



Online article and related content  
current as of June 3, 2010.

## Monitoring Sedation Status Over Time in ICU Patients: Reliability and Validity of the Richmond Agitation-Sedation Scale (RASS)

E. Wesley Ely; Brenda Truman; Ayumi Shintani; et al.

*JAMA*. 2003;289(22):2983-2991 (doi:10.1001/jama.289.22.2983)

<http://jama.ama-assn.org/cgi/content/full/289/22/2983>

Correction

[Contact me if this article is corrected.](#)

Citations

This article has been cited 159 times.  
[Contact me when this article is cited.](#)

Topic collections

Critical Care/ Intensive Care Medicine; Adult Critical Care; Anesthesia  
[Contact me when new articles are published in these topic areas.](#)

Subscribe

<http://jama.com/subscribe>

Permissions

[permissions@ama-assn.org](mailto:permissions@ama-assn.org)  
<http://pubs.ama-assn.org/misc/permissions.dtl>

Email Alerts

<http://jamaarchives.com/alerts>

Reprints/E-prints

[reprints@ama-assn.org](mailto:reprints@ama-assn.org)

# Monitoring Sedation Status Over Time in ICU Patients

## Reliability and Validity of the Richmond Agitation-Sedation Scale (RASS)

E. Wesley Ely, MD, MPH

Brenda Truman, RN, MSN

Ayumi Shintani, PhD, MPH

Jason W. W. Thomason, MD

Arthur P. Wheeler, MD

Sharon Gordon, PsyD

Joseph Francis, MD, MPH

Theodore Speroff, PhD

Shiva Gautam, PhD

Richard Margolin, MD

Curtis N. Sessler, MD

Robert S. Dittus, MD, MPH

Gordon R. Bernard, MD

INCREASED SCRUTINY HAS RECENTLY been placed on appropriate titration of sedative and analgesic medications in critically ill patients, especially those being treated with mechanical ventilation.<sup>1-5</sup> Patient comfort should be a primary goal in the intensive care unit (ICU), including adequate pain control,<sup>6,7</sup> anxiety, and prevention and treatment of delirium.<sup>8</sup> However, achieving an appropriate balance of sedation and analgesia is challenging.<sup>9-11</sup> Without rational and agreed on target levels of sedation, different members of the health care team will have disparate treatment goals,<sup>12-14</sup> increasing the chance for iatrogenic complications and potentially impeding recovery.

The clinical practice guidelines of the Society of Critical Care Medicine emphasize the need for goal-directed delivery of psychoactive medications.<sup>8,15</sup> Al-

**Context** Goal-directed delivery of sedative and analgesic medications is recommended as standard care in intensive care units (ICUs) because of the impact these medications have on ventilator weaning and ICU length of stay, but few of the available sedation scales have been appropriately tested for reliability and validity.

**Objective** To test the reliability and validity of the Richmond Agitation-Sedation Scale (RASS).

**Design** Prospective cohort study.

**Setting** Adult medical and coronary ICUs of a university-based medical center.

**Participants** Thirty-eight medical ICU patients enrolled for reliability testing (46% receiving mechanical ventilation) from July 21, 1999, to September 7, 1999, and an independent cohort of 275 patients receiving mechanical ventilation were enrolled for validity testing from February 1, 2000, to May 3, 2001.

**Main Outcome Measures** Interrater reliability of the RASS, Glasgow Coma Scale (GCS), and Ramsay Scale (RS); validity of the RASS correlated with reference standard ratings, assessments of content of consciousness, GCS scores, doses of sedatives and analgesics, and bispectral electroencephalography.

**Results** In 290-paired observations by nurses, results of both the RASS and RS demonstrated excellent interrater reliability (weighted  $\kappa$ , 0.91 and 0.94, respectively), which were both superior to the GCS (weighted  $\kappa$ , 0.64;  $P < .001$  for both comparisons). Criterion validity was tested in 411-paired observations in the first 96 patients of the validation cohort, in whom the RASS showed significant differences between levels of consciousness ( $P < .001$  for all) and correctly identified fluctuations within patients over time ( $P < .001$ ). In addition, 5 methods were used to test the construct validity of the RASS, including correlation with an attention screening examination ( $r = 0.78$ ,  $P < .001$ ), GCS scores ( $r = 0.91$ ,  $P < .001$ ), quantity of different psychoactive medication dosages 8 hours prior to assessment (eg, lorazepam:  $r = -0.31$ ,  $P < .001$ ), successful extubation ( $P = .07$ ), and bispectral electroencephalography ( $r = 0.63$ ,  $P < .001$ ). Face validity was demonstrated via a survey of 26 critical care nurses, which the results showed that 92% agreed or strongly agreed with the RASS scoring scheme, and 81% agreed or strongly agreed that the instrument provided a consensus for goal-directed delivery of medications.

**Conclusions** The RASS demonstrated excellent interrater reliability and criterion, construct, and face validity. This is the first sedation scale to be validated for its ability to detect changes in sedation status over consecutive days of ICU care, against constructs of level of consciousness and delirium, and correlated with the administered dose of sedative and analgesic medications.

JAMA. 2003;289:2983-2991

www.jama.com

**Author Affiliations** are listed at the end of this article.  
**Corresponding Author and Reprints:** E. Wesley Ely, MD, MPH, Department of Medicine, Division of Allergy, Pulmonary, and Critical Care Medicine, Center for Health Services Research of Vanderbilt University, Tennessee Valley Veteran's Affairs Healthcare System, Geriatric Research Education and Clinical

Center (GRECC), Nashville, TN 37232 (e-mail: wes.ely@vanderbilt.edu).

**Caring for the Critically Ill Patient Section Editor:** Deborah J. Cook, MD, Consulting Editor, JAMA.

**Advisory Board:** David Bihari, MD; Christian Brun-Buisson, MD; Timothy Evans, MD; John Heffner, MD; Norman Paradis, MD; Adrienne Randolph, MD.

though the Ramsay Scale (RS)<sup>16</sup> was not originally intended for use as a clinical monitoring tool, it has been used for decades in both clinical practice and research. Still, most patients are not monitored with any scale to guide delivery of sedative medications. Objective, goal-directed sedation therapy is now the recommended standard to avoid overse- dation and to promote earlier extuba- tion.<sup>2,8,14,17-19</sup> As pointed out by others,<sup>1,20-22</sup> very few of the available sedation scales have been appropri- ately tested for reliability and validity. Even among available instruments that have been tested for reliability and validity,<sup>23-26</sup> none discretely separates verbal from physical stimulation (ie, the po- tency of the stimulus) in generating scores at pivotal levels of sedation. Two recent systematic reviews concluded that goal-directed sedative and analgesic ad- ministration would be enhanced if such instruments were shown to detect varia- tions in level of consciousness over time and according to delivery of psychoac- tive drugs.<sup>1,26</sup> Other investigators ob- served cardiac surgery patients over time, yet the duration of monitoring was only 6 hours, the sample size was small (only 14 patients followed up to extubation), and correlations with drug doses were not published.<sup>27</sup>

