

## ORIGINAL RESEARCH

## Comparison of Mental-Status Scales for Predicting Mortality on the General Wards

Frank J. Zdravec, MPH<sup>1</sup>, Linda Tien, MD<sup>2</sup>, Brian J. Robertson-Dick, MD<sup>3</sup>, Trevor C. Yuen, MS<sup>1</sup>, Nicole M. Twu, MS<sup>1</sup>, Matthew M. Churpek, MD, MPH, PhD<sup>4</sup>, Dana P. Edelson, MD, MS<sup>1\*</sup>

<sup>1</sup>Section of Hospital Medicine, University of Chicago, Chicago, Illinois; <sup>2</sup>Pritzker School of Medicine, University of Chicago, Chicago, Illinois; <sup>3</sup>Department of Physical Medicine and Rehabilitation, Medical College of Wisconsin, Milwaukee, Wisconsin; <sup>4</sup>Section of Pulmonary and Critical Care, University of Chicago, Chicago, Illinois.

**BACKGROUND:** Altered mental status is a significant predictor of mortality in inpatients. Several scales exist to characterize mental status, including the AVPU (Alert, responds to Voice, responds to Pain, Unresponsive) scale, which is used in many early-warning scores in the general-ward setting. The use of the Glasgow Coma Scale (GCS) and Richmond Agitation Sedation Scale (RASS) is not well established in this population.

**OBJECTIVE:** To compare the accuracies of AVPU, GCS, and RASS for predicting inpatient mortality.

**DESIGN:** Retrospective cohort study.

**SETTING:** Single, urban, academic medical center.

**PARTICIPANTS:** Adult inpatients on the general wards.

**MEASUREMENTS:** Nurses recorded GCS and RASS on consecutive adult hospitalizations. AVPU was extracted from the eye subscale of the GCS. We compared the accuracies of each scale for predicting in-hospital mortality

within 24 hours of a mental-status observation using area under the receiver operating characteristic curves (AUC).

**RESULTS:** There were 295,974 paired observations of GCS and RASS obtained from 26,873 admissions; 417 (1.6%) resulted in in-hospital death. GCS and RASS more accurately predicted mortality than AVPU (AUC 0.80 and 0.82, respectively, vs 0.73;  $P < 0.001$  for both comparisons). Simultaneous use of GCS and RASS produced an AUC of 0.85 (95% confidence interval: 0.82-0.87,  $P < 0.001$  when compared to all 3 scales).

**CONCLUSIONS:** In ward patients, both GCS and RASS were significantly more accurate predictors of mortality than AVPU. In addition, combining GCS and RASS was more accurate than any scale alone. Routine tracking of GCS and/or RASS on general wards may improve the accuracy of detecting clinical deterioration. *Journal of Hospital Medicine* 2015;10:658-663. © 2015 Society of Hospital Medicine

Altered mental status (AMS), characterized by abnormal changes in a patient's arousal and/or cognition, is a significant predictor of hospital mortality.<sup>1-3</sup> Yet despite its prevalence<sup>3-5</sup> and importance, up to three-quarters of AMS events go unrecognized by caregivers.<sup>6-8</sup> Acute changes in mental status, often caused by delirium in the hospitalized patient,<sup>3</sup> can present nonspecifically, making it difficult to detect and distinguish from other diagnoses such as depression or dementia.<sup>7,9</sup> Further complicating the recognition of AMS, numerous and imprecise qualitative descriptors such as "confused" and "alert and oriented" are used in clinical practice to describe the mental status of patients.<sup>10</sup> Thus, more objective measures may result in improved detection of altered mental status and in earlier diagnostic and therapeutic interventions.

In critically ill patients, several scales have been widely adopted for quantifying mental status. The Richmond Agitation and Sedation Scale (RASS) was created to optimize sedation.<sup>11</sup> The Glasgow Coma Scale (GCS) was developed for head-trauma patients<sup>12</sup> and is now a standardized assessment tool in intensive care units,<sup>13</sup> the emergency department,<sup>14</sup> and the pre-hospital setting.<sup>15</sup> In addition, a simplified scale, AVPU (Alert, responsive to Verbal stimuli, responsive to Painful stimuli, and Unresponsive) was initially used in the primary survey of trauma patients<sup>16</sup> but is now a common component of early-warning scores and rapid response activation criteria, such as the Modified Early Warning Score (MEWS).<sup>17,18</sup> In fact, in a systematic review of 72 distinct early-warning scores, 89% of the scores used AVPU as the measure of mentation.<sup>17</sup> However, the utility of these 3 scales is not well established in the general-ward setting. Our aim was therefore to compare the accuracies of AVPU, GCS, and RASS for predicting mortality in hospitalized general-ward patients to provide insight into the accuracy of these different scores for clinical deterioration.

