Routine use of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) by bedside nurses may underdiagnose delirium

Michael C Reade, Glenn M Eastwood, Leah Peck, Rinaldo Bellomo and Ian Baldwin

ABSTRACT

Background: The Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) is emerging as the most frequently used tool for identifying delirium among critically ill patients.

Objective: To determine whether the number of patients and nursing shifts in which delirium was diagnosed would increase after the introduction of the CAM-ICU in our unit.

Design: Before-and-after study. In a 30-day Phase 1, we asked bedside nurses to assess their ICU patients for delirium each shift. We then conducted intensive education on the CAM-ICU for 30 days, including lectures, bedside tutorials, and supervised practice. In Phase 2, for 30 days we asked bedside nurses to record the results of their CAM-ICU assessments.

Setting: 20-bed mixed medical and surgical ICU at the Austin Hospital, Melbourne.

Participants: All patients admitted to the ICU during each phase.

Main outcome measures: Diagnosis of delirium by bedside nurses using either the CAM-ICU or an unstructured clinical assessment, by patient and nursing shift.

Results: Compared with unstructured assessments, the CAM-ICU identified a significantly lower proportion of patients (36.7% v 21.3%; $P = 0.004$) and a significantly lower proportion of shifts (14.7% v 6.4% of shifts, $P = 0.002$) with delirium. When adjusted for differences in age, sex, Acute Physiology and Chronic Health Evaluation III risk of death and total length of stay between the two periods, assessment type remained a significant predictor of the diagnosis of delirium.

Conclusions: In our hospital, the CAM-ICU detected delirium less often than unstructured delirium assessments made by qualified intensive care nurses.

Crit Care Resusc 2011; 13: 217–224

Delirium is increasingly recognised as a common occurrence in patients in intensive care units.[1-3] Among intensive care patients, delirium predicts higher mortality,[4] longer hospital stay,[1,5] and increased likelihood of discharge to a destination other than home.[6] There are several promising new treatments for delirium in the ICU, such as quetiapine and dexmedetomidine.[7,8] Their effective use, however, requires accurate diagnosis of delirium. Although delirium with physical agitation is relatively easily identified, delirium without agitation ("hypoactive delirium") is difficult to detect unless actively sought.[9,10]

Two established methods for delirium assessment are the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)[11] and the Intensive Care Delirium Screening Checklist (ICDSC).[12] Both assessment methods reach similar conclusions in a range of patients,[13] although a recent study suggested the CAM-ICU was more sensitive, whereas the ICDSC was more specific.[14] Both were markedly better than the opinion of a psychiatrist, geriatrician or neurologist.[14] The CAM-ICU appears to be emerging as the more commonly used tool in the ICU.[15-17] When performed by trained research nurses, the CAM-ICU diagnosed more patients with delirium than did routine assessments by bedside nurses.[18,19] However, studies of the CAM-ICU have almost exclusively involved dedicated, trained assessors rather than bedside ICU nurses, who must incorporate delirium assessment into their many other clinical tasks. Further, direct comparison of the CAM-ICU with unstructured delirium assessments, both performed by bedside nurses, has not been reported.

The CAM-ICU is not routinely used in Australian ICUs,[20] and had not been used in our ICU before this study. We hypothesised that introduction of the CAM-ICU would increase the number of patients among whom, and nursing shifts in which, the diagnosis of delirium was made, compared with an earlier observation period in which unstructured assessments were made by bedside nurses.

Methods

The study was conducted at the Austin Hospital, a tertiary hospital that is affiliated with the University of Melbourne and admits 65 000 patients each year. In 2009, 2040 critically ill patients were admitted to our 20-bed general

medical and surgical ICU. At the time of the study, there were 174 nursing staff members filling 105 full-time-equivalent positions.

Our study involved a prospective assessment of delirium, as diagnosed by bedside ICU nurses, before and after the introduction of the CAM-ICU. For 30 days beginning in February 2010, without any special antecedent training in delirium diagnosis or management, we asked bedside nurses to record on an audit form whether they thought their patients had been delirious during any part of their shift. The audit form defined delirium as "fluctuating mental status with confused or disorganised thinking ± agitation", and the possible responses were "yes", "no", and "unable to assess". Whether the patients were intubated, had a tracheostomy or were breathing without an artificial airway was also recorded. All patients in all shifts were assessed.

Over the next 30 days, we conducted an intensive CAM-ICU education program with medical and nursing staff. We used the six techniques found to be effective in teaching the CAM-ICU, including explaining its rationale, detailed features, examples of situations that might cause confusion, critiqued practice assessments, clear documentation, and provision of an ongoing forum for discussion.[21]

For 30 days in April–May 2010, we distributed audit forms on which bedside nurses were asked to record the results of their CAM-ICU assessments. A CAM-ICU worksheet was provided.

