Delirium in Mechanically Ventilated Patients
Validity and Reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)

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

Delirium is a common problem in the intensive care unit (ICU). Accurate diagnosis is limited by the difficulty of communicating with mechanically ventilated patients and by lack of a validated delirium instrument for use in the ICU.

Objectives To validate a delirium assessment instrument that uses standardized non-verbal assessments for mechanically ventilated patients and to determine the occurrence rate of delirium in such patients.

Design and Setting Prospective cohort study testing the Confusion Assessment Method for ICU Patients (CAM-ICU) in the adult medical and coronary ICUs of a US university-based medical center.

Participants A total of 111 consecutive patients who were mechanically ventilated were enrolled from February 1, 2000, to July 15, 2000, of whom 96 (86.5%) were evaluable for the development of delirium and 15 (13.5%) were excluded because they remained comatose throughout the investigation.

Main Outcome Measures Occurrence rate of delirium and sensitivity, specificity, and interrater reliability of delirium assessments using the CAM-ICU, made daily by 2 critical care study nurses, compared with assessments by delirium experts using Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria.

Results A total of 471 daily paired evaluations were completed. Compared with the reference standard for diagnosing delirium, 2 study nurses using the CAM-ICU had sensitivities of 100% and 93%, specificities of 98% and 100%, and high interrater reliability (κ = 0.96; 95% confidence interval, 0.92-0.99). Interrater reliability measures across subgroup comparisons showed κ values of 0.92 for those aged 65 years or older, 0.99 for those with suspected dementia, or 0.94 for those with Acute Physiology and Chronic Health Evaluation II scores at or above the median value of 23 (all P < .001). Comparing sensitivity and specificity between patient subgroups according to age, suspected dementia, or severity of illness showed no significant differences. The mean (SD) CAM-ICU administration time was 2 (1) minutes. Reference standard diagnoses of delirium, stupor, and coma occurred in 25.2%, 21.3%, and 28.5% of all observations, respectively. Delirium occurred in 80 (83.3%) patients during their ICU stay for a mean (SD) of 2.4 (1.6) days. Delirium was even present in 39.5% of alert or easily aroused patient observations by the reference standard and persisted in 10.4% of patients at hospital discharge.

Conclusions Delirium, a complication not currently monitored in the ICU setting, is extremely common in mechanically ventilated patients. The CAM-ICU appears to be rapid, valid, and reliable for diagnosing delirium in the ICU setting and may be a useful instrument for both clinical and research purposes.

©2001 American Medical Association. All rights reserved.
The diagnosis of delirium requires the presence of acute onset of changes or fluctuations in the course of mental status, and inattention, and either disorganized thinking or an altered level of consciousness.

Delirium, as defined by the CAM,31 has 4 features: (1) an acute onset of changes or fluctuations in the course of mental status, (2) inattention, (3) disorganized thinking, and (4) an altered level of consciousness (ie, other than alert). The patient is determined to be delirious (ie, CAM positive) if he/she manifests both features 1 and 2, plus either feature 3 or 4 (FIGURE). The development of the CAM-ICU began by adapting the patient assessment (without modifying the actual CAM instrument) using nonverbal, objective tests derived through a comprehensive literature review and consultation with numerous delirium experts. This resulted in a pilot investigation that yielded promising results.34

Figure. Flow Diagram of Confusion Assessment Method for the ICU (CAM-ICU)

Delirium

Feature 4
Altered Level of Consciousness

AND EITHER

Feature 3
Disorganized Thinking

AND EITHER

Feature 2
Inattention

AND

Feature 1
Acute Onset of Changes or Fluctuations in the Course of Mental Status

Phases of Study Design

Development and Validation

Delirium, as defined by the CAM,31 has 4 features: (1) an acute onset of changes or fluctuations in the course of mental status, (2) inattention, (3) disorganized thinking, and (4) an altered level of consciousness (ie, other than alert). The patient is determined to be delirious (ie, CAM positive) if he/she manifests both features 1 and 2, plus either feature 3 or 4 (FIGURE). The development of the CAM-ICU began by adapting the patient assessment (without modifying the actual CAM instrument) using nonverbal, objective tests derived through a comprehensive literature review and consultation with numerous delirium experts. This resulted in a pilot investigation that yielded promising results.34