\*Address for correspondence and reprint requests: Dana P. Edelson, MD, Section of Hospital Medicine, University of Chicago Medical Center, 5841 S Maryland Avenue, MC 5000, Chicago, IL 60637; Telephone: 773-834-2191; Fax: 773-795-7398; E-mail: dperes@uchicago.edu

Additional Supporting Information may be found in the online version of this article.

Received: March 5, 2015; Revised: May 26, 2015; Accepted: June 9, 2015

2015 Society of Hospital Medicine DOI 10.1002/jhm.2415

Published online in Wiley Online Library (Wileyonlinelibrary.com).

### METHODS

#### Study Setting and Protocol

We conducted an observational cohort study of consecutive adult general-ward admissions from July

2011 through January 2013 at a 500-bed, urban US teaching hospital. During the study period, no early-warning scoring systems were in place on the hospital wards. Rapid response teams responding to altered mental status would do so without specific thresholds for activation. During this period, nurses on the general floors were expected to record each patient's GCS and RASS score in the electronic health record (EPIC Systems Corp., Verona, WI) as part of the routine patient assessment at least once every 12-hour shift. AVPU assessments were extracted from the eye component of the GCS. The letter "A" was assigned to a GCS Eye score of 4 (opens eyes spontaneously), "V" to a score of 3 (opens eyes in response to voice), "P" to a score of 2 (opens eyes in response to painful stimuli), and "U" to a score of 1 (does not open eyes). To avoid comparison of mental-status scores at different time points, only concurrent GCS and RASS scores, documented within 10 minutes of one another, were included in the analysis.

Location and time-stamped GCS and RASS scores, demographics, and in-hospital mortality data were obtained from the hospital's Clinical Research Data Warehouse, which is maintained by the Center for Research Informatics at The University of Chicago. The study protocol and data-collection mechanisms were approved by The University of Chicago Institutional Review Board (#16995A).

### Statistical Analysis

Baseline admission characteristics were described using proportions (%) and measures of central tendency (mean, standard deviations [SD]; median, interquartile ranges [IQR]). Patient severity of illness at first ward observation was calculated using the MEWS.<sup>19</sup> All mental-status observations during a patient's ward stay were included in the analysis. Odds ratios for 24-hour mortality following an abnormal mental-status score were calculated using generalized estimating equations, with an exchangeable correlation structure to account for the correlation of scores within the same patient, as more than 1 abnormal mental-status score may have been documented within the 24 hours preceding death. Spearman's rank correlation coefficients ( $\rho$ ) were used to estimate the correlation among AVPU, GCS, and RASS scores.

The predictive accuracies of AVPU, GCS, RASS, and the subscales of GCS were compared using the area under the receiver operating characteristic curve (AUC), with mortality within 24 hours of a mental-status observation as the primary outcome and the mental-status score as the predictor variable. Although AUCs are typically used as a measure of discriminative ability, this study used AUCs to summarize both sensitivity and specificity across a range of cutoffs, providing an overall measure of predictive accuracies across mental-status scales. To estimate AUCs, the AVPU, GCS, and GCS subscales were entered into a

**TABLE 1.** Baseline Characteristics of Hospital Admissions

Total no. of admissions	26,873
Total no. of unique patients	17,660
Age, y, mean (SD)	57 (17)
Female sex, n (%)	14,293 (53)
Race, n (%)	
White	10,516 (39)
Black	12,580 (47)
Other/unknown	3,777 (14)
Admission MEWS, median (IQR)	1 (1–1)
Days on ward, median (IQR)	5 (3–10)
Observations per person, per day, median (IQR)	3 (2–4)
Underwent surgery during hospitalization, n (%)	6,141 (23)
Deaths, n (%)	417 (1.6)

