**Implications of Objective vs Subjective Delirium Assessment in Surgical Intensive Care Patients**

By Ulf Guenther, MD, Jan Weykam, Ulrich Andorfer, Nils Theuerkauf, MD, Julius Popp, MD, E. Wesley Ely, MD, MPH, and Christian Putensen, MD

**Background** Delirium is an independent predictor of increased length of stay, mortality, and treatment costs in critical care patients. Its incidence may be underestimated or overestimated if delirium is assessed by using subjective clinical impression alone rather than an objective instrument.

**Objectives** To determine frequency of discrepancies between subjective and objective delirium monitoring.

**Methods** An observational cohort study was performed in a surgical-cardiosurgical 31-bed intensive care unit of a university hospital. Patients’ delirium status was rated daily by bedside nurses on the basis of subjective individual clinical impressions and by medical students on the basis of scores on the objective Confusion Assessment Method for the Intensive Care Unit.

**Results** Of 160 patients suitable for analysis, 38.8% (n = 62) had delirium according to objective criteria at some time during their stay in the intensive care unit. A total of 436 paired observations were analyzed. Delirium was diagnosed in 26.1% of observations (n = 114) with the objective method. This percentage included 6.4% (n = 28) in whom delirium was not recognized via subjective criteria. According to subjective criteria, delirium was present in 29.4% of paired observations (n = 128), including 9.6% (n = 42) with no objective indications of delirium. A total of 8 patients with no evidence of delirium according to the objective criteria were prescribed haloperidol and lorazepam because the subjective method indicated they had delirium.

**Conclusions** Use of objective criteria helped detect delirium in more patients and also identified patients mistakenly thought to have delirium who actually did not meet objective criteria for diagnosis of the condition. (*Am J Crit Care*. 2011;21:e12-e20)
Delirium is the most common neurological diagnosis among patients in intensive care units (ICUs). This brain dysfunction is an independent predictor of prolonged ICU length of stay; longer hospital length of stay; and increased treatment costs. After discharge from the hospital, patients who had delirium during their hospital stay have increased rates of cognitive deficits. Mortality rates are significantly higher in patients with delirium (34%, 6 months after ICU stay; 42% after 12 months) than in patients without delirium (15%). The incidence of delirium varies widely; the dysfunction occurs in 8% to 92% of ICU patients, depending on severity of illness, the number of patients who are or are not treated with mechanical ventilation, different populations of patients (eg, surgical vs medical), and the choice of the delirium assessment method.

Although an increasing number of delirium assessment tools have become available for hospitals, monitoring delirium routinely is still deemed too time consuming and a feature of care that most consider dispensable. In the absence of objective delirium screening tools, many physicians rely on nurses’ clinical impressions. Clinical impression without objective criteria has repeatedly been reported to provide marked underestimates of the incidence of delirium. On one hand, a clinical downside of subjective delirium assessment is that delirium may not be detected and thus treatment opportunities missed. On the other hand, subjective assessment may provide overestimates of the incidence of delirium. For example, patients who are uncooperative because they are in pain, desire to be extubated, or have a brief hallucination not accompanied by any actual major delirium criteria may have delirium diagnosed. Such a situation may lead to unnecessary medication and missed opportunities for treatment of the true cause of the condition.

Almost any study in which subjective assessments are compared with objective assessments will report subjective ratings as a subset of an amount assessed by using a reference method. Overestimation of the incidence of delirium (and reasons for the delirium) when a subjective method is used is thereby excluded. The study we report here differs from other studies mainly in the exploration of the subjective clinical impression. We sought to determine reasons and consequences for discrepancies between nurses’ subjective assessments of delirium and objective assessment with the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) in a surgical-cardiosurgical 31-bed ICU in Bonn University Hospital, Bonn, Germany.

Methods

This prospective, observational study was approved by the hospital’s ethics committee. Written informed consent for use of the data for the study was obtained from patients who were competent after discharge from the ICU. Patients admitted to the hospital from October 2007 through November 2007 had daily rating of the presence or absence of delirium by bedside nurses according to nurses’ subjective clinical impression. This type of assessment was routine then because no delirium monitoring test was available to physicians and nursing staff at the time of the study. The findings of the subjective clinical impression were recorded along with patients’ demographics and medication data. Patients were also assessed for delirium by 1 of 2 fourth-year medical students who used the CAM-ICU. The students were selected to avoid a bias originating from clinical information and subjective impressions. Only these medical students and 1 investigator (U.G.) were familiar with the CAM-ICU at the time of the study.