The age, sex, Acute Physiology and Chronic Health Evaluation (APACHE) III risk of death (incorporating information on diagnosis and comorbidity), and ICU and hospital length of stay for each patient was obtained from the Australian and New Zealand Intensive Care Society Adult Patient Database[22] data collected by our unit for the purpose of quality assurance. Each of these variables was entered into a multivariate regression model to explore potential confounders in the association of delirium diagnosis and assessment method. Finally, we assessed how the diagnosis of delirium

Table 1. Characteristics of patients admitted in each phase of the study

	Unstructured assessment ($n = 147$)	CAM-ICU ($n = 141$)	*P*
Median age in years (IQR)	60 (48–75)	59 (47–71)	0.61
Sex, % male	59.5%	66.1%	0.28
Median ICU LOS, days (IQR)	2 (1–4)	2 (1–5)	0.34
Median hospital LOS, days (IQR)	11 (7–22)	11 (7–23)	0.97
ICU mortality, % died	5.0%	8.2%	0.31
Hospital mortality, % died	12.7%	15.0%	0.61
Median APACHE III score (IQR)	51 (40–68)	55 (40–75)	0.44
Median APACHE III risk of death, % (IQR)	6% (1%–20%)	7% (2%–26%)	0.36

CAM-ICU = Confusion Assessment Method for the Intensive Care Unit. IQR = interquartile range. LOS = length of stay. APACHE = Acute Physiology and Chronic Health Evaluation.

Table 2. Diagnosis of delirium using unstructured assessments and the CAM-ICU

	Unstructured assessment ($n = 147$ patients; 1255 shifts)	CAM-ICU ($n = 141$ patients; 1113 shifts)	*P*
No. of patients with delirium in at least one shift (%)	54 (36.7%)	30 (21.3%)	0.004
No. of patients with delirium in two consecutive shifts (%)	34 (23.2%)	14 (9.9%)	0.003
No. of patients for whom nurses were unable to perform assessment in at least one shift (%)	54 (36.7%)	67 (47.6%)	0.06
Shifts in which delirium was diagnosed, % of total shifts for each patient			
Median (IQR)	0 (0–25%)	0 (0–0)	0.002
Mean (SD)	14.7% (25.8%)	6.4% (16.2%)	
Two consecutive shifts in which delirium was diagnosed, % of total shifts for each patient			
Median (IQR)	0 (0–0)	0 (0–0)	0.002
Mean (SD)	3.3% (8.8%)	1.1% (4.1%)	
Shifts in which an assessment was unable to be made, % of total shifts for each patient			
Median (IQR)	0 (0–29%)	0 (0–50%)	0.02
Mean (SD)	16.8% (27.8%)	26.7 (35.8%)	

CAM-ICU = Confusion Assessment Method for the Intensive Care Unit. IQR = interquartile range.

(at any point, and by proportion of the total number of 8–10-hour nursing shifts in which each patient was assessed) using each method predicted outcome (ICU and hospital length of stay and mortality).

Statistical analyses were performed using Stata, version 9.2 (StataCorp, College Station, Tex, USA), with differences between groups assessed using χ^2 tests and stepwise multivariate logistic regression (with $P < 0.1$ used for retention in the model) as appropriate. Data were visually inspected for normality and summarised using parametric or non-parametric statistics as required. Statistical significance was defined as $P < 0.05$. We calculated that, with a historical mean ICU stay of 2.85 days, in each 30-day study period there would be about 166 admissions and 1453 shifts for analysis. Estimating that patients would be too sedated for assessment in 20% of these shifts, and that the proportion with delirium as measured using traditional unstructured methods is 20%, the study would have an 80% chance of detecting (with 95% certainty) a significantly higher delirium prevalence with the CAM-ICU if this method found a prevalence of 25% or greater.

The Austin Hospital Human Research Ethics Committee approved the study (approval no. H2010/03889) and ruled that there was no requirement to obtain consent.