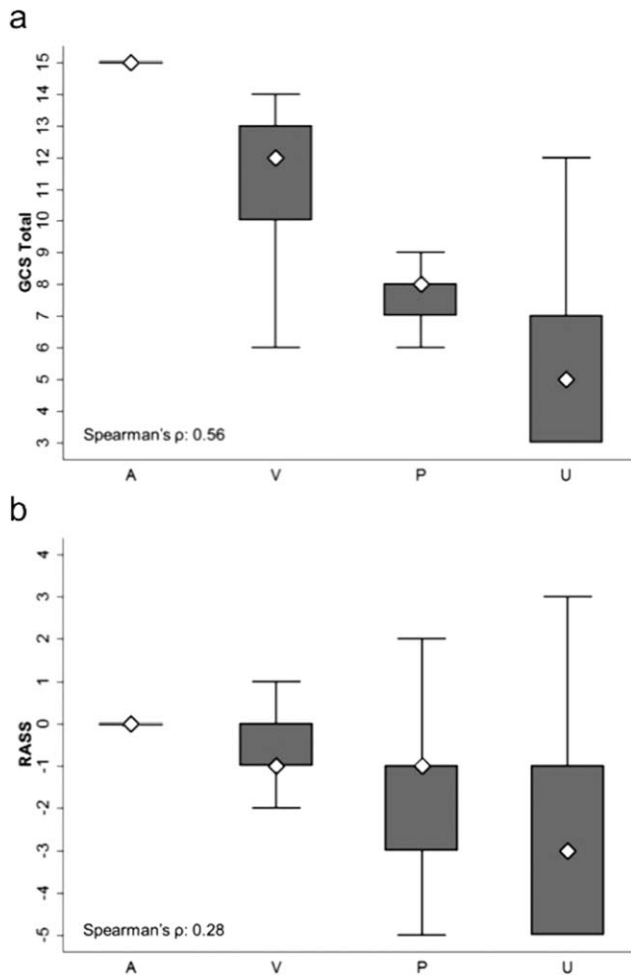
NOTE: Characteristics are stratified at the hospital admission level. Abbreviations: IQR, interquartile range; MEWS, Modified Early Warning Score; n, number of observations; SD, standard deviation.

logistic regression model as ordinal variables, whereas RASS was entered as a nominal variable due to its positive and negative components, and predicted probabilities were calculated. In addition, a combined model was fit where GCS and RASS were classified as categorical independent variables. AUCs were then calculated by utilizing the predicted probabilities from each logistic regression model using the trapezoidal rule.<sup>20</sup> A sensitivity analysis was performed to estimate the internal validity of the RASS model using 10-fold cross-validation.

Predefined subgroup analyses were performed that compared the accuracies of AVPU, GCS, and RASS for predicting 24-hour mortality in patients above and below the median age of the study population, and between patients who underwent surgery during their admission or not (surgical vs medical). All tests of significance used a 2-sided *P* value <0.05. All data analysis was performed using Stata version 13.0 (StataCorp, College Station, TX).

### RESULTS

During the study period, 313,577 complete GCS and 305,177 RASS scores were recorded in the electronic health record by nursing staff. A total of 26,806 (17,603 GCS and 9203 RASS) observations were excluded due to nonsimultaneous measurement of the other score, resulting in 295,974 paired mental-status observations. These observations were obtained from 26,873 admissions in 17,660 unique patients, with a median MEWS at ward admission of 1 (IQR 1–1). The mean patient age was 57 years (SD 17), and 23% were surgical patients (Table 1). Patients spent a median 63.9 hours (IQR 26.7–118.6) on the wards per admission and contributed a median of 3 paired observations (IQR 2–4) per day, with 91% of patients having at least 2 observations per day. A total of 417 (1.6%) general-ward admissions resulted in death during the hospitalization, with 354 mental-status observations occurring within 24 hours of a death. In



**FIG. 1.** Score correlations between (1a) AVPU and GCS total, and between (1b) AVPU and RASS. Boxes indicate interquartile range (25th to 75th percentiles), whiskers indicate 5th to 95th percentiles, and diamonds indicate median. Each correlation is significant at  $P < 0.001$ . Abbreviations: AVPU, Alert-Voice-Pain-Unresponsive; GCS, Glasgow Coma Scale; RASS, Richmond Agitation Sedation Scale.

addition, 26,618 (99.9%) admissions had at least 1 paired mental-status observation within the last 24 hours of their ward stay.

