Published in final edited form as:

Crit Care Med. 2016 January ; 44(1): 138–146. doi:10.1097/CCM.0000000000001375.

Validity of a Modified Sequential Organ Failure Assessment Score Using the Richmond Agitation-Sedation Scale

Eduard E. Vasilevskis, MD, MPH^{1,2,8}, Pratik P. Pandharipande, MD, MSCI^{4,9}, Amy J. Graves, MS, MPH⁵, Ayumi Shintani, PhD, MPH⁶, Ryosuke Tsuruta, MD⁷, E. Wesley Ely, MD, MPH^{1,3,8}, and Timothy D. Girard, MD, MSCI^{1,3,8}

¹Center for Health Services Research, Vanderbilt University School of Medicine, Nashville, TN

²Division of General Internal Medicine and Public Health, Vanderbilt University School of Medicine, Nashville, TN

³Division of Allergy, Pulmonary, and Critical Care Medicine in the Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN

⁴Division of Critical Care in the Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN

⁵Department of Urology at the Vanderbilt University School of Medicine, Nashville, TN

⁶Department of Emergency and Critical Care Medicine, Tokai University School of Medicine, Kanagawa, Japan

⁷Advanced Medical Emergency and Critical Care Center, Yamaguchi University Hospital, Yamaguchi, Japan

⁸Geriatric Research, Education and Clinical Center (GRECC) Service, Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville TN

⁹Anesthesia Service at the Department of Veterans Affairs Medical Center, Tennessee Valley Healthcare System, Nashville TN

Abstract

Objective—The Sequential Organ Failure Assessment (SOFA) and other severity of illness scales rely on the Glasgow Coma Scale (GCS) to measure acute neurologic dysfunction, but the GCS is unavailable or inconsistently applied in some institutions. The objective of this study was to assess the validity of a modified SOFA that uses the Richmond Agitation-Sedation Scale (RASS) instead of GCS.

Design—Prospective cohort study

Setting—Medical and surgical intensive care units within a large, tertiary care hospital.

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.

Author Disclosures: All the other authors report no financial conflicts of interest.

Copyright form disclosures: The remaining authors have disclosed that they do not have any potential conflicts of interest.

Patients—Critically ill medical/surgical ICU patients.

Interventions and measurements—We calculated daily SOFA scores using electronic medical record-derived data. Using bedside nurse-recorded GCS and RASS measures, we calculated neurologic SOFA scores using the original GCS-based approach (SOFA_{GCS}) and a novel RASS-based approach (SOFA_{RASS}), converting the 10-point RASS to a four-point neurologic SOFA score. We assessed construct validity of SOFA_{RASS} by analyzing correlations with established severity of illness constructs (APACHE II and SOFA_{GCS}) and predictive validity by using logistic regression to determine whether SOFA_{RASS} predicts ICU, hospital, and 1-year mortality. We assessed discriminative performance with C-statistics.

Main Results—Among 513 patients (5,199 patient-days), SOFA_{RASS} was strongly correlated with APACHE II acute physiology score at enrollment ($r = 0.583$, 95% CI 0.518, 0.642) and daily SOFA_{GCS} scores ($r = 0.963$, 95% CI 0.956 to 0.968). Mean SOFA_{RASS} scores predicted ICU mortality (AUC = 0.814)—as did mean SOFA_{GCS} (0.799)—as well as hospital and 1-year mortality. Admission SOFA scores, whether using RASS or GCS were less accurate predictors of mortality; AUCs for ICU mortality for SOFA_{RASS} and SOFA_{GCS}, for example, were 0.622 and 0.608, respectively.

Conclusions—A modified SOFA score that uses bedside RASS when GCS data are not available is a valid means of assessing daily severity of illness in the ICU and may be valuable for risk adjustment and benchmarking purposes.

Keywords

organ dysfunction scores; risk adjustment; severity of illness index; mortality; coma; sequential organ failure assessment scores

INTRODUCTION

Clinicians, investigators, and administrators frequently employ severity of illness scoring systems in the intensive care unit (ICU) to risk stratify patients for treatment, aid prognosis, risk adjust outcomes research, and benchmark performance across patients (1–3). The Sequential Organ Failure Assessment (SOFA) was one of the first severity of illness scoring systems to be validated for use over time in the ICU. The SOFA score, first shown to describe changing organ function in sepsis patients (2), has been since validated as a predictor of future organ dysfunction and mortality (4–6). Consisting of six individual organ dysfunction measures (i.e., respiratory, coagulation, liver, cardiovascular, neurologic, and renal), the SOFA is now widely used to assess the severity of illness in adult medical and surgical ICU patients (2, 4–7), as well as pediatric ICU patients (8).

