

UCSF

UC San Francisco Previously Published Works

Title

Developing a Cognition Endpoint for Traumatic Brain Injury Clinical Trials

Permalink

<https://escholarship.org/uc/item/8zv6g1fg>

Journal

Journal of Neurotrauma, 34(2)

ISSN

0897-7151

Authors

Silverberg, Noah D
Crane, Paul K
Dams-O'Connor, Kristen
[et al.](#)

Publication Date

2017-01-15

DOI

10.1089/neu.2016.4443

Peer reviewed

Developing a Cognition Endpoint for Traumatic Brain Injury Clinical Trials

Noah D. Silverberg,¹ Paul K. Crane,² Kristen Dams-O'Connor,³ James Holdnack,⁴ Brian J. Ivins,⁵ Rael T. Lange,⁶ Geoffrey T. Manley,⁷ Michael McCrea,⁸ and Grant L. Iverson⁹

Abstract

Cognitive impairment is a core clinical feature of traumatic brain injury (TBI). After TBI, cognition is a key determinant of post-injury productivity, outcome, and quality of life. As a final common pathway of diverse molecular and microstructural TBI mechanisms, cognition is an ideal endpoint in clinical trials involving many candidate drugs and nonpharmacological interventions. Cognition can be reliably measured with performance-based neuropsychological tests that have greater granularity than crude rating scales, such as the Glasgow Outcome Scale-Extended, which remain the standard for clinical trials. Remarkably, however, there is no well-defined, widely accepted, and validated cognition endpoint for TBI clinical trials. A single cognition endpoint that has excellent measurement precision across a wide functional range and is sensitive to the detection of small improvements (and declines) in cognitive functioning would enhance the power and precision of TBI clinical trials and accelerate drug development research. We outline methodologies for deriving a cognition composite score and a research program for validation. Finally, we discuss regulatory issues and the limitations of a cognition endpoint.

Keywords: clinical trials; cognition; craniocerebral trauma; neuropsychological tests; outcome assessment; psychometrics

Introduction

OVER THE PAST SEVERAL DECADES, every major clinical trial evaluating an acute treatment for traumatic brain injury (TBI) has failed to demonstrate significant therapeutic benefit.¹ The reasons for failure are undoubtedly multifaceted, but one universally recognized factor is inadequately accurate and sensitive outcome measurement.^{1–4} Crude measures of global functional outcome, most notably the Glasgow Outcome Scale-Extended (GOS-E),⁵ have been by far the most commonly studied primary endpoints in TBI trials.⁶

The GOS-E is an eightpoint ordinal rating scale, but it is typically dichotomized for analysis. It involves asking patients and/or their caregivers a series of questions (e.g., “Are you able to shop without assistance?”) to determine the overall degree of disability.

The International Mission on Prognosis and Clinical Trial Design in TBI⁷ and agencies such as the United States Department of Defense⁸ have recognized the limitations of these blunt tools, especially for mild to moderate TBI, and called for research to advance outcome measures for TBI clinical trials.

Although improved methods for measuring global functional outcome would be valuable, the thesis of this article is that developing and validating a cognition endpoint should also be a top research priority. Further, the ability to generate a cognition composite from differing neuropsychological batteries will enable interrogation of existing databases and thereby greatly accelerate the validation process. We outline the advantages of a cognition endpoint and sketch a road map for its development and validation.

Cognitive impairment is a core clinical feature of TBI. Depending on the severity of brain injury, cognition is often severely

¹Department of Medicine, Division of Physical Medicine and Rehabilitation, University of British Columbia, and GF Strong Rehab Centre, Vancouver, British Columbia, Canada, and Department of Physical Medicine and Rehabilitation, Harvard Medical School, and Red Sox Foundation and Massachusetts General Hospital Home Base Program, Boston, Massachusetts.

²Department of Medicine, University of Washington, Seattle, Washington.

³Department of Rehabilitation Medicine, Icahn School of Medicine at Mount Sinai, New York City, New York.

⁴Department of Physical Therapy, University of Delaware, Newark, Delaware.

⁵Defense and Veterans Brain Injury Center (DVBIC), Silver Spring, Maryland.

⁶Defense and Veterans Brain Injury Center (DVBIC), Walter Reed National Military Medical Center, and National Intrepid Center of Excellence, Bethesda, Maryland.

⁷Department of Neurological Surgery, University of California San Francisco, San Francisco, California.

⁸Department of Neurosurgery, Medical College of Wisconsin, Milwaukee, Wisconsin.

⁹Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, MassGeneral Hospital for Children Sports Concussion Program, and Red Sox Foundation and Massachusetts General Hospital Home Base Program, Boston, Massachusetts, and Defense and Veterans Brain Injury Center, National Intrepid Center of Excellence, Bethesda, Maryland.

affected in the acute post-injury stage and gradually improves over time, with a slowing trajectory.^{9–11} Even the mildest of TBIs at least transiently impacts cognition.¹² The risk of persistent or permanent cognitive impairment increases with the severity of injury.^{9–11,13–15} Patient characteristics such as age further modify the recovery curve.^{16–19} Cognition is associated with functional outcome in most studies, as measured by rehabilitation gains, functional independence, community reintegration, and employment.^{20–25} Given its prevalence and clinical significance, cognitive impairment is a key therapeutic target after TBI and requires sound measurement tools. It has been the primary outcome for efficacy in major TBI trials^{26,27} but far less frequently than the GOS-E and similar global outcome scales.⁶

Cognition has several distinct advantages as an endpoint for TBI trials. First, cognition can be reliably measured by performance-based neuropsychological tests, which have a rich scientific literature in general²⁸ and in TBI in particular.^{10,11,29} Second, a cognition endpoint can provide improved granularity over global functional outcome scales. In particular, neuropsychological tests may extend the dynamic range upward by detecting subtle differences among persons who score at the ceiling of the GOS-E.^{30,31} Third, compared with disability rating scales, cognition is more proximal to the TBI neuropathology such that it should be more responsive to treatments that alter neurophysiology while it should also be less influenced by non-TBI factors such as social support and comorbid bodily injuries.

