Towards a Better Measure of Brain Injury Outcome: New Measures or a New Metric?

Jacob Kean, PhD, CCC-SLP, a,b,c James F. Malec, PhD, ABPP-Cn, Rp b,d

From the aCenter for Health Information and Communication, Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service CIN 13-416, Richard L. Roudebush VA Medical Center, Indianapolis, IN; bDepartment of Physical Medicine and Rehabilitation, Indiana University School of Medicine, Indianapolis, IN; cRegenstrief Institute, Indianapolis, IN; and dRehabilitation Hospital of Indiana, Indianapolis, IN.

The Glasgow Outcome Scale (GOS), its successor the Glasgow Outcome Scale-Extended (GOSE), 1 and the Disability Rating Scale (DRS) 2 represent a desire for a single outcome score that can be used to track the progression of a patient with brain injury (BI) from coma to community. Although the GOS/GOSE and DRS have been rightly criticized for their psychometric and other limitations, 3-5 these scales span the major domains of the International Classification of Functioning, Disability and Health— that is, bodily functions, activities, and participation. 6 These early outcome scales also map directly onto the conceptual domains identified as representing outcome in 2010 by the Common Data Elements (CDE) Task Force (fig 1). Two of the areas targeted by the CDE Task Force but not explicitly mentioned by the GOS and DRS—behavioral function and neuropsychological impairment—may be considered more detailed and objective measurements of cognitive limitations that are clearly targeted as outcome indicators by both the GOSE and DRS. The additional domains identified by the CDE Task Force (health economic measures, patient-reported outcomes) represent, in the former case, an outcome indicator that does not reside within the person with BI, and in the latter case, a domain that overlaps with other domains.

The Traumatic Brain Injury Model Systems (TBIMS) study group has implicitly endorsed this conceptualization of outcome by retaining the GOS/GOSE and DRS throughout its 30-year history. The TBIMS group adopted the Satisfaction With Life Scale 1 to represent important content that was absent from the GOSE or DRS. Although included in both the GOSE and DRS, specific indicators of independent living and productive activity that appear basic to our idea of outcome are also recorded separately in the TBIMS longitudinal database. In short, the domains described in figure 1 appear to represent a conceptualization of outcome that has been informed by early outcome scales and has withstood the test of time, with ongoing and recent endorsement by the TBIMS study group and by the CDE Task Force.

Reflective outcome measurement

The GOSE and DRS are reflective measures. Reflective measures are composites of individual items (ie, observed variables) that are indicators believed to describe or reflect an underlying construct (ie, latent variable). 8 In contrast, formative measures, as we explain in more detail below, are composed of indicators believed to determine or form the construct. Socioeconomic status (SES) is an example of a formative measure. In reflective measurement, latent constructs are typically measured as the aggregation of reflective indicators (ie, items in a scale) because the indirectness of the indicators requires that several aspects be considered together to reasonably approximate a construct. The ability of a reflective measure to differentiate ability is related to the number of items and their interrelatedness. Because good measures reliably differentiate performance and better measures do so with greater precision, the recipe for creating the best measures is to develop a large number of interrelated items that reflect the full range of ability on the construct of interest.

However, improving reflective composites of global outcome such as GOSE and DRS using this approach is difficult because outcome is so broad and likely multidimensional. In other words, completely capturing a broad conceptual domain along a continuum “from coma to community” can come at the cost of compromised measurement precision. Satisfaction with life, for example, which is perhaps as important as or more important to outcome than activities and participation, is a domain found by...
Corrigan and Bogner as well as in our own unpublished work to be statistically orthogonal (ie, uncorrelated) to other outcome domains. Other outcome domains depicted in figure 1 correlate to varying degrees with one another. The complex relationships among these various outcome domains pose a significant challenge to the establishment of a single outcome scale to track recovery.

One approach to improving on available but coarse reflective measures is the use of advanced multidimensional and hierarchical psychometric models to account for the complex relationships among latent traits. One might, for example, collate extant items from available reflective measures, such as those listed in figure 1. We are currently collaborating on a TBIMS module project of this type. A challenge for efforts such as this is that the psychometric qualities of these available measures vary considerably. Developing new multidimensional reflective measures from the ground up may yield more precise outcomes but is time and resource intensive, requiring domain mapping, qualitative item review or development, item library construction, cognitive interviewing, psychometric evaluation and calibration using advanced models, and large sample sizes (eg, 2000+). These advanced reflective models may more realistically approximate the latent construct of interest. However, the resulting measures can be difficult to interpret since they consist of multiple distinct but correlated unidimensional subscales.