A major strength of the SOFA is that most of the data used to calculate the score are highly reliable and routinely available in medical records (9). The neurologic component, however, is based upon the Glasgow Coma Scale (GCS) (10), which has been shown to be less reliably measured (11). In addition, the GCS is often not recorded in a structured format in medical files, is miscalculated (e.g., when scoring intubated patients), is simply missing, and has been supplanted in many institutions by instruments such as the Richmond Agitation-Sedation Scale (RASS) for daily neurological monitoring of medical and surgical ICU

patients (or any population admitted to the ICU for reasons other than primary structural neurological injury) (9, 12–15). As a result, the SOFA has been commonly modified to exclude the GCS-based neurologic component (12, 14). Such modifications are less than ideal since they remove or attenuate neurologic end-organ information that is predictive of mortality (5, 16) and critical when there is interest in changing neurologic function.

Like the GCS, the RASS (17, 18) is an ordinal score that objectively measures level of consciousness (see Figure E1 electronic supplement). Importantly, the RASS has been shown to be valid and highly reliable for arousal assessment over time of critically ill patients, whether mechanically ventilated or not (17, 18), and in both research and clinical settings (19, 20). We therefore hypothesized that a modified SOFA score, which replaces the GCS-based with a RASS-based neurologic component, would meet objective measures of validity and provide equivalent predictive performance of mortality compared with the original GCS-based SOFA score.

MATERIALS AND METHODS

Study Design and Population

We assessed the construct and predictive validity of the RASS-based SOFA scoring system within the BRAIN-ICU study (21), a prospective cohort study examining neurologic and other short- and long-term outcomes in a broad population of critically ill patients. A detailed description of the parent study was previously published (21). Briefly, in BRAIN-ICU we enrolled patients admitted to medical or surgical ICUs with respiratory failure and/or shock after excluding those who could not be assessed for delirium by study staff during the entirety of the hospital stay (e.g., due to severe acute or advanced chronic neurologic conditions [e.g., advanced dementia only], persistent moribund state, blindness, or inability to understand English). Patients admitted with coma (e.g., due to medication or encephalopathy) with an expectation for recovery, however, were not excluded. Patients were additionally excluded from the parent study if they had a recent ICU visit to eliminate the potential of confounding of this recent exposure with long-term cognitive outcomes assessed as part of that study. We restricted eligibility for the current analyses to BRAIN-ICU subjects enrolled at Vanderbilt University Hospital, where bedside arousal assessments had been previously validated (19). In addition, GCS assessments are routinely available in the medical and surgical ICUs at Vanderbilt, allowing for a rigorous comparison of a GCS-based SOFA with a RASS-based SOFA. The institutional review board at Vanderbilt University approved this study.

Data Collection: Clinical Characteristics, SOFA Data, and Outcomes

Research personnel used electronic medical records to record demographic, comorbidity, and severity of illness data at ICU admission, including age, Acute Physiology and Chronic Health Evaluation (APACHE) II score at admission (1), use of mechanical ventilation, sepsis at admission (22, 23), and the Charlson comorbidity index (24). In addition, research personnel used the electronic medical record to obtain data, including GCS and RASS, to calculate the SOFA score for each specific organ domain. Both the GCS and RASS were assessed and recorded daily as part of routine care by bedside nurses in all ICUs that

participated in this study and were thus not necessarily concurrently performed. Specifically, nurses recorded the GCS twice daily and the RASS at least every 4 hours.

GCS and RASS measures were used to determine the SOFA neurologic component scores based on either GCS (SOFA-Neuro_{GCS}) (10) or RASS (SOFA-Neuro_{RASS}) scores using the *a priori*-specified methods shown in Table 1. RASS conversions to SOFA-Neuro_{RASS} scores were informed by prior validation research that demonstrated a strong relationship between the total GCS and the RASS (18). Similar to this prior work, due to the extremely low number of RASS assessments that were greater than 0 (i.e., only 0.6% of all assessments were in the positive RASS range indicating agitation), we considered any RASS assessment with a value greater than zero in the same fashion as RASS = 0 (i.e., translating to SOFA-Neuro_{RASS} = 0).

We then calculated SOFA scores using the original method, which we designate SOFA_{GCS}, and a RASS-based method, which we designate SOFA_{RASS}. SOFA scores included the cardiovascular, coagulation, liver, renal, and respiratory SOFA scores using original methods and either SOFA-Neuro_{GCS} or SOFA-Neuro_{RASS} to calculate a SOFA_{GCS} or SOFA_{RASS} score. SOFA_{GCS} or SOFA_{RASS} were calculated for the first 24 hours of ICU admission as well as at study enrollment, and daily thereafter until ICU discharge or death. In the case of multiple, repeated SOFA measures on the same day, the highest (worst) SOFA score was used. If data required to calculate a daily SOFA value were missing, we imputed using last observation carried forward. Finally, we computed summary SOFA severity of illness measures, including admission, maximum, mean, and 48-hour change, according to the methods of Ferreira et al. (6). Whereas the admission SOFA may guide point-of-care risk assessment by clinicians, the mean SOFA score is commonly used by researchers to reflect the cumulative severity of illness when risk-adjusting for outcomes that follow ICU discharge (21). See the online data supplement for further details regarding data collection methods.