Table 3. Diagnosis of delirium using unstructured assessments and the CAM–ICU in strata defined by whether a patient had ever been intubated

	Unstructured assessment	CAM–ICU	*P*
Intubated at some point, *n*	77 patients; 925 shifts	60 patients; 693 shifts	
No. of patients with delirium in at least one shift (%)	35 (45.5%)	14 (23.3%)	0.007
No. of patients with delirium in two consecutive shifts (%)	23 (29.9%)	6 (10.0%)	0.005
No. of patients for whom nurses were unable to perform assessment in at least one shift (%)	49 (63.6%)	51 (85.0%)	0.005
Shifts in which delirium was diagnosed, % of total shifts intubated for each patient			
Median (IQR)	0 (0–25%)	0 (0–0)	0.002
Mean (SD)	13.4% (20.2%)	4.1% (9.7%)	
Two consecutive shifts in which delirium was diagnosed, % of total shifts intubated for each patient			
Median (IQR)	0 (0–3%)	0 (0–0)	0.004
Mean (SD)	2.7% (5.6%)	0.6% (2.3%)	
Shifts in which an assessment was unable to be made, % of total shifts intubated for each patient			
Median (IQR)	20% (0–50%)	52% (25%–92%)	< 0.001
Mean (SD)	29.1% (31.5%)	54.1% (35.8%)	
Never intubated, *n*	70 patients; 330 shifts	81 patients; 420 shifts	
No. of patients with delirium in at least one shift (%)	19 (27.1%)	16 (19.8%)	0.28
No. of patients with delirium in two consecutive shifts (%)	11 (15.7%)	8 (9.9%)	0.28
No. of patients for whom nurses were unable to perform assessment in at least one shift (%)	5 (7.1%)	16 (19.8%)	0.03
Shifts in which delirium was diagnosed, % of total shifts not intubated for each patient			
Median (IQR)	0 (0–17%)	0 (0–0)	0.20
Mean (SD)	16.2% (30.9%)	8.1% (19.6%)	
Two consecutive shifts in which delirium was diagnosed, % of total shifts not intubated for each patient			
Median (IQR)	0 (0–0)	0 (0–0)	0.24
Mean (SD)	4.1% (11.3%)	1.4% (5.1%)	
Shifts in which an assessment was unable to be made, % of total shifts not intubated for each patient			
Median (IQR)	0 (0–0)	0 (0–0)	0.03
Mean (SD)	3.2% (13.8%)	6.4% (17.7%)	

CAM-ICU = Confusion Assessment Method for the Intensive Care Unit. IQR = interquartile range.

Table 4. Diagnosis of delirium using unstructured assessments and the CAM–ICU in strata defined by whether a patient was intubated at the time of assessment

	Unstructured assessment	CAM–ICU	*P*
Intubated at the time of assessment, *n*	77 patients; 365 shifts	60 patients; 344 shifts	
No. of patients with delirium in at least one shift (%)	21 (27.3%)	4 (6.7%)	0.002
No. of patients for whom nurses were unable to perform assessment in at least one shift (%)	48 (62.3%)	49 (81.7%)	0.02
Shifts in which delirium was diagnosed, % of total shifts intubated for each patient			
Median (IQR)	0 (0–7%)	0 (0–0)	0.002
Mean (SD)	7.8 (21.0%)	2.0 (13.0%)	
Shifts in which an assessment was unable to be made, % of total shifts intubated for each patient			
Median (IQR)	9% (0–33%)	20% (5%–50%)	0.04
Mean (SD)	23.1% (31.9%)	32.3% (34.9%)	
Not intubated at the time of assessment, *n*	70 patients; 569 shifts	81 patients; 514 shifts	
No. of patients with delirium in at least one shift (%)	36 (51.4%)	22 (27.2%)	0.003
No. of patients for whom nurses were unable to perform assessment in at least one shift (%)	7 (10.0%)	20 (24.7%)	0.02
Shifts in which delirium was diagnosed, % of total shifts not intubated for each patient			
Median (IQR)	0 (0–9%)	0 (0–0)	0.06
Mean (SD)	7.8% (16.2%)	4.6% (13.1%)	
Shifts in which an assessment was unable to be made, % of total shifts not intubated for each patient			
Median (IQR)	0 (0–0)	0 (0–0)	0.006
Mean (SD)	2.1% (11.0%)	4.8% (15.2%)	

CAM-ICU = Confusion Assessment Method for the Intensive Care Unit. IQR = interquartile range.

Results

Characteristics of the 147 patients admitted to the unit during the first (unstructured assessment) month of the study and the 141 patients in the second (CAM-ICU) month were very similar, as shown in Table 1.

Almost all the nurses (151; 87%) employed in the ICU attended at least one of the CAM-ICU education sessions. All the nursing shift supervisors attended an education session, and as these nurses act as clinical advisors to our bedside nurses, we were confident that all nursing staff had access to a resource person trained in the CAM-ICU.

Unstructured assessments, as compared with the CAM-ICU, found a significantly higher proportion of patients (36.7% v 21.3% of patients; $P = 0.004$) and shifts (14.7% v 6.4% of shifts; $P = 0.002$) in which the diagnosis of delirium was made (Table 2). When adjusted for age, sex, APACHE III risk of death and total ICU length of stay, assessment type remained a significant predictor of the diagnosis of delirium in patients ($P = 0.001$) and shifts ($P = 0.001$). Using the CAM-ICU, nurses were more often unable to make a delirium assessment (16.8% v 26.7% of all shifts; $P = 0.024$), although this difference was no longer significant when adjusted for age, sex, and APACHE III risk of death.