Methods

Patients

The study population, none of whom have been included in previous investigations, included mechanically ventilated adult medical and coronary ICU patients admitted to Vanderbilt University’s 641-bed academic medical center with 96 adult ICU beds. The institutional review board approved this study, and informed consent was obtained from the patient or surrogate. The study interval was from February 1, 2000, to July 15, 2000. During this time, 158 mechanically ventilated patients were admitted to the ICU, of whom 111 (70.3%) were enrolled and 47 (29.7%) met exclusion criteria. A priori exclusion criteria included a history of psychosis or neurologic disease (eg, cerebrovascular accident) that would confound the diagnosis of delirium (n = 16), inability to communicate with assessors (ie, did not speak or understand English or was deaf) (n = 5), admitted to the ICU but extubated before study nurses’ assessments (n = 10), previously enrolled in the study (n = 5), patient or family refusal to participate (n = 6), or died before study nurses’ assessments (n = 5). In addition, 15 (13.5%) of the 111 enrolled patients remained comatose throughout the investigation and were excluded from further analysis. The final sample size was 96.

Structure of Validation

Study Procedures

Two critical care study nurses enrolled patients and performed daily, independent CAM-ICU ratings during the patients’ ICU stay. The reference standard evaluations were performed independently by the delirium experts (described below), who applied DSM-IV criteria for delirium. The delirium experts demonstrated excellent reliability (κ = 0.97) using standardized methods to complete the DSM-IV ratings during prestudy training period. All cognitive assessments by the nurses and delirium experts were conducted independently in a blinded fashion in the ICU between 1 and 5 PM. None of the raters had access to any of the other’s evaluations or ratings. From these ratings, reliability and validity measures were calculated as described below.
Based on the pilot study, the CAM-ICU was refined to improve face validity and further standardize the evaluation, as described below.

**Standardized Assessment to Rate the CAM-ICU**

The CAM-ICU was completed by the study nurses as described in Table I. To aid in the assessment of acute onset or fluctuation of mental status changes, patients were followed up daily with the Glasgow Coma Scale and an agitation/sedation scale called the Richmond Agitation Sedation Scale (RASS) (See Table 1 footnotes). Any fluctuation during the prior 24-hour period in the patient’s Glasgow Coma Scale or RASS scores was used by the study nurse to indicate a positive mental status, and as such these scales became a standardized part of the rating of the CAM-ICU by the nurses.

To aid the assessment of inattention, we used an Attention Screening Examination (ASE), which included a visual recognition component and an auditory recognition component. The ASE visual was validated previously by Hart et al in the Cognitive Test for Delirium and the ASE auditory was validated as the Vigilance A random letter test. Both of these tools are presented in the pilot publication of the CAM-ICU, but during refinement it was decided to administer both visual and auditory components of the ASE and limit the number of letters in the sequence of auditory testing to only 10 letters. Data from the pilot study were used to establish the cut point for presence of inattention, which was rated as present if the patient scored less than 8 correct answers on either the visual or auditory components of the ASE.

Assessment of disorganized thinking was refined after the pilot study by adding the standardized questions and commands included in Table I. Disorganized thought was said to be present if the patient was unable to answer at least 3 of the 4 questions correctly and could not complete commands such as those described in Table I. An altered level of consciousness, was rated as being vigilant or lethargic or in a stupor or coma using the standardized definitions in Table 1.

