



# HHS Public Access

Author manuscript

*J Crit Care.* Author manuscript; available in PMC 2017 August 01.

Published in final edited form as:

*J Crit Care.* 2016 August ; 34: 12–16. doi:10.1016/j.jcrc.2016.03.008.

## Identification of Acute Brain Failure using electronic medical records

**Dereddi Raja Shekar Reddy, MD<sup>1,2</sup>, Tarun D Singh, MBBS<sup>2,4</sup>, Pramod K Guru, MD<sup>1,2</sup>, Amra Sakusic, MD<sup>1</sup>, Ognjen Gajic, MD, MSc<sup>1,2</sup>, John C O'Horo, MD, MPH<sup>1,2,3</sup>, and Alejandro A Rabinstein, MD<sup>2,4</sup>**

Dereddi Raja Shekar Reddy: drraja.reddy2005@gmail.com; Tarun D Singh: tarundsingh6@gmail.com; Pramod K Guru: guru.pramod@mayo.edu; Amra Sakusic: sakusic.amra@mayo.edu; Ognjen Gajic: gajic.ognjen@mayo.edu; Alejandro A Rabinstein: rabinstein.alejandro@mayo.edu

<sup>1</sup>Department of Medicine, Division of Pulmonary and critical Care Medicine, Mayo Clinic, Rochester, MN

<sup>2</sup>Department of Medicine, Multidisciplinary Epidemiology and Translational Research in Intensive Care, Emergency and Perioperative Medicine (METRIC), Mayo Clinic, Rochester, MN

<sup>3</sup>Department of Medicine, Division of Infectious Diseases, Mayo Clinic, Rochester, MN

<sup>4</sup>Department of Neurology, Division of Critical Care Neurology, Mayo, Rochester, MN

### Abstract

**Purpose**—Up to 80 % of critically ill patients suffer from acute neurological dysfunction syndromes. We evaluated inter-rater reliability between the examination by the investigator and the charted assessment by the nurse, since the accuracy and reliability of detailed datasets extracted from the EMR represents a keystone for creating EMR based definitions.

**Materials and Methods**—We conducted a prospective observational study of ICU patients to assess the reliability of charted Confusion Assessment Method for the ICU (CAM-ICU), Glasgow Coma Score(GSC), Full Outline of Unresponsiveness(FOUR) and Richmond Agitation Sedation Scale (RASS) scores, and a composite measure of ABF defined as new onset coma or delirium. Trained investigator blinded to nursing assessments performed the neurological evaluations that were compared with nursing documentation.

**Results**—202 observations were performed in 55 ICU patients. Excellent correlation was noted for GCS and FOUR scores on Bland-Altman plots (Pearson correlation 0.87 and 0.92 respectively). Correlation for CAM-ICU was also high ( $k=0.86$ , 95% confidence interval [CI] 0.70-1.01). RASS had good agreement when scores were dichotomized as over-sedated ( $<-2$ ) vs. not over-sedated, with  $k=0.76$  (95% CI 0.54-0.98). Investigator assessment and nurse charting were highly concordant ( $k=0.84$ , 95% CI 0.71-0.99).

---

Correspondence to: John C. O'Horo, MD, MPH, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, Phone: (507)284-3309, Fax: (507) 255-7767, OHoro.John@mayo.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conclusion**—Neurological assessments documented on the EMR are reliable.

## Keywords

delirium; acute brain failure; ICU; coma; electronic digital signature

---

## 1. Introduction

Delirium affects up to 80% of critically ill patients and negatively impacts prognosis.<sup>1</sup> It is associated with both poor short and long term outcomes, such as increased length of intensive care unit (ICU) stay, prolonged invasive mechanical ventilation, increased mortality and costs, and long term cognitive impairment.<sup>2-5</sup> The recently published Society of Critical Care Medicine Pain, Agitation and Delirium (PAD) guidelines recommend routine monitoring of delirium in adult ICU patients using validated bedside instruments.<sup>6</sup>

There are several validated tools to identify delirium, most notably the CAM-ICU.<sup>7</sup> Yet, evaluation of delirium requires assessment of thought content and therefore its recognition is confounded in patients with depressed level of consciousness and those who are deeply sedated. As a result, delirium is both over- and under diagnosed.<sup>8-11</sup> Reduced level of consciousness can be reliably defined using the Glasgow Coma Score (GCS) or the Full Outline of Unresponsiveness (FOUR score), both of which have been extensively validated in the ICU.<sup>12, 13</sup> The Richmond Agitation Sedation Score (RASS) can be effectively used to determine the level of sedation. Since brain dysfunction in patients with critical illness can manifest with alterations in the level and content of consciousness, delirium does not encompass the entire spectrum of cerebral disorders in these patients. An endpoint that includes delirium (i.e. alteration in the content of consciousness) and diminished level of consciousness (drowsiness, stupor, or coma) is necessary to capture the spectrum of acute brain failure (ABF)