AVPU was moderately correlated with GCS (Spearman's  $\rho = 0.56$ ) (Figure 1a) and weakly correlated with RASS (Spearman's  $\rho = 0.28$ ) (Figure 1b). GCS scores were also weakly correlated to RASS (Spearman's  $\rho = 0.13$ ,  $P < 0.001$ ). Notably, AVPU mapped to distinct levels of GCS, with "Alert" associated with a median GCS total score of 15, "Voice" a score of 12, "Pain" a score of 8, and "Unresponsive" a score of 5. Abnormal mental-status scores on any scale were associated with significantly higher odds of death within 24 hours than normal mental-status scores (Table 2). This association was consistent within the 3 subscales of GCS and for scores in both the sedation ( $< 0$ ) and agitation ( $> 0$ ) ranges of RASS.

AVPU was the least accurate predictor of mortality (AUC 0.73 [95% confidence interval {CI}: 0.71–0.76]), whereas simultaneous use of GCS and RASS was the most accurate predictor (AUC 0.85 [95% CI:

**TABLE 2.** Odds of Mortality Within 24 Hours of an Abnormal Mental-Status Score

Mental-status Score	Observations, n (%)	Odds Ratio for Mortality (95% CI)
<b>GCS Eye (AVPU)</b>		
4 (alert)	289,857 (98)	Reference
<4 (not alert)	6,117 (2)	33.8 (23.9–47.9)
<b>GCS Verbal</b>		
5	277,862 (94)	Reference
4	11,258 (4)	4.7 (2.8–7.9)
<4	6,854 (2)	52.7 (38.0–73.2)
<b>GCS Motor</b>		
6	287,441 (97)	Reference
<6	8,533 (3)	41.8 (30.7–56.9)
<b>GCS total</b>		
15	276,042 (93)	Reference
13, 14	12,437 (4)	5.2 (3.3–8.3)
<13	7,495 (3)	55.5 (40.0–77.1)
<b>RASS</b>		
>0	6,867 (2)	8.5 (5.6–13.0)
0	275,708 (93)	Reference
<0	13,339 (5)	25.8 (19.2–34.6)

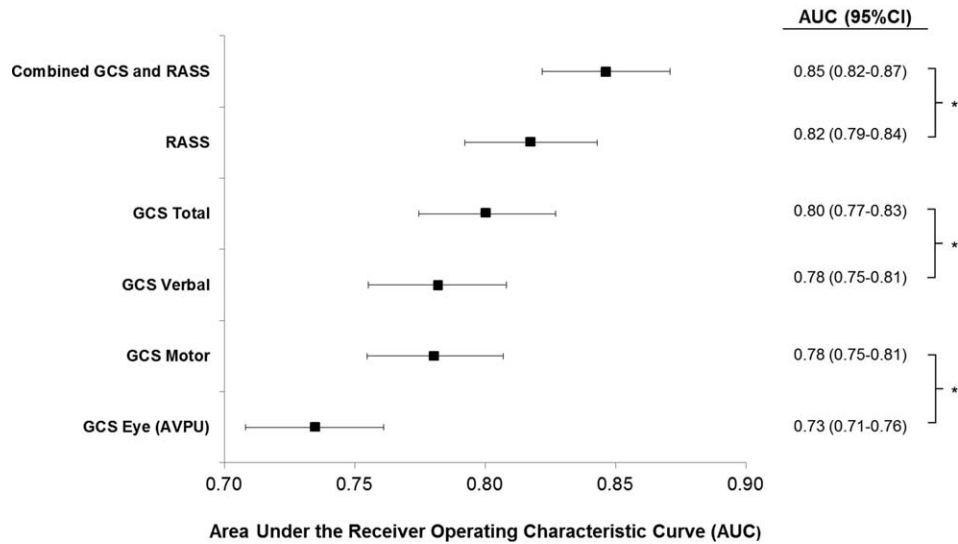
NOTE: Odds ratios, with 95% CIs, comparing the probability of mortality within 24 hours of an abnormal mental-status score to the probability of mortality within 24 hours of a normal mental-status score (Reference). All calculations control for clustering of observations within the same admission. All odds ratios were significant at  $P < 0.001$ . Abbreviations: AVPU, Alert-Voice-Pain-Unresponsive; CI, confidence interval; GCS, Glasgow Coma Scale; n, number of observations; RASS, Richmond Agitation Sedation Scale.