About the Authors

Ulf Guenther and Nils Theuerkauf are attending physicians in critical care, Christian Putensen is a professor in anesthesiology and critical care, and Jan Weykam and Ulrich Andorfer are medical students, Clinic of Anesthesiology and Intensive Care, Bonn University Hospital, Bonn, Germany. Julius Popp is an attending physician in psychiatry, Department of Psychiatry, University Hospital of Lausanne, Lausanne, Switzerland. E. Wesley Ely is a professor, Division of Allergy, Pulmonary, and Critical Care, Center of Health Services Research, Vanderbilt University School of Medicine and Geriatric Research, Education and Clinical Center for the Tennessee Valley Veteran’s Affairs Healthcare System, Nashville, Tennessee.

Corresponding author: Dr U. Guenther, Universitätssklinikum Bonn, Sigmund Freud Strasse 25, D-53105 Bonn, Germany (e-mail: u.guenther@uni-bonn.de).
rates 180 days after surgery were obtained by mail and telephone interviews.

**Delirium Assessment**

An algorithm-type variant of the CAM-ICU, the CAM-ICU Flowsheet, was used to assess delirium. The variant is a reliable screening tool with high sensitivity and specificity and is easy not only for nurses and physicians but also for medical students to use.24 (For further educational information on delirium screening, see the CAM-ICU training manual at http://www.icudelirium.org.)

Delirium was classified into motoric subtypes according to the Richmond Agitation Sedation Scale (RASS).25,26 The RASS is a 10-level scale used to assess degree of arousal and agitation. Scores range from -5 (unarousable) to +4 (combative). The RASS or other validated arousal-sedation monitoring tools are used as “sister instruments” with the CAM-ICU to assess feature 3 (altered level of consciousness). Patients were classified as having hypoactive delirium if they had delirium according to the CAM-ICU and had RASS scores of -3 to 0 and as having hyperactive delirium if they had delirium according to the CAM-ICU and had RASS scores of +1 to +4. Patients with a RASS score of -4 or -5 are considered comatose and by convention (because they do not respond to verbal stimulation) are classified as unable to assess.

**Statistical Analysis**

Prism5 Software for Macintosh (GraphPad Software Inc, San Diego, California) was used for statistical analysis. Patients’ characteristics and outcome parameters, including age, height, weight, Charlson Comorbidity Index, white blood cell count, serum creatinine level, TISS score, SAPS, number of days of mechanical ventilation, ICU length of stay, and hospital length of stay were analyzed by using the Mann-Whitney test. Except for mean length of survival, results are given as median and interquartile range. The number of patients receiving mechanical ventilation and patients’ age and sex were analyzed by using the Fisher exact test. The incidence of delirium (CAM-ICU vs subjective clinical impression) was compared by using the Wilcoxon signed rank test. RASS scores were compared with a hypothetical value (zero) by using the Wilcoxon signed rank test. Mean survival was analyzed by using the log rank test. P < .05 was considered significant.

**Results**

**Patients**

Each medical student received bedside training in use of the CAM-ICU in five 15-minute sessions before the start of the study (total of 75 minutes). During the study, the students discussed occasional discrepancies between subjective delirium findings and CAM-ICU ratings with Dr Guenther, who is one of the original group who translated the CAM-ICU into German. A total of 147 nurses in the ICU with at least 5 years of job experience participated in the study. None of them had formal knowledge of delirium assessment; delirium assessment was not and still is not a routine part of nurses’ training in Germany. Nurses and CAM-ICU raters saw patients once a day, independently, and both assessments were done within 30 minutes’ time. All of the nurses and raters were unaware of each other’s judgment. No other delirium screening was performed.

Patients were excluded if they had severe pre-existing neurological impairments (e.g., acute stroke, dementia, aphasia) and if they did not speak German. Demographic data, including age, weight, height, sex, Charlson Comorbidity Index, white blood cell count, serum creatinine level, TISS score, SAPS, number of days of mechanical ventilation, ICU length of stay, and hospital length of stay were analyzed by using the Mann-Whitney test. Except for mean length of survival, results are given as median and interquartile range. The number of patients receiving mechanical ventilation and patients’ age and sex were analyzed by using the Fisher exact test. The incidence of delirium (CAM-ICU vs subjective clinical impression) was compared by using the Wilcoxon signed rank test. RASS scores were compared with a hypothetical value (zero) by using the Wilcoxon signed rank test. Mean survival was analyzed by using the log rank test. P < .05 was considered significant.