The widespread adoption of electronic medical records (EMR) allows for novel research into common conditions using electronic search strategies.<sup>14, 15</sup> We hypothesized that ABF and its components could be reliably identified using EMR queries and “big data” research methods. Any sort of EMR query, however, is contingent upon the accuracy of the data being entered into the patient records. Therefore, it was necessary to first validate neurologic assessment documentation before making an electronic search algorithm.

This is a validation study in which we assessed the accuracy of nurse-determined neurological scoring (Confusion Assessment Method ICU (CAM-ICU), Glasgow Coma Score (GCS), Full Outline of Unresponsiveness (FOUR score) and Richmond Agitation Sedation Scale (RASS)).

## 2. Material and Methods

### 2.1. Subjects

Since some patients had altered mental status, we were waived from obtaining informed consent from subjects, but required to obtain a delayed consent from a Legally Authorized

Representative (LAR). Subjects were recruited from the Mayo Clinic's medical, surgical, cardiac and trauma ICUs.

## 2.2. Study Population

This study was approved as minimal risk research by the Mayo Clinic Institutional Review Board (IRB). Subjects were adult patients (Age>18 years) admitted to the Mayo Clinic medical, cardiac, or surgical ICUs between July and December 2014. Patients admitted with primary neurological disease (e.g. stroke, head trauma) were excluded.

## 2.3. Study Methods and Personnel

The research team consisted of two critical care fellows trained in neurologic assessments; (DRR and PKG). The team used standard printed reference cards for each neurologic assessment. While conducting this prospective study and randomizing patients, we preferentially examined patients with abnormal exam so that we can better assess the differences in the abnormal scores between the physician and nurses. Both researchers were assessed for competence by an expert neurocritical care physician (AAR) before the start of the study. An additional researcher (TS) was assigned to alert the clinical researchers when a subject was due for an exam to keep the examiners blinded to the charted values. Thus, one researcher would randomly identify patients with normal or abnormal neurologic scores across participating ICUs, and instruct another researcher which subjects needed to be examined. The researcher performing the examination was thus blinded to the nurse assessment.

The tools assessed in the present study included GCS, FOUR score, CAM-ICU and RASS. GCS and FOUR scores are measured every four hours on every patient. RASS is performed on initiation of sedation and at least hourly until the sedation goal is reached. CAM-ICU is evaluated at least twice daily, and frequency increased if a patient has an acute change in mental status. The increase can be to every 4, 2 or 1 hour. This is specifically ordered by the medical team on a case by case basis. All evaluations are performed with the assistance of a computerized scoring guide and charted directly into the EMR. Unit staff was not made aware that research staff would randomly perform prospective neurologic assessments following nursing assessments to minimize Hawthorne effect. The time delay in performing these assessments between nurses and research study fellows was usually around 30 minutes and never exceeded 60 minutes. The results of the prospective examinations were considered the gold standard and compared to the recorded nursing assessments on the EMR to determine the reliability of the latter.

## 2.4. Acute Brain Failure (ABF)

Our preliminary algorithm for identifying ABF had three components; a measurement for confounding sedation (RASS<-2), a level of consciousness component (GCS or FOUR score less than maximum achievable for that patient, accounting for intubated status) and a thought content component (CAM-ICU positive), as explained in figure 1. For this reliability study, we only evaluated inter-rater reliability between the examination by the investigator and the charted assessment by the nurse. The details are shown in figure 1.

## 2.5. Statistical Analysis

Bland-Altman plots were used to evaluate the presence of bias and inter-observer variability between nursing and researcher assessments on continuous and categorical data.<sup>16</sup> The Kappa coefficient was used to evaluate inter-observer variability for binary data. For binary data, values were dichotomized into “abnormal” vs “not abnormal.” The thresholds for “abnormal” were GCS < 15 for non-intubated patients, GCS<11 for intubated patients, FOUR score <16 for non-intubated patients, FOUR score <13 for intubated patients. The threshold for coma was GCS equal to or less than 8. Delirium was defined by a positive CAM-ICU. Deep sedation was defined as a RASS -3 or lower. Deeply sedated patients could not be further assessed. ABF was considered present when GCS or FOUR scores were abnormal or the CAM-ICU was positive. Analyses were performed using JMP 10 software (SAS Institute, Inc., Cary, NC).

