

# UCSF

## UC San Francisco Previously Published Works

### Title

Functional Status Examination Yields Higher Measurement Precision of Functional Limitations after Traumatic Injury than the Glasgow Outcome Scale-Extended: A Preliminary Study

### Permalink

<https://escholarship.org/uc/item/17976462>

### Journal

Journal of Neurotrauma, 37(4)

### ISSN

0897-7151

### Authors

Nelson, Lindsay D  
Brett, Benjamin L  
Magnus, Brooke E  
et al.

### Publication Date

2020-02-15

### DOI

10.1089/neu.2019.6719

Peer reviewed

# Functional Status Examination Yields Higher Measurement Precision of Functional Limitations after Traumatic Injury than the Glasgow Outcome Scale-Extended: A Preliminary Study

Lindsay D. Nelson,<sup>1</sup> Benjamin L. Brett,<sup>1</sup> Brooke E. Magnus,<sup>2</sup> Steve Balsis,<sup>3</sup> Michael A. McCrea,<sup>1</sup> Geoffrey T. Manley,<sup>4</sup> Nancy Temkin,<sup>5</sup> and Sureyya Dikmen<sup>6</sup>

## Abstract

The Glasgow Outcome Scale-Extended (GOSE) is one of the most widely used measures of functional limitations after traumatic brain injury (TBI), and is the primary outcome measure used in clinical trials of acute TBI treatment. However, the GOSE appears insensitive to the full spectrum of TBI-related functional limitations, which may limit its potential to capture treatment effects or correlate with other variables that impact outcome. The Functional Status Examination (FSE) was designed to improve on the assessment of injury-related functional limitations using a standardized assessment and wider possible score range. The aim of this pilot study was to employ item response theory (IRT) to test the hypothesis that the FSE yields more precise estimation of functional outcome than the GOSE. Traumatically injured patients ( $n = 100$ , 77 TBI, 23 orthopedic injuries) were interviewed at 3 months post-injury using both the GOSE and FSE structured interviews. IRT was used to quantify and compare the tests' information functions, which reflect the degree to which each instrument precisely measures functional limitations across the severity spectrum. Findings were consistent with predictions: the FSE yielded stronger measurement of functional limitations (i.e., higher test information) across a wider range of severity than the GOSE, whether scoring the GOSE from all interview items or using the traditional GOSE overall score. Although the FSE appears to be a promising alternative measure to the GOSE, further research is needed to cross-validate these findings in a larger sample and understand how to best deploy it in clinical and translational research.

**Keywords:** FSE; GOSE; IRT; outcome measurement; TBI

**T**RAUMATIC BRAIN INJURY (TBI) is a prevalent and sometimes disabling problem with almost no empirically supported treatments.<sup>1</sup> Despite several clinical trials for acute TBI treatment with strong pre-clinical and phase I–II data, no pharmacotherapy has significantly reduced TBI-related morbidity.<sup>2–4</sup> Researchers have suggested that imprecise outcome measurement contributed in part to the limited success of prior clinical trials.<sup>2,5</sup> The most commonly used outcome measure in TBI clinical trials—the Glasgow Outcome Scale (GOS)/GOS-Extended (GOSE)—has limitations that may contribute to imprecise estimation of outcome and, in turn, diminished estimates of treatment effects.<sup>6</sup> Among these factors are the lack of a single, sufficiently standardized administration protocol, which can lead to substantial rates of misclassification,<sup>7</sup> as well as a crude

ordinal rating scale focused on severe disability, which has limited sensitivity to individual differences (especially mild functional limitations).<sup>4,6</sup>

The Functional Status Examination (FSE)<sup>8</sup> is another interview-based measure developed to yield more precise estimates of injury-related functional limitations. Using a structured interview form and accompanying manual, examiners rate the degree to which an injury has affected patients' degree of participation across up to seven domains\* of life—Personal Care, Mobility/Ambulation,

\*Three additional domains available in earlier versions of the FSE—Executive Functioning/Cognitive Competency, Financial Independence, and Standard of Living—were eliminated from the instrument and are not discussed.

<sup>1</sup>Departments of Neurosurgery and Neurology, Medical College of Wisconsin, Milwaukee, Wisconsin.

