ICU raters to the delirium expert rating of cognitive status by using the DSM-IV criterion as the reference standard. To determine reliability and validity, we used the first-alert or lethargic comparison evaluation for each patient, as determined by the reference standard delirium expert, to include only interactive patient evaluations (i.e., avoiding comatose evaluations because they lack characteristic delirium features). This also allowed us to avoid repeat observer bias because patients had variable numbers of evaluations. The cells of the 2 × 2 tables comparing the reference standard ratings with the CAM-ICU ratings included the number of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). Sensitivity = TP/(TP + FN); specificity = TN/(FP + TN); accuracy = (TP + TN)/(TP + FP + FN + TN); and likelihood ratio = sensitivity/(1 − specificity). Ninety-five percent CIs for these test characteristics were calculated by using Stata software (College Station, TX). Subgroup analyses were designed and conducted for patients who would likely pose the greatest challenge in delirium assessment: those receiving mechanical ventilation, those >65 yrs old, and those with suspected dementia as defined in Study Variables. Statistical significance was defined as $p \leq .05$. Statistical analyses were performed by using the SAS software program version 8 (SAS Institute, Cary, NC).

RESULTS

Patient Characteristics. The reference standard delirium experts, two critical care nurses, and the intensivists completed evaluations in 38 patients over 293 patient evaluation days. Patient characteristics at the time of enrollment for the study population are presented in Table 1. The mean age of the study population was 60 ± 19 (SD), and 22 (58%) were mechanically ventilated on enrollment. The admission diagnoses capture the broad spectrum of these medical ICU patients. Although exclusion criteria included a documented history of overt dementia, 11 patients (29%) with suspected dementia were identified (see Study Variables) who served as an appropriate subgroup in whom to analyze the performance of the CAM-ICU.

Cognitive Reference Standard Findings. The delirium experts found that 69% of their 293 patient evaluations were abnormal during the investigation: 27% delirious, 15% stuporous, and 27% comatose. Of the 38 patients, 33 (87%) developed delirium at some point in their ICU stay, with average onset during the 2nd day (±1.7 days) and mean duration of delirium of 4.2 ± 1.7 days. At the time of hospital discharge, the MMSE examination was completed in 31 patients and was abnormal (<24 of possible 30 points) in 18 of the patients (58%). After we excluded two of these patients because they were among those patients with suspected dementia at baseline, 16 of 31 (52%) were cognitively impaired by the

Table 1. Patient characteristics at enrollment

Characteristics	Frequency (Total n = 38)
Age, mean ± SD	60 ± 19
Male, n (%)	23 (60)
Race	
White, %	84
African-American, %	14
Hispanic, %	2
APACHE II score, mean ± SD[a]	17.1 ± 8.7
Mechanical ventilation, n (%)	22 (58)
Visual or hearing deficit by patient or family report, %	71
Blessed Dementia Rating Scale, mean ± SD[b]	0.76 ± 1.26
ICU admission diagnosis, n (%)	
Acute respiratory distress syndrome	11 (29)
Myocardial infarction or arrhythmia	6 (16)
Congestive heart failure	6 (16)
Hepatic or renal failure	5 (13)
Chronic obstructive pulmonary disease	4 (11)
Gastrointestinal bleeding	3 (8)
Malignancy	2 (5)
Drug overdose	1 (2)

ICU, intensive care unit.

[a]APACHE II denotes Acute Physiology and Chronic Health Evaluation II score (32), an assessment of severity of illness; [b]the Blessed Dementia Rating Scale is an instrument to measure a patient's baseline likelihood of dementia using surrogate interviews, scores of ≥4 indicate likely dementia (35). For our subgroup analysis presented in Table 4 of patients with suspected dementia, we used an even lower score (≥3) to increase sensitivity.

MMSE examination at discharge. In addition, the CAM-ICU performed by the study nurses was found to be abnormal in ten patients (32%) within 36 hrs before hospital discharge.

Interrater Reliability of the CAM-ICU. One of these 38 patients was evaluated by only one of the study nurses, allowing 37 paired comparisons between nurses and 26 between the intensivists and nurses for interrater reliability. Of the 37 paired assessments between nurse 1 and nurse 2 by using the first-alert or lethargic evaluation for each patient as determined by the reference standard delirium expert (see statistical analysis in Methods), the CAM-ICU was completed with excellent reliability ($\kappa = 0.95$) (Table 2). The two nurses' CAM-ICU ratings compared to the intensivist's ratings also had substantial reliability ($\kappa = 0.84$ and 0.79, respectively).