Fourth, cognition is functionally relevant. It underlies TBI-related challenges with daily functioning, or in the World Health Organization's International Classification of Functioning framework, cognition links body structure and activity/participation.³² Fifth, a cognition endpoint could facilitate the transition from pre-clinical drug development to human clinical trials. Cognition, typically measured in animals by the Morris Water Maze task or variant, is one of the most widely used endpoints in rodent models of TBI.^{33,34} A discordance between the endpoints in pre-clinical studies (cognition) and Phase III clinical trials (global disability) might partly explain why promising pre-clinical evidence fails to translate into positive human trials.³⁵

Sixth, cognition has value as a prognostic marker. Cognition can be assessed early, before hospital discharge.²⁰ The strongest predictors of outcome from moderate to severe TBI (e.g., pupil response, computed tomography (CT) findings, and Glasgow Coma Scale)³⁶ lose their prognostic power when applied to patients with mild TBI.³⁷ Early neuropsychological testing, on the other hand, robustly predicts long-term outcome across the range of TBI severity.^{20,25,38,39} Therefore, a cognition outcome can be used to enrich clinical trials (e.g., enrolling only patients who are likely to have persistent cognitive impairment) or enhance the power of clinical trials through risk stratification or covariate adjustment.⁴⁰

Finally, cognition is a final common pathway of diverse TBI mechanisms, such as focal contusions, hematomas, diffuse axonal injury, edema, cellular dysfunction (e.g., excitotoxicity, calcium overload, oxidative stress, mitochondrial dysfunction, and inflammation), impaired synaptic transmission, cell death (necrosis or apoptosis), and axonal degeneration.⁴¹ Cognition is therefore well suited to measure benefit from therapeutics that influence multiple neurophysiological processes. For example, citicoline was recently evaluated in a large TBI trial, where it was hypothesized to exert neuroprotective effects and promote neurorecovery through numerous mechanisms.²⁶ Combination therapies, which are increasingly studied in TBI, by nature have diverse mechanisms of action and so require a downstream endpoint.⁴² Most nonpharmacological interventions also have complex neurophysiological mechanisms

(e.g., hyperbaric oxygen therapy)⁴³ or they target cognition directly (e.g., attention training).⁴⁴

Remarkably, there is no well-defined, widely accepted, and validated cognition endpoint for TBI. There is incomplete to no overlap in the test batteries used in past and ongoing clinical trials and observational studies. Further complicating matters, neuropsychological test data have been scored and analyzed in several different ways, reflecting a lack of agreement as to how multidimensional outcomes in clinical trials should be handled. Selecting a cognition endpoint and analytic approach are not trivial decisions. They directly influence sample size requirements, stopping rules for futility, and conclusions about treatment efficacy. Before initiating efforts to validate a cognition endpoint, it is essential to decide which tests to use and how to combine information from those tests. Only then will we have a single “measure” to validate.

Which Tests?

No single neuropsychological test captures the diversity of cognitive impairments seen after TBI. Adequate measurement of cognitive outcome from TBI requires administering a battery of tests that, at minimum, cover the domains of attention, memory, processing speed, and executive functioning.^{45,46} A large number of neuropsychological tests are available to measure these functions.²⁸

Over the past 25 years, there have been several attempts to endorse and promote particular neuropsychological tests for TBI clinical trials. The National Institutes of Health/National Institute of Neurologic Disorders and Stroke (NIH/NINDS) sponsored a conference in 1991 to identify appropriate outcome measures for TBI clinical trials.⁴⁷ The NINDS Head Injury Centers subcommittee met the following year and advised a broader battery to use when a richer characterization of cognition is desired.⁴⁸

The National Institute of Child Health and Development-sponsored TBI Clinical Trials Network was established in 2008 to support multicenter clinical trials.⁴⁶ They selected a battery of neuropsychological tests for the Citicoline Brain Injury Treatment Trial²⁶ and presumably others to follow. More recently, the NINDS Common Data Elements (CDE) project was undertaken to standardize outcome measurement for TBI research. One aim of this project was to improve the precision and consistency of measuring treatment effects.⁴⁵ In 2010, the NINDS-CDE TBI Outcomes Workgroup recommended a set of core measures across multiple domains, including neuropsychological functioning, with a revised set (version 2.0) in 2012.⁴⁹ The specific tests and group endorsing them are summarized in Table 1.

The NIH Blueprint for Neuroscience Research (www.neuroscienceblueprint.nih.gov) called for standardized outcome measurement, not just in TBI research, but across neurological conditions. Rather than attempt consensus for the adoption a common set of existing measures, the Blueprint advanced the development of new measures. The NIH Toolbox for the Assessment of Neurological and Behavioral Function⁵⁰ was released in 2012, after more than 8 years in development. The end product is a set of tests, standardized and normed across the life span, that measure different aspects of cognitive, emotional, motor, and sensory health and functioning. The Cognition Battery⁵¹ comprises performance-based neuropsychological tests measuring attention, working memory, language, processing speed, and executive functioning that can be administered in approximately 30 min. Validation studies in TBI are under way.

Another option, rather than settling on a particular set of tests, is to instead validate a composite scoring system that can be used for

TABLE 1. NEUROPSYCHOLOGICAL TESTS FOR ADULTS ENDORSED BY EXPERT CONSENSUS GROUPS

	<i>NIH/NINDS (Clifton, 1991)</i>	<i>NINDS Head Injury Centers (Hannay, 1996)</i>	<i>TBI Clinical Trials Network (Bagiella, 2010)</i>	<i>NINDS- CDE 1.0 (Wilde 2010)</i>	<i>NINDS- CDE 2.0 (Hicks 2013)**</i>
Ruff 2 & 7		X			
BVMT-R				X*	
CVLT-2			X		
COWA	X	X	X	X*	
Finger tapping		X			
Grooved pegboard	X	X		X*	
Naming (MAE)		X			
PASAT	X [^]	X			
Rey complex figure	X	X			
Reaction time		X			
RAVLT				X	X
Selective reminding test	X	X			
SDMT		X			
Stroop/CWIT			X	X*	
Token test (MAE)		X			
Trail making test	X	X	X	X	X
Visual form Discrimination		X			
WAIS digit span			X	X*	
WAIS digit symbol	X		X	X	X
WAIS L-N sequencing				X*	
WAIS symbol search			X	X	X
WCST	X [^]	X			

NIH/NINDS, National Institutes of Health/National Institute of Neurologic Disorders and Stroke; TBI, traumatic brain injury; NINDS-CDE, NINDS Common Data Elements; BVMT-R, Brief Visuospatial Memory Test-Revised; CVLT, California Verbal Learning Test-2; COWA, Controlled Oral Word Association test; MAE, Multilingual Aphasia Examination; PASAT, Paced Auditory Serial Addition Test; RAVLT, Rey Auditory Verbal Learning Test; SDMT, Symbol Digit Modalities Test; CWIT, Color-Word Interference Test; WAIS, Wechsler Adult Intelligence Scale; WCST, Wisconsin Card Sorting Test.