<sup>2</sup>Department of Psychology, Marquette University, Milwaukee, Wisconsin.

<sup>3</sup>Department of Psychology, University of Massachusetts Lowell, Boston, Massachusetts.

<sup>4</sup>Department of Neurological Surgery, University of California San Francisco, Zuckerberg San Francisco General Hospital and Trauma Center, and the Brain and Spinal Injury Center, University of California San Francisco, San Francisco, California.

Departments of <sup>5</sup>Neurological Surgery and Biostatistics and <sup>6</sup>Rehabilitation Medicine, University of Washington, Seattle, Washington.

Mobility/Travel, Major Activity (Work or School, if applicable), Home Management, Leisure and Recreation, and Social Integration. Each domain is scored on a 0–3 ordinal scale, with domain ratings summed to achieve a more granular overall index. Preliminary work has suggested that the FSE has good test–retest reliability and correspondence between patient and significant others as well as fewer ceiling effects, more sensitivity to change, and stronger relationships with other variables (e.g., markers of injury severity) than the GOSE.<sup>8–10</sup>

However, comparisons between the GOSE and FSE have been limited to relatively simple descriptive analyses, predominantly of samples with higher TBI severity. Modern psychometric analytic tools bring additional rigor in improving outcome measurement for health conditions.<sup>11,12</sup> For example, item response theory (IRT) provides a formal strategy for determining how well items within a test, or a test as a whole, reflect a hypothetical continuum such as injury-related functional limitations. In particular, IRT allows one to quantify the association between one's level on a latent continuum (e.g., mild in functional limitations) and the probability of making a particular item response. We recently used IRT to investigate the GOSE structured interview, confirming its low sensitivity to mild functional limitations, while also providing feedback about specific aspects of the GOSE examination that perform most and least strongly for assessing different levels of severity within the dimension of functional limitations.<sup>6</sup>

The aim of this pilot study was to test the hypothesis that the FSE yields more precise estimation of functional limitations after injury than the GOSE. The sample ( $n=100$ ) was recruited for a pilot study aimed at comparing outcome measures relevant to the hospital trauma population. Enrollment included patients with TBI ( $n=77$ ) and orthopedically injured trauma patients ( $n=23$ ) to extend findings across the continuum of injury-related functional limitations (Table 1). To obtain a broad sample of patients likely to vary in outcomes, we recruited patients released from the emergency department (ED) as well as those admitted to the hospital (level 1 trauma center). Consistent with the TBI population, the majority of the sample had “mild” TBI (i.e., Glasgow Coma Scale score  $\geq 13$ ). Inclusion criteria were being  $\geq 18$  years old, English speaking, able to give informed consent, and not a prisoner. Participants in the TBI group were required to have sustained head trauma or rapid head movement leading to at least transient altered mental status.<sup>13</sup> Participants enrolled without TBI were required to have other traumatically induced bodily injury and none of the above-listed characteristics of TBI, and were selected to match the TBI group in mean age and overall distribution of participants by gender, race, cause of injury, and level of care (ED or inpatient). Inclusion criteria were confirmed by review of medical charts and completion of a semi-structured interview with each participant.

Participants completed an in-person clinical outcome assessment at 3 months post-injury (mean=92 days, range 85–111) comprising assessment of day-to-day functioning, injury-related symptoms, and quality of life. We administered a structured interview of the GOSE<sup>14</sup> to elicit self-report of the overall impact of all injuries sustained in the same accident on six areas of day-to-day functioning: Independence in the home, independence outside the home (i.e., shopping and travel), work, social/leisure activities, psychological problems that impact family and friend relationships, and other symptoms or problems that affect daily life. (For details on interview questions and response options, see Ranson et al.<sup>6</sup>) To avoid item-level missingness, all domains were assessed for every

participant. For this study, we aimed to retain as much granularity in GOSE item responses as possible, but needed to combine responses to the shopping and travel items together to avoid estimation issues caused by their high correlation. This resulted in six GOSE items comprising two two-level, two three-level, and two four-level items. As stated earlier, the FSE<sup>8</sup> structured interview assessed the effects of injury on seven domains of life, each coded on a four point severity scale. To maximize comparability of the two instruments' performance, data were collected by a small team of research assistants working under the supervision of the study principal investigator (L.D.N.), without direct input from the authors of either instrument.