Criterion Validity of the CAM-ICU. Patients were evaluated during their entire hospital stay and had a total of 293 paired observations. As in the reliability comparisons, we used the first-alert or lethargic evaluation for each patient to evaluate the test performance of the CAM-ICU in Table 3. The two critical care nurses' and the intensivist's sensitivities when using the CAM-ICU compared with the reference standard were 95%, 96%, and 100%, respectively, whereas their specificities were 93%, 93%, and 89%, respectively. The likelihood ratios for the CAM-ICU for the two nurses and intensivists were 14, 14, and 9, respectively.

Individual Components of the CAM-ICU. Patients were able to complete the ASE on 149 of 293 (51%) attempted evaluations, which yielded useful information regarding inattention and disorganized thinking (i.e., the ability to follow commands). All 38 patients were eventually able to complete the ASE but with varying degrees of overall correctness. By using the first-alert or lethargic evaluation for each patient, we have plotted the results of the picture recognition component of the ASE in differentiating the presence or absence of delirium in Figure 1. (Note: the Vigilance A random letter test was performed only selectively in visually impaired patients, and individual data for this component of the ASE are not presented.) This analysis of the ASE showed that at a cut point of eight or more correct answers, the sensitivity was 86% (95% CI, 64–97%), specificity was 77% (95% CI, 50–93%), and the likelihood ratio was 3.5 (95% CI, 1.8–7.1). The ASE significantly differentiated between the presence and absence of delirium ($p < .0001$).

Subgroup Analyses of the CAM-ICU. To evaluate the performance of the CAM-ICU in different patient groups that pose particular challenges in delirium assessment, we conducted three subgroup analyses including mechanically ventilated patients, those >65 yrs old, and those patients with suspected dementia as defined in Methods. In mechanically ventilated patients, the sensitivity of nurse 1, nurse 2, and the intensivist when using the CAM-ICU were 93%, 100%, and 100%, respectively; the specificities were 100%, 100%, and 88%, respectively (Table 4).

Table 2. Interrater reliability of CAM-ICU by using the first-alert or lethargic evaluation for each patient

Rater	κ (95% CI)	Percentage agreement (95% CI)
Intensivist vs. nurse 1 (n = 26)	0.84 (0.63–0.99)	97 (86–100)
Intensivist vs. nurse 2 (n = 26)	0.79 (0.64–0.95)	89 (70–98)
Nurse 1 vs. nurse 2 (n = 37)	0.95 (0.84–1.00)	93 (78–99)

CAM-ICU, Confusion Assessment Method for the Intensive Care Unit; CI, confidence interval.

$p < .0001$ for all comparisons. By using all paired observations (n = 293), the kappa ranged from 0.81 to 0.95. This table represents only one evaluation per patient (the first-alert or lethargic) to avoid possible repeat observer bias.

Table 3. Validity of the CAM-ICU for delirium by using the first-alert or lethargic evaluation for each patient

Rater	Sensitivity	Specificity	Accuracy	Likelihood Ratio
Nurse 1 (n = 37)	95 (77–100)	93 (68–100)	95 (86–100)	14 (3–56)
Nurse 2 (n = 38)	96 (78–100)	93 (68–100)	95 (86–100)	14 (3–39)
Intensivist (n = 26)	100 (80–100)	89 (51–100)	96 (80–100)	9 (2–99)

After excluding comatose evaluations, Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) comparisons were made to reference standard evaluation by a delirium expert who used criteria from the *Diagnostic and Statistical Manual for Mental Disorders* (fourth edition). Values are shown with 95% confidence intervals in parentheses.