0.82–0.87] (Figure 2). The accuracies of GCS and RASS were not significantly different from one another in the total study population (AUC 0.80 [95% CI: 0.77–0.83] and 0.82 [0.79–0.84], respectively,  $P = 0.13$ ). Ten-fold cross-validation to estimate the internal validity of the RASS model resulted in a lower AUC (0.78 [95% CI: 0.75–0.81]) for RASS as a predictor of 24-hour mortality. Subgroup analysis indicated that RASS was more accurate than GCS in younger patients ( $< 57$  years old) and in surgical patients (Figure 3).

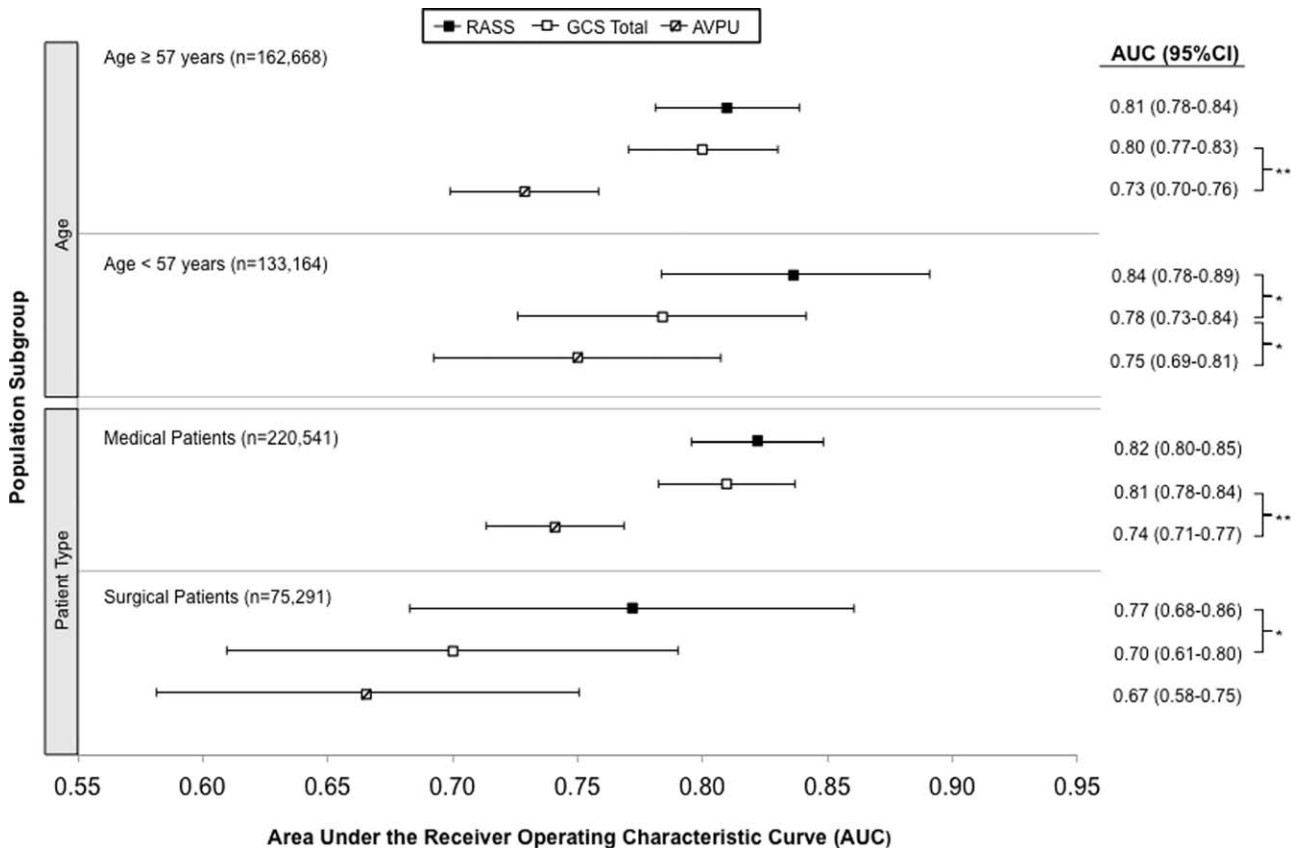
Removal of the 255 admissions missing a paired mental-status observation within the last 24 hours of their ward stay resulted in no change in the AUC values. A sensitivity analysis for prediction of a combined secondary outcome of 24-hour intensive care unit ICU transfer or cardiac arrest yielded lower AUCs for each mental-status scale, with no change in the association among scales.