The Richmond Agitation-Sedation Scale (RASS)<sup>28,29</sup> was developed by a mul- tidisciplinary team at Virginia Common- wealth University in Richmond. It is a 10- point scale that can be rated briefly using 3 clearly defined steps and that has dis- crete criteria for levels of sedation and agitation. A unique feature of the RASS is that it uses the duration of eye con- tact following verbal stimulation as the principal means of titrating sedation. Hence, this scale's validation could be linked to both arousal and content of thought—the 2 components of con- sciousness.<sup>30</sup> We determined that the du- ration of eye contact could be easily mea- sured with minimal training, allowing reproducibility and increased acceptabil- ity of the instrument by bedside physi- cians, nurses, and researchers alike. The RASS has been demonstrated to have ex- cellent interrater reliability in a broad

range of adult medical and surgical ICU patients and to have excellent validity when compared with a visual analogue scale and selected sedation scales.<sup>29</sup>

The current investigation was de- signed to extend the reliability and va- lidity testing of the RASS in novel ways to include assessment of sedation over time, correlation of the RASS with in- dependent neuropsychiatric experts' measures of level of consciousness and formally measured content of conscious- ness (ie, inattention and delirium), doses of sedative and analgesic medications, and objective measurement of brain function using bispectral array electro- encephalography.

## METHODS

This investigation was conducted in the adult medical and coronary ICUs at Vanderbilt University Medical Center, a 641-bed tertiary-care, academic medi- cal center. The institutional review board approved the study, and written in- formed consent was obtained from the patients or proxies. Reliability and va- lidity testing was performed in 2 phases. We enrolled patients into the reliabil- ity testing cohort from July 21, 1999, to September 7, 1999, and into the valid- ity testing cohort from February 1, 2000, to May 3, 2001.

While none of the data in this re- port have been previously published, other data from this cohort of patients have been published as would be ex- pected from prospective cohort inves- tigations that have the capacity to ad- dress different issues. Specifically, the 38 patients from the RASS reliability testing cohort were those reported in the first Confusion Assessment Method for the Intensive Care Unit (CAM- ICU) study,<sup>31</sup> and 96 of the patients from this RASS validity testing were those reported in the second CAM- ICU study cohort.<sup>32</sup>

Patients receiving and not receiving mechanical ventilation were screened during reliability testing to ensure reli- ability in verbal and nonverbal patients. Any adult admitted to the ICU who did not meet the following a priori exclu- sion criteria was eligible for enroll-

ment: a history of severe dementia, psy- chosis, or neurologic disease (n=12); patient or family refusal to participate (n=8); and admission to the ICU after the predefined cap of 10 study patients per day (because of research staffing limi- tations) had been reached (n=18). Con- secutive patients receiving mechanical ventilation were enrolled into the valid- ity testing cohort (and followed up un- til ICU discharge), because the primary challenge and use for sedation scales in the ICU are for patients who are non- verbal and are intubated. Patients were excluded if they had a history of psycho- sis or neurologic disease (n=16), were non-English-speaking or deaf (n=5), were extubated or had died before nurses' screen (n=15), were previously en- rolled in reliability cohort (n=5), or be- cause of patient or family refusal to par- ticipate (n=9).

## Psychoactive Medications

During this investigation, no protocol to guide analgesia, sedation, or neuromus- cular blockade existed in our ICU, and no objective target levels of sedation were routinely identified according to dis- ease state or ventilator settings. All doses of narcotics, benzodiazepines, propo- fol, and neuromuscular blocking agents were recorded prospectively in 8-hour intervals throughout the investigation. Administered narcotics were either mor- phine or fentanyl. Administered benzo- diazepines were either lorazepam or mid- azolam (the midazolam dose was converted to lorazepam equivalents by dividing by 3 to achieve equipotent dose<sup>33</sup>). If neuromuscular blocking agents had been given within 8 hours of RASS assessment, patient observations were excluded from our pharmacologi- cal analyses.

## Performing the RASS

Prior to this investigation, raters had no experience with the RASS. Sessler et al<sup>29</sup> from the Virginia Commonwealth Uni- versity provided a 1-page handout with the RASS description and the Proce- dure for RASS assessment (TABLE 1) to the investigators. No other formal train- ing was received or required. All raters

performed the RASS using the same sequence of 3 steps as outlined in Table 1. If the patients were alert or agitated prior to stimulation, they were scored 0 to +4 accordingly. If patients were not spontaneously alert, they were then called by name to look at the rater, with the duration of eye contact measured, at which time a positive response was scored accordingly as -1 to -3. If the patients did not respond to verbal stimulation, they were then physically stimulated (ie, shoulder shake and/or sternal rub) and scored according to their response as -4 or -5. If calm and not alert prior to verbal and physical stimulation, patients were then rated as -1 to -5 (as standardized in Table 1) even if they became agitated on stimulation. Assessments required less than 20 seconds.

**Structure of Reliability Study Procedures**

Two critical care study nurses, an intensivist, and a neuropsychiatric expert performed daily, independent RASS ratings during each patient’s ICU stay. Interrater reliability assessments were conducted in the afternoon, and none of the raters had access to the others’ scores at any time. The study nurses performed simultaneous ratings (one interacted with and rated the patient while the other observed and rated the same patient), while the intensivists and neuropsychiatric experts performed their independent RASS ratings within 4 hours of the nurses’ ratings. In addition, both nurses performed independent RS<sup>16</sup> ratings and Glasgow Coma Scale (GCS)<sup>34</sup> ratings for each patient.

**Structure of the Validation Study Procedures**

Validity, the extent to which the instrument measured what it was intended to measure, was tested in 3 ways according to standard definitions.<sup>35,36</sup> *Criterion validity*, the extent to which a measure relates to a set of externally derived criteria, was tested on the first 96 patients of the validation cohort by comparing patients’ RASS levels against reference standard evaluations performed by neuropsychiatric experts (ie, a geri-

**Table 1.** The Richmond Agitation-Sedation Scale (RASS)

Score	Term	Description	
+4	Combative	Overtly combative, violent, immediate danger to staff	
+3	Very agitated	Pulls or removes tube(s) or catheter(s); aggressive	
+2	Agitated	Frequent nonpurposeful movement, fights ventilator	
+1	Restless	Anxious but movements not aggressive or vigorous	
0	Alert and calm		
-1	Drowsy	Not fully alert, but has sustained awakening (eye opening/eye contact) to voice (>10 seconds)	Verbal stimulation
-2	Light sedation	Briefly awakens with eye contact to voice (<10 seconds)	
-3	Moderate sedation	Movement or eye opening to voice (but no eye contact)	
-4	Deep sedation	No response to voice, but movement or eye opening to physical stimulation	Physical stimulation
-5	Unarousable	No response to voice or physical stimulation	

Procedure for RASS Assessment		
1. Observe patient		
• Patient is alert, restless, or agitated.		Score 0 to +4
2. If not alert, state patient’s name and say to open eyes and look at speaker.		
• Patient awakens with sustained eye opening and eye contact.		Score -1
• Patient awakens with eye opening and eye contact, but not sustained.		Score -2
• Patient has any movement in response to voice but no eye contact.		Score -3
3. When no response to verbal stimulation, physically stimulate patient by shaking shoulder and/or rubbing sternum.		
• Patient has any movement to physical stimulation.		Score -4
• Patient has no response to any stimulation.		Score -5

Adapted with permission.<sup>29</sup>

atric psychiatrist and a geriatric neuropsychologist) who rated patients’ levels of consciousness as normal, delirious, stuporous, or comatose using standardized definitions that did not incorporate the RASS in any way.<sup>31,32,37</sup> The neuropsychiatric experts were blinded to the study nurses’ RASS ratings and the 2 ratings were performed within 4 hours of one another. To address 2 areas of sedation monitoring that have not been studied adequately to date,<sup>21,26</sup> we analyzed data for patients who were and were not receiving mechanical ventilation and the ability of the RASS to identify changes in level of consciousness over time (as rated by the neuropsychiatric expert).