## 3. Results

202 observations were performed in 55 patients. Patient characteristics are shown in table 1. No patients were lost due to withdrawal of consent. One patient was omitted from analysis because of a drastic change in clinical status between the time of nurse and researcher assessments which had required administration of large sedative doses. The details are presented in table 1 and figure 2.

### 3.1. GCS Score

GCS scores were obtained on all 55 patients and 30 of them had positive ones. GCS scores showed a positive Pearson correlation at 0.87, with a mean difference of 0.35 (95% CI -0.80-0.11) and no evidence of systematic bias ( $p=0.13$ ) (figure 3). When treated as dichotomous data, this also correlated well, with excellent agreement with Kappa coefficient 0.96 (0.89-1.03). the details are presented in figure 3.

### 3.2. FOUR Score

FOUR scores were obtained in 49 subjects while 25 of them had positive scores. FOUR scores were also highly correlated with a Pearson coefficient of 0.92. Bland-Altman plotting (figure 2) shows a mean difference of 0.21 (95% CI -0.55-0.14) with no systematic evidence of bias ( $p=0.10$ ). Dichotomized, agreement was quite good with Kappa coefficient of 0.95 (0.87-1.04). The details are shown in figure 4.

### 3.3. CAM-ICU

CAM-ICU scores were measured on 44 patients, while 18 of them had positive scores (40,9%). Scores were positively correlated, with only three disagreements and a Kappa coefficient of 0.86 (0.70-1.01).

### 3.4. RASS

RASS was scored on 55 subjects. Overall agreement was good, with Pearson correlation of 0.73. Minimal bias was seen in the Bland-Altman plot (figure 5), with a mean difference of 0.33 between researcher and nurse assessments ( $P=0.04$ ). However, agreement was good

when dichotomized as over-sedated ( $<-2$ ) vs. not over-sedated (all other scores), with Kappa 0.76 (0.54-0.98). The details are shown in the figure 5.

### 3.5. ABF

Among 55 patients 34 were diagnosed with ABF. Overall agreement on the presence or absence of the aggregate outcome of acute brain failure was excellent, with a kappa coefficient of 0.84 (0.71-0.99). Using investigator scores as the gold standard, EMR had 83.3% sensitivity and 100% specificity for the recognition of ABF.

## 4. Discussion

In this prospective observational study, there was good correlation between investigators neurologic exam and nurse charting on the EMR for GCS, FOUR score, RASS CAM-ICU and a composite outcome of ABF.

Despite extensive evidence showing the hazards and poor outcomes associated with delirium, there are several gaps and inconsistencies in our understanding of the problem. It's the true prevalence of acute brain dysfunction in critically ill patients is probably underestimated. Cheung *et al.* studied the semantic diagnostic classification of ICU delirium among Canadian intensivists and what they called cognitive and perceptive abnormalities in ICU patients. They found that there was a wide variability in the use of the term "delirium" to diagnose cognitive abnormalities in ICU patients.<sup>10, 17, 18</sup> In a commentary on the work by van Eijk and colleagues, Patel and Kress indicated that the components of CAM-ICU maybe improperly performed in non-research settings or insufficient to adequately identify a significant proportion of delirious patients. They also raised an interesting question whether there is a difference in both short and long term outcomes among patients with "gold standard" delirium versus delirium diagnosed by CAM-ICU.<sup>9</sup> Furthermore, delirium truly should reflect alterations in the content of consciousness, but may fail to identify patients in whom the main problem is depression of the level of consciousness and therefore cannot be reliably assessed solely using the CAM-ICU or similar delirium scores. To overcome some of these difficulties we focused on a composite outcome of acute brain failure (ABF) that will not only capture alterations in the content but also in the level of consciousness of ICU patients.

The results of our study demonstrate adequate correlation between prospective examinations by trained investigators and the nursing documentation on the EMR, This is the necessary first step to be confident that the data on the EMR can be used to construct a digital signature for ABF that will allow us to examine its incidence, its factors and its impact on short and long term clinical outcomes. To the best of our knowledge, this is the first study that has validated the reliability of routine neurological assessments documented in the EMR in a mixed ICU population.

