Published in final edited form as:
J Am Geriatr Soc. 2011 November ; 59(Suppl 2): S249–S255. doi:10.1111/j.1532-5415.2011.03673.x.

Delirium and Sedation Recognition Using Validated Instruments: Reliability of Bedside ICU Nursing Assessments from 2007 to 2010

Eduard E. Vasilevskis, MD[*,†,‖,¶,#], **Alessandro Morandi, MD**[**], **Leanne Boehm, MSN., RN, ACNS-BC**[§§], **Pratik P. Pandharipande, MD, MSCI**[§,††], **Timothy D. Girard, MD, MSCI**[*,‖,¶,**], **James C Jackson, PsyD**[**], **Jennifer Thompson, MPH**[‡‡], **Ayumi Shintani, PhD, MPH**[‡‡], **Sharon Gordon, PsyD**[*,‡], **Brenda T. Pun, RN, MSN., ACNP**[§§], and **E. Wesley Ely, MD, MPH**[*,‖,¶,**]

[*]Geriatric Research, Education and Clinical Center (GRECC) Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, TN

[†]Clinical Research Training Center of Excellence (CRCoE), Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, TN

[‡]Mental Health Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, TN

[§]Anesthesia Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville, TN

[‖]Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN

[¶]Center for Health Services Research, Vanderbilt University School of Medicine, Nashville, TN

[#]Division of General Internal Medicine and Public Health, Vanderbilt University School of Medicine, Nashville, TN

[**]Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University School of Medicine, Nashville, TN

[††]Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN

Corresponding Author: Eduard E. Vasilevskis, M.D., Vanderbilt University Medical Center; 1215 21st Ave, S., 6006 Medical Center East, NT; Nashville, TN 37232-8300; Voice 615-936-1935, Fax 615-936-1269, eduard.vasilevskis@vanderbilt.edu. Alternate Corresponding Author: E. Wes Ely, wes.ely@vanderbilt.edu.

Disclaimer: The views expressed in this article are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs.

Author Contributions:
Study Concept: Eduard E. Vasilevskis, M.D, Alessandro Morandi, Pratik P. Pandharipande, Timothy D. Girard, James C. Jackson, Sharon Gordon, E Wes Ely
Study Design: Eduard E. Vasilevskis, M.D, Alessandro Morandi, Pratik P. Pandharipande, Timothy D. Girard, James C Jackson, Jennifer Thompson, Ayumi Shintani, Sharon Gordon, Brenda T. Pun, E Wes Ely
Acquisition of Data: Eduard E. Vasilevskis, M.D, Alessandro Morandi, Leanne Boehm, Pratik P. Pandharipande, Timothy D. Girard, James C Jackson, Jennifer Thompson, Ayumi Shintani, Brenda T. Pun, E Wes Ely
Data Analysis and/or Interpretation: Eduard E. Vasilevskis, M.D, Alessandro Morandi, Leanne Boehm, Pratik P. Pandharipande, Timothy D. Girard, James C Jackson, Jennifer Thompson, Ayumi Shintani, Sharon Gordon, Brenda T. Pun, E Wes Ely
Preparation of Manuscript: Eduard E. Vasilevskis, M.D, Alessandro Morandi, Leanne Boehm, Pratik P. Pandharipande, Timothy D. Girard, James C Jackson, Jennifer Thompson, Ayumi Shintani, Sharon Gordon, Brenda T. Pun, E Wes Ely

Conflicts of Interest: Dr. Pandharipande has received honoraria from Hospira, Inc and GlaxoSmithKline. Dr. Girard has received honoraria from Hospira, Inc. Dr. Ely has received grant support and honoraria from Eli Lilly, Pfizer, Hospira, Aspect Medical Systems, and GlaxoSmithKline. Ms. Brenda T. Pun has received honoraria from Hospira. Ms. Leanne Boehm has received honoraria from Hospira.

‡‡Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, TN

§§Vanderbilt University School of Nursing, Nashville, TN

Abstract

OBJECTIVES—The aim of this study was to describe the reliability and sustainability of delirium and sedation measurements by bedside intensive care unit (ICU) nurses.

DESIGN—Prospective cohort study.

SETTING—A tertiary care academic medical center.

PARTICIPANTS—510 ICU patients from 2007 to 2010. 627 bedside nurses.

MEASUREMENTS—Delirium and sedation levels were independently measured in routine care by bedside nurses and well-trained reference-rater research nurses. Bedside nurses were instructed to use the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) every 12 hours and the Richmond Agitation-Sedation Scale (RASS) every 4 hours to measure delirium and sedation, respectively. CAM-ICU and RASS assessment agreement were computed using weighted kappa statistics across the entire population and subgroups (e.g., ICU type). Sensitivity and specificity of bedside nurse identification of delirium were calculated to understand sources of discordance.