**Reference Standard Evaluations**

Reference standard evaluations were performed by a geriatrician delirium expert (J.F.), a board certified geriatric consult-liaison psychiatrist (R.M.), or a neuropsychologist (S.G.). The delirium experts, basing evaluation on interviews with family members, the patient’s nurse, and chart review for lab data

---

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

<table>
<thead>
<tr>
<th>Features and Descriptions</th>
<th>Absent</th>
<th>Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Acute onset or fluctuating course*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Is there evidence of an acute change in mental status from the baseline?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Or, did the (abnormal) behavior fluctuate during the past 24 hours, that is, tend to come and go or increase and decrease in severity as evidenced by fluctuations on the Richmond Agitation Sedation Scale (RASS) or the Glasgow Coma Scale?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Inattention†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the patient have difficulty focusing attention as evidenced by a score of less than 8 correct answers on either the visual or auditory components of the Attention Screening Examination (ASE)?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Disorganized thinking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there evidence of disorganized or incoherent thinking as evidenced by incorrect answers to 3 or more of the 4 questions and inability to follow the commands?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Will a stone float on water?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Are there fish in the sea?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Does 1 pound weigh more than 2 pounds?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Can you use a hammer to pound a nail?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Are you having unclear thinking?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Hold up this many fingers. (Examiner holds 2 fingers in front of the patient.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Now do the same thing with the other hand (without holding the 2 fingers in front of the patient).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(If the patient is already extubated from the ventilator, determine whether the patient’s thinking is disorganized or incoherent, such as rambling or irrelevant conversation, unclear or illogical flow of ideas, or unpredictable switching from subject to subject.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV. Altered level of consciousness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the patient’s level of consciousness anytbing other than alert, such as being vigilant or lethargic or in a stupor, or coma?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alert: spontaneously fully aware of environment and interacts appropriately</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vigilant: hyperalert</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lethargic: drowsy but easily aroused, unaware of some elements in the environment or not spontaneously interacting with the interviewer; becomes fully aware and appropriate interactive when prodded minimally</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stupor: difficult to arouse, unaware of some or all elements in the environment or not spontaneously interacting with the interviewer; becomes incompletely aware when prodded strongly; can be aroused only by vigorous and repeated stimuli and as soon as the stimuli ceases, stuporous subject lapses back into unresponsive state</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coma: unarousable, unaware of all elements in the environment with no spontaneous interaction or awareness of the interviewer so that the interview is impossible even with maximal prodding</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Overall CAM-ICU Assessment (Features 1 and 2 and either Feature 3 or 4): Yes__ No__**

*The scores included in the 10-point RASS range from a high of 4 (comatose) to a low of –5 (deeply comatose and unresponsive). Under the RASS system, patients who were spontaneously alert, calm, and not agitated were scored at 0 (neutral zone). Anxious or agitated patients received a range of scores depending on their level of anxiety: 1 for anxious, 2 for agitated (fighting ventilator), 3 for very agitated (pulling on or removing catheters), or 4 for combative (violent and a danger to staff). The scores –1 to –5 were assigned for patients with varying degrees of sedation based on their ability to maintain eye contact: –1 for more than 10 seconds, –2 for less than 10 seconds, and –3 for eye opening but no eye contact. If physical stimulation was required, then the patient was scored as either –4 for eye opening or movement with physical or painful stimulation or –5 for no response to physical or painful stimulation. The RASS has excellent interrater reliability and intraclass correlation coefficients of 0.95 and 0.97, respectively, and has been validated against visual analog scale and geropsychiatric diagnoses in 2 ICU studies. Those completing the visual ASE, the patients were shown 5 simple pictures (previously published) at 3-second intervals and asked to remember them. They were then immediately shown 10 subsequent pictures and asked to nod “yes” or “no” to indicate whether they had or had not just seen each of the pictures. Since 5 pictures had been shown to them already, for which the correct response was to nod “yes,” and 5 others were new, for which the correct response was to shake their heads “no,” patients scored perfectly if they achieved 10 correct responses. Scoring accounted for either errors of omission (indicating “no” for a previously shown picture) or for errors of commission (indicating “yes” for a picture not previously shown). In completing the auditory ASE, patients were asked to squeeze the rater’s hand whenever they heard the letter A during the recitation of a series of 10 letters. The rater then read 10 letters from the following list in a normal tone at a rate of 1 letter per second: S, A, H, E, V, A, A, R, A, T. A scoring method similar to that of the visual ASE was used for the auditory ASE testing. This table may be reproduced without permission for clinical use only (Ely EW et al. JAMA. 2001;286:2707-2710).
and nursing notes, standardized their approach to DSM-IV delirium ratings during the pilot study.18 Patients were rated as either normal, delirious, stuporous, or comatose using the DSM-IV criteria for delirium or standardized definitions for stupor and coma.35 These experts also rated the patient for the presence or absence of dementia using standard DSM-IV criteria. To increase our sensitivity for detection of dementia, the experts could rate a patient as demented if at any time during his/her course it became clear (eg, through interviews with the patient, family members, the patient’s nurse, or chart review) that the patient’s baseline cognitive state met DSM-IV criteria. For cases in which the experts rated their own confidence level of DSM-IV ratings for either dementia or delirium as low or intermediate rather than high, 2 raters would form a consensus opinion that day about the rating in question.