Patients were analysed in groups defined by whether they had been intubated at some point in their ICU stay (Table 3). Among patients who had been intubated, the unstructured assessments diagnosed delirium in at least one shift significantly more often than did the CAM-ICU (45.5% v 23.3% of patients; $P = 0.007$). There was a similar trend among patients who had never been intubated, but this did not reach significance (27.1% v 19.8% of patients with delirium in at least one shift, $P = 0.28$). The median proportion of shifts in which an assessment was unable to be made was zero for patients who had not been intubated, but for patients intubated at some time, the CAM-ICU was unable to be performed in more than twice as many shifts (52.4% v 20.0%, $P < 0.001$).

Analogous analyses were performed on assessments of patients who had been, and who had not been, intubated at the time of their assessment (Table 4). These analyses reached the same conclusions, with the exception that the trend for unstructured assessments to diagnose a higher

rate of delirium in at least one shift became statistically significant (51.4% v 27.2% of patients with delirium in at least one shift, $P = 0.003$).

Using unstructured assessments, delirium tended to be more common among patients who had been intubated than among those who had not (45.5% of intubated patients v 27.1% of patients never intubated were delirious at least once [$P = 0.11$], and 29.8% v 15.7% were delirious in two consecutive shifts [$P = 0.04$]). The CAM-ICU did not uncover this distinction between intubated and never intubated patients (delirious in at least one shift, 23.3% v 19.8% [$P = 0.68$]; delirious in two consecutive shifts, 10.0% v 9.9% [$P = 0.98$]).

In multivariate regression analyses (Table 5), the diagnosis of delirium by unstructured assessment in at least one shift was a significant predictor of ICU but not hospital mortality, independent of age, sex and APACHE III risk of death. In contrast, the CAM-ICU diagnosis of delirium in at least one shift did not independently predict either ICU or hospital mortality. ICU and hospital length of stay were independently predicted by the presence of delirium in at least one shift, for delirium assessed using either method. The proportion of shifts when delirium was present, regardless of assessment method, was a significant predictor of ICU and hospital length of stay in univariate analyses, but after adjustment this remained true only for delirium diagnosed using the CAM-ICU.

Discussion

We found the CAM-ICU identified delirium significantly less frequently than unstructured assessments when both methods of patient examination were made by trained ICU bedside nurses in the process of caring for their patients. The difference was most prominent among patients who had been intubated at some point in their ICU stay, and among patients who were intubated at the time of assessment. Delirium only independently predicted ICU mortality when diagnosed using unstructured assessments. The association between the diagnosis of delirium and length of stay was similar regardless of the assessment method used, but whether this relationship was cause, effect or both was unclear.

Around 70% of ICU patients are said to have delirium (diagnosed by researchers using the CAM-ICU) at some point in their ICU stay.[2,10,23] We were initially not surprised that when making unstructured assessments, we found only one-third of our patients were delirious, as we expected on the basis of other studies[18,24] that unstructured assessments would be relatively insensitive. However, subsequently we found that the CAM-ICU diagnosed delirium in an even smaller proportion of patients and shifts. Without a "gold-standard" test, it is uncertain whether this means the CAM-ICU was less sensitive or whether it produced fewer false positives, but given the strikingly lower incidence of delirium in this study compared with other reports, the most likely explanation is that the CAM-ICU is less sensitive when performed by a trained bedside nurse in the course of their other clinical tasks than when performed by dedicated research nurses.

There have been no direct comparisons between the results of the CAM-ICU as applied by bedside nurses in the "real world" and the clinical impression gained through their unstructured routine assessments (as opposed to in a research study, when a dedicated investigator can focus exclusively on their assessment[18,19]). One before-and-after study of 1742 patients found that introducing the CAM-ICU led to more patients being treated with haloperidol, indirectly suggesting that the CAM-ICU was diagnosing more people with delirium, but perhaps only highlighting delirium as a problem worth treating.[24]

The only study to examine specifically whether real-life implementation of the CAM-ICU was an improvement on routine clinical assessment when both approaches were used in the same patients ($n = 711$) was performed in two American hospitals between 2001 and 2003.[25] Interrater reliability of delirium assessment using traditional methods was very low ($\kappa = 0.20$ and 0.03 at the two hospitals involved), but this improved (to $\kappa = 0.92$ and 0.75) with the introduction of the CAM-ICU. The rate of delirium diagnosed using unstructured assessments was not reported, so although this study suggests the CAM-ICU is more repeatable, it provides no information about the relative rates of diagnosis. Performed in an era before delirium had become topical, and in an environment with a nurse–patient ratio of 1 : 2, this study is not necessarily applicable to modern Australian practice.