RESULTS—6198 CAM-ICU and 6880 RASS measurement pairs obtained on 3846 patient-days. For CAM-ICU measurements, agreement between bedside and research nurses was substantial (weighted kappa = 0.67, 95% CI 0.66 to 0.70) and stable over 3 years of data collection. RASS measures also demonstrated substantial agreement (weighted kappa = 0.66, 95% CI 0.64 to 0.68), which was stable across all years of data collection. The sensitivity and specificity of delirium nurse assessments was 0.81 (95% CI 0.78 to 0.83) and 0.81 (95% CI 0.78 to 0.85), respectively.

CONCLUSION—Bedside nurse measurements of delirium and sedation are sustainable and reliable sources of information. These measures can be used for clinical decision making, quality improvement and measurement activities.

Keywords

Delirium; Sedation; Measurement; Critical Care; Nursing

INTRODUCTION

Acute brain dysfunction in the form of delirium or coma affects up to 80% of critically ill patients and is extremely hazardous to the affected patient.[1–4] Those suffering from delirium experience longer hospitalizations,[5;6] three-fold increase in mortality that persists up to one year after discharge,[7–9] and long-term brain dysfunction.[10–12]

As the body of evidence demonstrating the harmful effects of acute brain dysfunction has increased, so has awareness among providers.[13] Moreover, the ability to measure acute brain dysfunction in the ICU has substantially improved with the development of instruments that accurately and reliably measure delirium[2;3;14;15] and sedation.[16–19] Routine use of these tools is recommended in authoritative guidelines,[20] is a central component to structured approaches for delirium prevention,[21–23] and is necessary to uncover delirium that would otherwise go undetected in up to 72% of patients without the use of a validated instrument.[24;25] The dangers of potent sedative drugs[26;27] demand that providers monitor patients for acute brain dysfunction.

Despite this, concerns about the time and expertise needed to use validated delirium and sedation assessments and the perceived usefulness of information remain barriers to adoption.[28;29] In order to improve adoption of validated measurement instruments,[29–32] it is important to demonstrate both the feasibility and reliability of implementing these tools in clinical practice. To date, studies have demonstrated short-term feasibility and reliability of implementing delirium and sedation instruments. Devlin et al.,[33;34] for example, demonstrated improvements in detection of delirium and sedation following a scenario-based educational intervention for providers. Three additional studies showed high compliance (at least 92%) in achieving at least one bedside Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) measure per nursing shift in the 2-months following implementation.[35–37] One of these also reported outstanding reliability of bedside measures (kappa range 0.78 to 0.89).[37] Finally, Pun et al., performed 6- and 12-month implementation trials at two institutions.[28] At 3-month intervals, compliance and reliability of delirium and sedation measures was outstanding.

Initial implementation efforts demonstrated short-term feasibility and reliability of delirium and sedation monitoring. Importantly, there are no data to indicate how these measures perform outside the context of an active implementation campaign and over prolonged periods of time. This information is critical since implementation campaigns bring increased resources, attention, and leadership. It is important to understand measure performance under "usual care," where this information may be used for clinical decisions, quality improvement, and surveillance of acute brain dysfunction.

The primary aim of this study was to examine the recognition of delirium and sedation by ICU bedside nurses using validated tools during routine clinical care compared with concurrent measurements made by well-trained research personnel over a three year period.

METHODS

Patients

This study includes critically ill patients admitted to the 34-bed medical ICU (MICU), 27-bed cardiovascular ICU (CVICU), or 34-bed surgical ICU (SICU) between March 2007 and May 2010 at Vanderbilt University Medical Center (VUMC), a 900-bed teaching hospital. Patients with states preventing administration of the English version CAM-ICU were excluded: severe baseline cognitive impairment, anoxic brain injury, moribund state, and inability to understand English. Patients for whom follow-up (done for a separate study) was anticipated to be low or impossible were also excluded: lack of informed consent, blindness, active substance abuse, psychotic disorder, or recent suicidal gesture. The institutional review board at VUMC approved the study.

Delirium and Sedation Monitoring Instruments

All ICUs at VUMC included in this study have implemented monitoring for delirium (CAM-ICU[2;3]) and sedation (RASS[16]) as part of routine care. The CAM-ICU is a valid delirium assessment tool designed for bedside health care providers.[2;3] CAM-ICU results are documented as positive, negative, or unable to assess (i.e., coma). The RASS is a ten-level sedation scale ranging from −5 (unarousable) to +4 (combative).[16] These tools were chosen for use at VUMC over alternate instruments because of feasibility, excellent testing characteristics, reproducibility, and local expertise.