We determined vital status until hospital discharge using electronic medical records and evaluated one-year mortality using phone follow-up and the Social Security Administration's (SSA) Death Master File (25). Data in the Death Master File, which provides a valid, publicly available source of mortality data widely used in outcomes research, were linked to patient records by social security number or, when that was not available, via a matching algorithm that utilizes other variables.

Statistical Analysis

We assessed the construct validity of the newly created SOFA-Neuro_{RASS} component as a measure of neurologic dysfunction by assessing the distribution of the raw measures and correlations between RASS and GCS and between SOFA-Neuro_{RASS} and SOFA-Neuro_{GCS}. We also assessed the construct validity of SOFA_{RASS} as a measure of severity of illness in several ways. First, we calculated the correlation between SOFA_{RASS} and APACHE II scores at ICU admission (1). We also evaluated the correlation between individual daily SOFA_{RASS} and daily SOFA_{GCS}. Finally, we assessed correlations between four summary SOFA_{RASS}-derived severity of illness measures (i.e., admission, maximum, mean, 48-hour change) and corresponding SOFA_{GCS} measures. We used Spearman's rank correlation to

analyze all correlations. For all analyses where there were multiple measurements per individual, the bootstrapped confidence intervals take into account non-independence of measurements within each patient.

We assessed the predictive validity of the four summary SOFA-Neuro_{RASS}-derived neurologic component measures (i.e., admission, maximum, mean, 48-hour change) as well as the four summary SOFA_{RASS}-derived severity of illness measures by analyzing the associations between these measures and ICU, hospital, and 1-year mortality. Specifically, we used bivariable logistic regression to calculate the odds ratio and Wald 95% confidence interval for each predictor. We evaluated the discriminative performance of each score by calculating the areas under the curve (AUC) derived from the receiver-operating characteristic (ROC). Finally, we compared the AUCs of SOFA_{GCS} predicting ICU, hospital, and 1-year mortality with the AUCs of the best performing SOFA_{RASS} to assess for any statistically significant differences ($p < 0.05$) between AUCs using the bootstrap method using R package “pROC” (26).

We used R version 2.12 (www.r-project.org) to perform all statistical analyses. Some results have previously been published in abstract form (27), and data from these subjects pertaining to delirium and long-term outcomes have previously been published (21). Because some patients had missing SOFA data between ICU admission and study enrollment (day 2 of the ICU stay for these patients) we additionally performed a complete case analysis to examine for any potential bias introduced from imputation of missing data ($N = 235$). Additional details, including those regarding sensitivity analyses, are described in the electronic supplement.

RESULTS

Patient Characteristics

A total of 3,556 patients were screened at Vanderbilt; 3,035 of these were excluded, most often because of a recent ICU stay (Figure 1). Of the 521 patients enrolled, we excluded 3 patients who withdrew consent. An additional 5 patients did not have clinical RASS data available or had less than 24 hours of data. The characteristics and outcomes of the 513 patients included in this study are shown in Table 2. The mean SOFA_{GCS} at ICU admission and study enrollment indicated a high severity of illness. As expected, enrollment SOFA_{GCS} was lower, indicating lower severity of illness, among those who survived to ICU discharge than those who died in the ICU (9.7 ± 3.6 vs. 11.3 ± 3.8).

GCS and RASS were used to assess the 513 participants on a total of 5,199 patient-days. Though all three methods of converting RASS to SOFA-Neuro_{RASS} scores (see Table 1) yielded consistent results (see electronic supplement), the score based on ‘method C’ had the best predictive validity based on its association with mortality. Thus, the following results represent this conversion, whereas all results obtained using the alternate conversions are described in the electronic supplement.

Construct Validity

GCS and RASS were strongly correlated (Spearman $\rho = 0.806$, 95% CI 0.785 to 0.825) with each other across all daily values (Figure E2 in supplement shows distribution). Similarly, SOFA-Neuro_{GCS} and SOFA-Neuro_{RASS} neurologic component scores were also strongly correlated when considered on a daily basis or as the mean score during the ICU stay (Table 3), and moderately correlated for admission, maximum, or 48-hour change scores (Table 3).

Total SOFA_{RASS} correlated with established measures of severity of illness. SOFA_{RASS} scores at enrollment, for example, were moderately correlated with APACHE acute physiology scores at enrollment ($r = 0.58$, 95% CI 0.52, 0.64). In addition, total SOFA_{RASS} scores were strongly correlated with SOFA_{GCS} scores, regardless of whether they were examined at admission, on a daily basis, or as the maximum, mean, or 48-hour change during the ICU stay (Table 3).