**Formative outcome measurement**

Another option is to create a formative, rather than reflective, composite. Whereas a reflective measure is composed of observed indicators believed to reflect the manifestation of an underlying latent construct, a formative measure is composed of observable indicators believed to determine the construct. In a formative model, the indicators (ie, items) do not reflect the latent construct because it makes more conceptual sense to view the construct as the reflection of the items (rather than vice versa as in reflective measures). Bollen and Lennox offer the example of SES as a construct that is realized as the aggregation of 4 formative indicators: education, occupational prestige, income, and neighborhood. In this case, the direction of determination flows from the 4 indicators to the concept of SES rather than the reverse, as would be the case for a reflective measure. A formative measure may combine various types of indicators (eg, categorical, ordinal, continuous) that may or may not be correlated. For example, the components of SES are not highly correlated, but each informs the assessment of an individual’s SES. That is, a college professor may be of relatively high SES on the basis of education and occupational prestige despite relatively low income, whereas a lottery winner may achieve greater SES based on income alone. A formative measure, in simplest form, is based on a regression model in which the dependent variable is the construct of interest and the explanatory (predictive) variables are observable indicators. For this reason, including variables in the model that are correlated with the latent trait, in this case outcome, but not with each other, should improve the model.

Table 1 shows criteria offered by Jarvis et al to help researchers consider whether the relationship between measures and the construct of interest is reflective or formative. The first set of questions relates to the direction of determination between the construct and its indicators, the second to the interchangeability of the indicators, the third to the covariance among indicators, and the fourth to the nomothetic net, or antecedents and consequences, of the indicators. Despite these criteria, the general advice is that most constructs can be measured reflectively and that researchers should opt for reflective measurement, if possible. Nevertheless, for constructs such as SES, the consensus seems to be that formative measurement is a better option than a reflective measure with indicators such as “How high are you up the social ladder?” (as suggested in Borsboom et al).

Despite the potential advantages, many have been critical of formative measurement. One issue is the challenge of determining how to aggregate and weight components, which has been shown to have pronounced effects on the ordering of cases of interest. A related issue is the need to include reflective measures in the measurement model as determinants of the formative construct to achieve identification. In other words, without also modeling reflective measures “caused” directly or indirectly by the formative construct, there are an infinite number of solutions to a formative measurement model. However, the choice of reflective measures

**Fig 1** Brain injury outcome domains and recommended measures. Abbreviations: CHART-SF, The Craig Handicap Assessment and Reporting Technique—Short Form; PART, Participation Assessment with Recombined Tools; TBI, traumatic brain injury.

**List of abbreviations:**
- BI: brain injury
- CDE: Common Data Elements
- DRS: Disability Rating Scale
- GOS: Glasgow Outcome Scale
- GOSE: Glasgow Outcome Scale—Extended
- SES: socioeconomic status
- TBIMS: Traumatic Brain Injury Model Systems
can affect the loadings of the formative elements on the composite, which arguably affects the meaning of the construct itself.16 Unlike reflective measures, too great of a relationship among indicators (ie, multicollinearity) poses a computational problem for the statistical model. In addition, there is no psychometric theory to guide validity testing of formative measures. Finally, some consider it inappropriate to test theory using formative measures because formative measurement models assume that the theoretical construct is created by rather than reflected in its measures.

Nonetheless, a formative measure may make conceptual sense for BI outcome, as it does for SES. The statistical orthogonality between satisfaction with life and activities and participation after BI indicates no relationship between these outcome domains. This dissociation is difficult to handle in a reflective model, which expects indicators to covary. A formative outcome measure, in contrast, would account for the independent influence of life satisfaction as well as activity and accommodation of complex relationships among measures. A formative measurement allows “mixing apples and oranges” (ie, uncorrelated indicators) in situations where the measurement goal is analogous to recording servings of fruit consumed each day. With its diversity of composite indicators, outcome may be such a mix of apples and oranges.

More concretely, a formative measure might be constructed using available measures and indicators. Assuming that underlying neuropsychological, psychological and behavioral impairment would be captured by indicators of activity limitations, a measure that combines the GOSE, DRS, FIM subscales, Satisfaction With Life Scale, and employment and independent living status may sufficiently cover the outcome domains identified historically by the field (see fig 1). Our research group is currently engaged in examining the practicality and viability of such a measure. As mentioned previously, the greatest challenges in developing such a measure may be achieving model identification and verifying its external validity.

Although challenges to the development of a formative outcome measurement for BI are significant and the topic of formative measurement is a controversial one, developing precision reflective outcome measures de novo is no less challenging and certainly more resource intensive. A formative outcome measure for BI is achievable in a shorter period since it is based on available data and has the practical advantages of a single score and accommodation of complex relationships among measures. Treating outcome as a construct determined by indicators may also make more conceptual sense than modeling outcome as a latent construct. For these and other reasons, formative measurement may deserve consideration in physical medicine and rehabilitation clinical and research communities.

### Corresponding author

James F. Malec, PhD, ABPP-Cn, Rp, Professor and Research Director, Department of Physical Medicine and Rehabilitation, Indiana University School of Medicine and Rehabilitation Hospital of Indiana, 4141 Shore Dr, Indianapolis, IN 46254. E-mail address: jmalec@rhin.com.
References