*Construct validity*, the extent to which a measure relates to the other measures that would theoretically support the concept (or construct) being measured, also should be measured whenever no universally accepted criterion exists.<sup>35,38,39</sup> Thus, construct validity was tested 5 ways in patients with normal and abnormal consciousness: (1) RASS scores were compared with abnormali-

ties in the “content of consciousness” as rated by the nurse using a screening examination for attention,<sup>31</sup> which served as a measure of content of consciousness since this is the pivotal feature of delirium<sup>32</sup>; (2) comparisons were made against ratings of the GCS, a standard instrument widely used in neurologic monitoring throughout the world; (3) RASS scores were correlated with the quantity of sedative and analgesic drugs administered to patients in the 8- and 24-hour periods prior to their assessments; (4) outcomes of planned extubation were compared with concurrent RASS scores; and (5) bispectral-XP electroencephalography (BIS-XP EEG) (Aspect Medical Systems Inc, Newton, Mass) was measured and correlated with RASS scores as described below. Both the GCS and BIS-XP EEG measurements were recorded at the same time as the RASS ratings, while the ratings by the neuropsychiatric experts occurred within 4 hours of the study nurses’ ratings. The decision to extubate was made by the attending physician according to

the results of spontaneous breathing trials with no knowledge of the RASS scores and without a predetermined neurologic status. Extubations were considered successful if the patient remained extubated and did not receive mechanical ventilation for 48 hours. Detailed reasons for unsuccessful extubation were not recorded.

**Face validity**, the extent to which an instrument appears to measure what it is intended to measure, was tested by surveying critical care nurses about whether or not the RASS was an appropriate and clinically useful measure of agitation and sedation. The survey included 4 relevant questions each worded 2 different ways and embedded within other questions regarding patient management. The survey asked nurses to rate the following 4 statements using a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]): (1) RASS levels for agitation (+1 to +4) are clinically relevant and easy to score; (2) it makes sense clinically that the RASS assessment begins with verbal stimulation (-1 to -3) and then moves to physical stimulation (-4 to -5); (3) use of the RASS improves communication among the health care team; and (4) use of the RASS provides a team consensus for target-level sedation.

### BIS-XP EEG

Bispectral index uses a nonlinear signal processor to measure brain wave activity in the form of raw electroencephalography (EEG) and to create a score ranging from 100 (awake) to 0 (no cortical activity). Both raw EEG and BIS-XP EEG data were recorded using a fronto-temporal montage with disposable sensors that were connected to a portable EEG monitor (A1050, Aspect Medical Systems Inc).<sup>40,41</sup> As opposed to earlier versions of the BIS-XP EEG used in other investigations designed for the anesthetized surgical patients in the operating room,<sup>42</sup> this version has been specifically designed for use in ICU patients who are sedated and receiving mechanical ventilation to reduce electromyographic interference.<sup>43,44</sup> Electrode impedance values were acceptable if they

were no greater than 5 k $\Omega$  and if the threshold for acceptable signal quality index was greater than 80%. Raw EEG data was sampled at 128 samples per second and recorded continuously in real time; processed variables were downloaded and recorded every 5 seconds. The data sampling rate was 256 times per second with filter settings of 70 Hz for the high frequency (70 Hz) and 2 Hz for the low frequency. The study nurses were blinded to both the raw EEG and the BIS-XP EEG data. The continuous EEG recording was later reviewed by a separate investigator blinded to the clinical assessment ratings to identify the reported baseline value (ie, a stable portion of the tracing that the EEG would return to after the sensor was attached).

### Statistical Analysis

Patients' baseline characteristics were presented using means and SDs for continuous variables, and frequencies and proportions for dichotomous variables. Interrater reliability was determined for the RASS, RS, and GCS by comparing ratings between raters using weighted  $\kappa$  indices and 95% confidence intervals. For criterion validity, RASS scores were compared with the neuropsychiatric expert rating using Wilcoxon rank sum tests. As part of criterion validity, and to account for dependency among observations within an individual patient, proportional odds regression analysis with generalized estimating equations was performed to assess the neuropsychiatric expert ratings vs RASS ratings, and odds ratios were presented. Similar analysis using Wilcoxon rank sum tests assessed the difference between observations within a level of consciousness of patients receiving and not receiving mechanical ventilation. For construct validity testing, the Spearman correlation coefficient ( $r$ ) was calculated for the associations between RASS scores and onset of inattention, GCS score, daily dosage of sedative drugs, and BIS-XP EEG results. Wilcoxon rank sum tests were used to assess significant differences in GCS scores according to RASS levels, and proportional odds models were used with the generalized estimating equations to

analyze the relationship over time among the RASS score, GCS score, and daily dosage of sedative drugs. Statistical analyses were performed using SAS version 8.02 (SAS Institute Inc, Cary NC) and STATA version 7.0 (STATA Corp, College Station, Tex).

## RESULTS

### Patient Characteristics

Eighty-six patients were admitted to the medical and coronary ICUs during the reliability testing phase, of whom 38 (44%) were enrolled. During the validity testing phase, 325 consecutive patients receiving mechanical ventilation were admitted to the ICU, of whom 275 (84.6%) were enrolled. Baseline characteristics from all 313 patients are presented in TABLE 2. At the time of enrollment, 22 (46%) of the reliability cohort received mechanical ventilation, while all 275 of the validation cohort received mechanical ventilation. The 2 cohorts had a high baseline severity of illness as measured by Acute Physiology and Chronic Health Evaluation II (APACHE II) scores (mean [SD], 17.1 [8.7] and 25.0 [8.0]), and had a wide variety of medical diagnoses.

### RASS Interrater Reliability Testing

Patients were evaluated on multiple occasions during their ICU stay. The mean (SD) RASS scores for each rater were as follows: nurse 1, -1.60 (2.16); nurse 2, -1.88 (2.20); intensivist, -1.61 (2.17); and neuropsychiatric expert, -1.50 (2.25). In 290 paired observations by nurses, both the RASS and the RS demonstrated excellent interrater reliability (weighted  $\kappa$ , 0.91 and 0.94, respectively), which were superior to GCS (weighted  $\kappa$ , 0.64;  $P < .001$  for both comparisons). Using only the first observation for each patient ( $n = 38$ ), the weighted  $\kappa$  values for the RASS, RS, and GCS were unchanged at 0.95, 0.95, and 0.65, respectively. The interrater reliability of the RASS was very high across nurses, intensivists, and neuropsychiatric experts (TABLE 3). Because reliability testing was expected to be most challenging for patients receiving mechanical ventilation, we conducted an-



other analysis of interrater reliability restricted to the 22 patients who were intubated. The weighted  $\kappa$  for the RASS between the 2 nurse ratings for patients who were intubated was 0.88 (95% confidence interval, 0.78-0.97).