[^]Not recommended for severe TBI.

*Designated as "supplementary."

**A large number of additional tests (not shown in this table) were designated as "supplementary."

differing or even nonoverlapping test batteries, provided they meet some minimal criteria (e.g., coverage of cognitive domains, reliability, and adequate normative data). Recent research supports that a cognition composite based on different tests that measure the same construct (e.g., Rey Auditory Verbal Learning Test vs. California Verbal Learning Test) perform very similarly.⁵² Batteries that are composed of different tests cannot be expected to perform comparably to one another when the differences in battery composition become extreme. For example, a composite score derived from a brief battery of mostly speeded tests would likely have different properties than one derived from a more comprehensive battery with balanced coverage of processing speed, attention, memory, and executive functioning.

Further empirical work will be necessary to identify the limits of flexibility with respect to the number and type of tests. A composite that can be derived from multiple neuropsychological batteries will permit harmonization of these disparate measures across existing databases, thereby permitting interrogation of measures with more highly powered analyses, thus accelerating the validation process.

How to Combine Information from the Tests in the Battery?

Neuropsychological test data have been analyzed in numerous ways, reflecting a lack of agreement as to how to handle multidimensional outcomes in clinical trials. Each of these methods has unique limitations. The most common approach has been to perform group comparisons for each neuropsychological test in the

battery. For example, High and associates⁵³ compared patients receiving growth hormone replacement versus placebo on a comprehensive battery of neuropsychological tests. They found an "encouraging" (pg. 1573) treatment effect on total correct responses on the Wisconsin Card Sorting Test, but no significant group differences on another 27 measures. This approach risks mixed findings or chance positive findings (Type I error). Corrections for the family-wise error rate can guard against Type I error, but at the cost of sacrificing statistical efficiency.

Multivariate techniques such as multivariate analysis of variance/multivariate analysis of covariance (MANOVA/MANCOVA) are also widely used, such as in a study of prescribed rest for sport-related concussion.⁵⁴ Methods for translating the variance-accounted-for effect size (e.g., eta-squared) yielded by MANOVA/MANCOVA into a clinically meaningful metric are lacking.⁵⁵ Multivariate techniques can also provide misleading, biologically implausible results.⁴⁶ For example, a large improvement on one cognitive test and small decline on several other cognitive tests in the battery could lead to rejection of the null hypothesis.

Another solution implemented in at least one TBI trial (evaluating citicoline)²⁶ is a global test procedure based on the logistic model. This approach does not grade the severity of impairment on each test and assumes a common treatment effect across all tests in the battery.

Finally, there are nonparametric alternatives. For example, a study investigating the side effects of valproate in TBI⁵⁶ rank ordered neuropsychological test scores within the sample, averaged the ranks across the tests in the battery, and performed nonparametric statistics on the resulting mean rank score.

There are several advantages to analyzing a cognition composite endpoint instead of a set of neuropsychological test scores. First, having a single prospectively defined primary outcome is required by the Food and Drug Administration (FDA) to license new drugs.⁵⁷ Second, a single endpoint is also more readily subjected to validation studies such that the endpoint can be qualified for drug development research. Third, more flexible and powerful statistical approaches can be used to analyze a single endpoint. For example, generalized linear modeling and extensions can be used for an endpoint with any distribution that was measured across any number of time points. Fourth, clinical trials with a single endpoint can be readily meta-analyzed to more definitively establish treatment efficacy. Finally, a clinical trial with a single endpoint produces findings that will be easier for patients and their caregivers to understand,⁵⁸ empowering them to make informed treatment decisions.

Although a cognition composite endpoint integrates information from multiple tests, it can be constructed in a manner that facilitates straightforward interpretation, with higher scores indicating better (or worse) cognition. This is in contrast to truly multidimensional endpoints that integrate diverse outcomes, such as mortality and functional, neuropsychological, and patient-reported measures.⁵⁹

Approaches to Creating a Cognition Endpoint

The simplest and most widely used contemporary method for creating a composite score is to place the individual test scores on a common metric (e.g., norm-referenced *z*, *T*, scaled score, percentile, or other standard scores) and average them, producing what can be called the Overall Test Battery Mean (OTBM).⁶⁰ The OTBM has been shown to discriminate between levels of TBI severity,²⁹ and it has been used for clinical trials in TBI.⁴³

An alternative approach is to use Item Response Theory (IRT)-based methods to model latent ability based on a set of neuropsychological test scores. A model is built using confirmatory factor analysis in which cognitive test scores serve as indicators of the latent trait, and the model that fits the available data and is concordant with cognitive theory is selected.⁶¹ The latent ability model captures variance across cognitive tests and can also account for covariance attributable to methods effects or theoretical similarities across cognitive subtests.⁶²

A latent ability composite score is then computed for each participant in the dataset; the scale is typically and arbitrarily anchored to a mean of 0 and variance of 1. The latent ability composite score can be validated by comparing it with constituent cognitive test scores or other relevant data with respect to strength of association with markers of cognition or disease (e.g., imaging or other biomarkers) or ability to detect change over time. This approach has proven valuable for dementia research^{63,64} but it has not yet been applied to TBI.

The primary advantages of a latent ability composite are that it has linear measurement properties and accounts for differences among tests in the battery with regard to precision and relationships with demographic variables. Linear measurement refers to the notion that differences between points on a test or scale are equal across the range of the scale. Linear scaling has the potential to improve precision across a wider range of cognitive ability and reduce bias in estimating changes in cognition over time.⁶⁵

When constituent cognitive tests vary in difficulty and precision (e.g., one test measures high cognitive functioning with good precision and another precisely measures low cognitive functioning),

the latent ability score can measure the underlying trait well across the full range of ability. A further advantage of this approach to modeling is its ability to quantify and account for differences in test performance across demographic groups that persist after controlling for the underlying ability measured by all of the tests (differential item functioning).⁶⁶

Each test that contributes to the latent ability composite can be evaluated for differential item functioning effects; that is, systematic bias for people with certain demographic profiles. It is possible to model the direct effect of pre-injury variables (e.g., pre-morbid intelligence) on latent ability, as well as their relationship with the test scores independent of latent ability. Latent ability composite scores have been created for Alzheimer disease^{63,64} but not yet for TBI.