A total of 170 patients were screened. Of these, 10 were excluded from further analysis (Figure 1).
Table 1 displays preoperative demographic data and the admitting surgical specialities. A total of 4 patients never had surgery because they were referred to our ICU with the primary diagnosis of adult respiratory distress syndrome. The baseline characteristics of the 160 patients remaining for analysis indicate statistically nonsignificant trends of higher serum levels of creatinine and bilirubin in the group in whom delirium developed. Compared with patients in whom delirium did not develop in the ICU, patients who experienced delirium in the ICU had higher TISS scores and SAPS values at the time of admission to the unit. Patients who had delirium received more opioids and benzodiazepines, were treated with mechanical ventilation more often and longer, and had longer ICU and hospital stays than did patients without delirium (Table 2). One patient fractured his femoral neck in a fall in the hospital during a hypoactive episode of delirium. The mean 180-day survival rate after surgery was 143 days (95% CI, 128-159) in the group with delirium and 169 days (range, 162-176) in the group without delirium ($P= .002$; hazards ratio, 3.5; 95% CI, 1.6-7.5).

### Subjective Clinical Impression vs CAM-ICU

According to the CAM-ICU, delirium developed in 62 of the 160 patients (38.8%) at some time during their ICU stay. Only a minority had hyperactive ($n = 12$) or mixed type ($n = 10$) delirium. The majority ($n = 40$) had hypoactive delirium; most appeared calm, quiet, or drowsy. In 160 patients, 597 paired observations were completed. Patients were classified as unable to assess with the CAM-ICU 161 times because of RASS scores of -4 or -5.
but was significantly more often deemed present by subjective clinical impression (29.4%; n = 128; \( P = .047 \); Figure 2). Two subgroups accounted for these disparate findings in 16% of patients. First, in 9.6% of paired observations (n = 42), delirium was deemed present according to subjective clinical impression, although according to the CAM-ICU, the abnormality was not present. Second, delirium was not deemed present by subjective clinical impression in 6.4% (n = 28) of those who met CAM-ICU criteria for the abnormality. A total of 4 patients who had delirium according to the CAM-ICU during their stay in the ICU were transferred to a general unit without ever being classified as delirious on the basis of subjective clinical impression.

Agreement between CAM-ICU indications of delirium and subjective clinical impression differed depending on RASS scores. Agreement was high (93%) for patients who were delirious according to the CAM-ICU and agitated according to RASS scores (score >0). It was also high (90%) for patients who were not delirious according to the CAM-ICU and were calm and alert according to RASS scores (score = 0). Agreement was much lower for patients who had delirium according to the CAM-ICU and were calm, drowsy, or sedated (score <0, 73%; score = 0, 63%). Likewise, agreement was low for patients who did not have delirium according to the CAM-ICU but were drowsy (score <0, 74%) or agitated (score >0, 62%) according to RASS scores.

The manifestations leading to the subjective clinical impression of delirium are given in Table 3. They were grouped into positive and negative motoric signs and other findings not related to psychomotoric...
manifestations. Findings not related to psychomotoric manifestations included hallucinations, delusions, anxiousness, and indications of impaired memory, such as word-finding difficulties. Positive motoric signs included agitation, aggressive behavior, and unwillingness to cooperate. Negative motoric signs included disorientation, sleepiness, and slow communication.

**Discussion**

We found that subjective clinical impressions indicated delirium more often than did objective assessment with the CAM-ICU as the operational reference standard. Overall, findings between the subjective and objective approaches differed in 16% of patients: subjective clinical impression did not indicate delirium in 6.4% of paired observations for which the CAM-ICU did and also indicated delirium in 9.6% of observations for which the CAM-ICU did not. Agreement between subjective clinical impression and the results of the CAM-ICU was high (>90%) for delirious, agitated patients and for nondelirious, calm and alert patients. Rates of agreement were lower (<75%) for delirious, but seemingly calm and alert patients and for nondelirious patients who were either drowsy or agitated. Patients with delirium had higher SAPS and TISS scores at the time of ICU admission, received more opioids and benzodiazepines, had longer ICU and hospital stays, and had lower survival rates than did patients without delirium.