**RASS Criterion, Construct, and Face Validity Testing**

The relative frequency that the validation cohort spent in each major arousal category for 1833 observations in 275 patients indicated that the patients who received mechanical ventilation spent about one third of their time either unarousable or in a deeply sedated state (RASS score, -5 or -4; n=548 observations), one third in a moderate to light sedation state (RASS score, -3 to -1; n=625 observations), and one third in an alert and calm state (RASS score, 0; n=619 observations). Spontaneous agitation (RASS score, +1 to +4; n=41 observations), which was rated prior to stimulation, was an uncommon state (<5%) found by the study nurses.

Criterion validity was tested in 411 paired observations in the validation cohort for the first 96 patients with a median of 3 observations per patient. The results of the RASS showed excellent discrimination between levels of consciousness as rated using the neuropsychiatric expert reference standard ( $P<.001$  for all) (FIGURE 1). Furthermore, as the neuropsychiatric expert raters and RASS raters independently tracked level of consciousness within patients over successive days of ICU care, RASS scores continued to correlate with expert raters' evaluations despite fluctuations in consciousness ( $P<.001$  for all) (TABLE 4). When comparing patients over the course of their ICU stay, the reference standard ratings of abnormal levels of consciousness (ie, delirium, stupor, and coma) compared similarly with RASS ratings regardless of intubation status: delirium (-2 median RASS score if intubated vs -2 median RASS score if extubated,  $P=.18$ ), stupor (-3 vs -3,  $P=.92$ ), or coma (-5 vs -5,  $P=.62$ ) (Figure 1).

Five methods were used to test construct validity: (1) RASS was correlated

**Table 2. Patient Characteristics at Enrollment**

Characteristic	Reliability Phase (n = 38)	Validation Phase (n = 275)
Age, mean (SD), y	60.0 (19.0)	55.7 (16.5)
Men, No. (%)	23 (60.5)	140 (50.9)
Race, No. (%)		
White	32 (8.2)	222 (80.7)
African American	5 (13.2)	53 (19.3)
Hispanic	1 (2.6)	0
APACHE II score, mean (SD)*	17.1 (8.7)	25.0 (8.0)
Mechanical ventilation, No. (%)	22 (45.8)	275 (100)
ICU admission diagnosis, No. (%)		
Sepsis/acute respiratory distress syndrome	11 (28.9)	143 (52.0)
Pneumonia	0	24 (8.7)
Myocardial infarction or arrhythmia	6 (15.8)	8 (2.9)
Congestive heart failure	6 (15.8)	16 (5.8)
Hepatic or renal failure	5 (13.2)	9 (3.3)
Chronic obstructive pulmonary disease	4 (10.5)	21 (7.6)
Gastrointestinal bleeding	3 (7.9)	10 (3.6)
Hematologic/oncologic disease	2 (5.3)	23 (8.4)
Drug overdose	1 (2.6)	13 (4.7)
Other	0	8 (2.9)

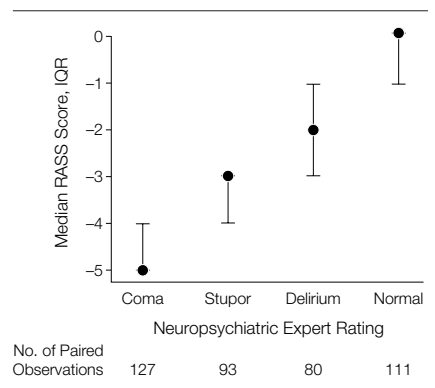
Abbreviations: APACHE II, Acute Physiology and Chronic Health Evaluation II; ICU, intensive care unit.  
 \*An assessment of severity of illness that was calculated using patients' most abnormal values during the first 24 hours following ICU admission.<sup>45</sup>

**Table 3. Interrater Reliability of the Richmond Agitation-Sedation Scale**

Rater	No. of Paired Observations	Weighted $\kappa$ (95% Confidence Interval)
Nurse 1 vs nurse 2	290	0.91 (0.86-0.95)
Nurse 1 vs intensivist	127	0.79 (0.61-0.88)
Nurse 1 vs neuropsychiatric expert	150	0.82 (0.67-0.90)
Nurse 2 vs intensivist	128	0.88 (0.76-0.94)
Nurse 2 vs neuropsychiatric expert	151	0.84 (0.71-0.92)
Neuropsychiatric expert vs intensivist	129	0.91 (0.82-0.96)

with onset of inattention using an attention-screening examination, the pivotal criterion for delirium ( $r=0.78$ ,  $P<.001$ ). (2) To compare the RASS to the GCS, 1360 paired observations (among 275 patients with a median of 3 observations per patient) showed excellent correlation and discrimination ( $r=0.91$ ,  $P<.001$ ) (FIGURE 2). The RASS also correlated with the GCS over time ( $P<.001$ ), and the odds ratio of having higher RASS scores with greater GCS scores was 1.39 ( $P<.001$ ). (3) We compared RASS scores with the cumulative lorazepam, propofol, fentanyl, and morphine dose over the 8-hour and 24-hour periods prior to RASS assessments (TABLE 5). As described in the "Methods" section, midazolam dose was converted into lorazepam equivalents for the purposes of these analyses. As an example of the dose-response relation-

**Figure 1. Criterion Validity: Richmond Agitation-Sedation Scale (RASS) Scores vs Reference Standard Evaluations Performed by a Neuropsychiatric Expert**



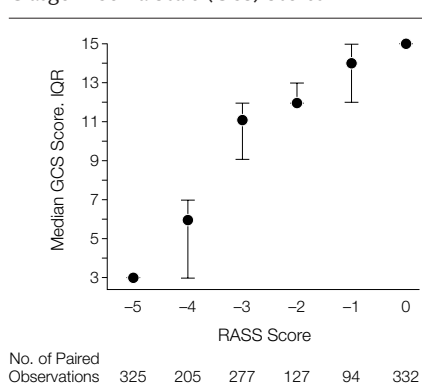
Neuropsychiatric experts (ie, a geriatric psychiatrist or a geriatric neuropsychologist) rated patients' levels of consciousness as normal, delirium, stuporous, or comatose. Data (median [interquartile range][IQR]) are from 411 paired observations in the first 96 patients in the validation cohort and demonstrated significant discrimination between each level of consciousness (all  $P<.001$ ).

**Table 4.** Neuropsychiatric Expert Rating vs Richmond Agitation-Sedation Scale (RASS) Rating: Tracking Level of Consciousness Over Time Within Patients\*

Reference Standard Neuropsychiatric Rating	Odds Ratios of Higher RASS Score Over Time at Given Level of Consciousness	P Value
Normal	>100	<.001
Delirium	36	<.001
Stupor	9	<.001
Coma	1 (reference)	...