**Criterion Validity and Interrater Reliability**

Criterion validity was determined by comparing the 2 CAM-ICU raters to the delirium expert rating of delirium using the DSM-IV criterion as the reference standard. Interrater reliability was determined by comparing the CAM-ICU ratings of nurse 1 with those of nurse 2 using the $\kappa$ coefficient. To determine validity and reliability, we used the first alert or lethargic evaluation of each patient (as rated by the reference standard delirium expert) for the comparison evaluation. This allowed us to avoid repeat observer bias since patients had variable numbers of evaluations. For these evaluations the patient had to be aroused with verbal stimulation, demonstrate eye contact, and follow some commands.38 The a priori decision to include this level of consciousness as the first comparison for validation and reliability testing was made to include only interactive patient evaluations, avoiding comatose and stuporous evaluations because they lack characteristic delirium features and their relative ease of rating might falsely elevate the test characteristics of the CAM-ICU.

**Subgroup Analyses**

Subgroup analyses were conducted for patients who would likely pose the greatest challenge in delirium assessment: those 65 years or older, those with possible dementia, and those with more severe illness (an Acute Physiology and Chronic Health Evaluation II [APACHE II] score above the median). The first 2 subgroups (older age and possible dementia) were prospectively determined while the subgroup based on higher level of illness severity was not.

**Ease of Use**

Some training is required for optimal use of the CAM-ICU. As an indicator of ease of use, we measured time for completion of the test, and completion rates in patients with and without delirium.

**Sample Size Calculations**

In the pilot study of the CAM-ICU,33 the test sensitivity of mechanically ventilated patients was 95% averaging across raters and test specificity was 88%. A group of 12 intensivists agreed a priori that the instrument’s sensitivity was the critical feature so as not to miss the diagnosis, stating that it would be important to confirm the lower limit of the 95% confidence interval (CI) to be 85% or higher while an acceptable specificity would be 75% or higher. Sample size was calculated to ensure the appropriate number of patients necessary to achieve the expected lower limit of the 95% CI for the CAM-ICU test sensitivity and specificity in alert or lethargic patients. Assuming an incidence of delirium of 50%, it would require 50 mechanically ventilated patients to ensure that the lower range of the CI for sensitivity in the entire study population was 85% and another 45 patients to ensure that specificity was at least 75%. Thus, the planned total sample size for the investigation was 95 patients. The study was not powered for subgroup analyses.

**Study Variables**

Patient demographics and the severity of illness measurement using the APACHE II27 were collected at time of enrolment.17,19 Baseline visual or auditory deficits were recorded if patients wore corrective lenses (glasses, bifocals, or contacts) or had a hearing aid, as well as if the family reported that the patient had any documented impairment in vision or hearing. The modified Blessed Dementia Rating scale (mBDRS),60 originally validated against brain pathological specimens, was used to screen for dementia using family interviews (an additional surrogate question asked them to rate on a 5-point scale whether they believed the patient was demented). Because baseline dementia could serve as a confounder in rating the CAM-ICU, we chose to increase our sensitivity for detecting dementia by defining patients as having suspected dementia if they met any of the following 3 criteria: (1) delirium expert rated them as having dementia, (2) modified Blessed Dementia Rating scale score of at least 3, or (3) rating by the surrogate of at least 3 out of 5 as possibly having dementia. At the time of hospital discharge, the patients completed a Folstein Mini-Mental State Examination (MMSE)47 and CAM-ICU rating. The MMSE is a standard method of screening for dementia or related cognitive impairment with a scale ranging from 0 to 30 points. Scores below 24 indicate cognitive impairment.