The only study to attempt to characterise the sensitivity and specificity of the CAM-ICU when applied by bedside nurses (using the presumed gold standard of psychiatrist, neurologist or geriatrician diagnosis) involved 282 patients, 181 of whom were sufficiently awake to be assessed.[26] The sensitivity of the CAM-ICU was 47%, with a gold-standard delirium prevalence of 43% among patients able to be assessed. If the sensitivity of the real-life CAM-ICU is indeed 47%, our unstructured assessments, which detected delirium about twice as often as the CAM-ICU, diagnose delirium at a strikingly similar rate to these reported detailed neuropsychiatric assessments.

For patients who were intubated, our nurses thought they could not assess delirium using the CAM-ICU more frequently than when they used unstructured assessments. Inability to assess patients was (anecdotally, and logically) almost always because the patient was too sedated. It is not

Table 5. Predictors of intensive care unit and hospital mortality and LOS, assessed using univariate and stepwise multivariate regression modelling*

Unstructured assessments	Univariate odds ratio	Univariate *P*	Multivariate odds ratio	Multivariate *P*	CAM-ICU	Univariate odds ratio	Univariate *P*	Multivariate odds ratio	Multivariate *P*
Hospital mortality					Hospital mortality				
Age	1.00	0.98	—	—	Age	0.96	0.04	0.94	0.026
Sex	1.38	0.58	—	—	Sex	2.56	0.16	—	—
APACHE III risk of death	0.04	0.001	0.04	0.001	APACHE III risk of death	0.00	< 0.001	0.00	< 0.001
Presence of delirium in any shift	1.17	0.79	—	—	Presence of delirium in any shift	1.61	0.55	—	—
Proportion of shifts with delirium	3.88	0.4	—	—	Proportion of shifts with delirium	0.65	0.82	—	—
Proportion of shifts unable to assess	0.03	< 0.001	0.04	0.001	Proportion of shifts unable to assess	0.09	0.001	—	—
ICU mortality					ICU mortality				
Age	1.02	0.37	1.13	0.046	Age	0.99	0.84	—	—
Sex	0.64	0.6	—	—	Sex	1.15	0.85	—	—
APACHE III risk of death	0.01	< 0.001	0.00	0.005	APACHE III risk of death	0.00	< 0.001	0.00	< 0.001
Presence of delirium in any shift	3.10	0.31	53.73	0.04	Presence of delirium in any shift	—	—	—	—
Proportion of shifts with delirium	30.60	0.35	—	—	Proportion of shifts with delirium	—	—	—	—
Proportion of shifts unable to assess	0.05	0.01	—	—	Proportion of shifts unable to assess	0.01	< 0.001	—	—
Hospital LOS					Hospital LOS				
Age	0.28	< 0.001	—	—	Age	0.28	< 0.001	0.14	0.1
Sex	11.80	< 0.001	—	—	Sex	11.57	< 0.001	—	—
APACHE III risk of death	53.65	< 0.001	—	—	APACHE III risk of death	46.14	< 0.001	10.74	0.1
Presence of delirium in any shift	25.67	< 0.001	11.60	0.002	Presence of delirium in any shift	27.16	< 0.001	11.86	0.002
Proportion of shifts with delirium	47.75	< 0.001	—	—	Proportion of shifts with delirium	68.74	< 0.001	34.32	0.002
Proportion of shifts unable to assess	33.36	< 0.001	—	—	Proportion of shifts unable to assess	23.00	< 0.001	– 8.2	0.1
ICU LOS					ICU LOS				
Age	0.06	< 0.001	—	—	Age	0.07	< 0.001	—	—
Sex	2.66	< 0.001	—	—	Sex	3.14	< 0.001	—	—
APACHE III risk of death	11.87	< 0.001	3.57	0.048	APACHE III risk of death	15.64	< 0.001	8.26	< 0. 001
Presence of delirium in any shift	6.84	< 0.001	4.10	< 0.001	Presence of delirium in any shift	9.32	< 0.001	5.74	< 0.001
Proportion of shifts with delirium	9.65	< 0.001	—	—	Proportion of shifts with delirium	22.11	< 0.001	12.43	0.001
Proportion of shifts unable to assess	9.97	< 0.001	4.19	0.007	Proportion of shifts unable to assess	9.54	< 0.001	—	—