Our study has several limitations. First, it is single-center design within a quaternary care setting and a sample of mixed medical-surgical population, factors that limit inferences we can make about the reliability of neurological scores in other subpopulations. Second, the performance of neurologic assessments between study fellows and nursing staff could not be

done simultaneously, which may have affected our results (though the expected impact of this limitation would have been exaggerating the differences between the two assessments and therefore it should not question the validity of our main conclusion). Additionally, this study was performed on a small and relatively ill-defined patient population. Results clearly show that neurological assessments documented on the EMR are reliable but it is however, dependent on the teaching and auditing of nurses to demonstrate quality of assessment. Moreover, we always have to account on stuff changing and other similar technical issues. Furthermore, we have some missing data on the patients with regards to some tests. There was a minimal time gap between the assessments done by the nurses and the physicians. As we mentioned in the methods before, that the bedside staff was not aware of the research study being conducted, some of these patients were not available for assessments as they were taken to the OR, for performing some tests or procedures or were sedated and their RASS was too low to be tested. In addition, there was a delay to the first assessment in some of the patients. Since most of the patients were admitted to the ICUs, they were suffering from critical conditions and we didn't wanted to interrupt their active care. We waited for the patients to stabilize so that they could answer questions related to CAM-ICU and other assessments, but they were being continuously monitored by their primary team.

A final limitation is that we did not demonstrate the superiority of ABF over existing concepts of delirium and coma. However, the combination of the two into a single metric has some “face validity,” and given the reliability of assessments observed here, a larger validation study is merited to determine the link between ABF and patient outcomes.

## Conclusion

Our study shows that neurological assessments documented on the EMR are reliable, and accurately and reproducibly defines acute brain failure in critically ill patients. With the widespread EMR implementation a pragmatic EMR-based definition of ABF will facilitate large scale quality improvement and outcome research efforts. The present definition is reliable and logical; however, further validation is needed to link it to clinical outcomes.

## Acknowledgments

CTSA Grant Number UL1 TR000135 made this publication possible from the National Center for Advancing Translational Sciences (NCATS), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

This publication was made possible, in part by funding from the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, and the Mayo Critical Care Research Committee.

## References

1. Salluh JIWH, Schneider EB, Nagaraja N, Yenokyan G, Damluji A, Serafim RB, Stevens RD. Outcome of delirium in critically ill patients: systematic review and meta-analysis. *BMJ (Clinical research ed)*. 2015; 350:2538.
2. Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *Jama*. 2004; 291:1753–1762. [PubMed: 15082703]
3. Ely EW, Gautam S, Margolin R, et al. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med*. 2001; 27:1892–1900. [PubMed: 11797025]

4. Girard TD, Jackson JC, Pandharipande PP, et al. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Crit Care Med*. 2010; 38:1513–1520. [PubMed: 20473145]
5. Hopkins RO, Jackson JC. Long-term neurocognitive function after critical illness. *Chest*. 2006; 130:869–878. [PubMed: 16963688]
6. 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: executive summary. *Am J Health Syst Pharm*. 2013; 70:53–58. [PubMed: 23261901]
7. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med*. 2001; 29:1370–1379. [PubMed: 11445689]
8. Pun BT, Devlin JW. Delirium monitoring in the ICU: strategies for initiating and sustaining screening efforts. *Semin Respir Crit Care Med*. 2013; 34:179–188. [PubMed: 23716309]
9. Patel SB, Kress JP. Accurate identification of delirium in the ICU: problems with translating the evidence in the real-life setting. *Am J Respir Crit Care Med*. 2011; 184:287–288. [PubMed: 21804118]
10. Spronk PE, Riekerk B, Hofhuis J, et al. Occurrence of delirium is severely underestimated in the ICU during daily care. *Intensive Care Med*. 2009; 35:1276–1280. [PubMed: 19350214]
11. Poston JTG, Pohlman AS, et al. Assessment of delirium relative to daily sedative interruption. *Am J Respir Crit Care Med*. 2010
12. Wijdicks EF, Bamlet WR, Maramattom BV, et al. Validation of a new coma scale: The FOUR score. *Ann Neurol*. 2005; 58:585–593. [PubMed: 16178024]
13. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale *Lancet*. 1974; 2:81–84. [PubMed: 4136544]
14. Herasevich V, Yilmaz M, Khan H, et al. Validation of an electronic surveillance system for acute lung injury. *Intensive Care Med*. 2009; 35:1018–1023. [PubMed: 19280175]
15. Herasevich V, Pickering BW, Dong Y, et al. Informatics infrastructure for syndrome surveillance, decision support, reporting, and modeling of critical illness. *Mayo Clin Proc*. 2010; 85:247–254. [PubMed: 20194152]
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1:307–310. [PubMed: 2868172]
17. Vasilevskis EE, Ely EW, Speroff T, et al. Reducing iatrogenic risks: ICU-acquired delirium and weakness--crossing the quality chasm. *Chest*. 2010; 138:1224–1233. [PubMed: 21051398]
18. Cheung CZ, Alibhai SM, Robinson M, et al. Recognition and labeling of delirium symptoms by intensivists: does it matter? *Intensive Care Med*. 2008; 34:437–446. [PubMed: 18087692]