**Statistical Analysis**

To compare demographic variables and other baseline characteristics between enrolled and excluded patients, the $t$ test was used. To compare proportions and rates, $\chi^2$ tests were used when sample sizes were large and Fisher exact tests when appropriate.48 The performance test characteristics for the CAM-ICU were calculated using standard definitions: sensitivity, specificity, predictive values, overall accuracy (true positives + true negatives/true positives + false positives + true negatives), and likelihood ratios (sensitivity/1–specificity) were estimated from simple $2 \times 2$ tables. Exact 95% CIs for these test characteristics were calculated using SAS software version 8.02 (SAS Institute Inc, Cary, NC). The CIs for the likelihood ratios were calculated using theboot-
strapping technique using STATA 7.0 (STATA Corp, College Station, Tex). All statistical tests were 2 sided, and a P value of <.05 was considered statistically significant.

RESULTS

Patient Characteristics

The reference standard delirium experts and 2 critical care study nurses completed evaluations in 111 patients. Comparison of the demographic variables of the 111 enrolled patients with the 47 patients excluded based on criteria designated a priori showed no significant differences in age, sex, race, severity of illness, or other baseline characteristics (all P>.25). After excluding 15 enrolled patients because they remained comatose throughout the investigation, 96 (86.5%) of 111 enrolled patients remained and comprised the final study population. Characteristics of these 96 patients at the time of enrollment are presented in TABLE 2. The mean (SD) age of the study population was 55.3 (17.4) years (range, 18-92 years). The study patients, all of whom were mechanically ventilated, had high severity of illness with mean (SD) APACHE II scores of 22.9 (7.2) (median, 23; interquartile range [IQR], 18-29), and a wide spectrum of illnesses. The presence of visual or auditory deficits were reported at baseline in 61.5% of patients, supporting the need to have both visual and auditory means to assess cognitive function as part of the CAM-ICU.

Criterion Validity of the CAM-ICU

Patients were evaluated during their entire hospital stay and had a total of 471 paired observations between the nurses and delirium experts. Using the first alert (n=91) or lethargic (n=93) paired evaluation of each patient, the test performance of the CAM-ICU was determined (TABLE 3). The 2 nurses’ sensitivities using the CAM-ICU compared with the reference standard were 100% for nurse 1 and 93.5% for nurse 2. Their specificities were 97.8% and 100%, respectively. The likelihood ratios for the CAM-ICU for the 2 nurses were 50 and 80, respectively. The likelihood ratios for the specificities were 97.8% and 100%, respectively. Their agreement with the reference standard were 100% for nurse 1 and 93.5% for nurse 2. Their performance of the CAM-ICU compared with the reference standard rating, and (2) a lapse of more than 3 hours between ratings.

Interrater Reliability of the CAM-ICU

In the 84-paired assessments of the first alert or lethargic evaluation for each patient, the CAM-ICU was completed with excellent interrater reliability between nurse 1 and nurse 2 (κ=0.96, 95% CI, 0.92-0.99).

Subgroup Performance of the CAM-ICU

We conducted 3 subgroup analyses (TABLE 4) to evaluate the performance of the CAM-ICU in patient groups that could pose particular challenges in delirium assessment: (1) those 65 years or older, (2) those with suspected dementia as defined in the “Methods” section, and (3) those with higher severity of illness. Interrater reliability measures across comparisons showed high agreement with κ values: 0.92 for those 65 or older, 0.99 for those with suspected dementia, and 0.94 for those with APACHE II scores at or above the median value of 23 (all P<.001). Comparing sensitivity and specificity between patient subgroups according to age, suspected dementia, and severity of illness showed no significant differences (all P>.56).