CAM-ICU = Confusion Assessment Method for the Intensive Care Unit. LOS = length of stay. APACHE = Acute Physiology and Chronic Health Evaluation.
— = Not significant ($P > 0.1$) in stepwise multivariable regression.
* With age, sex, APACHE III risk of death, and *either* presence of delirium in any shift *or* proportion of shifts with delirium and proportion of shifts in which assessment was not possible as the full model. $P < 0.1$ used for retention. Multivariate odds ratios for age, sex and APACHE III risk of death are presented for the models containing presence of delirium in any shift.

surprising that this affected the CAM-ICU assessments more, as the CAM-ICU must be performed at a single time point, whereas the unstructured assessments are likely to be a composite over an entire nursing shift. More frequent inability to assess delirium using the CAM-ICU might partly explain the lower rates of CAM-ICU delirium diagnosis. However, the difference is probably too small to account for all of the discrepancy. Another likely explanation is that the single time point CAM-ICU may have missed periods of delirium that would have been captured by continuous assessments made over the entire shifts.

We found a trend to more delirium diagnosis with unstructured assessments than with the CAM-ICU among patients who had never been intubated. However, of assessments made while patients were not intubated, unstructured assessments diagnosed significantly more delirium than the CAM-ICU. Inability to assess non-intubated patients with either technique was infrequent in comparison to when patients were intubated. Others have also found the CAM-ICU relatively insensitive among non-intubated patients — for example, in a non-critically ill cohort of oncology patients, the sensitivity of the CAM-ICU was only 18%.[27]

Knowing that a patient is delirious is useful only if this precipitates effective therapy or predicts outcome. Like other studies,[1,28,29] we found delirium identified by both assessment methods predicted longer ICU and hospital stay. When performed by dedicated research personnel, delirium detected by the CAM-ICU predicted higher mortality,[1,4,30] but in our study, this was true of only the unstructured assessments, suggesting these may be the more clinically useful technique in our setting.

Our study has several limitations. We did not assess the same patients with both unstructured examinations and the CAM-ICU, as we were interested in the real-world performance of these techniques by bedside ICU nurses and we felt the results of the unstructured assessment might be influenced by that of the CAM-ICU performed in the same shift by the same nurse. If the incidence or prevalence of delirium was different in the two periods, we will have made an unfair comparison. However, the demographic characteristics, severity of illness mortality and length of stay were similar in the two groups, suggesting the true rates of delirium might also have been similar. We cannot report the sensitivity or specificity of either of the assessment techniques we used, as we have no gold-standard assessment for delirium.

An important impediment to all ICU delirium research is the lack of a suitable gold standard. Other studies have used analysis of the medical record,[31] agitation on a sedation scale and a doctor's unstructured assessment,[32] and the mini-mental state examination.[33] The most commonly used criteria are those of the *Diagnostic and statistical manual of mental disorders*, fourth edition, text revision (DSM-IV-TR),[34] performed by anyone from interns[35] to "delirium experts".[11] The DSM-IV-TR is known to have low interrater reliability for many disorders,[36-39] in part because of dependence on the experience of the assessors.[40] The criteria for delirium in the DSM-IV-TR (disturbance of consciousness with reduced ability to focus or shift attention, a change in cognition or development of a perceptual disturbance, and development over a short period and with a fluctuating course) are open to interpretation, and not tailored to critically ill patients. For example, how should an intubated patient be assessed for reduced ability to shift attention? How is a fluctuating course assessed in patients receiving varying rates of sedatives or boluses of sedating analgesics? The unstructured technique used in this study, while apparently more sensitive (ie, diagnosing more delirium), may have been less specific than the CAM-ICU. However, delirium diagnosed by unstructured assessments had greater prognostic significance, suggesting greater validity. Furthermore, the overall incidence of delirium diagnosed by unstructured assessments was half that reported in other studies, which also argues against low specificity.

Our nurses may not have performed the CAM-ICU satisfactorily. The CAM-ICU study period followed only 30 days of intensive education and practice. Others have found the ability to reproducibly perform the CAM-ICU improves with practice.[24] We did not assess interrater reliability for each assessment technique. However, the diagnosis of delirium in two consecutive shifts is a reasonable surrogate, albeit one that underestimates reliability, as there will be discordance if the patient's condition changes. The CAM-ICU was one-third as likely to diagnose delirium in two consecutive shifts as the unstructured assessment method. It is possible that, had we persisted with the CAM-ICU, its accuracy may have improved.

We made no distinction between hyper- and hypoactive delirium in our unstructured assessments. However, neither does the CAM-ICU, unless the Richmond Agitation Sedation Scale[41] is simultaneously reported. Hypoactive delirium is more common among ICU patients[10,23,42] and is associated with a worse prognosis.[42] Although there are effective treatments for hyperactive delirium,[7,8] therapy was no better than placebo in a population with mixed delirium subtypes.[43] Hyper- and hypoactive delirium may therefore represent different disease processes, and future studies should make this simple distinction.