Education and Training in Delirium and Sedation Monitoring Instruments

Routine use of the CAM-ICU and RASS was implemented in the MICU at VUMC in Jan 2002 as part of a one year implementation project.[28] SICUs implemented use of the same

tools more recently, in 2006. The CVICU has used validated measures of delirium and sedation since prior to 2007, but did not undergo a formal intensive educational intervention until 2010. Since 2007, graduate nurses and those with <1 year of experience in nursing are hired into Vanderbilt's Nurse Residency Program, in which there is a lecture dedicated to ICU delirium and its identification. Specifically, nurses are provided with instruction for performing the CAM-ICU and RASS assessments. In addition, all newly hired MICU and SICU nurses receive training in CAM-ICU and RASS assessments, followed by competency assessment. SICU nurses additionally complete an online educational module covering pain, anxiety, and delirium that incorporate the use of CAM-ICU and RASS tools. Assessments of CAM-ICU and RASS competency and reliability are not routinely performed across all units. Repeat training and competency assessments occur if deficiencies are identified by those charged with competency assessment and are continued until competency is achieved. VUMC bedside nurses are instructed to assess for delirium once every 12 hours and for sedation every 4 hours, though nurses are able to document additional assessments when performed. Results are entered into the electronic medical record nursing flowsheet every 4 hours, or sooner for more frequent assessments.

Research personnel (N = 6) received thorough training from the CAM-ICU and RASS creators, including training manual review, videos, role playing, and a manual of standard operating procedures. These personnel are now all trainers who educate others in the administration of the tools.

Patient Demographics and Monitoring of Reliability

Demographic data including APACHE II severity of illness score,[38] admission diagnoses, type of ICU admission (medical vs. surgical) and pre-existing comorbidities[39] including depression and cognitive impairment[40] were recorded at admission by research personnel. The bedside nurse's documentation of the CAM-ICU and the RASS was obtained from the patients' electronic medical record. Research personnel assessed CAM-ICU and RASS twice per day for each day during a patient's ICU stay.

Statistical Analysis

Patients' demographic and clinical variables are presented using means and standard deviations (SD) or medians and interquartile ranges (IQR) for continuous variables, and proportions for categorical variables. Assessment frequency was calculated as the total number and mean number of clinical CAM-ICU or RASS assessments for each patient day. To determine agreement between bedside and research nurses' assessments of the CAM-ICU and RASS, we paired assessments that were no more than 4 hours apart and calculated weighted kappa statistics.[41;42] When multiple bedside assessments were available within the given time frame, we examined the beside nurse assessment carried out closest to the time of the research assessment. Because patients could be assessed multiple times over their ICU stay by both clinical bedside nurses and research personnel, we used bootstrapped confidence intervals to account for the correlation between these repeated measurements. Specifically, the kappa statistics obtained from original data were used as the point estimates of each comparison. We then used random sampling with replacement to select a new cohort from the original cohort and recalculated the kappa statistic for that set of patients. This process of resampling was repeated 1,000 times, each one generating another kappa statistic; the 2.5th and 97.5th percentiles of these 1,000 kappa statistics were considered the lower and upper confidence limits, respectively. In keeping with widely accepted standards, interpretation of kappa statistics were as follows: values < 0 indicate no agreement, 0 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1.0 as almost perfect agreement.[43]

Whereas agreement quantifies the inter-rater reliability, after correcting for chance agreement, this metric does not describe the details of disagreement. The kappa statistic provides no "directionality" of the disagreements, i.e., whether bedside nurses under-diagnose (false negatives) or over-diagnose (false positives) delirium when compared to research nurses (when used as reference standards). To describe the nature of disagreement, we calculated sensitivity and specificity of bedside nurse assessments of delirium as compared with the research nurse assessments, which we considered as a reference standard. As with analyses of agreement, we paired assessments that were no more than 4 hours apart. For all analyses, MICU and CVICU delirium and sedation measurements are reported together and are grouped under "Medical." All statistical analyses were performed with R version 2.12 (www.r-project.org).

RESULTS

Patient Characteristics

A total of 510 patients were included in this investigation. Their median age was 58 (IQR, 47 to 67), 12% had cognitive impairment prior to ICU admission, a large percentage (91%) were mechanically ventilated, and the median APACHE II score at enrollment was 27 (IQR, 21 to 33), indicating a high severity of illness (Table 1).