**Underrating of Delirium**

Most patients in whom delirium was detected by using the CAM-ICU and not by using subjective clinical impression had a RASS score of 0 or less, meaning that they were calm and alert or appeared drowsy or sedated. Hence, hypoactive delirium was apt to remain undiagnosed because of its seemingly calm clinical manifestations.\(^\text{17,27}\) Spronk et al\(^\text{16}\) reported that ICU delirium (mostly hypoactive motoric subtype) is missed in 75% of assessments when delirium is not actively monitored by using an objective instrument. This finding is clinically important because compared with other subtypes, the hypoactive motoric subtype of delirium is particularly associated with prolonged hospital length of stay\(^\text{10}\) and a higher incidence of decubitus ulcers.\(^\text{28}\)

In our study, in 2 paired observations, even agitated patients with delirium were not delirious according to subjective clinical impression, because the patients were oriented to person and place. Orientation to person and place, however, does not exclude delirium according to delirium criteria in the *Diagnostic and Statistical Manual of Mental Diseases* (Fourth Edition, Text Revision, *DSM-IV-TR*).\(^\text{29}\) Indeed, the CAM-ICU evaluation indicated that these patients were inattentive and had disorganized thinking, cardinal features of delirium. This finding underlines the importance of acknowledging different motoric subtypes of delirium. Our results confirm that even patients who appear alert and calm may have delirium.

**Possible Overrating of Delirium With Subjective Assessment**

Most validation studies of instruments used to measure delirium have compared the results obtained with an assessment tool with results obtained by using a reference standard. Delirium is then reported as a fractional amount or a ratio of the reference standard, and false-positive results and the reasons for the results are thus excluded. We did not perform our study to determine validation criteria in comparison with a reference rater applying the delirium criteria of the *DSM-IV-TR*. We simply sought reasons for the discrepancies between subjective and objective approaches.

---

**Table 3**

Clinical manifestations cited by nurses that led to subjective impression of delirium

<table>
<thead>
<tr>
<th>Positive motoric signs</th>
<th>Negative motoric signs</th>
<th>Manifestations not related to psychomotor domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wants repetitively to get out of bed</td>
<td>Sleepy, not fully alert</td>
<td>Not oriented to person, place, or time</td>
</tr>
<tr>
<td>Restlessness, fiddling about (with sheets, catheters)</td>
<td>Slow reaction looks “through” you</td>
<td>Indications of impaired memory</td>
</tr>
<tr>
<td>Aggressive behavior (harmful to staff, pulls drains, tubes, etc)</td>
<td>Slurred speech</td>
<td>Word-finding difficulty</td>
</tr>
<tr>
<td>Logorrhea</td>
<td>Slow/inappropriate communication</td>
<td>Anxiousness without reason</td>
</tr>
<tr>
<td>Works against nursing activities</td>
<td></td>
<td>Hallucinations</td>
</tr>
<tr>
<td>Restlessness</td>
<td></td>
<td>Delusions</td>
</tr>
<tr>
<td>Agitation</td>
<td></td>
<td>Suspicion of everything</td>
</tr>
</tbody>
</table>

---

Delirium was diagnosed with the CAM-ICU less often than by using subjective clinical impression.
for discrepancies between subjective assessments of delirium and the CAM-ICU; the CAM-ICU was the a priori determined operational reference standard. A major finding was that subjective assessment indicated a higher rate of delirium than the CAM-ICU did. Possibly, delirium was not detected in some patients by using the CAM-ICU because the patients had a low severity of illness or because extended periods of time had passed since training in use of the instrument occurred; other investigators\textsuperscript{30} have reported low sensitivity of the CAM-ICU for detection of delirium. A specificity of almost 100\% has been reported\textsuperscript{14} for the CAM-ICU Flowsheet used in our study.

Discrepancies in which patients were delirious according to subjective assessment and not delirious according to the CAM-ICU occurred mostly if the patients appeared sedated or agitated. Unwillingness to cooperate, hyperactive states, and drowsiness were reasons given for subjectively rating patients as delirious. This finding suggests that patients were rated delirious because of their behavior, which could be deliberate if a patient were frustrated or had painful sensations. The finding also emphasizes that conspicuous states of consciousness (agitation or drowsiness) per se are not specific to delirium.

The difference between frequencies of delirium measured by using the CAM-ICU (26.1\%) and by using subjective clinical impression (29.4\%) may appear small, but it corresponds to 8 paired observations in patients who were given antipsychotic medications or sedatives though they were not delirious. For patients’ safety, minimizing the use of any unnecessary medications is important, especially drugs that might be arrhythmogenic via prolongation of the QTC or might cause unwanted side effects such as extrapyramidal signs and symptoms.