We completed preliminary factor and IRT analyses in two phases. In Model A, we included all GOSE and FSE structured interview items in a single model, to yield comparable estimates of item and, by extension, test performance. In Model B, we included all FSE items alongside the GOSE overall score in a model. Whereas the former model maximally advantaged the GOSE by using all item-level data available in the interview, the latter was more consistent with the typical scoring and usage of the GOSE as a single ordinal variable. For each model, we first tested an assumption of the unidimensional IRT model to be employed that all of the items inputted reflected a single common dimension of injury-related functional limitations. This was achieved through exploratory factor analysis (EFA) and one-factor confirmatory factor analysis (CFA) using Mplus version 8.2.<sup>15</sup> Based on prior recommendations,<sup>11</sup> unidimensionality of the modified GOSE and FSE items (Model A) was established based on good fit via practical CFA fit indices (root mean square error of approximation [RMSEA]=0.05, comparative fit index [CFI]=0.95, Tucker–Lewis Index [TLI]=0.94), robust standardized factor loadings (0.61–0.84), and first to second eigenvalue ratio  $>4$  (4.56). In Model B, although RMSEA was not suggestive of good fit (RMSEA=0.12), we proceeded with the IRT model based on sufficient evidence for unidimensionality of the model reflected by other fit indices (CFI=0.96, TLI=0.95), robust standardized factor loadings (0.67–0.82), and first to second eigenvalue ratio of 6.53.

Second, we estimated a unidimensional two-parameter/graded response hybrid IRT model (i.e., one allowing inclusion of both binary and polytomous items) for Model A and Model B separately using Item Response Theory for Patient-Reported Outcomes (IRTPRO) v. 4.2.<sup>16</sup> The local independence assumption of IRT was supported by all local dependence  $\chi^2$  statistics being well under the recommended threshold of 10.<sup>17</sup> This model yielded feedback about the level of the latent continuum at which items could be expected to be endorsed (i.e., item difficulty) and the strength of the relationship between item responses and the latent continuum of functional limitations (i.e., discrimination). Because of the preliminary nature of these findings and our main objective of comparing overall information derived from each measure, we report only test-level (vs. item-level) findings. In particular, we computed test information for the GOSE and FSE separately by summing each instrument's item information values from the combined IRT model. *Test information* reflects the degree to which the test precisely measures a given level of the underlying continuum of functional limitations ( $\theta$ ). Advancing on classical test theory, IRT allows one to separately estimate the degree to which a test measures different levels of the continuum, acknowledging that some items may be informative primarily, for example, for high-severity limitations (e.g., basic activities of daily living), whereas others may be more informative for milder limitations. Second, test information is useful for estimating the

TABLE 1. SAMPLE CHARACTERISTICS

	Full sample (N=100)	TBI (n=77)	Trauma control (n=23)
	Mean (SD) or n (%)		
<b>Demographics</b>			
Age (years)	45.0 (16.3)	44.9 (15.7)	45.3 (18.5)
Sex (male)	56 (56%)	43 (56%)	13 (57%)
<b>Race</b>			
Black	39 (39%)	32 (42%)	7 (30%)
White	54 (54%)	39 (51%)	15 (65%)
Unknown/not reported	7 (7%)	6 (8%)	1 (4%)
<b>Education</b>			
Below high school	15 (15%)	12 (15%)	3 (13%)
High school graduate	32 (32%)	25 (33%)	17 (30%)
College or more	53 (53%)	40 (52%)	13 (57%)
<b>Injury characteristic</b>			
<b>Cause of injury</b>			
Motor vehicle accident	63 (63%)	49 (64%)	14 (61%)
Fall	25 (25%)	17 (22%)	8 (35%)
Assault	6 (6%)	6 (8%)	0 (0%)
Other	6 (6%)	5 (6%)	1 (4%)
Loss of consciousness	46 (46%)	46 (60%)	0 (0%)
Post-traumatic amnesia	48 (48%)	48 (62%)	0 (0%)
Retrograde amnesia	21 (21%)	21 (27%)	0 (0%)
<b>TBI severity group</b>			
GCS 3-8	3 (4%)	3 (4%)	N/A
GCS 9-12	0 (0%)	0 (0%)	N/A
GCS 13-15 CT positive	18 (23%)	18 (23%)	N/A
GCS 13-15 CT negative	56 (73%)	56 (73%)	N/A
<b>Highest level of care</b>			
Inpatient	48 (48%)	34 (44%)	14 (61%)
Emergency department	52 (52%)	43 (56%)	9 (39%)
<b>GOSE overall score</b>			
Median (IQR)	6 (5-7)	6 (5-7)	6 (5-7)
M (SD)	6.2 (1.3)	6.2 (1.3)	6.2 (1.3)
3, n (%)	3 (3%)	3 (4%)	0
4, n (%)	4 (4%)	2 (3%)	2 (9%)
5, n (%)	26 (26%)	21 (27%)	5 (22%)
6, n (%)	28 (28%)	20 (26%)	8 (35%)
7, n (%)	18 (18%)	15 (20%)	3 (13%)
8, n (%)	21 (21%)	16 (21%)	5 (22%)
<b>FSE overall score</b>			
Median (IQR)	7 (3-12)	7 (3-12)	6 (4-10)
Mean (SD)	7.4 (5.0)	7.6 (5.1)	6.7 (4.9)