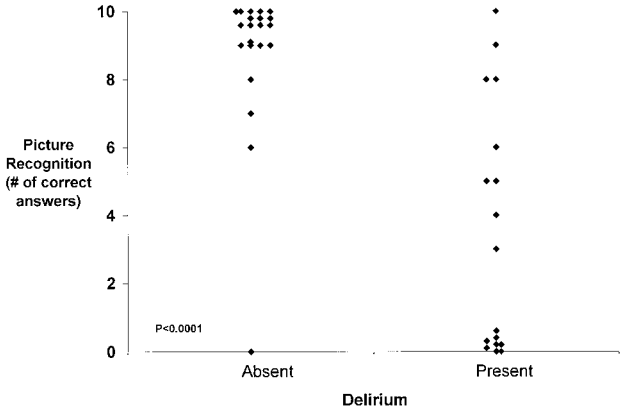


Fig. 1. Attention Screening Examination (ASE). The picture recognition component of the ASE, which is outlined in Methods and presented in Appendix 2, was used to help determine the presence or absence of inattention (Feature 2 of the Confusion Assessment Method for the Intensive Care Unit, or CAM-ICU). This scatter plot presents data from the first awake or lethargic observation of the patients on the y-axis and the presence or absence of delirium as rated by the reference standard psychiatric expert on the x-axis. By using a cutoff of eight or more correct answers (out of a possible ten), the instrument differentiated between the presence or absence of delirium as rated by the reference standard criteria (*Diagnostic and Statistical Manual of Mental Disorders*, fourth edition) with a sensitivity and specificity of 86% and 77%, respectively ($p < .0001$).

Table 4. Subgroup analyses of the CAM-ICU in ventilated, elderly, and suspected dementia patients

	Patients Mechanically Ventilated (n = 22)		Patients >65 Yrs Old (n = 18)		Patients With Suspected Dementia (n = 11)	
Rater	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Nurse 1	92 (66–99)	100 (63–100)	100 (69–100)	88 (47–99)	100 (56–100)	100 (56–100)
Nurse 2	100 (77–100)	88 (47–99)	90 (55–100)	100 (63–100)	100 (56–100)	100 (56–100)
Intensivist	100 (75–100)	88 (47–99)	100 (69–100)	83 (35–98)	100 (56–100)	100 (56–100)

After excluding comatose evaluations, Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) comparisons were made to reference standard evaluation by a delirium expert who used criteria from the *Diagnostic and Statistical Manual for Mental Disorders* (fourth edition). Values shown with 95% confidence intervals in parentheses. Interrater reliability measures across all comparisons showed kappa from 0.88 to 1.0 ($p < .0001$).

In patients >65 yrs old, the sensitivities for the CAM-ICU were 100%, 90%, and 100%, whereas the specificities were 88%, 100%, and 83%. Among the 11 patients with suspected dementia, sensitivity and specificity were 100% for all raters. For all of the comparisons of interrater reliability in these subgroups, the kappa statistic ranged from 0.88 to 1.0 ($p < .0001$).

Ease of Use. The CAM-ICU instrument could be completed in 2–3 mins and was easy for the study nurses and physicians to incorporate into their daily routine. As mentioned previously, the ASE could not be completed (i.e., the patient could not follow commands adequately to complete the evaluation) in half of the patient assessments because of severely impaired cognitive function. Once alert or lethargic, however, 89% of the patients could be evaluated by using the ASE, whereas 11% remained unable to follow commands necessary to complete an ASE.

DISCUSSION

There is currently no validated method by which critical care nurses and intensivists can monitor ICU patients for delirium while they are receiving mechanical ventilation. Yet it is important for these clinicians to be able to readily identify delirium, which is well documented to be associated with numerous adverse outcomes. Patients being supported by a ventilator receive numerous psychoactive medications for sedation and analgesia to increase their comfort and tolerance of treatment modalities during the ICU stay. These medications, along with patients' underlying comorbidities and advancing age, place ICU patients at extremely high risk for delirium, prolonged mechanical ventilation, and complicated hospital stays (41–44). Because of the difficulty of their assessment, however, mechanically ventilated patients have been excluded from nearly all delirium studies. The purpose of this investigation was to develop an instrument to be used by ICU personnel to assess delirium. We have modified the CAM for use in critically ill, nonverbal (i.e., intubated) patients and named it the CAM-ICU. We have shown the CAM-ICU to have high interrater reliability and a high sensitivity and specificity compared with the reference standard DSM-IV ratings for delirium.