## DISCUSSION

To our knowledge, this study is the first to compare the accuracies of AVPU, GCS, and RASS for predicting mortality in the general-ward setting. Similar to McNarry and Goldhill, we demonstrated that AVPU scores mapped to distinct levels of GCS. Although our study reports the same median GCS scores of 15 and 8 for AVPU levels of Alert and Pain, respectively, we indicate slightly lower corresponding median GCS scores for AVPU scores of Voice (12 vs 13) and Unresponsive (5 vs 6) than their previous work.<sup>21</sup> We found that AVPU was the least accurate predictor of mortality within 24 hours of an observation, and the



**FIG. 2.** Predictive accuracies of mental-status scales (and GCS subscales) for mortality within 24 hours of a mental-status observation (\* $P < 0.001$ ). AUC with whiskers indicating 95% confidence intervals for predicting mortality occurring within 24 hours of a mental-status observation. AUCs are shown for each mental-status scale, for the combination of GCS and RASS, and for the 3 subscales of the GCS. Abbreviations: 95% CI, 95% confidence interval; AUC, area under the receiver operating characteristic curve; AVPU, Alert-Voice-Pain-Unresponsive; GCS, Glasgow Coma Scale; RASS, Richmond Agitation Sedation Scale.



**FIG. 3.** Predictive accuracies of AVPU, GCS, and RASS for mortality within 24 hours of a mental-status observation. Subgroup analysis is based on age and surgical status (\* $P < 0.05$ , \*\* $P < 0.001$ ). AUC with whiskers indicating 95% CI for predicting mortality occurring within 24 hours of a mental-status observation, analyzed at the observation level, and stratified by patient age (below or greater than or equal to the median age of 57 years) and surgical status (patient with surgery during hospitalization or medical patient only). Abbreviations: 95% CI, 95% confidence interval; AUC, area under the receiver operating characteristic curve; AVPU, Alert-Voice-Pain-Unresponsive; GCS, Glasgow Coma Scale; RASS, Richmond Agitation Sedation Scale.

combination of GCS and RASS was the most accurate. RASS was at least as accurate a predictor for 24-hour mortality in comparison to GCS total in the

overall study population. However, the RASS score was the most accurate individual score in surgical and younger patients. These findings suggest that changing

from the commonly used AVPU scale to the RASS and/or GCS would improve the prognostic ability of mental-status assessments on the general wards.

Buist and colleagues have previously demonstrated altered mental status to be one of the strongest predictors of death on the wards. In that study, a GCS score of 3 and a decrease in GCS score by more than 2 points were independently associated with mortality (odds ratio 6.1 [95% CI: 3.1–11.8] and 5.5 [95% CI: 2.6–11.9], respectively).<sup>22</sup> We have also previously shown that after adjusting for vital signs, being unresponsive to pain was associated with a 4.5-fold increase in the odds of death within 24 hours,<sup>23</sup> whereas Subbe and colleagues showed a relative risk ratio of 5.2 (95% CI: 1.5–18.1) for the combined endpoint of cardiac arrest, death at 60 days, or admission to the intensive care/high dependency unit.<sup>19</sup> In the current study, the magnitude of these associations was even stronger, with a GCS score <13 correlating with a 55-fold increase in the odds of death, compared to a normal GCS, and not being alert being associated with a 33.8-fold increase in the odds of death. This difference in magnitude is likely a product of the univariate nature of the current analysis, compared to both the Buist et al. and Churpek et al. studies, which adjusted for vital signs, thereby lessening the impact of any single predictor. Because this study was designed to compare mental-status variables to one another for future model inclusion, and all the analyses were paired, confounding by additional predictors of death was not a concern.