CT, computed tomography findings of acute intracranial abnormalities; FSE, Functional Status Examination; GCS, Glasgow Coma Scale; GOSE, Glasgow Outcome Scale—Extended; IQR, interquartile range; SD, standard deviation; TBI, traumatic brain injury.

precision around estimates of individual patients’ degree of functional limitations in that information corresponds inversely, and non-linearly, with the standard error (SE) around patients’ estimated scores conditional on their level of  $\theta$  as follows:  $SE = 1/\sqrt{(Information|\theta)}$ .

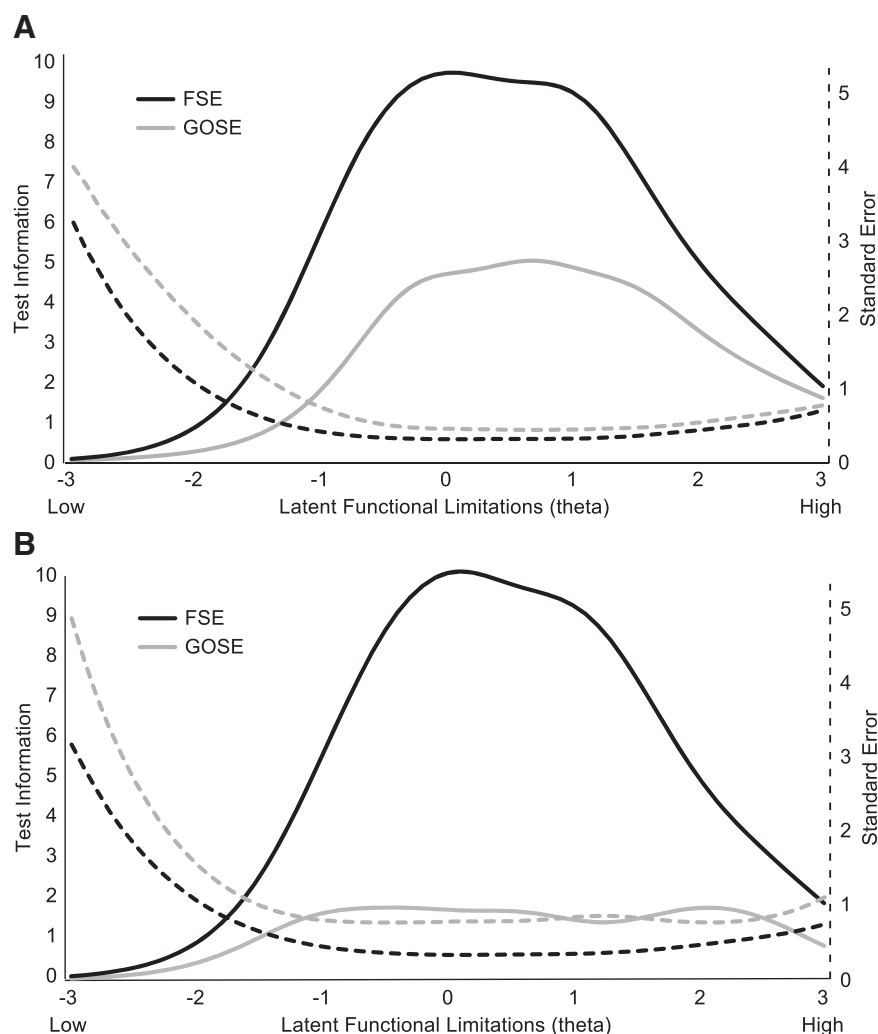
In other words, more information reflects more precise measurement.