## List of Abbreviations Used

<b>ABF</b>	Acute brain failure
<b>EMR</b>	Electronic medical records
<b>ICU</b>	Intensive care unit
<b>CAM-ICU</b>	Confusion Assessment Method for the ICU
<b>GSC</b>	Glasgow Coma Score
<b>FOUR score</b>	Full Outline of Unresponsiveness
<b>RASS scores</b>	Richmond Agitation Sedation Scale
<b>PAD</b>	Society of Critical Care Medicine Pain, Agitation and Delirium (PAD)

**LAR**

Legally Authorized Representative

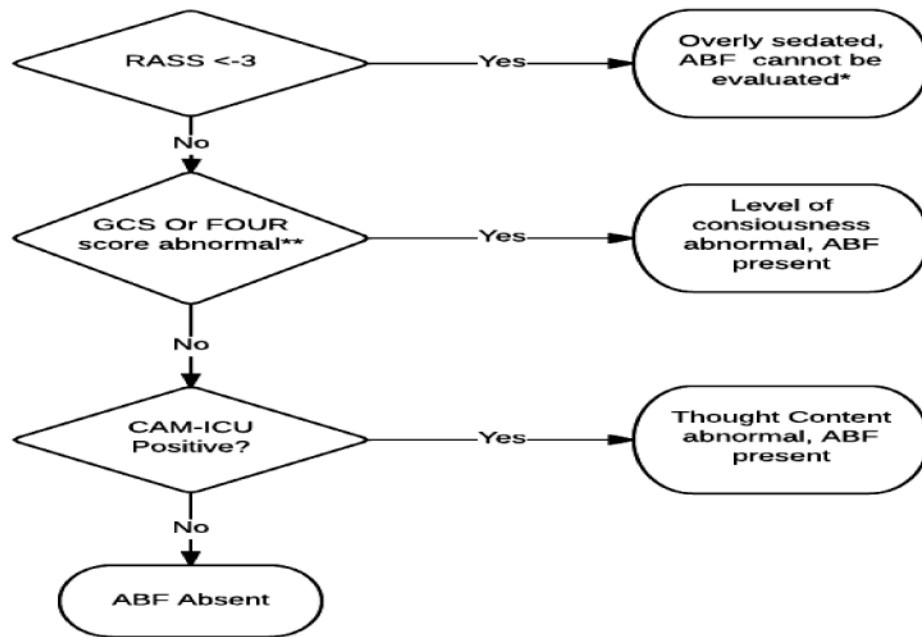
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript





**Figure 1. ABF Identification algorithm**

\*When data is insufficient to categorize it as ABF, it is scored as “not present.” \*\*Abnormal GCS or FOUR score depends on patient status. For intubated patients who cannot be assessed for verbal components, <13 FOUR score or <12 GCS is abnormal. For all others, GCS<15 and FOUR score <16 is scored as abnormal.