Predictive Validity

Mean SOFA-Neuro_{RASS} and SOFA_{RASS} scores in the ICU showed excellent discrimination for ICU and hospital mortality (Table 4). Of the SOFA-Neuro_{RASS} and SOFA_{RASS} summary measures examined, mean and maximum scores consistently demonstrated the highest discrimination, whereas the admission score demonstrated the lowest. Though all versions of SOFA_{RASS} significantly predicted mortality at all time points, discrimination was highest for ICU mortality and lowest for 1-year mortality. Upon comparison of AUCs, mean SOFA_{RASS} was a statistically significantly better predictor of ICU ($p = 0.002$) and hospital mortality ($p = 0.03$) than mean SOFA_{GCS} (Figure 2), though the differences between the two measures are small and unlikely to be clinically relevant. The AUCs for 1-year mortality were similar ($p = 0.31$).

Neurologic function alone, as assessed by the SOFA-Neuro_{RASS}, was predictive of mortality, with the odds of death in the ICU increased 5.08-fold (95% CI 3.33 – 7.72) for each 1-point increase in mean SOFA-Neuro_{RASS}. For the complete score, the odds of death in the ICU increased 1.52-fold (95% CI 1.38 to 1.68) for each 1-point increase in mean SOFA_{RASS}, and the odds of death in the ICU increased 1.13-fold (95% CI 1.06 to 1.21) for each point increase in admission SOFA_{RASS} (Table E1).

Sensitivity Analysis

There were minimal differences in predictive validity with the use of alternate RASS conversions (Tables E2 and E3). The relationship between mean SOFA-Neuro_{RASS} scores and ICU and hospital mortality were nonlinear (p -value test for linearity < 0.05), though allowing nonlinearity in the regression models had minimal effect on corresponding AUCs (Tables E4–E6). There were no significant nonlinear associations for the SOFA_{RASS} models (eTables E4–E6). Complete case analyses ($N = 278$) of both neurologic and total SOFA scores produced similar findings in terms of OR, although there was no longer statistically significant differences between the AUCs of the SOFA_{GCS} and SOFA_{RASS} in prediction of ICU ($p = 0.06$) and hospital mortality ($p = 0.19$) (Tables E7–9).

DISCUSSION

In this prospective cohort study of medical and surgical ICU patients, we found that the RASS, a commonly performed bedside arousal scale, can be successfully converted to a neurologic SOFA score, and that following conversion, the RASS-based neurologic and total SOFA scores exhibit construct and predictive validity. Whether using RASS or GCS, admission SOFA scores were less accurate than mean SOFA scores when predicting mortality, suggesting a more limited clinical role for risk assessment at the time of ICU admission. Mean SOFA scores, however, had excellent predictive accuracy and may serve as a risk-adjustment tool for post-ICU outcomes research or benchmarking, especially in ICUs where GCS is unavailable or inconsistently scored.

Our data suggest that the RASS is useful as a tool to quantify severity of acute brain dysfunction during critical illness. Sessler et al. (17) and Ely et al. (18) previously described high correlations between RASS and GCS, and we converted RASS to a neurologic SOFA using the GCS-to-RASS conversions reported by Ely et al. (18). We found moderate correlations between the original RASS and GCS scores, with higher correlations when multiple data points were used or averaged across days. Since five of the six SOFA components were unchanged, it was unsurprising that correlations between the two versions of SOFA were high. We subsequently found that a RASS-based SOFA score consistently had predictive performance for mortality similar to that of the GCS-based SOFA score, supporting the validity of RASS as an appropriate measure of acute neurologic dysfunction. Although the predictive performance of the three unique SOFA-Neuro_{RASS} scores (described in Table E1) were only modestly different from each other, the conversion presented has the benefit of being numerically aligned with the SOFA scoring system (i.e., 0 : 0, -1 : 1, -2 : 2, -3 : 3, < -4 : 4). Because of this alignment with SOFA, this conversion may be the easiest to apply.

It is unclear whether the effect of sedation should be accounted for when assessing the severity of neurologic organ dysfunction. Some have suggested that sedation may falsely elevate patients' extent of neurologic injury and subsequent risk of mortality (2), whereas recent studies have shown deep sedation to be associated with significantly higher mortality risk (28, 29). GCS and RASS assessments in the current study were performed at the bedside by clinical nurses as part of routine care of patients; the results, therefore, incorporate the effects of sedation, which tends to be heaviest early during the ICU stay. As a result, mortality prediction in this manner can provide prognostic information but should not be used for benchmarking of outcomes that occur prior to ICU discharge given the inclusion of care processes, e.g., sedation. In addition, use of clinical measures may, in part, explain the less accurate predictions of ICU mortality when using either the SOFA-Neuro_{GCS} and SOFA-Neuro_{RASS} obtained at ICU admission.