Figure 1 depicts the test information (solid lines) and SE (dashed lines) functions for the GOSE (gray) and FSE (black). Consistent with our hypothesis, the FSE displayed higher information across a wider range of the continuum of functional limitations than the GOSE, both in comparison with the model that used all item-level GOSE interview responses (A) and the model that used the GOSE overall score (B). The area under the curve (AUC) of the test information function

for the FSE was 1.98\* and 4.15 times that of the GOSE for Models A and B, respectively. As is depicted in Figure 1, the FSE yields estimates of functional limitations with less measurement error (SE), particularly for individuals with lower-severity limitations.

Strengths of this study include the application of rigorous analyses in a sample with more mild injuries than most published samples that performed the FSE, providing strong preliminary data

\*Because the decision to collapse Shopping and Travel items resulted in one fewer item for the GOSE than for the FSE, we also computed the AUC ratio for mean item information. This FSE/GOSE ratio fell to 1.70 when computed from mean item rather than test information. Although weaker than the 1.98 ratio reported, this analysis achieved a similar conclusion about the relative superiority of the FSE for precisely measuring functional limitations.



**FIG. 1.** Test information functions for the Glasgow Outcome Scale—Extended (GOSE; gray solid line) and Functional Status Examination (FSE; black solid line) derived from a two-parameter/graded response hybrid item response theory (IRT) model. Model A used all items from the structured interview for both the GOSE and FSE. Model B compares the “one-item” GOSE overall score to the FSE scored from all interview items, which more closely reflects the two instruments as they are typically used in practice. Test information reflects the degree of precision with which individual scores on the latent continuum of injury-related functional limitations ( $\theta$ ) can be estimated across the continuum, which is inversely associated with the standard error of measurement (dashed lines) around  $\theta$ . The ratio of area under the curve (AUC) for the FSE versus GOSE was 1.98 (Model A) and 4.15 (Model B).

about the psychometric performance of these instruments, even in a sample for whom functional limitations 3 months post-injury are relatively mild. Further, these preliminary data suggest that the lower sensitivity of the GOSE as compared with the FSE is unlikely to be overcome by leveraging a more powerful scoring approach that uses the information available in all of the GOSE structured interview questions. It is of note that this study also included a small sample size that aggregated patients with TBI and non-brain injuries. Advantages of this include an ability to generalize the findings across a broader spectrum of trauma patients. That said, given the small sample size, it is possible that IRT parameter estimates were instable, which underscores the need to replicate these findings in an independent, larger sample. Finally, although GOSE scores are often dichotomized in practice, this analysis was restricted to investigating the ordinal GOSE overall score and the individual GOSE interview questions. Follow-up work should also examine the performance of these outcome measures in distinct TBI strata (e.g., cohorts with milder vs. more severe TBI), consider how different scoring

approaches affect measurement precision, and investigate how and when it is advantageous to parse the contributions of brain from those of non-brain injuries to ratings on these instruments.

The findings support the FSE as a promising alternative measure to the GOSE when more fine-grained measurement of injury-related functional limitations is desired. This may be especially valuable for patients with Glasgow Coma Scale scores of 13–15 (e.g., “mild” TBI/concussion), where residual deficits can be quite subtle. We demonstrated that, when assessed along a continuum, functional outcome after traumatic injury is measured more precisely across a wider spectrum of severity by the FSE than the GOSE. In offering more nuanced differentiation among individuals, the FSE may more accurately represent the reality of TBI outcomes, which are more heterogeneous than can be captured by a 5/8-level ordinal variable such as the GOS/GOSE. This could increase statistical power to detect relationships among other variables of interest or to demonstrate differences among different treatment groups. On the other hand, both the FSE and GOSE