A third possibility is to aggregate information about the number and severity of low scores in a neuropsychological battery. This may be particularly important for TBI because the neuropathology and resulting neuropsychological deficits after TBI are heterogeneous.^{67,68} As well, the pathophysiological mechanisms underlying cognitive impairment may be different during the acute versus chronic phases. Information from a neuropsychological battery should be combined in a manner that does not “wash out” individual differences. In a clinical trial, prominent treatment effects on certain tests could be averaged out by relatively stable performances in other areas.

We refer to a cognition composite score created by weighting the number and severity of low performances as a Neuropsychological Deficit Score (NDS). The NDS can be derived by assigning a pre-specified weight to the norm-referenced percentile score for each test (where greater weights are assigned as observed scores deviate farther from normative expectations), summing the weights, and then dividing the sum by the number of tests in the battery. The NDS has a legacy in neuropsychology, with the Halstead Impairment Index (circa 1955), Average Impairment Rating (circa 1970), General Neuropsychological Deficit Scale (circa 1988), and Global Deficit Scale (circa 1994).^{69,70}

An NDS for TBI could extend these methods by (1) creating finer gradations, because responsiveness (sensitivity to change) is a highly desirable property of clinical trial endpoints; (2) raising the ceiling, to potentially better detect impairment in patients with milder injuries and/or high pre-morbid ability; and (3) differentially weighting deficit levels informed by multivariate base rates in healthy persons. In analyzing multivariate base rates, we have previously shown^{71–75} that (1) a substantial percentage of healthy persons obtain one or more low scores when administered a battery of neuropsychological tests; (2) the more tests that are administered, the more likely it is for a healthy person to obtain a low score; and (3) there are differences in the prevalence of low scores in healthy persons associated with level of education, race, and level of intellectual ability. Therefore, these factors need to be considered in the development of an algorithm for the NDS.

Validating a Cognition Endpoint

Extensive research is required before a cognition endpoint will be ready for clinical trials, and more broadly, for use as an outcome measure in TBI research. The three approaches to creating a cognition endpoint outlined above (OTBM, latent ability modeling, NDS) should be evaluated and compared on multiple aspects of validity, including (1) diagnostic validity, i.e., the ability of the composite scores to discriminate between patients with TBI versus with bodily injuries not involving the head (controls) throughout

the TBI severity spectrum; (2) concurrent validity, i.e., the strength of the relationships between the composite scores and other TBI outcomes such as trauma-related intracranial abnormalities on magnetic resonance imaging and patient-reported outcomes; (3) responsiveness to TBI recovery, i.e., sensitivity of the composite scores to change with time since TBI; (4) prognostic validity, i.e., the ability of early neuropsychological testing to independently predict long-term TBI outcomes.

It may be possible to further improve a cognition composite derived from any of the above approaches (OTBM, latent ability modeling, and NDS) by accounting for pre-injury individual differences. Age, sex, race, education level, and pre-morbid intelligence explain at least 40% of the variability in cognitive performance before injury.⁷⁶ Neuropsychological tests may therefore systematically over- or underestimate cognitive impairment after TBI in patients with certain demographic profiles. TBI “signal” could theoretically be better isolated by removing the “noise” variance associated with pre-morbid functioning. Demographic characteristics, however, appear to contribute to TBI outcome through multiple pathways⁷⁷ such that attempting to remove their contribution to a cognition composite might weaken, not strengthen, outcome prediction.⁶⁶ This needs to be examined carefully through applied research studies with TBI patients across the spectrum of concussion to lengthy coma.

The cognition endpoint emerging from head-to-head validity testing will ideally further demonstrate that it can capture small but important intra-individual changes. Its Minimal Clinically Important Difference could be derived through distribution-based methods (e.g., Reliable Change Index) and by anchoring to external criteria (e.g., change from one level of the GOS-E to the next among persons below the ceiling of the GOS-E).⁷⁸ Reanalyzing clinical trial results or simulating treatment effects in cohort databases could help contrast the responsiveness of a cognition composite with traditional approaches to analyzing neuropsychological outcomes (e.g., group mean comparisons on individual tests).

Finally, if not tied to a fixed battery of neuropsychological tests, the context in which the cognition endpoint will be used must be further refined. The endpoint cannot be expected to perform similarly for all possible neuropsychological test batteries. It will be important to examine the equivalence of partially overlapping and nonoverlapping test batteries, of variable size (e.g., 4 through N tests) and coverage of cognitive domains, with simulation studies using healthy normative samples as well as TBI samples. This will inform the requirements for a minimally adequate set of tests on which to derive the composite.

Much of this work can be accomplished using existing databases. Responding to the US Department of Defense Psychological Health and Traumatic Brain Injury Research Program’s January 2014 call for proposals to address the problem of insensitive outcome measurement, the TBI Endpoints Development (TED) Initiative received a \$17 million 5-year award to improve clinical trial methodology by advancing outcome measurement. The investigators have compiled the TED metadataset, which integrates data from eight TBI trials involving more than 3500 patients. More patients and studies are expected to be added.

The TED metadataset contains longitudinal demographic, clinical, biomarker, and neuropsychological data from civilian, sport, and military cohorts (<https://tbiendpoints.ucsf.edu>). The International Initiative for Traumatic Brain Injury Research⁷⁹ is also working to compile data for secondary analyses. These metadatasets will be a tremendous resource for endpoint development and validation.

Regulatory Requirements

An important goal of TBI clinical trials is to achieve market-ready therapeutics that improve patient outcomes. Qualification of a primary endpoint through the FDA’s Drug Development Tool program can facilitate FDA approval for new therapeutics. To be considered for qualification, the FDA requires that a Clinical Outcome Assessment be supported by evidence of being “well-defined and reliable” in measuring a Concept of Interest (e.g., a symptom, such as fatigue) in a specified Context of Use (COU).⁸⁰

The FDA defines COU as a comprehensive statement that fully and clearly describes the way an outcome measure is to be used and the drug development-related purpose of the use (e.g., enriching a clinical trial in mild to moderate TBI by selecting only patients with early neuropsychological impairment). The COU delineates the boundaries within which the available data adequately justify use of an outcome measure and describes important criteria regarding the circumstances under which the measure is qualified.