It is likely that mean and maximum measures are more accurate because they take into account more data. Whereas Khwannimit et al. (16) found an AUC of 0.840 for admission SOFA-Neuro_{GCS} as a predictor of ICU mortality, we found an AUC of only 0.569. A wide range of discrimination for SOFA scores has been reported that reflect different populations and evolving ICU practices over time (30). In addition, AUCs are predictably lower when

there is a more limited range of risk represented in the population (our population was restricted to patients with sepsis and/or respiratory failure) coupled with a relatively frequent outcome (31). In contrast, the mean SOFA-Neuro_{RASS} score in our study yielded excellent predictive ability for ICU and hospital mortality. The greatest predictive value of the mean score followed by the maximum score observed in this study corresponds to the original work by Ferreira et al. (6), which similarly found mean values to be the most predictive.

Although “removing” the effect of sedation may be feasible in a research setting, it runs counter to nursing and physician practice in the clinical setting. The goal of the RASS is to assess a patient’s level of arousal in light of everything that may be affecting their brain, including sedation, so that adjustments in treatments can be made to meet sedation goals (15, 32). We made no efforts to account for sedation in our analyses, and yet the SOFA-Neuro_{RASS} and SOFA_{RASS} scores both strongly predicted mortality. Although in some instances there was a statistically significant difference in discrimination between SOFA_{RASS} and SOFA_{GCS}, the clinical difference is minimal. Importantly, the similar predictive value suggests that the use of “actual” rather than “assumed” (i.e., if sedation were not being used) level of consciousness continues to accurately measure the severity of neurologic dysfunction, especially as the ICU stay progresses. This may reflect either the importance of a severe underlying illness that requires the deeper levels of sedation, and/or harmful effects of sedative agents or burst suppression (33).

Whereas alternative versions of the SOFA that do not include the neurologic status have been proposed (9, 14), we demonstrated that the replacement of the GCS with the RASS is a valid method by which to calculate the complete SOFA. A variety of attributes makes the RASS an appealing choice for neurologic component of SOFA. First, daily sedation assessments that include the RASS are recommended by critical care consensus guidelines (15). Second, the RASS is easily measured with minimal training, allowing reproducibility and increased acceptance of the instrument (18–20). Although our study did not examine whether other sedation scales (e.g., Ramsay (34), Sedation-Agitation Scale (35), etc.) could also be used to calculate the neurologic component of the SOFA, other studies have suggested these measures may be interchangeable (36).

The SOFA_{RASS} correlated with two validated severity of illness measures (i.e., APACHE II and SOFA_{GCS}), and accurately predicted mortality. The observed decrease in accuracy for long-term outcomes is expected since additional unobserved and accumulating risk factors have increasing influence on longer-term mortality. Of the various SOFA_{RASS}-derived severity of illness measures, the mean SOFA_{RASS} was the most predictive for each mortality outcome followed by the maximum SOFA_{RASS} score. This result was consistent with those reported by Ferreira et al. (6) using SOFA_{GCS} and confirms the importance of considering the cumulative effect of organ dysfunction over the course of the ICU stay. In fact, Moreno et al. (5) found that maximum dysfunction of renal, coagulation, and hepatic organs typically does not occur until several days into an ICU stay, such that SOFA scores at admission or during the first 48 hours may not reflect the full extent of organ dysfunction. These factors may limit the utility of the SOFA score for clinicians looking to use the admission score to guide treatment decisions and for risk assessment. The mean SOFA

score, however, may serve an important risk-adjustment function for researchers or administrators investigating or benchmarking post-ICU outcomes.

Several important limitations of this study deserve attention. First, the setting was a single, large, academic teaching hospital, and the study was performed within a selected cohort of critically ill patients who were managed with practices, e.g., light sedation, that may be specific to this hospital. Although this institution treats a large population of medical and surgical ICU patients, our findings may not generalize to all settings or to all patients. Second, we calculated all SOFA scores using data obtained in the course of usual care; some data needed were therefore not available. It is possible that the last observation carried forward method, which we used to account for missing data, may underestimate the risk of death in patients whose organ failure was not assessed until it was more advanced. A complete case analysis among patients without missing data, however, revealed almost identical results with the primary analysis. Third, we treated all RASS values greater than 0 (indicating agitation) equivalent to a RASS equal to 0 (normal). Although there are clinical differences between agitation and an alert/calm state, agitation as indicated by RASS >0 accounted for a very small fraction of measurements (0.6%) in our study and may be even less common in a less severely ill population. Our approach to analyzing agitation is therefore unlikely to affect the findings. Finally, clinical RASS data may be less accurate than that obtained by trained research assistants. Research-derived data may generate SOFA scores with better discrimination, but this “real-world” limitation produced results that are generalizable to clinical practice.