appear to measure moderate-severe functional limitations better than they can measure than milder limitations, and it is unknown whether alternative measures would more precisely measure mild limitations. Additionally, we caution readers from inferring from these data that a continuous outcome measure is most efficient in all contexts. As has been pointed out by others, the statistical efficiency of binary, ordinal, or interval-level measures depends on sample- and context-specific factors.<sup>3,18–22</sup> Nevertheless, in providing a formal quantitative strategy for modeling the continuum of functional limitations, mapping items onto that continuum, and deriving estimates of outcomes using all available information, IRT may facilitate the measurement of complex behavioral and neuropsychological outcomes important to patients with TBI.

### Acknowledgments

We thank Amy Nader for coordinating the study, and Hannah Bartels, Robyn Furger, Alex Kirk, Georgia Ristow, Alexa Wild, and the MCW ED Enroller Program for contributing to recruitment, data collection, and entry.

### Funding Information

This pilot study was funded by the Medical College of Wisconsin Center for Advancing Population Science. Data analyses and manuscript preparation were supported by National Institutes of Health (NIH) grant R01NS110856 and United States Army Medical Research Acquisition Activity (USAMRRA) grant W81XWH-13-1-0441. Study data were collected and managed using REDCap electronic data capture tools hosted at the Medical College of Wisconsin, which was supported by National Institutes of Health grant UL1TR001436. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Army and do not necessarily represent the official views of the NIH.

### Author Disclosure Statement

Drs. Dikmen and Temkin developed the FSE but receive no financial incentives for its use. The other authors have no conflicts of interest to disclose.