Estimated Prevalence With Subjective vs Objective Assessments

At first sight, our results seem in contrast to those of other researchers\textsuperscript{10,16,31} who reported that the incidence of delirium was severely underestimated in the ICU. Van den Boogaard et al\textsuperscript{19} reported that the incidence of diagnosed delirium and the number of patients who were prescribed haloperidol increased after implementation of a delirium screening tool. According to these authors,\textsuperscript{19} the increase was due to the introduction of the delirium assessment tool, which led to a higher rate of detection of delirium. Interestingly, in the same study,\textsuperscript{19} the amount of prescribed haloperidol per patient decreased during the study period. Van den Boogaard et al reported that this finding was the result of earlier discontinuation of medication due to earlier recognition of patients who no longer “required” haloperidol, suggesting that the objective assessments were helping avoid “overdiagnosis” of delirium. Another possible mechanism could be a reduction in the duration of delirium due to earlier onset of treatment, as has been suggested by others.\textsuperscript{32}

Limitations

Some limitations of our study warrant comment. First, the CAM-ICU is not a gold standard for diagnosis of delirium, and it should not be a substitute for a delirium expert such as a psychiatrist. In recent validation studies,\textsuperscript{24,33,34} the sensitivity of the instrument ranged from 79\% to 92\%. Most likely our CAM-ICU raters underestimated the true incidence of delirium to some extent. In a recent Dutch multicenter study,\textsuperscript{30} the CAM-ICU had a sensitivity of only 47\%. The authors\textsuperscript{30} attributed this finding to their multicenter design; many clinical staff members involved were not as familiar with the CAM-ICU as the more intensely trained and experienced investigators in single-center studies. The lower sensitivity could also reflect a study in which the severity of illness was lower, and thus delirium in some patients might have been less overt and might have required a more extensive test for core features such as inattention.

Second, our study was performed in a single center in surgical patients, and thus might reflect only a segment of possible discrepancies of subjective vs objective rating of delirium. The target of our research was the possible misconceptions that would be encountered when the CAM-ICU was introduced on a larger scale to our ICU staff. A multicenter study and a larger number of patients, including nonsurgical patients, will extend the generalizability of the results and lead to delirium monitoring tools suitable for use on a large scale.

Third, we did not recruit ICU nurses to perform the CAM-ICU testing. To minimize possible bias, we opted for medical students who were not familiar with the nursing staff or the patients’ medical history. We did not intend to assess the performance of nursing staff with the CAM-ICU; this topic has been covered in other studies.\textsuperscript{35-38} Rather, we wished to elucidate the manifestations that led nurses to rate patients as having delirium although the patients
were not delirious according to the CAM-ICU. Because nurses and physicians’ perceptions of delirium features may differ significantly, objective tools for monitoring delirium provide the foundation and common language for monitoring and treating patients with delirium throughout the patients’ clinical course.

Conclusion

Subjective assessment did not indicate delirium in a worrisome number of patients who had delirium according to the CAM-ICU. At the same time, delirium was possibly overestimated in patients who did not fulfill delirium criteria such as inattention or disorganized thinking as assessed by the operational reference standard. Routine delirium monitoring with the CAM-ICU Flowsheet is an easy way to help detect delirium so that agreed-upon nonpharmacological and drug-cessation approaches could be applied before administration of new and potentially harmful medications is started. Such objective monitoring may also increase the recognition of patients who are at risk for delirium-specific complications such as falls and pressure ulcers. Cross-talk among members of the interdisciplinary ICU team is limited when only subjective impressions are available for monitoring delirium because differences of opinion are difficult to articulate and contrast. Future quality improvement projects may shed more light on ways to facilitate the new culture of ICU monitoring and safety management with regard to delirium, which is one of the most common types of acute organ dysfunction.

FINANCIAL DISCLOSURES

U. Guenther received honoraria from GlaxoSmithKline, Hamburg, Germany, and Orion Pharma, Hamburg, Germany. E. W. Ely received grants and honoraria from Pfizer Inc, New York, New York; Hospira, Inc, Lake Forest, Illinois; GlaxoSmithKline, Philadelphia, Pennsylvania; Aspect Software, Inc, Chelmsford, Massachusetts; and Masimo Corp, Irvine, California.

eLetters

Now that you’ve read the article, create or contribute to an online discussion on this topic. Visit www.ajcconline.org and click “Submit a response” in either the full-text or PDF view of the article.

REFERENCES