Ease of Use of the CAM-ICU

The CAM-ICU instrument and rating were completed in a mean (SD) of 2(1) minutes, demonstrating the relative ease with which it could be incorporated into nurses’ daily routine. When patients were either alert or lethargic, they were able to complete the visual ASE in 167 (69.6%) of 240 cases and able to complete the auditory ASE in 176 (73.3%) of 240 cases. Delirious patients were unable to complete both ASE tests and were considered nonresponsive in 82% of cases while nondelirious patients could complete both ASE tests in 91% of testing attempts.

Cognitive Reference Standard Findings

Delirium occurred in 80 patients (83.3%) during their ICU stay for a mean (SD) duration of 2.4 (1.6) days (median, 2; IQR, 1-3 days). Of 471 daily evaluations completed during this study, a diagnosis by the reference standard of delirium was made in 25.2%, stupor in 21.3%, and coma in 28.5% of all observations (with the remaining 25.0% rated as normal). The level of confidence of the delirium experts was recorded for each evaluation. Overall, the level of confidence was 100% of the time.

©2001 American Medical Association. All rights reserved.
DELIRIUM IN MECHANICALLY VENTILATED PATIENTS

Table 3. Validity of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) for Delirium Using the First Alert or Lethargic Evaluation for Each Patient*

<table>
<thead>
<tr>
<th>Rater</th>
<th>No. of Patients</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive Predictive Value (95% CI)</th>
<th>Negative Predictive Value (95% CI)</th>
<th>Likelihood Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurse 1</td>
<td>91</td>
<td>100 (90 to 100)</td>
<td>98 (91 to 100)</td>
<td>92 (77 to 99)</td>
<td>100 (93 to 100)</td>
<td>50 (20 to 77)</td>
</tr>
<tr>
<td>Nurse 2</td>
<td>92</td>
<td>99 (82 to 99)</td>
<td>100 (93 to 100)</td>
<td>100 (79 to 100)</td>
<td>98 (90 to 99)</td>
<td>&gt;100 (21 to 100)</td>
</tr>
</tbody>
</table>

*After excluding stuporous and comatose evaluations, CAM-ICU comparisons were made to reference standard evaluations by delirium experts using Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria. Interrater reliability measures across 84 paired comparison showed k of 0.96 (95% confidence interval [CI], 0.92-0.99; P<.001).

Table 4. Subgroup Analyses of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)*

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. of Patients</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nurse 1</td>
<td>Nurse 2</td>
<td>Nurse 1</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>68</td>
<td>97 (82-99)</td>
<td>100 (88-100)</td>
</tr>
<tr>
<td>≥65</td>
<td>28</td>
<td>100 (86-100)</td>
<td>100 (85-100)</td>
</tr>
<tr>
<td>Dementia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspected</td>
<td>12</td>
<td>100 (66-100)</td>
<td>100 (63-100)</td>
</tr>
<tr>
<td>Not suspected</td>
<td>84</td>
<td>98 (88-99)</td>
<td>100 (92-100)</td>
</tr>
<tr>
<td>APACHE II, score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;Median</td>
<td>45</td>
<td>100 (72-100)</td>
<td>100 (72-100)</td>
</tr>
<tr>
<td>≥Median</td>
<td>51</td>
<td>92 (64-100)</td>
<td>100 (75-100)</td>
</tr>
</tbody>
</table>

*After excluding stuporous and comatose evaluations, CAM-ICU comparisons were made to reference standard evaluations by delirium experts using Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria. Interrater reliability measures across comparisons showed k values of 0.92 for those aged 65 years or older, 0.99 for those with suspected dementia, and 0.94 for those with Acute and Chronic Health Evaluation II (APACHE II) scores at or above the median value of 23 (P<.001). Comparing sensitivity and specificity among patient subgroups by age, suspected dementia, and severity of illness showed no significant difference (all, P> .50). Patients were defined as having suspected dementia if they met any of the following 3 criteria: if a delirium expert rated them as having dementia, if they had a modified Blessed Dementia Rating Scale score of at least 3, or if they had a surrogate rating of at least 1 of 5 as possibly having dementia.