### References

- Centers for Disease Control and Prevention (2019). Surveillance Report of Traumatic Brain Injury-related Emergency Department Visits, Hospitalizations, and Deaths—United States, 2014. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. [http://www.cdc.gov/traumaticbraininjury/get\\_the\\_facts.html](http://www.cdc.gov/traumaticbraininjury/get_the_facts.html) (last accessed September 20, 2019).
- Maas A.I., Roozenbeek B., and Manley G.T. (2010). Clinical trials in traumatic brain injury: past experience and current developments. *Neurotherapeutics* 7, 115–126.
- Maas A.I., Steyerberg E.W., Marmarou A., McHugh G.S., Lingsma H.F., Butcher I., Lu J., Weir J., Roozenbeek B., and Murray G.D. (2010). IMPACT recommendations for improving the design and analysis of clinical trials in moderate to severe traumatic brain injury. *Neurotherapeutics* 7, 127–134.
- Narayan R.K., Michel M.E., Ansell B., Baethmann A., Biegon A., Bracken M.B., Bullock M.R., Choi S.C., Clifton G.L., Contant C.F., Coplin W.M., Dietrich W.D., Ghajar J., Grady S.M., Grossman R.G., Hall E.D., Heetderks W., Hovda D.A., Jallo J., Katz R.L., Knoller N., Kochanek P.M., Maas A.I., Majde J., Marion D.W., Marmarou A., Marshall L.F., McIntosh T.K., Miller E., Mohberg N., Muizelaar J.P., Pitts L.H., Quinn P., Riesenfeld G., Robertson C.S., Strauss K.I., Teasdale G., Temkin N., Tuma R., Wade C., Walker M.D., Weinrich M., Whyte J., Wilberger J., Young A.B., and Yurkewicz L. (2002). Clinical trials in head injury. *J. Neurotrauma* 19, 503–557.
- Manley G.T., MacDonald C.L., Markowitz A., Stephenson D., Robbins A., Gardner R.C., Winkler E.A., Bodien Y., Taylor S., Yue J.K., Kannan L., Kumar A., McCrea M., and Wang K.K.W. (2017). The Traumatic Brain Injury Endpoints Development (TED) Initiative: progress on a public-private regulatory collaboration to accelerate diagnosis and treatment of traumatic brain injury. *J. Neurotrauma* 34, 2721–2730.
- Ranson J., Magnus B.E., Temkin N., Dikmen S., Giacino J.T., Okonkwo D.O., Valadka A.B., Manley G.T., Nelson L.D., and the TRACK-TBI Investigators (2019). Diagnosing the GOSE: structural and psychometric properties using item response theory, a TRACK-TBI study. *J. Neurotrauma* 36, 2493–2505.
- Wilson J.T., Sliker F.J., Legrand V., Murray G., Stocchetti N., and Maas A.I. (2007). Observer variation in the assessment of outcome in traumatic brain injury: Experience from a multicenter, international randomized clinical trial. *Neurosurgery* 61, 123–129.
- Dikmen S., Machamer J., Miller B., Doctor J., and Temkin N. (2001). Functional status examination: a new instrument for assessing outcome in traumatic brain injury. *J. Neurotrauma* 18, 127–140.
- Hudak A.M., Caesar R.R., Frol A.B., Krueger K., Harper C.R., Temkin N.R., Dikmen S. S., Carlisle M., Madden C., and Diaz-Arrastia R. (2005). Functional outcome scales in traumatic brain injury: a comparison of the Glasgow Outcome Scale (Extended) and the Functional Status Examination. *J. Neurotrauma* 22, 1319–1326.
- Dikmen S., Machamer J., Manley G.T., Yuh E.L., Nelson L.D., Temkin N.R., and the TRACK-TBI Investigators. (2019). Functional Status Examination versus Glasgow Outcome Scale Extended as outcome measures in traumatic brain injuries: how do they compare? *J. Neurotrauma*. [Epub ahead of print; doi: 10.1089/neu.2018.6198.]
- Hays R.D., Morales L.S., and Reise S.P. (2000). Item response theory and health outcomes measurement in the 21st century. *Med. Care* 38, I128–42.
- Balsis S., Choudhury T.K., Geraci L., Bengt J.F., and Patrick C.J. (2018). Alzheimer's disease assessment: a review and illustrations focusing on item response theory techniques. *Assessment* 25, 360–373.
- Kay T.H.D., Adams, R., Anderson, T., Berrol, S., Cicerone, K., Dahlberg, C., Gerber, D., Goka, R., Harley, P., Hilt, J., Horn, L., Lehmkuhl, D., Malec, J. (1993). Definition of mild traumatic brain injury. *J. Head Trauma Rehabil.* 8, 86–87.
- Wilson J.T., Pettigrew L.E., and Teasdale G.M. (1998). Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. *J. Neurotrauma* 15, 573–585.
- Muthén L.K., and Muthén B.O. (1998–2017). *Mplus User's Guide*, 8th ed. Muthén & Muthén: Los Angeles.
- IRTPRO for Windows*, Version 4.2. (2017). Scientific Software International: Skokie, IL.
- Chen W.H., and Thissen D. (1997). Local dependence indexes for item pairs using item response theory. *J. Educ. Behav. Stat.* 22, 265–289.
- Alali A.S., Vavrek D., Barber J., Dikmen S., Nathens A.B., and Temkin N.R. (2015). Comparative study of outcome measures and analysis methods for traumatic brain injury trials. *J. Neurotrauma* 32, 581–589.
- Bagiella E., Novack T.A., Ansell B., Diaz-Arrastia R., Dikmen S., Hart T., and Temkin N. (2010). Measuring outcome in traumatic brain injury treatment trials: Recommendations from the traumatic brain injury clinical trials network. *J. Head Trauma Rehabil.* 25, 375–382.
- Naggara O., Raymond J., Guilbert F., Roy D., Weill A., and Altman D.G. (2011). Analysis by categorizing or dichotomizing continuous variables is inadvisable: an example from the natural history of unruptured aneurysms. *AJNR. Am. J. Neuroradiol.* 32, 437–440.
- Altman D. G. (1991). *Practical Statistics for Medical Research*. Chapman and Hall: London.
- Choi S.C., Clifton G.L., Marmarou A., and Miller E.R. (2002). Misclassification and treatment effect on primary outcome measures in clinical trials of severe neurotrauma. *J. Neurotrauma* 19, 17–22.

Address correspondence to:

Lindsay D. Nelson, PhD

Department of Neurosurgery

Medical College of Wisconsin

8701 West Watertown Plank Road

Milwaukee, WI 53226

E-mail: linelson@mcw.edu