The FDA recognizes different types of clinical outcome measures. The cognition composites described above are performance outcomes for cognition (the construct of interest). They are derived from a series of standardized tasks performed by a patient according to instructions from a trained professional.

Like all endpoints, a cognition composite will only be valid in a defined COU. We anticipate that a cognition composite will be appropriate as a primary or secondary endpoint for studying interventions that target cognition directly (e.g., attention training) or underlying neurophysiological processes that support brain function (e.g., with a neuroprotective drug). It is not appropriate as the sole endpoint for interventions designed to promote survival, but might enhance such a trial with further understanding of the implications of those interventions on cognition.

For interventions with a known specific mechanism, a more proximal outcome would be preferable (e.g., slowed intracranial hemorrhage growth on serial CT for tranexamic acid), at least in the early stages of drug development.⁸¹ For interventions that target a narrow aspect of cognition (e.g., mental imagery for prose memory),⁸² a specific cognitive domain score may be more appropriate than a global cognition composite.

Neuropsychological tests are not designed to capture the transition from coma to minimally conscious state. An aspirational goal is to develop a cognition composite that will be suitable for patients with mild to severe TBI who are in the post-acute stage of recovery, excluding those with a persistent disorder of consciousness. Studies are needed to determine whether the cognition composite retains psychometric robustness (1) across the full range of pre-morbid ability, TBI severity, and time since injury, and (2) with variations in the specific tests that comprise it. Empirically based consensus recommendations on specific measures for TBI clinical trials can then provide guidance on the COU (i.e., what measures to use based on acuity, population, setting, and form of intervention under study).

Limitations of Cognition Endpoint

We advocate for a cognition composite score as a primary endpoint for clinical trials that target cognition directly and/or multiple neurophysiological processes upstream to cognition. A number of other outcome assessments should be developed and validated alongside, however. Most neuropsychological tests are insufficiently process-pure for mechanistic targeting.⁸³ Cognition is also influenced by medical and psychiatric comorbidities (e.g., mood disturbance, sleep disorders, and chronic pain) commonly

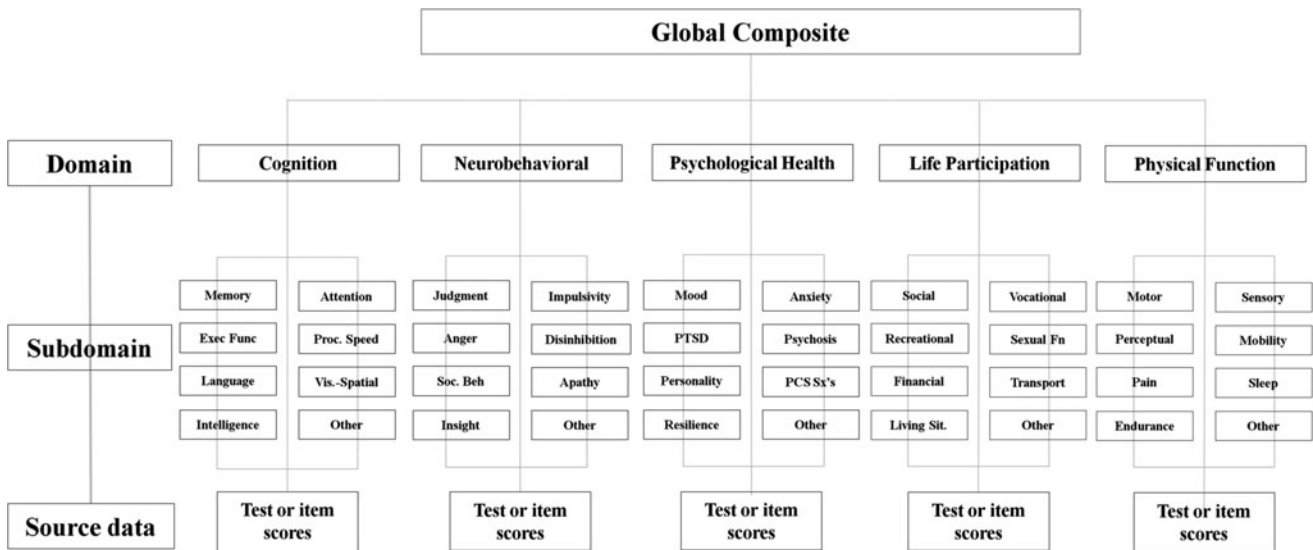


FIG. 1. Framework for multidimensional outcome assessment for traumatic brain injury clinical trials.

Note: The concept for this framework and a more detailed version of this figure were developed through a collaborative effort between the Transforming Research and Clinical Knowledge (TRACK-TBI) and TBI Endpoint Development (TED) teams. In this framework, cognition is one of five key domains that contribute to a comprehensive clinical picture of TBI outcome. Test scores or item responses from relevant subdomains are combined into domain composite scores.

Exec, Executive; Func, Functioning; Soc., Social; Beh, Behavior; Proc., Processing; Vis., Visual; PCS., Post-Concussion' Sx's, Symptoms; Sit., Situation; PTSD, Post-traumatic Stress Disorder; Sit., Situation.

encountered after TBI. Especially in the proof-of-concept and dosing phases of drug development, it will be important to have intermediate biomarkers that quantify the drug mechanisms (e.g., reducing inflammation, edema, ischemia, or oxidative stress) and link the mechanism(s) to changes in cognition.³ We expect that genotyping,⁸⁴ serum proteins,⁸⁵ neuroimaging,⁸⁶ and other biomarkers obtained soon after TBI will improve outcome prediction above and beyond neuropsychological testing, further refining risk stratification and enhancing the power of clinical trials.

A comprehensive picture of the therapeutic benefits of a new intervention requires supplementing neuropsychological testing with measures of neurobehavioral symptoms, psychological health, physical functioning, and life participation.^{45,48} Some researchers have proposed that a multidimensional endpoint that captures several or all of these domains is most appropriate for TBI clinical trials.^{7,59,87} For example, a battery of neuropsychological tests supplemented with functional, mood, and behavioral outcome measures could be analyzed using statistical methods that accommodate multiple dependent variables. Figure 1 illustrates this model.