CONCLUSIONS

The SOFA is a common measure of severity of illness used in intensive care units. Use of the RASS in SOFA offers an opportunity for wider implementation of this valuable risk-prediction tool, especially for research and administrative purposes in ICUs where GCS is unavailable or inconsistently scored. Measures of arousal, including the RASS, are an important component to targeted sedation and also provide valuable information predicting short-and long-term mortality. RASS measures may play a role in severity of illness measurement including SOFA and should be evaluated in other severity of illness measures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Source: This study was supported by the National Institutes of Health (AG027472). In addition, Drs. Vasilevskis, Pandharipande, Ely, and Girard received support from the National Institutes of Health (AG040157, HL111111, AG035117, and AG034257). Drs. Pandharipande and Ely also received support from the VA Clinical Science Research and Development Service (VA Career Development Award and VA Merit Review Award, respectively) and Drs. Vasilevskis, Ely, and Girard received support from the Veterans Affairs Tennessee Valley Geriatric Research, Education and Clinical Center (GRECC).

Dr. Pandharipande has received honoraria from Hospira, Inc. and Orion Pharma. Dr. Ely has received honoraria from Pfizer, Inc., Eli Lilly and Company, Hospira, Inc., Abbott Laboratories, and Orion Pharma. Dr. Girard has received honoraria from Hospira, Inc.

Dr. Vasilevskis received support for article research from the National Institutes of Health (NIH). His institution received grant support from the NIH. Dr. Pandharipande received support for article research from the NIH. His institution received grant support from Hospira Inc. Dr. Shintani received support for article research from the NIH. Dr. Ely received support for article research from the NIH and VA Funding and VA Merit Awards. He received funding from Orion, Abbott Laboratories, Hospira. His institution received grant support from the NIH (support for several ongoing studies: BRAIN ICU, EXD Study, MENDS2 Study and MIND-USA Study) and from the VA (MIND-ICU, PTSD and Predictors of cognitive impairment in survivors of critical illness). His institution received funding from the NIH, VA Funding and VA Merit Awards. Dr. Girard received funding from Hospira and received support for article research from the NIH. His institution received funding from the NIH.

REFERENCES

1. 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]
2. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996; 22:707–710. [PubMed: 8844239]
3. Marshall JC, Cook DJ, Christou NV, et al. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
4. Vincent JL, Mendonca Ad, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: Results of a multicenter, prospective study. *Crit Care Med.* 1998; 26:1793–1800. [PubMed: 9824069]
5. Moreno R, Vincent JL, Matos R, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. Working Group on Sepsis related Problems of the ESICM. *Intensive Care Med.* 1999; 25:686–696. [PubMed: 10470572]
6. Ferreira FL, Bota DP, Bross A, et al. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA.* 2001; 286:1754–1758. [PubMed: 11594901]
7. Badreldin A, Elsobky S, Lehmann T, et al. Daily-Mean-SOFA, a new derivative to increase accuracy of mortality prediction in cardiac surgical intensive care units. *The Thoracic and cardiovascular surgeon.* 2012; 60:43–50. [PubMed: 22215501]
8. Shime N, Kageyama K, Ashida H, et al. Application of modified sequential organ failure assessment score in children after cardiac surgery. *J Cardiothorac Vasc Anesth.* 2001; 15:463–468. [PubMed: 11505351]
9. Nates JL, Cardenas-Turanzas M, Wakefield C, et al. Automating and simplifying the SOFA score in critically ill patients with cancer. *Health Informatics J.* 2010; 16:35–47. [PubMed: 20413411]
10. Bastos PG, Sun X, Wagner DP, et al. Glasgow Coma Scale score in the evaluation of outcome in the intensive care unit: findings from the Acute Physiology and Chronic Health Evaluation III study. *Crit Care Med.* 1993; 21:1459–1465. [PubMed: 8403953]
11. Chen LM, Martin CM, Morrison TL, et al. Interobserver variability in data collection of the APACHE II score in teaching and community hospitals. *Crit Care Med.* 1999; 27:1999–2004. [PubMed: 10507631]
12. Junger A, Engel J, Benson M, et al. Discriminative power on mortality of a modified Sequential Organ Failure Assessment score for complete automatic computation in an operative intensive care unit. *Crit Care Med.* 2002; 30:338–342. [PubMed: 11889305]
13. Baykara N, Gokduman K, Hosten T, et al. Comparison of sequential organ failure assessment (SOFA) scoring between nurses and residents. *Journal of anesthesia.* 2011; 25:839–844. [PubMed: 21931987]
14. Nates JL, Cardenas-Turanzas M, Ensor J, et al. Cross-validation of a modified score to predict mortality in cancer patients admitted to the intensive care unit. *J Crit Care.* 2011; 26:388–394. [PubMed: 21195582]
15. Barr J, Fraser GL, Puntillo K, et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med.* 2013; 41:263–306. [PubMed: 23269131]