\*To understand correlations between repeated observations over time, this generalized estimating equations analysis would be interpreted as follows: using coma as a reference point (ie, odds ratio of 1), patients rated by the neuropsychiatric experts as stuporous, delirious, or normal were 9, 36, and >100 times more likely, respectively, to have higher RASS scores across time compared with patients whose ratings remained comatose (eg, comparing 2 initially comatose patients with RASS scores of -5, one of whom transitioned to delirium on day 2 while the other remained comatose, the odds of the delirious patient having a higher RASS score would be 36 times greater).

**Figure 2.** Construct Validity: Richmond Agitation-Sedation Scale (RASS) Scores vs Glasgow Coma Scale (GCS) Scores



Data (median [interquartile range] [IQR]) are from a total of 1360 observations and demonstrate excellent correlation ( $r=0.91$ ,  $P<.001$ ) and discrimination (all pairs with  $P<.001$ ). Median GCS for RASS+1 to +4 ( $n=27$ ) was 13 (IQR 10-14) (data not shown). Generalized estimating equations analysis was used to independently evaluate association within patients over successive days of intensive care and yielded a similar result (odds ratio, 2.3;  $P<.001$ ).

ship between psychoactive medication and RASS scores, correlative data for lorazepam equivalents over 8 hours prior to RASS assessment are shown in FIGURE 3. In contrast to the benzodiazepines, correlations between RASS and fentanyl and morphine were performed and presented as separate analyses because of significantly different results for these 2 agents (Table 5). (4) Of 185 planned extubations in the 275 validation cohort patients, 19 (10.3%) patients required reintubation within 48 hours. Of these planned extubations, 137 had RASS scores available during the shift prior to extubation, of which 13 (9.5%) were unsuccessful. The median (interquartile range) RASS score for success-

ful extubation was -2 (-3 to 0) and for unsuccessful extubation was -3 (-3 to -2) ( $P=.07$ ). (5) From the validation cohort, a random sampling of 124 patients was monitored over 321 days in the ICU with the BIS-XP EEG. The RASS scores correlated with the BIS-XP EEG results over the range of levels of consciousness ( $r=0.64$ ,  $P<.001$  for all) (FIGURE 4). For patients in 3 different, clinically relevant states (ie, spontaneously awake and alert [RASS score, 0], arousable with verbal stimulation [RASS score, -1 to -3], and arousable only with physical stimulation or not at all [RASS score, -4 or -5]), the median (interquartile range) EEG values were 96.8 (90.0-97.6), 69.0 (57.6-87.6), and 57.4 (46.4-66.3), respectively.

Face validity was demonstrated via a survey of 26 bedside critical care nurses. According to the results of the survey, 77% of the nurses agreed or strongly agreed that the RASS levels for agitation were clinically relevant and easy to score. Regarding the construct of the sedation assessment, 92% agreed or strongly agreed with the "verbal followed by physical stimulation" scoring scheme, 69% agreed or strongly agreed that the RASS improved communication among the managing team, and 81% agreed or strongly agreed this instrument provided a consensus target for goal-directed delivery of sedative and analgesic medications.

**COMMENT**

This investigation was designed to test a very brief yet structured approach to assessment of patient sedation in the ICU for reliability and validity, using a new,

broader, and more rigorous set of validation procedures than those previously studied. The RASS demonstrated strong interrater reliability and criterion, construct, and face validity. In previous work, Sessler et al<sup>29</sup> demonstrated strong interrater reliability of RASS between 5 nurse, pharmacist, and physician investigators in 192 consecutive patients who did or did not receive mechanical ventilation in surgical, neuroscience, and medical ICUs. The authors further documented strong reliability between a nurse educator and 27 bedside nurses in more than 100 patient observations. Building on these data, the present study included the use of neuropsychiatric experts as reference standard raters of patients' levels of consciousness to demonstrate criterion validity, 5 methods of confirming construct validity, and incorporating views of bedside critical care nurses for face validity.

In keeping with the stated priorities of recent critical appraisals of sedation scales,<sup>1,20-22,26</sup> the present report shows the RASS to be valid over successive days in individual patients, to correlate with the administered dose of sedative and analgesic drugs as well as brain wave activity, and confirms that it is a reliable and valid measurement for patients who were and were not receiving mechanical ventilation.

This report and that by Sessler et al<sup>29</sup> are complementary and constitute an evaluation of the RASS in more than 600 patients, both having avoided selection bias present in other investigations by enrolling consecutive patients. The large sizes of these 2 investigations, combined with the strong reliability of the instrument and the scope of our validity testing, lend strong credibility for the use of the RASS in patient management.

**Strengths of the RASS**

The RASS itself has several important strengths that warrant comment. Unlike other recently validated instruments,<sup>23,25</sup> the RASS separates verbal from physical stimulation so that the patient's level of arousal may be graded ac-

**Table 5.** Correlation Coefficients Between RASS Scores and Sedative and Analgesic Drug Equivalents\*

Sedative or Analgesic Equivalents	8 Hours Prior to RASS				24 Hours Prior to RASS			
	No. of Observations	Median Dose (IQR), mg	RASS vs Drug Dose		No. of Observations	Median Dose (IQR), mg	RASS vs Drug Dose	
			Spearman <i>r</i>	<i>P</i> Value			Spearman <i>r</i>	<i>P</i> Value
Lorazepam administration	294	4 (2-8)	-0.31	<.001	494	5 (2-12)	-0.25	<.001
Propofol administration	29	1269 (675-1896)	-0.32	.09	36	3315 (556-5814)	-0.40	.01
Fentanyl administration	247	1.6 (0.75-2.4)†	-0.25	<.001	295	3 (0.9-6.5)†	-0.22	<.001
Morphine administration	160	12 (0.75-2.4)	-0.13	.10	229	14 (0.9-6.5)	-0.12	.08

Abbreviations: IQR, interquartile range; RASS, Richmond Agitation-Sedation Scale.

\*Data are from the validation phase cohort. See Figure 3 for a detailed example of the dose-response relationship between RASS and 1 class of these agents (using lorazepam). Similar *P* values were observed using generalized estimating equations to adjust nonindependent observations within patients.

†Fentanyl is commonly reported to be 100 times more potent than morphine.<sup>33</sup> Therefore, the median morphine equivalent dose given to these patients (as fentanyl) would equate to 160 mg over 8 hours and 300 mg over 24 hours. While this mathematical conversion may be flawed or confounded in vivo, such large values are plausible considering its initially short duration of action,<sup>8</sup> the potential for rapid tolerance with fentanyl,<sup>46-48</sup> and that fentanyl was administered as a continuous infusion rather than an intermittent bolus.