Frequency and Sustainability of Assessments

There were 627 unique nurses that practiced within the MICU, CVICU, and SICU during the study period. Bedside nurses assessed individual patients for delirium (CAM-ICU) an average of 7.4 (SD = 2.9) times per patient-day, ranging from 7 (SD 3) to 7.8 (SD 3) over the study period. This corresponds to a bedside delirium measurement about every 3.2 hours. Bedside nurses assessed patient's level of sedation (RASS) an average of 8.6 (SD = 2.8) times per patient-day, ranging from 7.7 (SD 2.4) to 9.1 (SD 3) over the study period. This corresponds to a bedside sedation measurement about every 2.8 hours.

Beside Nurse-Research Nurse Agreement

Among 510 patients, there were 6198 CAM-ICU and 6880 RASS measurement pairs within four hours on 3846 patient-days; 98% of these paired assessments occurred ≤ 2 hours apart. Tables 2 and 3 display the inter-rater reliability of the independent CAM-ICU and RASS measures performed within 2 hours of each other. Across the entire population there was substantial agreement between bedside and research nurses on measures done within 2 hours of each other (CAM-ICU weighted kappa = 0.67, [95% CI 0.66 to 0.70] and RASS weighted kappa = 0.66, [95% CI 0.64 to 0.68]). Comparable agreement was seen for measures within 4 hours. Agreement was stable regardless of severity of illness, preexisting cognitive impairment, and study year. Agreement was slightly lower, but still substantial, in the surgical compared with the medical units, mechanically ventilated patients, and among patients 65 years and older compared with those < 65 years old.

Sensitivity and Specificity of Bedside Nurse Identification of Delirium

Among 3856 paired assessments for delirium (excluding comatose patients) within 2 hours, bedside nurse identified delirium with a sensitivity of 0.81 (95% CI 0.78 to 0.83) and a specificity of 0.81 (95% CI 0.78 to 0.85), compared to the research nurse reference standard (Table 4). These results correspond to a positive and negative likelihood ratio of 4.3 (95% CI 3.6, 5.2) and 0.2 (95% CI 0.2, 0.3), respectively. Specificity was highest for bed side nurse assessments in the medical ICU. Sensitivity was lowest for the least severely ill patients. Sensitivity was stable by age and year of data collection. There were fewer false

positives among younger patients as indicated by increased specificity. Results were unchanged for assessments compared up to 4 hours apart.

DISCUSSION

The main finding of this investigation is that delirium and sedation assessments in the ICU using validated instruments are reliable and sustainable in usual clinical practice. Despite prolonged lapses in time from initial implementation efforts, medical and surgical ICU nurses demonstrated substantial and stable agreement with research personnel in the assessment of both delirium (CAM-ICU weighted kappa = 0.67) and sedation (RASS weighted kappa = 0.66). Sensitivity and specificity of bedside delirium diagnosis were 0.81 and 0.81, suggesting balanced false positive and negative assessments. Critical care providers can confidently use bedside nurse measurements of delirium and sedation to assist with the diagnosis, timely medical decisions (e.g., sedation adjustments), and quality or research monitoring.

This study builds upon prior implementation efforts. First, Pun et al. demonstrated successful implementation of CAM-ICU and RASS tools across two medical ICUs (one of which was VUMC).[28] Following 6 to 12 months of staged training, nurses achieved outstanding measurement compliance (>95% for CAM-ICU and RASS) and agreement between bedside and reference rater nurses (CAM-ICU kappa 0.75 to 0.92, RASS kappa 0.77 to 0.89). Second, following brief education and training, Devlin et al. demonstrated substantial agreement in the diagnosis of delirium between practicing physicians and reference judges, with a post-training kappa of 0.6.[34] The agreement observed in these prior studies is generally equal to or higher than what we observed. This difference is not unexpected since the studies compared measurements between clinicians and trained researchers during or immediately after implementation efforts, which included implementation champions, visible marketing campaigns, and individualized detailing to drive change.

This study reflects a more common scenario of delirium and sedation training, which occurs as just two of many tasks and responsibilities ICU nurses must learn and prioritize as part of their regular clinical duties. The measurements in the MICU occurred a full 5 years after the initial implementation project, and 1 year following the SICU introduction, both representing a significant “wash-out” period between initial educational interventions at time of implementation. The CVICU had not had a formal implementation phase until after study completion. This study demonstrates that measures of delirium and sedation can be reliably performed in an environment where monitoring with validated tools is integrated into the usual course of clinical care. The overall compliance and sustainability of measurement was outstanding with more than seven delirium and sedation measurements per patient-day, more than was expected of nurses. Moreover, frequency of measurement and measure agreement remained stable over more than three years of the study period. It is remarkable to the extent that measurement was performed beyond that which was required. The dedication to frequent measurement may be explained by the clinical value that is perceived by nurses. As the first detectors of such an important clinical change, nurses may feel increased autonomy as well as advocacy for their patients in detecting brain dysfunction that may otherwise be missed without their assessment.