16. Khwannimit B. A comparison of three organ dysfunction scores: MODS, SOFA and LOD for predicting ICU mortality in critically ill patients. *J Med Assoc Thai*. 2007; 90:1074–1081. [PubMed: 17624199]
17. Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med*. 2002; 166:1338–1344. [PubMed: 12421743]
18. 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]
19. Vasilevskis EE, Morandi A, Boehm L, et al. Delirium and sedation recognition using validated instruments: reliability of bedside intensive care unit nursing assessments from 2007 to 2010. *J Am Geriatr Soc*. 2011; 59(Suppl 2):S249–S255. [PubMed: 22091569]
20. 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]
21. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. *N Engl J Med*. 2013; 369:1306–1316. [PubMed: 24088092]
22. Dellinger RP, Carlet JM, Masur H, et al. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. *Crit Care Med*. 2004; 32:858–873. [PubMed: 15090974]
23. Dellinger RP, Levy MM, Carlet JM, et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Crit Care Med*. 2008; 36:296–327. [PubMed: 18158437]
24. 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]
25. [cited 2015 May 6] Social Security Administration. Service to epidemiological researchers to provide vital status data on subjects of health research, OMB No. 0960-0701. Available from: <http://www.ssa.gov/policy/about/epidemiology.html>
26. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011; 12:77. [PubMed: 21414208]
27. Vasilevskis E, Pandharipande P, Graves A, et al. Validation of a modified Sequential Organ Failure Assessment (SOFA) score incorporating the Richmond Agitation-Sedation Score (RASS) to assess neurologic end-organ function. *Am J Respir Crit Care Med*. 2013; 187:A1551.
28. Shehabi Y, Riker RR, Bokesch PM, et al. Delirium duration and mortality in lightly sedated, mechanically ventilated intensive care patients. *Crit Care Med*. 2010; 38:2311–2318. [PubMed: 20838332]
29. Zimmerman JE, Kramer AA, McNair DS, et al. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med*. 2006; 34:1297–1310. [PubMed: 16540951]
30. Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: A systematic review. *Crit Care*. 2008; 12:R161. [PubMed: 19091120]
31. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007; 115:928–935. [PubMed: 17309939]
32. 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. [PubMed: 10628598]
33. Watson PL, Shintani AK, Tyson R, et al. Presence of electroencephalogram burst suppression in sedated, critically ill patients is associated with increased mortality. *Crit Care Med*. 2008; 36:3171–3177. [PubMed: 19020432]
34. Ramsay MA, Savege TM, Simpson BR, et al. Controlled sedation with alphaxalone-alphadolone. *Br Med J*. 1974; 2:656–659. [PubMed: 4835444]
35. Riker RR, Picard JT, Fraser GL. Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. *Crit Care Med*. 1999; 27:1325–1329. [PubMed: 10446827]

36. Khan BA, Guzman O, Campbell NL, et al. Comparison and agreement between the Richmond Agitation-Sedation Scale and the Riker Sedation-Agitation Scale in evaluating patients' eligibility for delirium assessment in the ICU. *Chest*. 2012; 142:48–54. [PubMed: 22539644]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.
Patient enrollment

Figure 2.

The figure represents the receiver operator characteristic (ROC) curve for the prediction of hospital mortality by either the mean SOFA_{GCS} (light gray line) or the mean SOFA_{RASS} (black line). The p-value represents the results in the test of difference ($p < 0.05$) between area under the curves using the bootstrap method using R package “pROC” (1).

Table 1

Methods used to convert GCS and RASS into the neurologic SOFA score.

Neurologic SOFA Score	GCS Score	RASS Method A *	RASS Method B *	RASS Method C *
0	15	0	0	0
1	13–14	–1	–1	–1
2	10–12	–2, –3	–2	–2
3	6–9	–4	–3, –4	–3
4	3–5	–5	–5	–4, –5

GCS – Glasgow Coma Scale, RASS – Richmond Agitation Sedation Scale, SOFA – Sequential Organ Failure Assessment.

* We converted RASS from a 10-point scale to a 4-point score using three different *a priori*-specified methods, with method A based on results of a RASS validation study that correlated the RASS with GCS (17) and methods B and C based on clinical judgment, i.e., after considering that alternate categories may more accurately reflect sequential severity of brain dysfunction and better predict outcomes.