According to the potency of the stimulus.<sup>28</sup> It has been common to consider sedation scales valid as long as they delineate levels of arousal (considered a surrogate of consciousness).<sup>20,23-25,49</sup> Consciousness, however, is classically defined as the combination of a person's level of arousal plus the content of consciousness (eg, delirium).<sup>30</sup> Only recently has the ICU community begun to focus on delirium as an essential element of patient comfort and outcome.<sup>8,50,51</sup> Importantly, a key feature of delirium is the presence or absence of inattention, which can be measured in part by the ability of a person to maintain eye contact. Indeed, this study demonstrated that RASS scores correlated with the onset of inattention and delirium. These 2 main strengths of the RASS assessment procedure (ie, completely distinguishing verbal from physical stimulation, and relying heavily on duration of eye contact) complement the recently developed delirium monitoring instrument, the CAM-ICU.<sup>31,32</sup> Using such a combined neurologic monitoring schema for patients receiving mechanical ventilation in future trials may offer a significant advance in our ability to measure short- and long-term cognitive outcomes of goal-directed delivery of sedative and analgesic medications or newer pharmacological agents with procognitive advantages over traditionally used agents in treating delirium.<sup>8,52,53</sup>

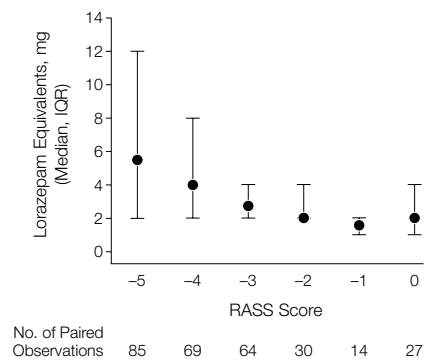
The κ values between nurses' and between physicians' paired assessments were both 0.91, and the κ values between any nurse and physician pair

ranged from 0.79 to 0.88. While all very high, the subtle differences in these values likely reflect that time elapsed between the nurse and physician assessments. The fact that the correlations were high despite elapsed time between the nurses' and physicians' assessments makes the data even more compelling.

**Study Limitations and Areas for Future Research**

We correlated RASS scores with quantity (ie, dose) of sedative and analgesic medications and found highly significant yet moderately to low correlation coefficients. Considering the broad distribution of drug dose variables and the numerous other covariates affecting consciousness (eg, underlying illness, electrolytes, hypoxemia, and other pharmacological agents), it would be unrealistic to expect RASS correlation with actual dose to be any higher. Similar to other reports on risk factors for delirium,<sup>47,54</sup> we found disparate correlation coefficients between RASS levels and different drugs. For example, the highly significant correlations between RASS levels vs lorazepam or fentanyl (both *P*<.001) contrast sharply with that of morphine (*P*=.10). Future investigations should attempt to determine the relative importance of other covariates contributing to level of consciousness, such as patients' age, sex, race, disease state, body mass index, the duration and cumulative drug levels of narcotics and analgesics over days of ICU care, and pharmacological interindividual variability based on drug metabolism and transport.

**Figure 3.** Richmond Agitation-Sedation Scale (RASS) Scores vs Cumulative Lorazepam Equivalent Dose

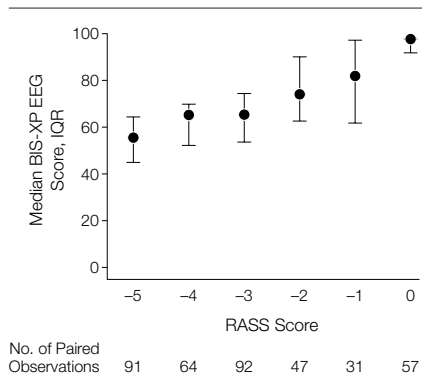


Benzodiazepine doses administered to patients over the 8-hour period prior to RASS assessment are shown (*r*=-0.31, *P*<.001). Benzodiazepines are expressed in lorazepam equivalents. Data on individual correlations between RASS vs lorazepam, propofol, fentanyl, and morphine for both 8-hour and 24-hour periods are shown in Table 5. Data on RASS scores greater than 0 were not shown because of insufficient sample size (*n*=5 for RASS > 0).

Most available evidence regarding sedatives and analgesics in ICU patients indicates that it may be less important which drugs are delivered than their proper titration using goal-directed delivery to optimize patient comfort while avoiding complications, such as prolonged mechanical ventilation or reintubation.<sup>1,3,8,55</sup> For example, recent data showed that deeper levels of sedation at the time of extubation, measured using the RASS, were associated with a higher likelihood of reintubation.<sup>56</sup> Future investigations should evaluate the usefulness of this tool in single or multicenter clinical trials, lo-



**Figure 4.** Richmond Agitation-Sedation Scale (RASS) Scores vs Bispectral-XP Electroencephalography (BIS-XP EEG)



The algorithm for BIS-XP EEG was used to measure brain wave activity in the fronto-temporal leads of 124 patients who received mechanical ventilation over 382 days. The RASS scores correlated with the electroencephalography (EEG) scores ( $r=0.64$ ) over a range of levels of arousal from awake and alert to coma ( $P<.001$ ). The (median [interquartile range] [IQR]) EEG values from 12 RASS evaluations of 9 agitated patients was 88.1 (79.5-97.3;  $P<.001$  for all).

cal or national quality-improvement collaboratives,<sup>57</sup> and different management strategies based on goal-directed delivery.

Some patients are sedated but subsequently demonstrate agitation following stimulation,<sup>5,20</sup> which represents another limitation of this investigation. Since we only recorded a single RASS score per patient assessment, we do not know the number of times patients were initially sedated and assigned negative RASS scores, only to become agitated minutes later. It may be surprising to some that so few patients in this investigation and that by Sessler et al<sup>29</sup> were found in the vigilant or agitated state, which we believe reflects a recurring theme of oversedation in critical care.<sup>2,10,17</sup> On the other hand, under-sedation and inadequate relief of symptoms could lead to problems such as unaddressed pain or anxiety,<sup>7,58</sup> and ultimately posttraumatic stress disorder.<sup>59-62</sup> The fact that the RASS has an expanded set of clinically relevant scores for tracking both agitation and sedation makes it well suited for future investigators to help better understand the clinical implications of such crossover states.

While the RASS demonstrated excellent face validity among our nurses, reports of ongoing large-scale implementation projects using the RASS will aid in our understanding of how to effect sustained change in the practice patterns of ICU nurses, therapists, pharmacists, and physicians using such instruments. Preliminary work in this area has already been reported from academic settings,<sup>63</sup> but should be forthcoming from community settings as well.

Lastly, as a barometer of brain wave activity, we used BIS-XP EEG monitoring in a method comparable with that used by other investigators.<sup>27,42,44</sup> The BIS-XP used in this investigation was an advance over that of earlier versions of BIS<sup>42</sup> because of improved screening of an electromyographic artifact.<sup>43,64</sup> However, the overlapping BIS values at different RASS levels (Figure 4) may be the result of the broad distribution of psychoactive drugs administered to this cohort and their interindividual effects on brain wave activity or merely a limitation in this emerging EEG technology.