In addition to stability over time, delirium and sedation measurements appear to be robust across key subgroups. The CAM-ICU was originally developed as a tool that would assist in the evaluation of critically ill patients who frequently require mechanical ventilation, an impediment to verbal communication.[2;3] In line with developer goals, there did not appear to be any significant changes in the reliability of either CAM-ICU or RASS measurements

across increasing severity of illness and remained substantial among mechanically ventilated patients. These findings are highly encouraging, and support the value of these tools in the critical care setting. In addition to illness severity, we examined assessments according to evidence of pre-ICU cognitive impairment and depression, two conditions that may hypothetically complicate diagnosis of delirium. Even among these two groups there were minimal clinically important differences. Of note, agreement and specificity of delirium appeared lower in older patients and specifically in those older than 65 years of age. This is an interesting finding since we did not detect significant difference according to presence of depression and cognitive impairment as one might have expected. Nonetheless, these data suggest health care providers should be more careful when assessing elderly patients for delirium since there may be increased potential for falsely over-diagnosing acute brain dysfunction in this group. Likewise, agreement and specificity appeared to be lower in the surgical units. A higher number of false positives in the surgical population may be explained by differences in training between units and/or patient differences (e.g., higher levels of pain, more limitations in mobility) that may be falsely identified as delirium or limit assessment.

There are several aspects of this study that strengthen the findings. First, the quality and consistency of the independent rater measurements for delirium and sedation allow for reliable comparisons over time. In addition, the large sample size and examination of important subgroups that may affect delirium and sedation measurement (e.g., mechanical ventilation status, premorbid cognitive status) demonstrated patient generalizability. This study is the first of its kind to examine bedside nursing measurements with validated instruments outside of the scope of an implementation trial and therefore represents effectiveness rather than efficacy. Finally, our ability to examine the measurements over more than three years provides substantial evidence of sustainability and validity of the assessments in a clinical setting where results can be relied upon for clinical decision-making, quality improvement, and surveillance over extended periods.

Several important limitations deserve attention as well. First, this study was performed at a large academic teaching hospital, and therefore findings may not generalize to all settings. Centers with lower nurse turnover or more intense ongoing training and audit procedures may see improved agreement. This single institution, however, represents a broad population of patients, across hundreds of individual nursing observation, and includes both medical and surgical ICUs. Second, measures were performed in the ICU and may not generalize to a ward setting. We believe this is still encouraging data, and further exploration among medical and surgical ward patients is warranted. An additional limitation is that patients with severe dementia or primary neurologic disease that would confound diagnosis of delirium were excluded, and results may not apply to these patients. Agreement may be decreased, in fact, in centers that care for large proportions of patients with baseline neurologic illness. Patients with mild or moderate dementia and pre-existing depression were, however, included in this study, providing important information on this subset of population. In addition, after initial training, educational procedures for clinical nurses are not standardized and somewhat heterogeneous. Once again, this reflects “real-world” limitations in an environment with multiple priorities and constrained resources. It is conceivable that with more intensive continuing education and training with standardized auditing procedures, agreement may increase to a level seen immediately following the implementation effort. Finally, paired measurements were not performed at the exact same time. As mental status may change quickly during critical illness, these results may underestimate the level of agreement that would be seen if measures were simultaneous.

In conclusion, the use of potent sedative agents[26;27] in the ICU, combined with evidence of poor mortality[7;8] and cognitive[12] outcomes among patients with delirium, demand that

critical care providers monitor brain function with equal attention to other vital signs. The ability of ICU nurses to monitor delirium and sedation with validated instruments has been a significant advance toward meeting this goal, and central to future prevention strategies.[21;22] We demonstrate that with commitment to their use, validated delirium and sedation measurement tools can be reliably and sustainably employed in the usual course of clinical care. By increasing confidence in the reliability and sustainability of such measures, we may further eliminate the barriers to more widespread adoption of delirium and sedation monitoring worldwide. Without the ability to monitor for acute brain dysfunction, treatment strategies cannot be appropriately targeted, evaluation of the effectiveness of prevention programs will not be possible, and measurement for quality improvement hindered. Future studies need to examine the effect of ongoing education and audit strategies to further maximize reliability and understand how we can best utilize this source of information to guide clinical decisions, monitor quality, and improve patient outcomes.

Acknowledgments

We acknowledge Christine Bass, MPH, Department of Medicine (General Internal Medicine and Public Health), Center for Health Services Research, Vanderbilt University, Nashville, TN, for her administrative and editorial assistance in the preparation of this manuscript.