We believe that a single pre-specified endpoint supplemented by a comprehensive battery of secondary outcomes has more advantages (summarized above) and less problematic limitations than multidimensional outcomes. The main advantages of a multidimensional endpoint are statistical efficiency and the ability to capture a range of important outcomes in a single metric. Disadvantages of multidimensional endpoints include having to make assumptions about the magnitude of the treatment effect across outcome domains, including domains that are less influenced by the treatment, obtaining stronger treatment effects for less important outcomes, susceptibility to *post hoc* “cherry-picking,” and that interpretation (explaining what the treatment effect means) can be challenging.^{58,88}

Several potential threats to the validity of a cognition endpoint are anticipated. First, a cognition endpoint may be systematically

biased in patients with diverse language and cultural backgrounds.⁸⁹ Second, interpreting neuropsychological performance as a measure of cognitive ability rests on the assumption that the examinee put forth good effort. We know that this assumption is sometimes violated, especially in TBI-related disability claim contexts but in other settings as well.^{90–92} Several relatively brief performance-based measures of response bias are now available, and could be included in a research battery.

Third, when a neuropsychological test battery is administered at baseline (pre-treatment) and again after treatment, some improvement is expected on the basis of previous exposure (e.g., via practice effects). Fortunately, this is only an issue for certain clinical trial designs (e.g., ABA) and can sometimes be mitigated by “washing out” practice effects with repeated baselines⁹³ or by using psychometrically comparable alternate forms. Finally, patients with severe TBI may be untestable at the first time point, creating “Missing Not at Random” data. Detailed completion codes may help attenuate or at least quantify this bias.^{46,94}

As mentioned above, patients who have a persistent disorder of consciousness will likely fall outside of the COU for a cognition endpoint because neuropsychological testing will be insensitive to clinically meaningful improvements in alertness and responsiveness. Alternative methods for measuring therapeutic benefit within this lower end of the ability spectrum, such as neurobehavioral rating scales, will be preferable.⁹⁵

Whereas cognitive impairment remains detectable many years after severe TBI,^{13–15,19} impairment is no longer detectable at the group level by 3 months after a mild TBI in most controlled studies.¹¹ There is a debate in the literature as to whether neuropsychological tests are inadequately sensitive to detect subtle deficits in post-acute mild TBI (i.e., the tests are too imprecise to detect the signal) or whether mild TBI does not result in lasting neurocognitive impairment (i.e., there is no signal to detect).⁹⁶ One possible explanation is that a minority of patients with mild TBI

have residual impairments and that this subgroup is masked by group comparisons.⁹⁷ Case-control studies that address selection and attrition bias are needed to further evaluate this hypothesis. An impaired subgroup argues strongly for having validated neuropsychological metrics for enriching clinical trials.

Conclusions

TBI intervention research will benefit greatly from a well-validated global cognition endpoint. Adequately measuring cognitive outcome from TBI requires a conceptually sound and psychometrically sophisticated method for integrating information from a battery of neuropsychological tests. A program of research is needed to develop, refine, and evaluate a cognition composite score for TBI in a manner that aligns with FDA regulatory requirements.

A starting point is to derive and compare candidate cognition composite scores created through both traditional and innovative methodologies. Preliminary validation studies will help identify the most promising cognition composite score and assess its readiness for TBI outcome research. Examining how the cognition composite performs across the full range of pre-morbid ability, TBI severity, and time since injury (acute, subacute, and chronic) will lead to evidence-based recommendations for how and when the cognition composite should be used as an endpoint for TBI clinical trials.

Author Disclosure Statement

GLI has been reimbursed by the government, professional scientific bodies, and commercial organizations for discussing or presenting research relating to MTBI and sport-related concussion at meetings, scientific conferences, and symposiums. He has a clinical practice in forensic neuropsychology involving individuals who have sustained mild TBIs. He has received honorariums for serving on research panels that provide scientific peer review of programs. He is a co-investigator, collaborator, or consultant on grants relating to mild TBI funded by several organizations. He has received grant funding from pharmaceutical companies to do psychometric research using neuropsychological tests. He has received research support from neuropsychological test publishing companies in the past (not in the past 5 years). For the remaining authors, no competing financial interests exist.

NDS receives salary support from a Vancouver Coastal Health Research Institute Clinician-Scientist Career Development Award. GLI acknowledges support from the INTRuST Posttraumatic Stress Disorder and Traumatic Brain Injury Clinical Consortium funded by the Department of Defense Psychological Health/Traumatic Brain Injury Research Program (X81XWH-07-CC-CSDoD), Harvard Integrated Program to Protect and Improve the Health of NFLPA Members, and the Mooney-Reed Charitable Foundation. GTM acknowledges support from NIH U01 NS086090-01, DOD USAMRAA W81XWH-13-1-0441, DOD W81XWH-14-2-0176, and One Mind. This work is related to the TBI Endpoints Development Initiative and a grant entitled Development and Validation of a Cognition Endpoint for Traumatic Brain Injury Clinical Trials.