Table 2**Patient Characteristics and Outcomes**

Variable	All Patients (n=513)
<i>Baseline characteristics</i>	
Age, mean (SD), years	57 (15)
Sex, N (%)	
Female	246 (48)
Male	267 (52)
Race, N (%)	
White	451 (88)
Non-white	62 (12)
Insurance Type, N(%)	
Medicare	70 (14)
Private	194 (38)
Private and Medicare	136 (27)
Medicare and Medicaid	69 (13)
Medicaid	19 (4)
Uninsured	25 (5)
ICU type, N (%)	
Medical	308 (60)
Surgical	205 (40)
SOFA at ICU admission, mean (SD)	9.7 (3.7)
<i>Enrollment characteristics</i>	
SOFA, mean (SD)	10.0 (3.6)
APACHE II,(1) mean (SD)	25.5 (8.5)
Charlson comorbidity index,(24) median (IQR)	2 (1 to 4)
Mechanically ventilated, N (%)	447 (87)
Septic	306 (60)
<i>Outcomes</i>	
Length of Stay - days, median (IQR)	
ICU	4.8 (2.1 to 10.1)
Hospital	10.1 (5.9 to 17.1)
Mortality, N (%)	
ICU	77 (15)
Hospital	90 (18)
1-Year	184 (36)

Abbreviations: APACHE II - Acute Physiology and Chronic Health Evaluation II, ICU - intensive care unit, IQCODE - Informant Questionnaire on Cognitive Decline in the Elderly, IQR – Interquartile Range, SD – standard deviation, SOFA – Sequential Organ Failure Assessment

Table 3

Correlation between GCS and RASS neurologic and total SOFA scores *

SOFA-derived severity of illness measures	Number of measures	GCS-based vs. RASS based Neurologic Component (SOFA-Neuro _{GCS} vs. SOFA-Neuro _{RASS})	Total SOFA (SOFA _{GCS} vs. SOFA _{RASS})
Admission ^a	513	0.599 (0.529 to 0.661)	0.944 (0.930 to 0.955)
Daily values ^b	5199	0.772 (0.748 to 0.795)	0.963 (0.956 to 0.968)
Maximum ^c	513	0.622 (0.558 to 0.683)	0.977 (0.971 to 0.981)
Mean ^d	513	0.845 (0.811 to 0.874)	0.983 (0.979 to 0.986)
48-hour change ^e	513	0.584 (0.519 to 0.643)	0.835 (0.796 to 0.867)

* Spearman rank correlation coefficients are calculated for the neurologic SOFA score components based on GCS (SOFA-Neuro_{GCS}) and RASS (SOFA-Neuro_{RASS}) measures (ranging from 0 to 4), as well as the total SOFA scores that include either GCS (SOFA_{GCS}) or RASS (SOFA_{RASS}) based neurologic components (ranging from 0 to 24).

^a Admission score = highest score in first 24 hours of ICU admission.

^b Daily values = Highest score for each 24 hour period in the ICU.

^c Maximum = Highest score recorded during any 24-hour period during the entire ICU stay.

^d Mean = Sum of daily scores divided by the number of 24-hour periods in the ICU.

^e 48-hour change score = Difference between the highest score on the score between 72 hours and 24 hours.

Table 4

Area under the receiver operating characteristic for prediction of mortality for GCS and RASS based neurologic and total SOFA scores (N = 513)*

Neurologic SOFA score	AUC -ICU mortality			AUC - Hospital mortality			AUC - I-year mortality		
	SOFA- NeuroGCS	SOFA- NeuroRASS	SOFA- NeuroGCS	SOFA- NeuroGCS	SOFA- NeuroRASS	SOFA- NeuroGCS	SOFA- NeuroGCS	SOFA- NeuroRASS	SOFA- NeuroRASS
Admission ^a	0.518	0.569	0.504	0.579	0.518	0.511			
Maximum ^b	0.637	0.693	0.630	0.683	0.569	0.594			
Mean ^c	0.784	0.819	0.761	0.784	0.636	0.635			
48-hour change ^d	0.636	0.621	0.629	0.600	0.603	0.584			
Total SOFA Score	SOFA _{GCS}	SOFA _{RASS}	SOFA _{GCS}	SOFA _{RASS}	SOFA _{GCS}	SOFA _{RASS}			
Admission ^a	0.608	0.622	0.604	0.623	0.596	0.599			
Maximum ^b	0.770	0.779	0.740	0.748	0.676	0.678			
Mean ^c	0.799	0.814	0.771	0.782	0.681	0.685			
48-hour change ^d	0.657	0.648	0.646	0.624	0.586	0.573			

* Area under the receiver operating characteristics are calculated for GCS based neurologic SOFA score (SOFA-NeuroGCS) and RASS based neurologic SOFA scores (SOFA-NeuroRASS), ranging from 0 to 4, as well as the total SOFAGCS and SOFARASS scores (ranging from 0 to 24).

All scores are based upon SOFA-NeuroRASS(C) to SOFA-NeuroGCS conversions.

^a Admission score = highest score in first 24 hours of ICU admission.

^b Maximum = Highest score recorded during any 24-hour period during the entire ICU stay.

^c Mean = Sum of daily scores divided by the number of 24-hour periods in the ICU.

^d 48-hour change score = Difference between the highest score on the score between 72 hours and 24 hours.