This investigation was designed to help explore the feasibility of accurate assessments of delirium in perhaps the most difficult to assess population—those whose medical treatments limit their ability to communicate through mechanical ventilation, restraints, psychoactive medications, or support tubes such as orogastric feeding tubes. In the three subgroups that we thought would pose the greatest challenge to the CAM-ICU (i.e., intubated patients, those >65 yrs, and those with suspected dementia), the tool still had excellent interrater reliability, sensitivity, and specificity (Table 4).

The commonly discussed "ICU psychosis" or "ICU syndrome" refers to delirious patients who demonstrate increased psychomotor activity and hallucinations, so-called hyperactive or "loud" delirium (10, 11, 45, 46). The more common type of hypoactive or "quiet" delirium, manifested as decreased mental activity and inattention, frequently is overlooked by physicians and nurses yet may be associated with a worse prognosis (3, 29, 31, 47). Such an oversight is likely to occur in the ICU setting where more "urgent" issues occupy the health care professionals' attention. It recently was postulated that the term "ICU syndrome" is a potentially dangerous misnomer, because it implies that delirium is an expected outcome in ICU patients (45). Indeed, 87% of this cohort developed delirium during their ICU stay, but we believe that the severity or duration of this delirium potentially could be modified by using appropriately designed interventions (2). Monitoring patients for delirium with an easy to use tool could both increase nurses' and physicians' awareness of the occurrence of delirium and also allow credible interventions to be tested to reduce its incidence or duration.

Although the original CAM (18) is used commonly to detect and monitor delirium, mechanically ventilated patients have been excluded from CAM investigations to date (2, 17, 18). In the original CAM investigation, the internists administered the MMSE examination during the structured CAM interview. Because this was not possible in our mechanically ventilated patients, we developed alternative methods of evaluating attentiveness and disorganized thinking (see Appendixes 1–3). Our ASE had both visual (picture recognition) and auditory (Vigilance A random letter test) assessment components, which are promising components for future investigation. Tools such as the ASE may aid nurses in completing delirium assessments but at this stage should not be considered as stand-alone instruments in delirium assessment.

Several limitations to this investigation deserve comment. In developing the CAM-ICU, we sought to develop a tool to

We have shown that ICU nurses and physicians who use the CAM-ICU can detect delirium reliably and with a high degree of sensitivity and specificity within minutes.

detect delirium, not dementia (48). Although we excluded severely demented patients by their medical history and family interview, it is common for mildly demented patients to be cared for in the ICU setting. Because such patients could pose a challenge for the CAM-ICU, we included patients with suspected dementia in our subgroup analysis. The findings of 100% sensitivity and specificity in this group are based on a small number of patients and deserve further analysis on a larger scale. Another limitation of this study is related to the fluctuating nature of delirium, particularly as influenced by the frequent use of sedative and analgesic medications in this population, which could have influenced the ratings by the nurses, intensivists, and delirium experts because they evaluated the patients at different times (although within 1.5 hrs of one another on average). In fact, in reviewing the three misclassified CAM-ICU ratings by the nurses and intensivists (of which there were two FPs and one FN), it became apparent that there were two likely etiologies to explain these discordant ratings: a) a dose of sedative or analgesic drug being given between the CAM-ICU rating and the DSM-IV rating, or b) >3 hrs in between ratings. Both of these explanations indicate that future studies which use the CAM-ICU may need to employ frequent, serial measurement of delirium in the ICU.

Future studies of delirium in the ICU setting should use larger populations of patients both during and after mechanical ventilation. Another focus for future investigations would include more intensive neuropsychologic testing of patients after their ICU stay (i.e., beyond the MMSE examination) considering that one in three were still delirious within 36 hrs of hospital discharge. This could help to delineate the prevalence and predictors of chronic cognitive deficits in this vulnerable population as well as to better define the prognostic importance of delirium among ICU patients. To date, the impact of delirium among mechanically ventilated patients on clinical outcomes such as reintubation or nosocomial pneumonia remains poorly understood, but this could be defined in future investigations that enroll larger numbers of elderly and mechanically ventilated patients by using the CAM-ICU.