References

- Stein, D. (2015). Embracing failure: what the Phase III progesterone studies can teach about TBI clinical trials. *Brain Inj.* 29, 1259–1272.
- Menon, D.K., and Maas, A.I. (2015). Traumatic brain injury in 2014. Progress, failures and new approaches for TBI research. *Nat. Rev. Neurol.* 11, 71–72.
- Maas, A.I., Menon, D.K., Lingsma, H.F., Pineda, J.A., Sandel, M.E., and Manley, G.T. (2012). Re-orientation of clinical research in traumatic brain injury: report of an international workshop on comparative effectiveness research. *J. Neurotrauma* 29, 32–46.
- Wright, D.W., Yeatts, S.D., Silbergleit, R., Palesch, Y.Y., Hertzberg, V.S., Frankel, M., Goldstein, F.C., Caveney, A.F., Howlett-Smith, H., Bengelink, E.M., Manley, G.T., Merck, L.H., Janis, L.S., and Barsan, W.G. (2014). Very early administration of progesterone for acute traumatic brain injury. *N. Engl. J. Med.* 371, 2457–2466.
- 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.
- Lu, J., Gary, K.W., Neimeier, J.P., Ward, J., and Lapane, K.L. (2012). Randomized controlled trials in adult traumatic brain injury. *Brain Inj.* 26, 1523–1548.
- 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.
- TBI Endpoints Development Award. Available at: http://cdmnp.army.mil/funding/pa/13phtbited_pa.pdf. Accessed: June 14, 2014.
- Christensen, B.K., Colella, B., Inness, E., Hebert, D., Monette, G., Bayley, M., and Green, R.E. (2008). Recovery of cognitive function after traumatic brain injury: a multilevel modeling analysis of Canadian outcomes. *Arch. Phys. Med. Rehabil.* 89, Suppl 12, S3–S15.
- Schretlen, D.J., and Shapiro, A.M. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *Int. Rev. Psychiatry* 15, 341–349.
- Karr, J.E., Areshenkoff, C.N., and Garcia-Barrera, M. (2014). The neuropsychological outcomes of concussion: a systematic review of meta-analyses on the cognitive sequelae of mild traumatic brain injury. *Neuropsychology* 28, 321–336.
- McCrea, M., Iverson, G.L., McAllister, T.W., Hammeke, T.A., Powell, M.R., Barr, W.B., and Kelly, J.P. (2009). An integrated review of recovery after mild traumatic brain injury (MTBI): implications for clinical management. *Clin. Neuropsychol.* 23, 1368–1390.
- Ruttan, L., Martin, K., Liu, A., Colella, B., and Green, R.E. (2008). Long-term cognitive outcome in moderate to severe traumatic brain injury: a meta-analysis examining timed and untimed tests at 1 and 4.5 or more years after injury. *Arch. Phys. Med. Rehabil.* 89, Suppl 12, S69–S76.
- Draper, K., and Ponsford, J. (2008). Cognitive functioning ten years following traumatic brain injury and rehabilitation. *Neuropsychology* 22, 618–625.
- Levin, H.S., Grossman, R.G., Rose, J.E., and Teasdale, G. (1979). Long-term neuropsychological outcome of closed head injury. *J. Neurosurg.* 50, 412–422.
- Green, R.E., Colella, B., Christensen, B., Johns, K., Frasca, D., Bayley, M., and Monette, G. (2008). Examining moderators of cognitive recovery trajectories after moderate to severe traumatic brain injury. *Arch. Phys. Med. Rehabil.* 89, Suppl 12, S16–S24.
- Dougan, B.K., Horswill, M.S., and Geffen, G.M. (2014). Athletes' age, sex, and years of education moderate the acute neuropsychological impact of sports-related concussion: a meta-analysis. *J. Int. Neuropsychol. Soc.* 20, 64–80.
- Dikmen, S., McLean, A., and Temkin, N. (1986). Neuropsychological and psychosocial consequences of minor head injury. *J. Neurol. Neurosurg. Psychiatry* 49, 1227–1232.
- Millis, S.R., Rosenthal, M., Novack, T.A., Sherer, M., Nick, T.G., Kreutzer, J.S., High, W.M. Jr., and Ricker, J.H. (2001). Long-term neuropsychological outcome after traumatic brain injury. *J. Head Trauma Rehabil.* 16, 343–355.
- Hanks, R.A., Millis, S.R., Ricker, J.H., Giacino, J.T., Nakese-Richardson, R., Frol, A.B., Novack, T.A., Kalmar, K., Sherer, M., and Gordon, W.A. (2008). The predictive validity of a brief inpatient neuropsychologic battery for persons with traumatic brain injury. *Arch. Phys. Med. Rehabil.* 89, 950–957.
- Rassovsky, Y., Satz, P., Alfano, M.S., Light, R.K., Zaucha, K., McArthur, D.L., and Hovda, D. (2006). Functional outcome in TBI I: neuropsychological, emotional, and behavioral mediators. *J. Clin. Exp. Neuropsychol.* 28, 567–580.
- Green, R.E., Colella, B., Hebert, D.A., Bayley, M., Kang, H.S., Till, C., and Monette, G. (2008). Prediction of return to productivity after severe traumatic brain injury: investigations of optimal neu-