Funding was provided by the National Institutes of Health (AG027472, AG034257, and AG031322), Veterans Affairs Clinical Research Center of Excellence, Veterans Affairs Merit Award, Veterans Affairs Career Development Award, and the Tennessee Valley Geriatric Research, Education and Clinical Center (GRECC).

Sponsor's Role: None of the funding agencies had any role in the conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript.

References

1. Dubois MJ, Bergeron N, Dumont M, et al. Delirium in an intensive care unit: A study of risk factors. Intensive Care Med. 2001; 27:1297–1304. [PubMed: 11511942]
2. 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; 286:2703–2710. [PubMed: 11730446]
3. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: Validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). Crit Care Med. 2001; 29:1370–1379. [PubMed: 11445689]
4. Pisani MA, Murphy TE, Van Ness PH, et al. Characteristics associated with delirium in older patients in a medical intensive care unit. Arch Intern Med. 2007; 167:1629–1634. [PubMed: 17698685]
5. 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; 27:1892–1900. [PubMed: 11797025]
6. 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–R381. [PubMed: 16137350]
7. Pisani MA, Kong SYJ, Kasl SV, et al. Days of delirium are associated with 1-year mortality in an older intensive care unit population. Am J Respir Crit Care Med. 2009; 180:1092–1097. [PubMed: 19745202]
8. 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; 291:1753–1762. [PubMed: 15082703]
9. Shehabi Y, Riker RR, Bokesch PM, et al. Delirium duration and mortality in lightly sedated, mechanically ventilated intensive care unit patients. Critical Care Medicine. 2010; 38:2311–2318. [PubMed: 20838332]
10. Jackson JC, Gordon SM, Hart RP, et al. The association between delirium and cognitive decline: A review of the empirical literature. Neuropsychol Rev. 2004; 14:87–98. [PubMed: 15264710]

11. Rudolph JL, Inouye SK, Jones RN, et al. Delirium: An independent predictor of functional decline after cardiac surgery. J Am Geriatr Soc. 2010; 58:643–649. [PubMed: 20345866]
12. Girard TD, Jackson JC, Pandharipande PP, et al. Duration of delirium as a predictor of long-term cognitive impairment in survivors of critical illness. Crit Care Med. 2010; 38:1513–1520. [PubMed: 20473145]
13. 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; 32:106–112. [PubMed: 14707567]
14. Bergeron N, Dubois MJ, Dumont M, et al. Intensive Care Delirium Screening Checklist: Evaluation of a new screening tool. Intensive Care Med. 2001; 27:859–864. [PubMed: 11430542]
15. Luetz A, Heymann A, Radtke FM, et al. Different assessment tools for intensive care unit delirium: Which score to use? Crit Care Med. 2010; 38:409–418. [PubMed: 20029345]
16. Ely EW, Truman B, Shintani A, et al. Monitoring sedation status over time in ICU patients: Reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA. 2003; 289:2983–2991. [PubMed: 12799407]
17. Ramsay MA, Savege TM, Simpson BR, et al. Controlled sedation with alphaxalone-alphadolone. BMJ. 1974; 2:656–659. [PubMed: 4835444]
18. Riker RR, Picard JT, Fraser GL. Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. Crit Care Med. 1999; 27:1325. [PubMed: 10446827]
19. Weinert C, McFarland L. The state of intubated ICU patients. Chest. 2004; 126:1883–1890. [PubMed: 15596688]
20. Jacobi J, Fraser GL, Coursin DB, et al. Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult. Crit Care Med. 2002; 30:119–141. [PubMed: 11902253]
21. Vasilevskis EE, Pandharipande PP, Girard TD, et al. A screening, prevention, and restoration model for saving the injured brain in intensive care unit survivors. Crit Care Med. 2010; 38:S683–S691. [PubMed: 21164415]
22. Vasilevskis EE, Ely EW, Speroff T, et al. ICU-acquired delirium and weakness - Crossing the quality chasm. Chest. 2010; 138:1224–1233. [PubMed: 21051398]
23. Skrobik Y, Ahern Sp, Leblanc M, et al. Protocolized intensive care unit management of analgesia, sedation, and delirium improves analgesia and subsyndromal delirium rates. Anesth Analg. 2010; 111:451–463. [PubMed: 20375300]
24. Spronk PE, Riekerk B, Hofhuis J, et al. Occurrence of delirium is severely underestimated in the ICU during daily care. Intensive Care Med. 2009; 35:1276–1280. [PubMed: 19350214]
25. 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; 37:1881–1885. [PubMed: 19384206]
26. Pandharipande P, Shintani A, Peterson J, et al. Lorazepam is an independent risk factor for transitioning to delirium in intensive care unit patients. Anesthesiology. 2006; 104:21–26. [PubMed: 16394685]
27. Pisani MA, Murphy TE, Araujo KLB, et al. Benzodiazepine and opioid use and the duration of intensive care unit delirium in an older population. Crit Care Med. 2009; 37:177–183. [PubMed: 19050611]
28. 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; 33:1199–1205. [PubMed: 15942331]
29. Tanios MA, de Wit M, Epstein SK, et al. Perceived barriers to the use of sedation protocols and daily sedation interruption: A multidisciplinary survey. J Crit Care. 2009; 24:66–73. [PubMed: 19272541]
30. Patel RPS, Gambrell MB, Speroff TP, et al. Delirium and sedation in the intensive care unit: Survey of behaviors and attitudes of 1384 healthcare professionals. Crit Care Med. 2009; 37:825–832. [PubMed: 19237884]
31. Mehta S, McCullagh I, Burry L. Current sedation practices: Lessons learned from international surveys. Crit Care Clin. 2009; 25:471–488. [PubMed: 19576525]