Conclusions

We conclude that, in our ICU, unstructured assessment diagnoses more patients with delirium than does the CAM-

ICU. Like others,[26] we sound a note of caution that the CAM-ICU in practice is not sufficiently sensitive. Outside research studies, where there is a greater need for objective, reproducible measures, in our hospital we have therefore decided against adoption of the CAM-ICU. However, if we had not asked nurses to make their unstructured assessments and record the results each shift, we may have identified fewer patients with delirium. We will therefore continue to require our nurses to actively seek delirium using their unstructured assessments, with the hope that application of one of the promising new treatments for ICU-acquired delirium will improve outcomes for such patients.

Competing interests

None declared.

Author details

Michael C Reade, Consultant Intensivist,[1] and Associate Professor;[2] currently, Professor of Military Medicine and Surgery, Australian Defence Force and Burns Trauma and Critical Care Research Centre, University of Queensland, Brisbane, QLD, Australia
Glenn M Eastwood, Research Manager[1]
Leah Peck, Research Coordinator[1]
Rinaldo Bellomo, Director of Research[1]
Ian Baldwin, Professor[1]
1 Intensive Care Unit, Austin Hospital, Melbourne, VIC, Australia.
2 University of Melbourne, Melbourne, VIC, Australia.
Correspondence: m.reade@uq.edu.au

References

1 Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA* 2004; 291: 1753-62.
2 Pandharipande P, Cotton BA, Shintani A, et al. Prevalence and risk factors for development of delirium in surgical and trauma intensive care unit patients. *J Trauma* 2008; 65: 34-41.
3 Ouimet S, Kavanagh BP, Gottfried SB, Skrobik Y. Incidence, risk factors and consequences of ICU delirium. *Intensive Care Med* 2007; 33: 66-73.
4 Shehabi Y, Riker RR, Bokesch PM, et al. Delirium duration and mortality in lightly sedated, mechanically ventilated intensive care patients. *Crit Care Med* 2010; 38: 2311-8.
5 Van Rompaey B, Schuurmans MJ, Shortridge-Baggett LM, et al. Long term outcome after delirium in the intensive care unit. *J Clin Nurs* 2009; 18: 3349-57.
6 Balas MC, Happ MB, Yang W, et al. Outcomes associated with delirium in older patients in surgical ICUs. *Chest* 2009; 135: 18-25.
7 Devlin JW, Roberts RJ, Fong JJ, et al. Efficacy and safety of quetiapine in critically ill patients with delirium: a prospective, multicenter, randomized, double-blind, placebo-controlled pilot study. *Crit Care Med* 2010; 38: 419-27.
8 Reade MC, O'Sullivan K, Bates S, et al. Dexmedetomidine vs. haloperidol in delirious, agitated, intubated patients: a randomised open-label trial. *Crit Care* 2009; 13: R75.
9 Klugkist M, Sedemund-Adib B, Schmidtke C, et al. [Confusion Assessment Method for the Intensive Care Unit (CAM-ICU): diagnosis of postoperative delirium in cardiac surgery] [German]. *Anaesthesist* 2008; 57: 464-74.
10 Pandharipande P, Cotton BA, Shintani A, et al. Motoric subtypes of delirium in mechanically ventilated surgical and trauma intensive care unit patients. *Intensive Care Med* 2007; 33: 1726-31.
11 Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA* 2001; 286: 2703-10.
12 Bergeron N, Dubois MJ, Dumont M, et al. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med* 2001; 27: 859-64.
13 Plaschke K, von Haken R, Scholz M, et al. Comparison of the confusion assessment method for the intensive care unit (CAM-ICU) with the Intensive Care Delirium Screening Checklist (ICDSC) for delirium in critical care patients gives high agreement rate(s). *Intensive Care Med* 2008; 34: 431-6.
14 van Eijk MM, van Marum RJ, Klijn IA, et al. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med* 2009; 37: 1881-5.
15 Ceraso DH, Duenas-Castel C, Raimondi N, et al. [Latin American survey on delirium in critical patients] [Spanish]. *Med Intensiva* 2010; 34: 495-505.
16 Ely EW, Stephens RK, Jackson JC, et al. Current opinions regarding the importance, diagnosis, and management of delirium in the intensive care unit: a survey of 912 healthcare professionals. *Crit Care Med* 2004; 32: 106-12.
17 Mac Sweeney R, Barber V, Page V, et al. A national survey of the management of delirium in UK intensive care units. *QJM* 2010; 103: 243-51.
18 Mistarz R, Eliott S, Whitfield A, Ernest D. Bedside nurse-patient interactions do not reliably detect delirium: An observational study. *Aust Crit Care* 2011; 24: 126-32.
19 Spronk PE, Riekerk B, Hofhuis J, et al. Occurrence of delirium is severely underestimated in the ICU during daily care. *Intensive Care Med* 2009; 35: 1276-80.
20 Shehabi Y, Botha JA, Boyle MS, et al. Sedation and delirium in the intensive care unit: an Australian and New Zealand perspective. *Anaesth Intensive Care* 2008; 36: 570-8.
21 Nelson LS. Teaching staff nurses the CAM-ICU for delirium screening. *Crit Care Nurs Q* 2009; 32: 137-43.
22 Stow PJ, Hart GK, Higlett T, et al. Development and implementation of a high-quality clinical database: the Australian and New Zealand Intensive Care Society Adult Patient Database. *J Crit Care* 2006; 21: 133-41.
23 Peterson JF, Pun BT, Dittus RS, et al. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc* 2006; 54: 479-84.
24 van den Boogaard M, Pickkers P, van der Hoeven H, et al. Implementation of a delirium assessment tool in the ICU can influence haloperidol use. *Crit Care* 2009; 13: R131.
25 Pun BT, Gordon SM, Peterson JF, et al. Large-scale implementation of sedation and delirium monitoring in the intensive care unit: a report from two medical centers. *Crit Care Med* 2005; 33: 1199-205.
26 van Eijk MM, van den Boogaard M, van Marum RJ, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *Am J Respir Crit Care Med* 2011; 184: 340-4.
27 Neufeld KJ, Hayat MJ, Coughlin JM, et al. Evaluation of two intensive care delirium screening tools for non-critically ill hospitalized patients. *Psychosomatics* 2011; 52: 133-40.
28 Thomason JW, Shintani A, Peterson JF, et al. Intensive care unit delirium is an independent predictor of longer hospital stay: a prospective analysis of 261 non-ventilated patients. *Crit Care* 2005; 9: R375-81.
29 Ely EW, Gautam S, Margolin R, et al. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med* 2001; 27: 1892-900.