In conclusion, we have shown that ICU nurses and physicians who use the CAM-ICU can detect delirium reliably and with a high degree of sensitivity and specificity. This instrument is easy to use, takes only 2–3 mins to perform, and requires relatively little training. For the first time, it appears likely that clinicians who are not psychiatrically trained will be able to monitor critically ill patients for the development and persistence of delirium, both during and after mechanical ventilation. The ease of use of this instrument, combined with the lack of need for special equipment or expensive testing, make it ideally suited for use by ICU nursing staff. We hope that the CAM-ICU will allow clinicians to begin monitoring delirium in the ICU and eventually intervening to reduce the incidence and impact of this potentially devastating complication.

REFERENCES

1. Inouye SK: The dilemma of delirium: Clinical and research controversies regarding diagnosis and evaluation of delirium in hospitalized elderly medical patients. *Am J Med* 1994; 97:278–288
2. Inouye SK, Bogardus ST, Charpentier PA, et al: A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med* 1999; 340:669–676
3. Inouye SK, Schlesinger MJ, Lyndon TJ: Delirium: A symptom of how hospital care is failing older persons and a window to improve quality of hospital care. *Am J Med* 1999:565–573
4. Francis J, Martin D, Kapoor WN: A prospective study of delirium in hospitalized elderly. *JAMA* 1990; 263:1097–1101
5. Levkoff SE, Evans DA, Liptzin B, et al: Delirum: The occurrence and persistence of symptoms among elderly hospitalized patients. *Arch Intern Med* 1992; 152:334–340
6. Williams-Russo P, Urquhart BL, Sharrock NE, et al: Post-operative delirium: Predictors and prognosis in elderly orthopedic patients. *J Am Geriatr Soc* 1992; 40:759–767
7. Legault SE, Joffe RT, Armstrong PW: Psychiatric morbidity during the early phase of coronary care for myocardial infarction: Association with cardiac diagnosis and outcome. *Can J Psychiatry* 1992; 37:316–325
8. Inouye SK, Rushing JT, Foreman MD, et al: Does delirium contribute to poor hospital outcomes? A three-site epidemiologic study. *J Gen Intern Med* 1998; 13:234–242
9. Hebert PC, Drummond AJ, Singer J, et al: A simple multiple system organ failure scoring system predicts mortality of patients who have sepsis syndrome. *Chest* 1993; 104: 230–235
10. Granberg A, Engberg B, Lundberg D. Intensive care syndrome. *Intensive Crit Care Nurs* 1996; 12:173–182
11. Geary SM: Intensive care unit psychosis revisited: Understanding and managing delirium in the critical care setting. *Crit Care Nursing* 1994; 17:51–63
12. Sloss EM, Solomon DH, Shekelle PG, et al: Selecting target conditions for quality of care improvement in vulnerable older adults. *J Am Geriatr Soc* 2000; 48:363–369
13. Ely EW, Evans GW, Haponik EF: Mechanical ventilation in a cohort of elderly patients admitted to an intensive care unit. *Ann Intern Med* 1999; 131:96–104
14. Chelluri L, Pinsky MR, Grenvik AN: Outcome of intensive care of the "oldest-old" critically ill patients. *Crit Care Med* 1992; 20:757–761
15. Ely EW, Wheeler A, Thompson T, et al: Age is an independent predictor of outcomes in ARDS. *Am J Respir Crit Care Med* 2000; 161:A210
16. Marcantonio ER, Goldman L, Mangione CM, et al: A clinical prediction rule for delirium after elective noncardiac surgery. *JAMA* 1994; 271:134–139
17. Marcantonio ER, Juarez G, Goldman L, et al: The relationship of postoperative delirium with psychoactive medications. *JAMA* 1994: 1518–1522
18. Inouye SK, van Dyck CH, Alessi CA, et al: Clarifying confusion: The confusion assessment method. *Ann Intern Med* 1990; 113: 941–948
19. Trzepacz PT, Baker RW, Greenhouse J: A symptom rating scale for delirium. *Psychiatry Res* 1987; 23:89–97
20. Trzepacz PT, Dew MA: Further analyses of the delirium rating scale. *Gen Hosp Psychiatry* 1995; 17:75–79
21. Pompei P, Foreman M, Cassel CK, et al: Detecting delirium among hospitalized older patients. *Arch Intern Med* 1995; 155:301–307
22. Foreman MD: A comparison of the reliability and validity of mental status questionnaires in hospitalized patients. *Nurs Res* 1987; 36: 216–220
23. Vermeersch PEH: The clinical assessment of confusion. *Appl Nurs Res* 1990; 3:128–133
24. Smith MJ, Breitbart WS, Meredith MP: A critique of instruments and methods to detect, diagnose, and rate delirium. *J Pain Symptom Manage* 1995; 10:35–77
25. Hart RP, Levenson JL, Sessler CN, et al:

Validation of a cognitive test for delirium in medical ICU patients. *Psychosomatics* 1996; 37:533–546
26. Hart RP, Best AM, Sessler CN, et al: Abbreviated cognitive test for delirium. *J Psychosom Res* 1997; 43:417–423
27. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fourth Edition. Washington, DC, American Psychiatric Association, 1994
28. Strub RL, Black FW: The mental status examination in neurology. Third Edition. Philadelphia, PA, FA Davis, 1993
29. Francis J, Kapoor WN: Delirium in hospitalized elderly. *J Gen Intern Med* 1990:65–79
30. Francis J, Kapoor WN: Prognosis after hospital discharge of older medical patients with delirium. *J Am Geriatr Soc* 1992:601–606
31. Francis J. Delirium in older patients. *J Am Geriatr Soc* 1992; 40:829–838.
32. Knaus WA, Draper EA, Wagner DP, et al: APACHE II: A severity of disease classification system. *Crit Care Med* 1985; 13: 818–829
33. Katz S, Ford AB, Moskowitz RW, et al: Studies of illness in the aged—the index of ADL: A standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 94–99
34. Inouye SK, Viscoli C, Horwitz RI, et al: A predictive model for delirium in hospitalized elderly medical patients based on admission characteristics. *Ann Intern Med* 1993: 474–481
35. Blessed G, Tomlinson BE, Roth M: The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiat* 1968:797–811
36. Folstein MF, Folstein SE, McHugh PR: Mini-mental state a practical method for grading the cognitive state of patients for the clinician. *J Psychiat Res* 1975; 12:189–198
37. Lyness JM, Noel TK, Cox C, et al: Screening for depression in elderly primary care patients. *Arch Intern Med* 1997; 157:449–454
38. Gandek B, Ware JE, Aaronson NK, et al: Cross-validation of item selection and scoring for the SF-12 health survey in nine countries: Results from the international quality of life assessment project. *J Clin Epidemiol* 1998; 51:1171–1178
39. Carone M, Bertolotti G, Anchisi F, et al: Analysis of factors that characterise health impairment in patients with chronic respiratory failure. *Eur Respir J* 1999; 13:1293–1300
40. Fleiss JL: Statistical Methods for Rates and Proportions. Second Edition. New York, John Wiley & Sons, 1981
41. Kress JP, Pohlman AS, O'Connor MF, et al: Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. *N Engl J Med* 2000; 342: 1471–1477
42. Heffner JE: A wake-up call in the intensive care unit. *N Engl J Med* 2000; 342: 1520–1522
43. Kollef MH, Levy NT, Ahrens T, et al: The use of continuous iv sedation is associated with prolongation of mechanical ventilation. *Chest* 1999; 114:541–548
44. Brook AD, Ahrens TS, Schaiff R, et al: Effect of a nursing implemented sedation protocol on the duration of mechanical ventilation. *Crit Care Med* 1999; 27:2609–2615
45. McGuire BE, Basten CJ, Ryan CJ, et al: Intensive care unit syndrome: A dangerous misnomer. *Arch Intern Med* 2000; 160: 906–909
46. Wilson LM: Intensive care delirium: The effect of outside deprivation in a windowless unit. *Arch Intern Med* 1972; 130:22–23
47. Meagher DJ, Hanlon DO, Mahony EO, et al: Relationship between symptoms and motoric subtype of delirium. *J Neuropsychiatry Clin Neurosci* 2000; 12:51–56
48. Trzepacz PT, Mulsant BH, Dew MA, et al: Is delirium different when it occurs in dementia? A study using the delirium rating scale. *J Neuropsychiatry* 1998; 10:199–204

Appendix 1. The Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)

Delirium is diagnosed when both Features 1 and 2 are positive, along with either Feature 3 or Feature 4.