32. Mehta SM, Burry LP, Fischer SM, et al. Canadian survey of the use of sedatives, analgesics, and neuromuscular blocking agents in critically ill patients. Crit Care Med. 2006; 34:374–380. [PubMed: 16424717]
33. Devlin J, Marquis F, Riker R, et al. Combined didactic and scenario-based education improves the ability of intensive care unit staff to recognize delirium at the bedside. Crit Care. 2008; 12:R19. [PubMed: 18291021]
34. Devlin JW, PharmD B, Fong JJ, et al. Use of a validated delirium assessment tool improves the ability of physicians to identify delirium in medical intensive care unit patients. Crit Care Med. 2007; 35:2721–2723. [PubMed: 18074477]
35. Riekerk B, Pen EJ, Hofhuis JGM, et al. Limitations and practicalities of CAM-ICU implementation, a delirium scoring system, in a Dutch intensive care unit. Intensive Crit Care Nurs. 2009; 25:242–249. [PubMed: 19540761]
36. Page V, Navarange S, Gama S, et al. Routine delirium monitoring in a UK critical care unit. Crit Care. 2009; 13:R16. [PubMed: 19203391]
37. van den Boogaard M, Peters S, van der Hoeven J, et al. The impact of delirium on the prediction of in-hospital mortality in intensive care patients. Crit Care. 2010; 14:R146. [PubMed: 20682037]
38. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. Crit Care Med. 1985; 13:818–829. [PubMed: 3928249]
39. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis. 1987; 40:373–383. [PubMed: 3558716]
40. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity and some norms. Psychol Med. 1989; 19:1015–1022. [PubMed: 2594878]
41. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960; 20:37–46.
42. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull. 1968; 70:213–220. [PubMed: 19673146]
43. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

Table 1

Baseline and Hospital Stay Characteristics, Patients with Clinical Assessments

Patient Characteristic	All Patients (n=510)
Age at enrollment, mean ± SD*	56 (±15)
Age group at enrollment, N (%)	
< 65	350 (69)
≥ 65	160 (31)
Sex, N (%)	
Female	246 (48)
Male	264 (52)
Race, N (%)	
White	448 (88)
Non-white	62 (12)
ICU† type, N (%)	
Medical	307 (60)
Surgical	203 (40)
APACHE‡ II at enrollment, mean ± SD	27 (± 9)
Charlson comorbidity index[39] at enrollment, median (IQR§)	2.0 (1.0 to 4.0)
Mechanical ventilation, N (%)	
Never on ventilator	45 (9)
Spent some time on ventilator	465 (91)
History of depression, N (%)	170 (33%)
Preexisting cognitive impairment, N (%)	
No cognitive impairment (IQCODE‖[40] ≤ 3.3)	447 (88)
Some cognitive impairment (IQCODE‖[40] > 3.3)	63 (12)

*SD – standard deviation,

†ICU - intensive care unit,

‡APACHE II - Acute Physiology and Chronic Health Evaluation II,

§IQR – Interquartile Range,

‖IQCODE - Informant Questionnaire on Cognitive Decline in the Elderly

Table 2

Agreement between Research Nurse Measurement and Bedside Nurse Measurement of Delirium Within Two Hours of Each Other