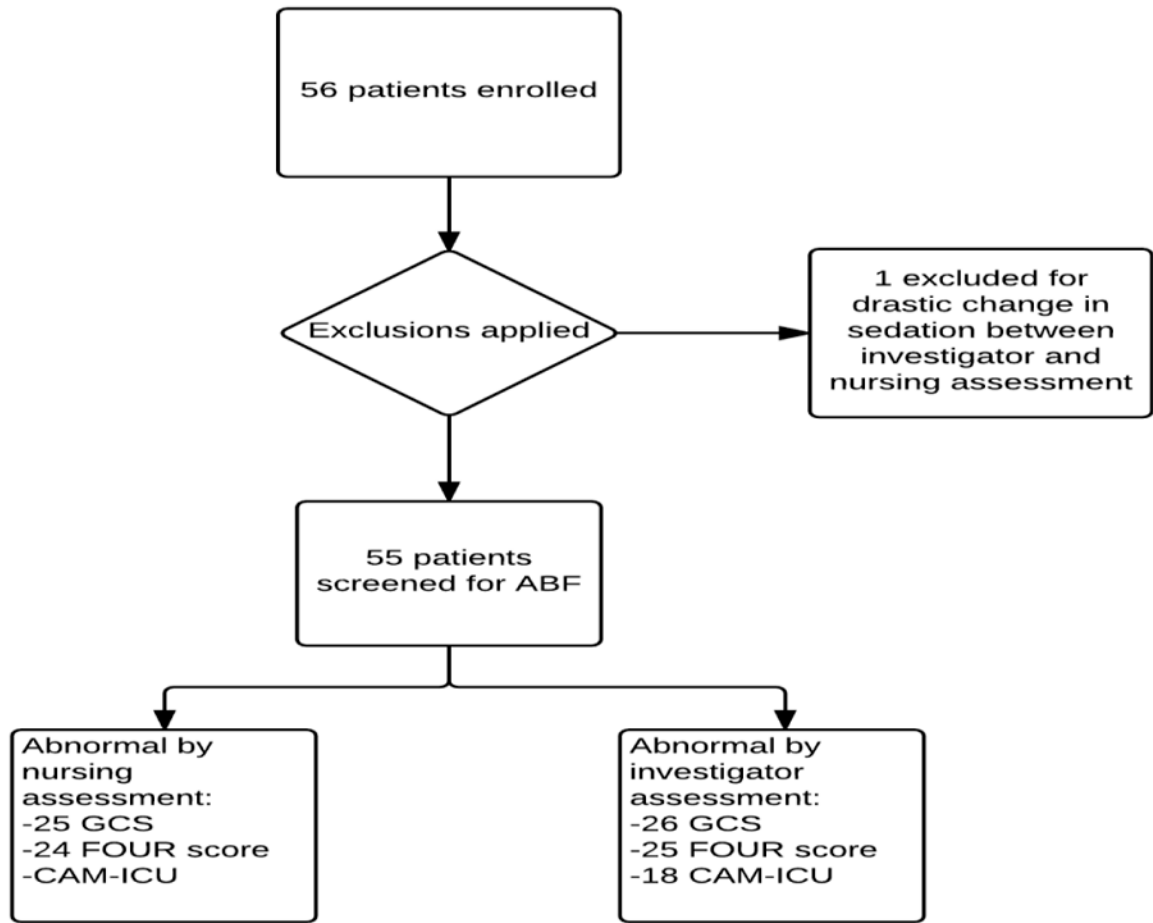
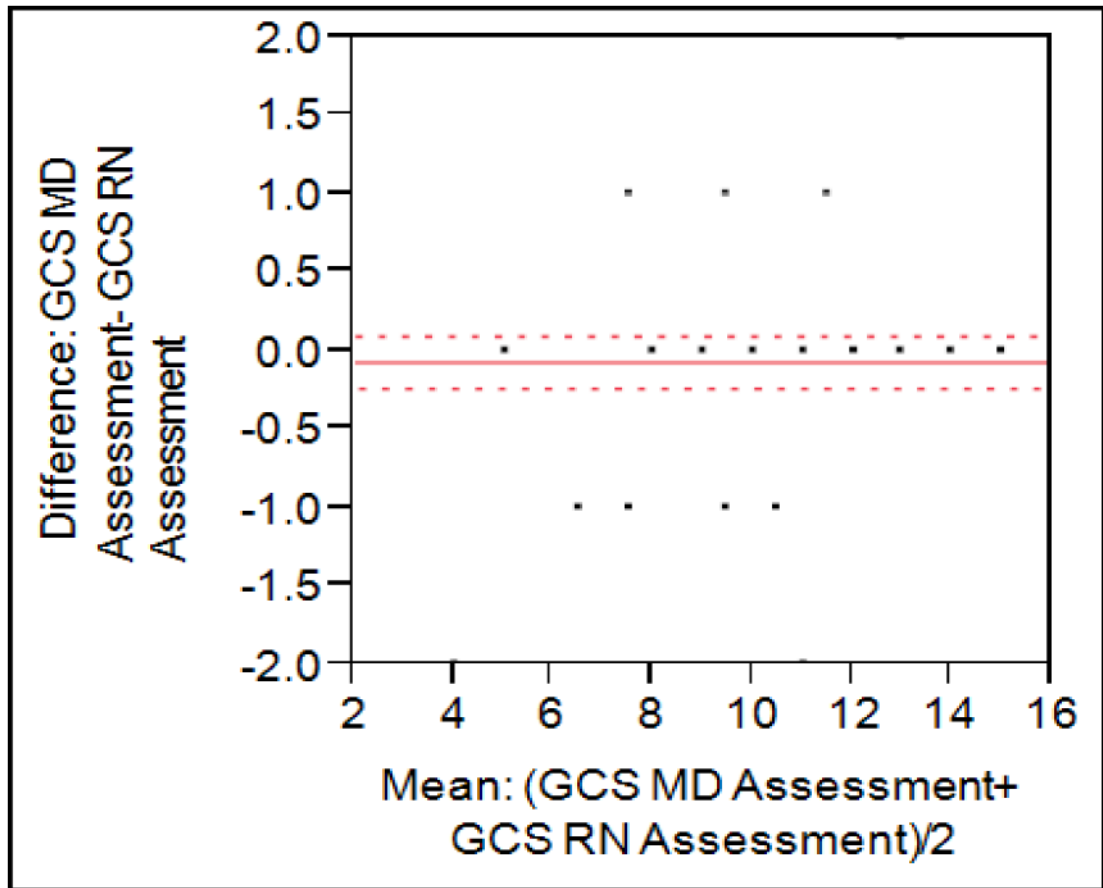