In this investigation, we have validated the CAM-ICU, a 2-minute assessment instrument, which demonstrated a sensitivity of 93% to 100%, a specificity of 98% to 100%, and high interrater reliability (k = 0.96) in the detection of delirium. In 96 consecutive mechanically ventilated patients, delirium occurred in 83.3% while they were in the ICU. In the 3 subgroups expected to pose the greatest challenges for the CAM-ICU (ie, those 65 years or older, those with suspected dementia, and those with the highest severity of illness), the instrument retained excellent sensitivity, specificity, and interrater reliability.

We simultaneously assessed ICU patients for level of sedation and delirium, finding that 40% of patients who were at a neutral level (neither agitated nor overly sedated) were delirious by both the reference standards’ (Diagnostic and Statistical Manual of Mental Disorders—IV) and nurses’ evaluations (CAM-ICU). Notably, 10% of patients met full delirium criteria, 20% met partial delirium criteria, and nearly 50% demonstrated substantial cognitive impairment (MMSE score, <24) at the time of hospital discharge.

The strengths of this study include the challenging study population of medically diverse but severely ill mechanically ventilated patients, the large number of patient evaluations, and the use of recognized delirium experts for the reference standard ratings. Another important strength of the study design was the use of a standardized, easily performed nursing assessment, which should allow the CAM-ICU to be readily implemented in both academic and community hospitals.
LIMITATIONS OF THIS INVESTIGATION WARRANT COMMENT. IN DEVELOPING THE CAM-ICU, WE Sought TO DEVELOP A TOOL FOR DETECTING DELIRIUM, NOT DEMENTIA. HOWEVER, IT IS COMMONPLACE FOR MILDELY DEMENTED PATIENTS TO BE CAREFUL FOR THE IN ICU SETTING. BECAUSE SUCH PATIENTS COULD POSE A CHALLENGE FOR THE CAM-ICU, WE USED LIBERAL INCLUSION OF PATIENTS WITH SUSPECTED DEMENTIA IN OUR SUBGROUP ANALYSIS TO VERIFY PERFORMANCE OF THE CAM-ICU IN PATIENTS WITH DEMENTIA. THE FINDINGS OF 100% SENSITIVITY AND SPECIFICITY IN THIS GROUP ARE BASED ON A SMALL NUMBER OF PATIENTS BUT ARE CONSISTENT WITH OUR DATA IN ANOTHER COHORT. THIS INVESTIGATION ALSO REPRESENTS A SELECTED POPULATION AT A SINGLE SITE, AND FUTURE STUDIES WILL NEED TO EVALUATE THE GENERALIZABILITY OF PERFORMANCE ACROSS OTHER PATIENT POPULATIONS INCLUDING THOSE WITH A LOWER PREVALENCE OF DELIRIUM.

THE VALIDATION OF A DELIRIUM INSTRUMENT FOR THE ICU OPENS A NEW FRONTIER FOR INVESTIGATION: TO EVALUATE THE IMPACT OF THIS IMPORTANT PROBLEM IN THE ICU, ESPECIALLY AS IT RELATES TO THE OUTCOMES OF OLDER, MECHANICALLY VENTILATED PATIENTS. IMPORTANT AREAS FOR FUTURE INVESTIGATION INCLUDE DETERMINATION OF RISK FACTORS FOR DELIRIUM IN THE ICU AND THE IMPACT OF DELIRIUM IN MECHANICALLY VENTILATED PATIENTS ON CLINICAL OUTCOMES SUCH AS REINTUBATION, nosocomial pneumonia, AS WELL AS BROADER OUTCOMES SUCH AS QUALITY OF LIFE. ALTHOUGH THE USE OF PSYCHOACTIVE MEDICATIONS SUCH AS SEDATIVES AND ANALGESICS IN MECHANICALLY VENTILATED PATIENTS IS INTENDED TO RELIEVE ANXIETY AND SUFFERING, RECENT STUDIES HAVE SUGGESTED THAT THESE MEDICATIONS MAY BE PRESCRIBED OVERZEALOUSLY. INTERVENTIONAL TRIALS DESIGNED TO REDUCE OVERUSE OF THESE MEDICATIONS AND THEIR ATTENDANT CONTRIBUTIONS TO DELIRIUM AND LONG-TERM COGNITIVE DEFICITS ARE GREATLY NEEDED.