One of the potential strengths of RASS over GCS and AVPU is its ability to measure agitation levels, in addition to depressed mentation, a feature that has been shown to be present in up to 60% of delirium episodes.<sup>24</sup> This may also explain why RASS was the most accurate predictor of mortality in our subset of younger patients and surgical patients, because hyperactive delirium is more common in younger and healthier patients, which surgical patients tend to be as compared to medical patients.<sup>25,26</sup> In this study, we found negative RASS scores portending a worse prognosis than positive ones, which supports previous findings that hypoactive delirium had a higher association with mortality than hyperactive delirium at 6 months (hazard ratio 1.90 vs 1.37) and at 1 year (hazard ratio 1.60 vs 1.30) in elderly patients at post-acute-care facilities in 2 separate studies.<sup>27,28</sup> However, a study of patients undergoing surgery for hip fracture found that patients with hyperactive delirium were more likely to die or be placed in a nursing home at 1 month follow-up when compared to patients with purely hypoactive delirium (79% vs 32%,  $P = 0.003$ ).<sup>29</sup>

We found the assessment of RASS and GCS by ward nurses to be highly feasible. During the study period, nurses assessed mental status with the GCS and RASS scales at least once per 12-hour shift in 91% of patients. GCS has been shown to be reliably

and accurately recorded by experienced nurses (reliability coefficient = 0.944 with 96.4% agreement with expert ratings).<sup>30</sup> RASS can take <30 seconds to administer, and in previous studies of the ICU setting has been shown to have over 94% nurse compliance for administration,<sup>31</sup> and good inter-rater reliability (weighted kappa 0.66 and 0.89, respectively).<sup>31,32</sup> Further, in a prior survey of 55 critical care nurses, 82% agreed that RASS was easy to score and clinically relevant.<sup>31</sup>

This study has several limitations. First, it was conducted in a single academic institution, which may limit generalizability to other hospitals. Second, baseline cognition and comorbidities were not available in the dataset, so we were unable to conduct additional subgroup analyses by these categories. However, we used age and hospital admission type as proxies. Third, the AVPU scores in this study were extracted from the Eye subset of the GCS scale, as AVPU was not directly assessed on our wards during the study period. Clinical assessment of mental status on the AVPU scale notes the presence of any active patient response (eg, eye opening, grunting, moaning, movement) to increasingly noxious stimuli. As such, our adaptation of AVPU using only eye-opening criteria may underestimate the true number of patients correctly classified as alert, or responding to vocal/painful stimuli. However, a sensitivity analysis comparing directly assessed AVPU during a 3-year period prior to the study implementation at our institution, and AVPU derived from the GCS Eye subscale for the study period, indicated no difference in predictive value for 24-hour mortality. Fourth, we did not perform trend analyses for change from baseline mental status or evolution of AMS, which may more accurately predict 24-hour mortality than discrete mental-status observations. Finally, the 3 scales we compared differ in length, which may bias the AUC against AVPU, a 4-point scale with a trapezoidal ROC curve compared to the smoother curve generated by the 15-point GCS scale, for example. However, the lack of discrimination of the AVPU is the likely source of its lesser accuracy.

## CONCLUSION

In the general-ward setting, routine collection of GCS and RASS is feasible, and both are significantly more accurate for predicting mortality than the more commonly used AVPU scale. In addition, the combination of GCS and RASS has greater accuracy than any of the 3 individual scales. RASS may be particularly beneficial in the assessment of younger and/or surgical patients. Routine documentation and tracking of GCS and/or RASS by nurses may improve the detection of clinical deterioration in general-ward patients. In addition, future early-warning scores may benefit from the inclusion of GCS and/or RASS in lieu of AVPU.

Disclosures: Drs. Churpek and Edelson have a patent pending (ARCD. P0535US.P2) for risk stratification algorithms for hospitalized patients. Dr. Churpek is supported by a career development award from the National Heart, Lung, and Blood Institute (K08 HL121080). Dr. Edelson has received research support from the National Heart, Lung, and Blood Institute (K23 HL097157), Philips (Andover, MA), the American Heart Association (Dallas, TX), Laerdal Medical (Stavanger, Norway), and Early Sense (Tel Aviv, Israel). She has ownership interest in Quant HC (Chicago, IL), which is developing products for risk stratification of hospitalized patients. All other authors report no conflicts of interest.