Patient Characteristic	N	Weighted Kappa (95% CI)
All Matched Assessments	6062	0.67 (0.66, 0.70)
Unit-Specific Assessments		
Medical	3182	0.73 (0.70, 0.75)
Surgical	2880	0.60 (0.57, 0.63)
Medical Subgroups		
Patient is Mechanically Ventilated	4346	0.60 (0.57, 0.62)
Cognitively Impaired at Baseline*	764	0.70 (0.66, 0.76)
History of Depression	2346	0.66 (0.62, 0.70)
APACHE† Acute Physiology Score, Quintile (Score)		
Quartile 1 (<12)	1178	0.68 (0.61, 0.73)
Quartile 2 (12–15)	1178	0.68 (0.61, 0.73)
Quartile 3 (16–20)	1720	0.66 (0.63, 0.70)
Quartile 4 (21 or greater)	1837	0.68 (0.64, 0.71)
Age Groups		
< 65	4031	0.71 (0.68, 0.73)
≥ 65	2031	0.59 (0.55, 0.63)
Year of Data Collection		
2007	1656	0.65 (0.60, 0.68)
2008	1640	0.70 (0.66, 0.74)
2009	2014	0.67 (0.62, 0.71)
2010	752	0.67 (0.60, 0.73)

*Cognitively impaired is defined as an IQCODE score of > 3.3

†APACHE: Acute Physiology and Chronic Health Evaluation

Table 3

Agreement between Research Measurement and Bedside Nurse Measurement of the Richmond Agitation-Sedation Scale (RASS) Within Two Hours of Each Other

Patient Characteristic	N	Weighted Kappa (95% CI)
All Matched Assessments	6731	0.66 (0.64, 0.68)
Unit-Specific Assessments		
Medical	3612	0.71 (0.67, 0.75)
Surgical	3119	0.60 (0.55, 0.63)
Medical Subgroups		
Patients Mechanically Ventilated	4747	0.60 (0.56, 0.62)
Cognitively Impaired at Baseline*	850	0.70 (0.66, 0.77)
History of Depression	2611	0.65 (0.61, 0.70)
APACHE† Acute Physiology Score, Quintile (Score)		
Quartile 1 (<12)	1295	0.64 (0.57, 0.68)
Quartile 2 (12–15)	1295	0.64 (0.57, 0.68)
Quartile 3 (16–20)	1891	0.64 (0.60, 0.68)
Quartile 4 (21 or greater)	2052	0.69 (0.65, 0.73)
Age Groups		
< 65	4450	0.69 (0.66, 0.72)
≥ 65	2281	0.58 (0.54, 0.63)
Year of Data Collection		
2007	1733	0.67 (0.61, 0.71)
2008	1903	0.65 (0.59, 0.70)
2009	2240	0.65 (0.61, 0.68)
2010	855	0.66 (0.60, 0.71)

*Cognitively impaired is defined as an IQCODE score of > 3.3

†APACHE: Acute Physiology and Chronic Health Evaluation

Table 4

Sensitivity and Specificity of Bedside Nurse Assessment of Delirium when Compared with Research Nurse Assessment Within Two Hours of Each Other

Patient Characteristic	N	Sensitivity (95% CI)	Specificity (95% CI)
All Matched Assessments	3856	0.81 (0.78, 0.83)	0.81 (0.78, 0.85)
Unit-Specific Assessments			
Medical	2044	0.80 (0.75, 0.84)	0.90 (0.87, 0.92)
Surgical	1812	0.82 (0.78, 0.85)	0.69 (0.63, 0.74)
Medical Subgroups			
Patients Mechanically Ventilated	2279	0.82 (0.79, 0.85)	0.77 (0.71, 0.82)
Cognitively Impaired at Enrollment*	453	0.85 (0.78, 0.90)	0.78 (0.67, 0.86)
History of Depression	1554	0.82 (0.78, 0.85)	0.80 (0.75, 0.85)
APACHE† Acute Physiology Score, Quintile (Score)			
Quartile 1 (<12)	845	0.75 (0.68, 0.81)	0.85 (0.78, 0.90)
Quartile 2 (12–15)	733	0.83 (0.75, 0.88)	0.84 (0.75, 0.91)
Quartile 3 (16–20)	1109	0.82 (0.75, 0.87)	0.80 (0.74, 0.86)
Quartile 4 (21 or greater)	1102	0.81 (0.77, 0.85)	0.78 (0.71, 0.84)
Age Groups			
< 65	2551	0.80 (0.76, 0.83)	0.84 (0.81, 0.88)
≥ 65	1305	0.82 (0.77, 0.86)	0.73 (0.64, 0.80)
Year of Data Collection			
2007	1016	0.80 (0.74, 0.84)	0.77 (0.68, 0.84)
2008	1095	0.78 (0.72, 0.84)	0.84 (0.78, 0.88)
2009	1290	0.84 (0.79, 0.88)	0.82 (0.75, 0.87)
2010	455	0.80 (0.73, 0.85)	0.83 (0.73, 0.89)

*Cognitively impaired is defined as an IQCODE score of > 3.3

†APACHE: Acute Physiology and Chronic Health Evaluation