COGNITIVE IMPAIRMENT IN THE ICU MAY BE INDEPENDENTLY RELATED TO PROLONGED NEUROPSYCHOLOGICAL DEFICITS, BUT STUDIES SPECIFICALLY ANALYZING THE INTERACTIONS BETWEEN DELIRIUM IN THE ICU AND LONG-TERM NEUROCOGNITIVE FUNCTION ARE LACKING. A SIGNIFICANT PERCENTAGE OF INDIVIDUALS DEVELOPING DELIRIUM IN THE HOSPITAL CONTINUE TO DEMONSTRATE SYMPTOMS OF DELIRIUM AFTER DISCHARGE. SUCH PATIENTS DEMONSTRATE INCREASED ACIVITY AND INCREASED COGNITIVE DETEORATION AND ARE MORE LIKELY TO DEVELOP DEMENTIA THAN PATIENTS WITHOUT DELIRIUM. FINALLY, PATIENTS WHO DEVELOP DELIRIUM HAVE A HIGHER RATE OF DEATH ON COGNITIVE TESTS THAN NONDELIRIOUS PATIENTS. IMPORTANTLY, FUTURE STUDIES ARE NEEDED TO DETERMINE THE PROGNOSTIC SIGNIFICANCE OF DELIRIUM IN THE ICU ON LONG-TERM COGNITIVE OUTCOMES.

In conclusion, delirium occurred in more than 8 of 10 mechanically ventilated adult medical ICU patients in this cohort, and it was present in 4 of 10 alert or easily aroused patients who are usually assumed to be cognitively intact by ICU personnel. Unfortunately, delirium is often not recognized by clinicians when it is noted in the ICU, it is often considered an “expected” occurrence attributed to ICU psychosocial. The most common type of delirium, hypoactive or quiet delirium, may be associated with a worse prognosis than hyperactive or agitated delirium. We documented that bedside nurses in the ICU with no formal psychiatric training can reliably detect delirium in mechanically ventilated patients with a high degree of sensitivity and specificity using the CAM-ICU. Requiring only a modest degree of training, this instrument is rapid and easy to use. Incorporation of the CAM-ICU into clinical practice and future investigations may lead to a more precise understanding of the incidence, predictors, and consequences of delirium among critically ill patients.

Author Affiliations: Department of Internal Medicine, Division of General Internal Medicine and Center for Health Services Research and the Veterans Affairs Tennessee Valley Health Care System Geriatric Research, Education and Clinical Center (Drs Ely, Gordon, Speroff, Gautam, and Dittus and Ms May and Truman), and Division of Allergy/Pulmonary/Critical Care Medicine (Drs Ely and Bernard), Department of Geriatric Psychiatry and Neurosciences (Drs Gordon and Margolin), Vanderbilt University School of Medicine, Nashville, Tenn; Department of Internal Medicine, Yale University School of Medicine, New Haven, Conn (Dr Inouye); Division of Quality and Data Management, St Vincent Hospital Health System, Indianapolis, Ind (Dr Francis); and the Department of Psychiatry, Virginia Commonwealth University Health Systems, Richmond (Dr Hart).

Author Contributions: Study concept and design: Ely, Bernard, Gordon, Francis, Truman, Speroff, Gautam, Marginol, Hart, Dittus.

©2001 American Medical Association. All rights reserved.
72. Sessler C, Gossnell M, Grant MP, et al. A new agi-