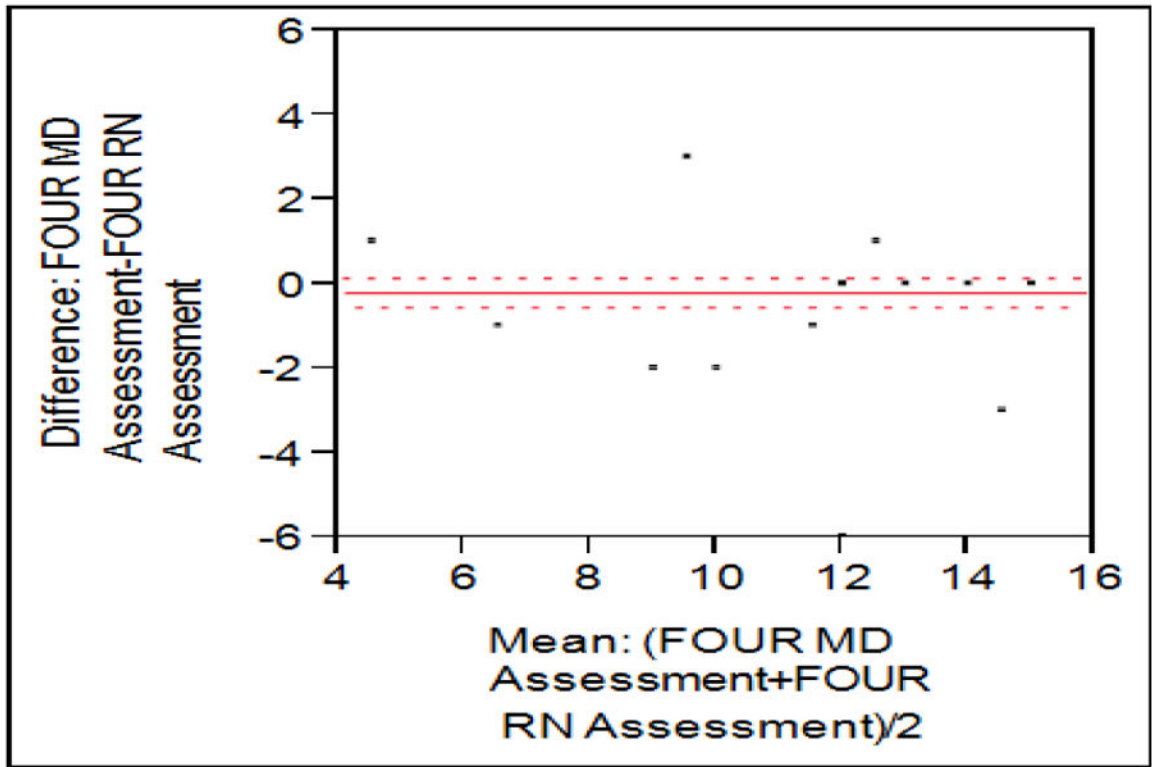