## References

- Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004;291(14):1753–1762.
- Pompei P, Foreman M, Ruddberg MA, Inouye SK, Braund V, Cassel CK. Delirium in hospitalized older persons: outcomes and predictors. *J Am Geriatr Soc*. 1994;42(8):809–815.
- Siddiqi N, House AO, Holmes JD. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. *Age Ageing*. 2006;35(4):350–364.
- Levkoff SE, Evans DA, Liptzin B, et al. Delirium. The occurrence and persistence of symptoms among elderly hospitalized patients. *Arch Intern Med*. 1992;152(2):334–340.
- Dyer CB, Ashton CM, Teasdale TA. Postoperative delirium. A review of 80 primary data-collection studies. *Arch Intern Med*. 1995;155(5):461–465.
- Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM Jr. Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. *Arch Intern Med*. 2001;161(20):2467–2473.
- Armstrong SC, Cozza KL, Watanabe KS. The misdiagnosis of delirium. *Psychosomatics*. 1997;38(5):433–439.
- Ely EW, Stephens RK, Jackson JC, et al. Current opinions regarding the importance, diagnosis, and management of delirium in the intensive care unit: a survey of 912 healthcare professionals. *Crit Care Med*. 2004;32(1):106–112.
- Farrell KR, Ganzini L. Misdiagnosing delirium as depression in medically ill elderly patients. *Arch Intern Med*. 1995;155(22):2459–2464.
- Simpson CJ. Doctors and nurses use of the word confused. *Br J Psychiatry*. 1984;145:441–443.
- 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(10):1338–1344.
- Teasdale G, Jennett B. Assessment and prognosis of coma after head injury. *Acta Neurochir (Wien)*. 1976;34(1–4):45–55.
- Bastos PG, Sun X, Wagner DP, Wu AW, Knaus WA. 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(10):1459–1465.
- Holdgate A, Ching N, Angonese L. Variability in agreement between physicians and nurses when measuring the Glasgow Coma Scale in the emergency department limits its clinical usefulness. *Emerg Med Australas*. 2006;18(4):379–384.
- Menegazzi JJ, Davis EA, Sucov AN, Paris PM. Reliability of the Glasgow Coma Scale when used by emergency physicians and paramedics. *J Trauma*. 1993;34(1):46–48.
- Alexander RH, Proctor HJ; American College of Surgeons. Committee on Trauma. Advanced Trauma Life Support Program For Physicians: ATLS. 5th ed. Chicago, IL: American College of Surgeons; 1993.
- Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation*. 2008;77(2):170–179.
- Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. *Resuscitation*. 2008;79(1):11–21.
- Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning score in medical admissions. *QJM*. 2001;94(10):521–526.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845.
- McNarry AF, Goldhill DR. Simple bedside assessment of level of consciousness: comparison of two simple assessment scales with the Glasgow Coma Scale. *Anaesthesia*. 2004;59(1):34–37.
- Buist M, Bernard S, Nguyen TV, Moore G, Anderson J. Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. *Resuscitation*. 2004;62(2):137–141.
- Churpek MM, Yuen TC, Edelson DP. Predicting clinical deterioration in the hospital: the impact of outcome selection. *Resuscitation*. 2013;84(5):564–568.
- Peterson JF, Pun BT, Dittus RS, et al. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc*. 2006;54(3):479–484.
- Angles EM, Robinson TN, Biffi WL, et al. Risk factors for delirium after major trauma. *Am J Surg*. 2008;196(6):864–869.
- Meagher DJ, O'Hanlon D, O'Mahony E, Casey PR, Trzepacz PT. Relationship between symptoms and motoric subtype of delirium. *J Neuropsychiatry Clin Neurosci*. 2000;12(1):51–56.
- Yang FM, Marcantonio ER, Inouye SK, et al. Phenomenological subtypes of delirium in older persons: patterns, prevalence, and prognosis. *Psychosomatics*. 2009;50(3):248–254.
- Kiely DK, Jones RN, Bergmann MA, Marcantonio ER. Association between psychomotor activity delirium subtypes and mortality among newly admitted post-acute facility patients. *J Gerontol A Biol Sci Med Sci*. 2007;62(2):174–179.
- Marcantonio E, Ta T, Duthie E, Resnick NM. Delirium severity and psychomotor types: their relationship with outcomes after hip fracture repair. *J Am Geriatr Soc*. 2002;50(5):850–857.
- Rowley G, Fielding K. Reliability and accuracy of the Glasgow Coma Scale with experienced and inexperienced users. *Lancet*. 1991;337(8740):535–538.
- 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(6):1199–1205.
- 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.