Feature 1. Acute Onset of Mental Status Changes or Fluctuating Course

- Is there evidence of an acute change in mental status from the baseline?
- Did the (abnormal) behavior fluctuate during the past 24 hrs, that is, tend to come and go or increase and decrease in severity?

Sources of information: Serial Glasgow Coma Scale or sedation score ratings over 24 hrs as well as readily available input from the patient's bedside critical care nurse or family.

Feature 2. Inattention

- Did the patient have difficulty focusing attention?
- Is there a reduced ability to maintain and shift attention?

Sources of information: Attention screening examinations by using either picture recognition or Vigilance A random letter test (see Methods and Appendix 2 for description of Attention Screening Examinations). Neither of these tests requires verbal response, and thus they are ideally suited for mechanically ventilated patients.

Feature 3. Disorganized Thinking

- Was the patient's thinking disorganized or incoherent, such as rambling or irrelevant conversation, unclear or illogical flow of ideas, or unpredictable switching from subject to subject?
- Was the patient able to follow questions and commands throughout the assessment?
  1. "Are you having any unclear thinking?"
  2. "Hold up this many fingers." (examiner holds two fingers in front of the patient)
  3. "Now, do the same thing with the other hand." (not repeating the number of fingers)

Feature 4. Altered Level of Consciousness

- Any level of consciousness other than "alert."
- Alert—normal, spontaneously fully aware of environment and interacts appropriately
- Vigilant—hyperalert
- Lethargic—drowsy but easily aroused, unaware of some elements in the environment, or not spontaneously interacting appropriately with the interviewer; becomes fully aware and appropriately interactive when prodded minimally
- Stupor—difficult to arouse, unaware of some or all elements in the environment, or not spontaneously interacting with the interviewer; becomes incompletely aware and inappropriately interactive when prodded strongly
- Coma—unarousable, unaware of all elements in the environment, with no spontaneous interaction or awareness of the interviewer, so that the interview is difficult or impossible even with maximal prodding

Appendix 2. The Attention Screening Examinations (ASE) for the Intensive Care Unit
(A) Picture Recognition ASE
(B) Vigilance A Random Letter Test

(A) Picture Recognition ASE
Step 1: Five pictures
Say to the patient: "Mr. or Mrs. __X__, I am going to show you pictures of some common objects. Watch carefully and try to remember each picture because I will ask what pictures you have seen." Then show step 1 of either form A or form B, alternating daily if repeat measures are taken.
Step 2: Ten pictures
Say to the patient: "Now I am going to show you some more pictures. Some of these you have already seen and some are new. Let me know whether or not you saw the picture before by nodding your head yes (demonstrate) or no (demonstrate)." Then show step 2 of form A or B, depending on which form was used in step 1.

Form A (step 1)
- key
- cup
- car
- table
- hammer

Form B (step 1)
- boot
- dog
- knife
- pants
- paint brush

Form A (step 2)
- key
- cup
- car
- table
- hammer
- glass
- lock
- truck
- chair
- saw

Form B (step 2)
- boot
- dog
- knife
- pants
- paint brush
- fork
- cat
- dress
- toothbrush
- shoe

This test is scored by the number of correct "yes" or "no" answers (out of a possible ten) during the second step. To improve the visibility for elderly patients, the images are printed on 6 × 10-inch buff-colored paper and laminated with a flat finish. As did Hart et al. (25, 26) we showed each image for 3 secs. When a patient had known visual impairment and no corrective lenses, we substituted the Vigilance A Random Letter Test (25, 26).

(B) Vigilance A Random Letter Test
Directions: Tell the patient: "I am going to read you a long series of letters. Whenever you hear the letter *A*, indicate by squeezing my hand." Read the following letter list in a normal tone at a rate of one letter per second.

L T P E A O A I C T D A L A A
A N I A B F S A M R Z E O A D
P A K L A U C J T O E A B A A
Z Y F M U S A H E V A A R A T

Scoring: Currently, only preliminary standardized norms exist for this test. The average person should complete the task without error. ($x = 0.2$); a sample of randomly selected brain-damaged patients made an average of ten errors. Examples of common organic errors are a) failure to indicate when the target letter has been presented (omission error); b) indication made when a nontarget letter has been presented (commission error); and c) failure to stop tapping with the presentation of subsequent nontarget letters (perseveration error) (28).