Figure 2. Study selection flow diagram



**Figure 3. Bland Altman Plot of GCS scores**  
Mean observed difference was 0.35, with a Pearson correlation of 0.87.



**Figure 4. Bland Altman plot of FOUR scores**  
Mean observed difference was -0.21, with a Pearson correlation of 0.92.

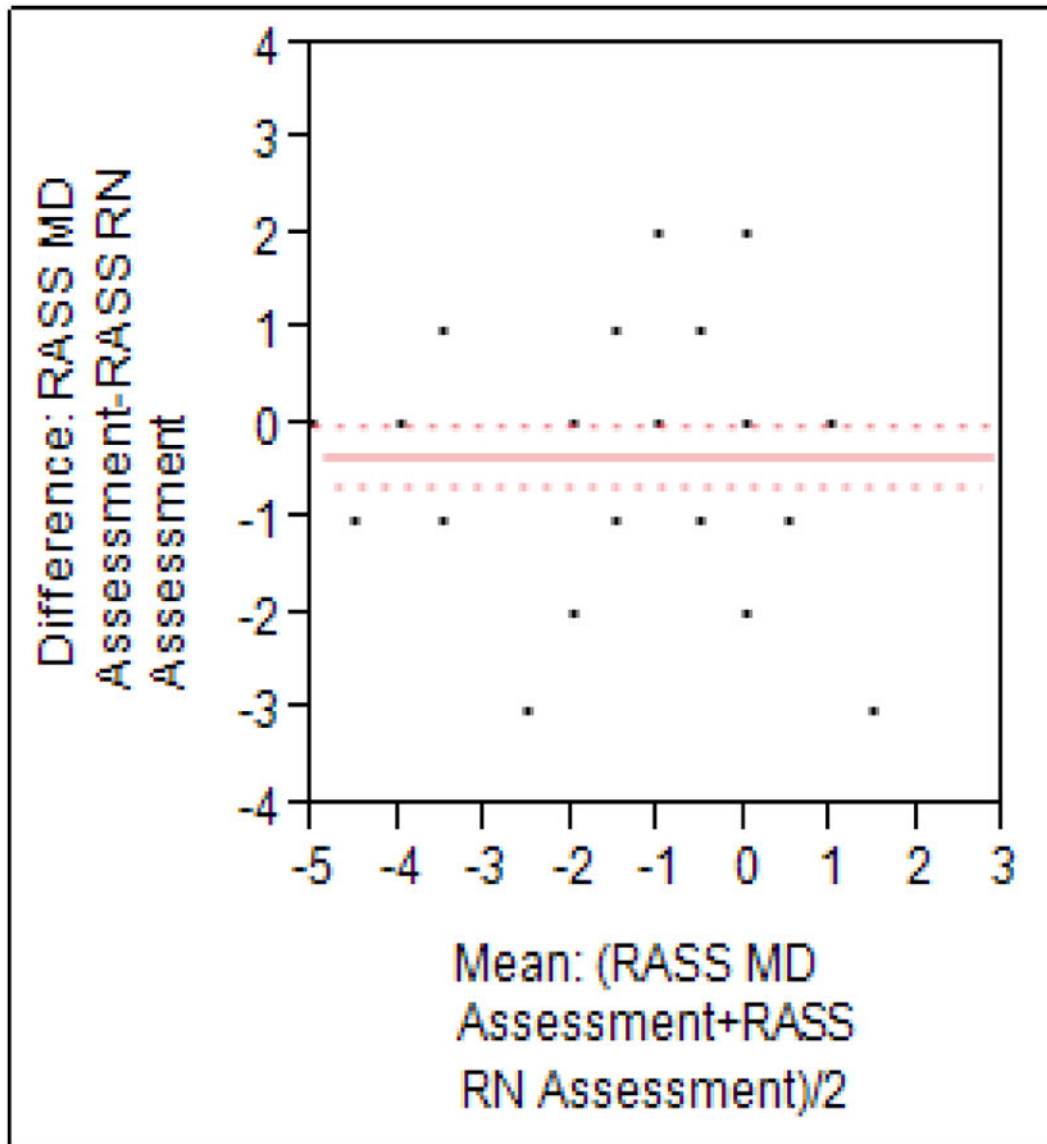


Figure 5. Bland Altman plot of RASS scores  
Mean observed difference was 0.33, with a Pearson correlation of 0.73.

**Table 1**

Baseline characteristics of included patients. Continuous variables are presented as median [interquartile range]. APACHE scores and SOFA scores are from the first 24 hours following ICU admission.

<i>Characteristic</i>	<i>Measurement (N=55)</i>
<i>Age</i>	67.1 [57.6-77.1]
<i>Gender (% male)</i>	49%
<i>Type of ICU admission (%)</i>	
<i>-Coronary care unit</i>	7.2%
<i>-Cardiovascular Surgery Intensive Care</i>	5.4%
<i>-Medical ICU</i>	72.3%
<i>-Surgical ICU</i>	14.5%
<i>SOFA score</i>	6 [3-10]
<i>APACHE III score</i>	69 [57-88]
<i>Abnormal GCS</i>	54.5%
<i>ABF</i>	61.8 %
<i>Interval between admission and assessment (in days)</i>	3.0 [0.7-3.1]
<i>Length of ICU Stay (in days)</i>	3.6 [2.2-7.8]
<i>Length of Hospital Stay (in days)</i>	12.1 [5.5-32.2]
<i>In-ICU mortality (%)</i>	7.2%
<i>In-Hospital mortality (%)</i>	12.7%