**Conclusion**

The RASS, which takes less than 20 seconds to perform and requires minimal training,<sup>29</sup> has been shown to be highly reliable among multiple types of health care professionals. The RASS has an expanded set of scores at pivotal levels of sedation that are determined by patients' response to verbal vs physical stimulation, which will help the clinician in titrating medications. This extensive body of new data, with a variety of unique approaches to assess an agitation-sedation scale, expands the usefulness of such instruments for patient care. In accordance with recent recommendations, health care professionals should use valid and reliable instruments such as the RASS to implement sedative and analgesic drug delivery protocols for patients receiving mechanical ventilation. The driving unmet need for goal-directed sedation practice has been met—now an instrument has been shown to detect variations in level of consciousness over time. Taken

together, advances in neurologic assessment provided by the RASS and the CAM-ICU should lead to better characterization of acute brain dysfunction as an organ failure, reductions in the random variation with which patients' sedatives are currently managed, and appropriate interventions aimed at prevention or reversal of acute brain dysfunction.

**Author Affiliations:** Department of Medicine, Division of Allergy, Pulmonary, and Critical Care Medicine (Drs Ely, Thomason, Wheeler, and Bernard, and Ms Truman), and Geriatric Psychiatry and Neuroscience (Drs Gordon and Margolin), Center for Health Services Research (Drs Ely, Shintani, Dittus, Gordon, Speroff, and Ms Truman) of Vanderbilt University School of Medicine, Nashville, Tenn; Geriatric Research Education and Clinical Center of the Veterans Administration Tennessee Valley Healthcare System, Nashville (Drs Ely, Gordon, Speroff, and Dittus); Department of Preventive Medicine, Division of Biostatistics (Dr Gautam), Division of Quality and Data Management (Dr Francis), St Vincent Hospital Health System, Indianapolis, Ind; Department of Internal Medicine, Virginia Commonwealth University, Richmond (Dr Sessler).

**Author Contributions:** Study concept and design: Ely, Truman, Wheeler, Gordon, Francis, Margolin, Sessler, Dittus, Bernard.

**Acquisition of data:** Ely, Truman, Wheeler, Gordon, Francis, Margolin, Bernard.

**Analysis and interpretation of data:** Ely, Truman, Shintani, Thomason, Wheeler, Speroff, Gautam, Margolin, Sessler, Dittus, Bernard.

**Drafting of the manuscript:** Ely, Truman, Shintani, Thomason.

**Critical revision of the manuscript for important intellectual content:** Ely, Truman, Shintani, Thomason, Wheeler, Gordon, Francis, Speroff, Gautam, Margolin, Sessler, Dittus, Bernard.

**Statistical expertise:** Ely, Shintani, Speroff, Gautam.

**Obtained funding:** Ely, Dittus, Bernard.

**Administrative, technical, or material support:** Ely, Truman, Thomason, Wheeler, Gordon, Francis, Margolin, Dittus, Bernard.

**Study supervision:** Ely, Wheeler, Dittus, Bernard.

**Funding/Support:** Dr Ely is a recipient of the AFAR Pharmacology in Aging Grant and the Paul Beeson Faculty Scholar Award from the Alliance for Aging Research, a recipient of a K23 from the National Institutes of Health (AG01023-01A1) and the Geriatric Research Education and Clinical Center (GRECC). Aspect Medical Systems Inc provided the necessary equipment and financial support to obtain the electroencephalographic measurements, but was not involved with patient assessments, data collection. Aspect did have access to and helped in analysis of the data for the BIS-XP-EEG figure.

**Acknowledgment:** We thank the tireless and devoted nursing staff and respiratory care practitioners in our ICU, who constantly strive to improve the delivery of care for our critically ill patients.

**REFERENCES**

- Ostermann ME, Keenan SP, Seiferling RA, Sibbald W. Sedation in the intensive care unit: a systematic review. *JAMA*. 2000;283:1451-1459.
- Kress JP, Pohlman AS, O'Connor MF, Hall JB. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. *N Engl J Med*. 2000;342:1471-1477.
- Young C, Knudsen N, Hilton A, Reves JG. Seda-