Appendix 2A
The Attention Screening Examination for the ICU – Picture Recognition ASE

- This is a test of attention, the ability of the patient to concentrate and demonstrate short term memory.
- This test, created by Hart and colleagues, was found to be a reliable and valid instrument for detecting delirium among mechanically ventilated patients.

Hart et al, Psychosomatics 1996; 37:533-546
Hart et al, J Psychosom Res 1997; 43:417-423

1st Step: 5 pictures

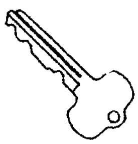
- Say to the patient: "Mr or Mrs. ___X____, I am going to show you pictures of some common objects. Watch carefully and try to remember each picture because I will ask what pictures you have seen."
- Then show Step 1 of Form A or Form B.
- Alternate daily between Form 1 and Form 2.

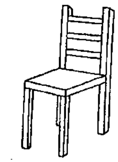
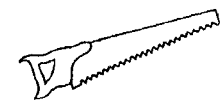
2nd Step: 10 pictures

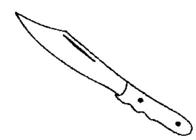
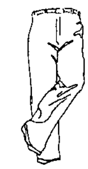
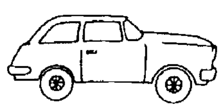
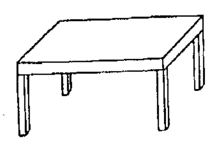
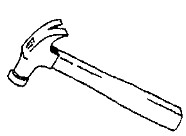
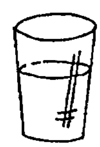
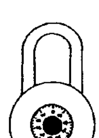
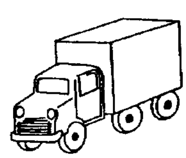
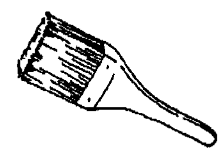
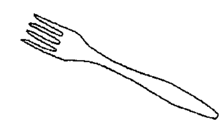
- Say to the patient: "Now I am going to show you some more pictures. Some of these you have already seen and some are new. Let me know whether or not you saw the picture before by nodding your head 'yes' (demonstrate) or 'no' (demonstrate)."
- Then show Step 2 of Form A or Form B, whichever form was used in Step 1.

Scoring the ASE

- This test is scored by the number of correct "yes" or "no" answers out of a possible 10 during the second step using Forms A or B.
- The pictures are listed in the following slides. In the ICU, they are shown on large, laminated, buff colored paper for ease of viewing by the elderly or visually impaired.





Appendix 3. Criteria Used by Delirium Experts

Criteria for Delirium from Diagnostic and Statistical Manual for Mental Disorders (fourth edition) (27)—Reference standard evaluations were performed by using all available information including patient examinations and interactions, nurse and family interviews, physicians' and nurses' notes, laboratory values, and any other chart data present.

A. Disturbance of consciousness (i.e., reduced clarity of awareness of the environment) with reduced ability to focus, sustain, or shift attention
B. A change in cognition (such as memory deficit, disorientation, language disturbance) or the development of a perceptual disturbance that is not better accounted for by a preexisting, established, or evolving dementia
C. Disturbance that develops over a short period of time (usually hours to days) and tends to fluctuate during the course of the day
D. Evidence from the history, physical examination, or laboratory findings that the disturbance is caused by one of the following:
   i. Direct physiological consequences of a general medical condition
   ii. Direct result of medication use or substance intoxication (substance intoxication delirium)
   iii. Direct result of a withdrawal syndrome (substance withdrawal delirium)
   iv. Direct result of more than one of the preceding etiologies (delirium due to multiple etiologies)

The diagnosis of cognitive impairment involved careful observations of the abilities of the patient and knowledge of the patient's former level of functioning. To identify all cases of cognitive impairment, we adopted the following measures: a) The preceding criteria from the *Diagnostic and Statistical Manual for Mental Disorders* and mental status definitions were employed consistently. b) A delirium expert evaluation was conducted to determine which of these criteria were met by the patient. This involved a bedside evaluation and screening for cognitive and attention deficits. c) Last, interviewing the family and nurse who provided the majority of patient care established baseline functioning and identified fluctuations (27).