30 van den Boogaard M, Peters SA, van der Hoeven JG, et al. The impact of delirium on the prediction of in-hospital mortality in intensive care patients. *Crit Care* 2010; 14: R146.
31 Pisani MA, Araujo KL, Van Ness PH, et al. A research algorithm to improve detection of delirium in the intensive care unit. *Crit Care* 2006; 10: R121.
32 Otter H, Martin J, Basell K, et al. Validity and reliability of the DDS for severity of delirium in the ICU. *Neurocrit Care* 2005; 2: 150-8.
33 McNicoll L, Pisani MA, Ely EW, et al. Detection of delirium in the intensive care unit: comparison of confusion assessment method for the intensive care unit with confusion assessment method ratings. *J Am Geriatr Soc* 2005; 53: 495-500.
34 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Text revision. Arlington, Va: APA, 2000.
35 Immers HE, Schuurmans MJ, van de Bijl JJ. Recognition of delirium in ICU patients: a diagnostic study of the NEECHAM confusion scale in ICU patients. *BMC Nurs* 2005; 4: 7.
36 Hogervorst E, Bandelow S, Combrinck M, et al. The validity and reliability of 6 sets of clinical criteria to classify Alzheimer's disease and vascular dementia in cases confirmed post-mortem: added value of a decision tree approach. *Dement Geriatr Cogn Disord* 2003; 16: 170-80.
37 Lipton AA, Simon FS. Psychiatric diagnosis in a state hospital: Manhattan state revisited. *Hosp Community Psychiatry* 1985; 36: 368-73.
38 Schmidt HO, Fonda CP. The reliability of psychiatric diagnosis; a new look. *J Abnorm Psychol* 1956; 52: 262-7.
39 Spitzer RL, Fleiss JL. A re-analysis of the reliability of psychiatric diagnosis. *Br J Psychiatry* 1974; 125: 341-7.
40 Klin A, Lang J, Cicchetti DV, Volkmar FR. Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: results of the DSM-IV autism field trial. *J Autism Dev Disord* 2000; 30: 163-7.
41 Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med* 2002; 166: 1338-44.
42 Robinson TN, Raeburn CD, Tran ZV, et al. Motor subtypes of postoperative delirium in older adults. *Arch Surg* 2011; 146: 295-300.
43 Girard TD, Pandharipande PP, Carson SS, et al. Feasibility, efficacy, and safety of antipsychotics for intensive care unit delirium: the MIND randomized, placebo-controlled trial. *Crit Care Med* 2010; 38: 428-37. ❑