- tion in the intensive care unit. *Crit Care Med.* 2000; 28:854-866.
4. Shapiro BA, Warren J, Egol A, et al. Society of Critical Care Medicine. Practice parameters for intravenous analgesia and sedation for adult patients in the intensive care unit: an executive summary. *Crit Care Med.* 1995;23:1596-1600.
  5. Fraser G, Prato BS, Riker R, Berthiaume D, Wilkins ML. Frequency, severity, and treatment of agitation in young versus elderly patients in the ICU. *Pharmacotherapy.* 2000;20:75-82.
  6. Joint Commission on Accreditation of Healthcare Organizations (JCAHO). *Comprehensive Accreditation Manual for Hospitals: The Official Handbook.* Oakbrook Terrace, Ill: JCAHO; 2003.
  7. Payne JP, Bru O, Bossou JL, et al. Assessing pain in critically ill sedated patients by using a behavioral pain scale. *Crit Care Med.* 2001;29:2258-2263.
  8. 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.
  9. Park G, Coursin D, Ely EW, et al. Balancing sedation and analgesia in the critically ill. *Crit Care Clin.* 2001; 17:1015-1027.
  10. Kollef MH, Levy NT, Ahrens T, et al. The use of continuous IV sedation is associated with prolongation of mechanical ventilation. *Chest.* 1999;114:541-548.
  11. Wheeler AP. Sedation, analgesia, and paralysis in the intensive care unit. *Chest.* 1993;104:566-577.
  12. Bair N, Bobek MB, Hoffman-Hogg L, et al. Introduction of sedative, analgesic, and neuromuscular blocking agent guidelines in a medical intensive care unit: physician and nurse adherence. *Crit Care Med.* 2000; 28:707-713.
  13. Slomka J, Hoffman-Hogg L, Mion LC, et al. Influence of clinicians' values and perceptions on use of clinical practice guidelines for sedation and neuromuscular blockade in patients receiving mechanical ventilation. *Am J Crit Care.* 2000;9:412-418.
  14. Mascia MF, Koch M, Medicis JJ. Pharmacoeconomic impact of rational use guidelines on the provision of analgesia, sedation, and neuromuscular blockade in critical care. *Crit Care Med.* 2000;28:2300-2306.
  15. Nasraway SA, Jacobi J, Murray MJ, Lumb PD. Sedation, analgesia, and neuromuscular blockade of the critically ill adult: revised clinical practice guidelines. *Crit Care Med.* 2002;30:117-118.
  16. Ramsay M, Savege TM, Simpson ER, Goodwin R. Controlled sedation with alphaxalone-alphadolone. *BMJ.* 1974;2:656-659.
  17. Brook AD, Ahrens TS, Schaiff R, et al. Effect of a nursing implemented sedation protocol on the duration of mechanical ventilation. *Crit Care Med.* 1999; 27:2609-2615.
  18. MacIntyre NR, Cook DJ, Ely EW, et al. Evidence-based guidelines for weaning and discontinuing ventilatory support. *Chest.* 2001;120(suppl 6):3755-3955.
  19. Ely EW, Meade M, Haponik EF, et al. Mechanical ventilator weaning protocols driven by non-physician health care professionals: evidence-based clinical practice guidelines. *Chest.* 2001;120(suppl):4545-4635.
  20. Hansen-Flaschen J, Cowen J, Poloman RC. Beyond the Ramsay scale: need for a validated measure of sedating drug efficacy in the intensive care unit. *Crit Care Med.* 1994;22:732-733.
  21. Lieberman J, Trember KK. Sedation: if you don't know where you are going, any road will get you there. *Crit Care Med.* 1999;27:1395-1396.
  22. Wittbrodt ET. The ideal sedation assessment tool: an elusive instrument. *Crit Care Med.* 1999;27:1384-1385.
  23. Devlin JW, Boleski G, Mlynarek M, et al. Motor activity assessment scale: a valid and reliable sedation scale for use with mechanically ventilated patients in an adult surgical intensive care unit. *Crit Care Med.* 1999;27:1271.
  24. Ambuel B, Hamlett KW, Marx CM, Blumer JL. Assessing distress in pediatric intensive care environments: the COMFORT scale. *J Pediatr Psychol.* 1992; 17:95-109.
  25. Riker R, Picard JT, Fraser G. Prospective evaluation of the sedation-agitation scale for adult critically ill patients. *Crit Care Med.* 1999;27:1325-1329.
  26. De Jonghe B, Cook D, Appere-De-Vecchi C, Guyatt G, Meade M, Outin H. Using and understanding sedation scoring systems: a systematic review. *Intensive Care Med.* 2000;26:275-285.
  27. Riker RR, Fraser GL, Simmons LE, Wilkins ML. Validating the sedation-agitation scale with the bispectral index and visual analog scale in adult ICU patients after cardiac surgery. *Intensive Care Med.* 2001; 27:853-858.
  28. Sessler CN, Grap MJ, Brophy GM. Multidisciplinary management of sedation and analgesia in critical care. *Semin Respir Crit Care Med.* 2001;22:211-225.
  29. Sessler CN, Gosnell M, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care patients. *Am J Respir Crit Care Med.* 2002;166:1338-1344.
  30. Plum F, Posner J. *The Diagnosis of Stupor and Coma.* 3rd ed. Philadelphia, Pa: FA Davis Co; 1980.
  31. 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.
  32. Ely EV, 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.
  33. Cammarano WB, Drasner K, Katz J. Pain control, sedation, and use of muscle relaxants. In: Hall JB, Schmidt GA, Wood LD, eds. *Principles of Critical Care Medicine.* New York, NY: McGraw Hill Publishers; 1998:90-97.
  34. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. *Lancet.* 1974; 1:81-84.
  35. Carmines EG, Zeller RA. *Reliability and Validity.* Newberry Park, Calif: Sage Publications; 1979.
  36. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: Wiley & Sons; 1981.
  37. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.* Washington, DC: American Psychiatric Association; 1994.
  38. Sackett DL, Haynes RB, Guyatt GH, et al. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* 2nd ed. Boston, Mass: Little Brown & Co; 1991.
  39. Parone SV, Tsuang MT. Measuring diagnostic accuracy in the absence of a "gold standard." *Am J Psychiatry.* 2002;151:650-657.
  40. Rosow C, Manberg PJ. Bispectral index monitoring. *Anesthesiol Clin North Am.* 2001;19:947-966.
  41. Johansen JW, Sebel PS. Development and clinical application of electroencephalographic bispectrum monitoring. *Anesthesiology.* 2002;2000:1336-1344.
  42. Nasraway SA, Wu EC, Kelleher RM, Yasuda CM, Donnelly AM. How reliable is the Bispectral Index in critically ill patients? a prospective, comparative, single-blinded observer study. *Crit Care Med.* 2002;30:1483-1487.
  43. Ely EW, Truman B, May L, et al. Validating the bispectral EEG for ventilated ICU patients. *Am J Respir Crit Care Med.* 2001;163:A899.
  44. Frenzel D, Greim CA, Sommer C, Bauerle K, Roewer N. Is the bispectral index appropriate for monitoring the sedation level of mechanically ventilated surgical ICU patients? *Intensive Care Med.* 2002;28:178-183.
  45. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med.* 1985;13:818-829.
  46. Trujillo KA, Akil H. Opiate tolerance and dependence: recent findings and synthesis. *New Biol.* 1991; 3:915-923.
  47. Vinik HR, Kissin I. Rapid development of tolerance to analgesia during remifentanyl infusion in humans. *Anesth Analg.* 1998;86:1307-1311.
  48. Eilers H, Philip LA, Bickler PE, McKay WR, Schumacher MA. The reversal of fentanyl-induced tolerance by administration of "small dose" ketamine. *Anesth Analg.* 2001;93:213-214.
  49. Eappen T, Wittbrodt E, Fuchs B, et al. Reliability of four sedation assessment scales. *Crit Care Med.* 1999; 27:A135.
  50. Bergeron N, Dubois MJ, Dumont M, Dial S, Skrobik Y. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med.* 2001; 27:859-864.
  51. Dubois MJ, Bergeron N, Dumont M, Dial S, Skrobik Y. Delirium in an intensive care unit: a study of risk factors. *Intensive Care Med.* 2001;27:1297-1304.
  52. Ely EW, Stephens RK, Jackson JC, et al. Current opinions regarding delirium in the ICU: a survey of 915 healthcare professionals. *Crit Care Med.* In press.
  53. Carnes M, Howell T, Rosenberg M, et al. Physicians vary in approaches to the clinical management of delirium. *J Am Geriatr Soc.* 2003;51:234-239.
  54. Marcantonio ER, Juarez G, Goldman L, et al. The relationship of postoperative delirium with psychoactive medications. *JAMA.* 1994;272:1518-1522.
  55. Hall RI, Sandham D, Cardinal P, et al. Propofol vs midazolam for ICU sedation: a Canadian multicenter randomized trial. *Chest.* 2001;119:1151-1159.
  56. Collier Z, Shintani A, Truman B, et al. The impact of delirium in the ICU on outcomes of mechanical ventilation. *Am J Respir Crit Care Med.* 2003;167:A970.
  57. Brattebo G, Hofoss D, Flaatten H, et al. Effect of a scoring system and protocol for sedation on duration of patients' need for ventilator support in a surgical intensive care unit. *BMJ.* 2002;324:1386-1389.
  58. Hall-Lord ML, Larsson G, Steen B, Sweden K. Pain and distress among elderly intensive care unit patients: comparison of patients' experiences and nurses' assessments. *Heart Lung.* 1998;27:123-132.
  59. Schelling G, Stoll C, Haller M, et al. Health-related quality of life and posttraumatic stress disorder in survivors of the acute respiratory distress syndrome. *Crit Care Med.* 1998;26:651-659.
  60. Michaels AJ, Michaels CE, Smith JS, et al. PTSD in critical care. *J Trauma.* 2000;48:841-850.
  61. Rotondi AJ, Chelluri L, Sirio C, et al. Patients recollections of stressful experiences while receiving prolonged mechanical ventilation in an intensive care unit. *Crit Care Med.* 2002;30:746-752.
  62. Scragg P, Jones A, Fauvel N. Psychological problems following ICU treatment. *Anaesthesia.* 2001;56: 9-14.
  63. Truman B, Shintani A, Jackson J, et al. Implementation of the SCCM guidelines for sedation and delirium monitoring in the ICU. *Am J Respir Crit Care Med.* 2003;167: A969.
  64. Riker RR, Fraser GL. Sedation in the intensive care unit: refining the models and defining the questions. *Crit Care Med.* 2002;30:1661-1663.