- ropsychological tests and timing of assessment. *Arch. Phys. Med. Rehabil.* 89, Suppl 12, S51–S60.
23. Chaytor, N., Temkin, N., Machamer, J., and Dikmen, S. (2007). The ecological validity of neuropsychological assessment and the role of depressive symptoms in moderate to severe traumatic brain injury. *J. Int. Neuropsychol. Soc.* 13, 377–385.
 24. Hanks, R.A., Rapport, L.J., Millis, S.R., and Deshpande, S.A. (1999). Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. *Arch. Phys. Med. Rehabil.* 80, 1030–1037.
 25. Ross, S.R., Millis, S.R., and Rosenthal, M. (1997). Neuropsychological prediction of psychosocial outcome after traumatic brain injury. *Appl. Neuropsychol.* 4, 165–170.
 26. Zafonte, R.D., Bagiella, E., Ansel, B.M., Novack, T.A., Friedewald, W.T., Hesdorffer, D.C., Timmons, S.D., Jallo, J., Eisenberg, H., Hart, T., Ricker, J.H., Diaz-Arrastia, R., Merchant, R.E., Temkin, N.R., Melton, S., and Dikmen, S.S. (2012). Effect of citicoline on functional and cognitive status among patients with traumatic brain injury: Citicoline Brain Injury Treatment Trial (COBRIT). *JAMA* 308, 1993–2000.
 27. Salazar, A.M., Warden, D.L., Schwab, K., Spector, J., Braverman, S., Walter, J., Cole, R., Rosner, M.M., Martin, E.M., Ecklund, J., and Ellenbogen, R.G. (2000). Cognitive rehabilitation for traumatic brain injury: a randomized trial. *Defense and Veterans Head Injury Program (DVHIP) Study Group. JAMA* 283, 3075–3081.
 28. Strauss, E.H., Sherman, E.M.S., and Spreen, O. (2006). *A Compendium of Neuropsychological Tests*, 3rd ed. Oxford University Press: Oxford, pp. 1216.
 29. Rohling, M.L., Meyers, J.E., and Millis, S.R. (2003). Neuropsychological impairment following traumatic brain injury: a dose-response analysis. *Clin. Neuropsychol.* 17, 289–302.
 30. Bazarian, J.J., Wong, T., Harris, M., Leahy, N., Mookerjee, S., and Dombrov, M. (1999). Epidemiology and predictors of post-concussive syndrome after minor head injury in an emergency population. *Brain Inj.* 13, 173–189.
 31. McCreary, M., Guskiewicz, K.M., Marshall, S.W., Barr, W., Randolph, C., Cantu, R.C., Onate, J.A., Yang, J., and Kelly, J.P. (2003). Acute effects and recovery time following concussion in collegiate football players: The NCAA Concussion Study. *JAMA* 290, 2556–2563.
 32. Whyte, J. (2014). Contributions of treatment theory and enablement theory to rehabilitation research and practice. *Arch. Phys. Med. Rehabil.* 95, Suppl 1, S17–S23.e2.
 33. Fujimoto, S.T., Longhi, L., Saatman, K.E., Conte, V., Stocchetti, N., and McIntosh, T.K. (2004). Motor and cognitive function evaluation following experimental traumatic brain injury. *Neurosci. Biobehav. Rev.* 28, 365–378.
 34. Wheaton, P., Mathias, J.L., and Vink, R. (2011). Impact of pharmacological treatments on outcome in adult rodents after traumatic brain injury: a meta-analysis. *J. Psychopharmacol.* 25, 1581–1599.
 35. Skolnick, B.E., Maas, A.I., Narayan, R.K., van der Hoop, R.G., MacAllister, T., Ward, J.D., Nelson, N.R., and Stocchetti, N.; SYNAPSE Trial Investigators. (2014). A clinical trial of progesterone for severe traumatic brain injury. *N. Engl. J. Med.* 371, 2467–2476.
 36. Perel, P., Edwards, P., Wentz, R., and Roberts, I. (2006). Systematic review of prognostic models in traumatic brain injury. *BMC Med. Inform. Decis. Mak.* 6, 38.
 37. Lingsma, H.F., Yue, J.K., Maas, A.I., Steyerberg, E.W., Manley, G.T.; TRACK-TBI Investigators. (2015). Outcome prediction after mild and complicated mild traumatic brain injury: external validation of existing models and identification of new predictors using the TRACK-TBI pilot study. *J. Neurotrauma* 32, 83–94.
 38. Sherer, M., Novack, T.A., Sander, A.M., Struchen, M.A., Alderson, A., and Thompson, R.N. (2002). Neuropsychological assessment and employment outcome after traumatic brain injury: a review. *Clin Neuropsychol.* 16, 157–178.
 39. Silverberg, N.D., Gardner, A., Brubacher, J.R., Panenka, W., Li, J.J., and Iverson, G.L. (2015). Systematic review of multivariable prognostic models for mild traumatic brain injury. *J. Neurotrauma* 32, 517–526.
 40. Hernandez, A.V., Steyerberg, E.W., Taylor, G.S., Marmarou, A., Habbema, J.D., and Maas, A.I. (2005). Subgroup analysis and covariate adjustment in randomized clinical trials of traumatic brain injury: A systematic review. *Neurosurgery* 57, 1244–1253.
 41. Walker, K.R., and Tesco, G. (2013). Molecular mechanisms of cognitive dysfunction following traumatic brain injury. *Front. Aging Neurosci.* 5, 29.
 42. Margulies, S., and Hicks, R. (2009). Combination therapies for traumatic brain injury: prospective considerations. *J. Neurotrauma* 26, 925–939.
 43. Boussi-Gross, R., Golan, H., Fishlev, G., Bechor, Y., Volkov, O., Bergan, J., Friedman, M., Hoofien, D., Shlamkovich, N., Ben-Jacob, E., and Efrati, S. (2013). Hyperbaric oxygen therapy can improve post concussion syndrome years after mild traumatic brain injury—randomized prospective trial. *PLoS One* 8, e79995.
 44. Cicerone, K.D., and Giacino, J. (1992). Remediation of executive function deficits after traumatic brain injury. *NeuroRehabilitation* 2, 12–22.
 45. Wilde, E., Whiteneck, G.G., Bogner, J., Bushnik, T., Cifu, D.X., Dikmen, S., French, L., Giacino, J.T., Hart, T., Malec, J.F., Millis, S.R., Novack, T.A., Sherer, M., Tulsy, D.S., Vanderploeg, R.D., and von Steinbuechel, N. (2010). Recommendations for the use of common outcome measures in traumatic brain injury research. *Arch. Phys. Med. Rehabil.* 91, 1650–1660.e17.
 46. Bagiella, E., Novack, T.A., Ansel, 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.
 47. Clifton, G.L., Hayes, R.L., Levin, H.S., Michel, M.E., and Choi, S.C. (1992). Outcome measures for clinical trials involving traumatically brain-injured patients: report of a conference. *Neurosurgery* 31, 975–978.
 48. Hannay, H.J., Ezrachi, O., Contant, C.F., and Levin, H.S. (1996). Outcome measures for patients with head injuries: report of the Outcome Measures Subcommittee. *J. Head Trauma Rehabil.* 11, 41–50.
 49. Hicks, R., Giacino, J., Harrison-Felix, C., Manley, G., Valadka, A., and Wilde, E.A. (2013). Progress in developing common data elements for traumatic brain injury research: version two – the end of the beginning. *J. Neurotrauma* 30, 1852–1861.
 50. Gershon, R.C., Wagster, M.V., Hendrie, H.C., Fox, N.A., Cook, K.F., and Nowinski, C.J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology* 80, Suppl 3, S2–S6.
 51. Weintraub, S., Dikmen, S.S., Heaton, R.K., Tulsky, D.S., Zelazo, P.D., Bauer, P.J., Carlozzi, N.E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N.A., Beaumont, J.L., Mungas, D., Nowinski, C.J., Richler, J., Deocampo, J.A., Anderson, J.E., Manly, J.J., Borosh, B., Havlik, R., Conway, K., Edwards, E., Freund, L., King, J.W., Moy, C., Witt, E., and Gershon, R.C. (2013). Cognition assessment using the NIH Toolbox. *Neurology* 80, Suppl 3, S54–S64.
 52. Rohling, M.L., Miller, R.M., Axelrod, B.N., Wall, J.R., Lee, A.J., and Kinikini, D.T. (2015). Is co-norming required? *Arch. Clin. Neuropsychol.* 30, 611–633.
 53. High, W.M. Jr., Briones-Galang, M., Clark, J.A., Gilkison, C., Mossberg, K.A., Zgaljardic, D.J., Masel, B.E., and Urban, R.J. (2010). Effect of growth hormone replacement therapy on cognition after traumatic brain injury. *J. Neurotrauma* 27, 1565–1575.
 54. Moser, R.S., Glatts, C., and Schatz, P. (2012). Efficacy of immediate and delayed cognitive and physical rest for treatment of sports-related concussion. *J. Pediatr.* 161, 922–926.
 55. Vacha-Haase, T., and Thompson, B. (2004). How to estimate and interpret various effect sizes. *J. Couns. Psychol.* 51, 473–481.
 56. Dikmen, S.S., Machamer, J.E., Winn, H.R., Anderson, G.D., and Temkin, N.R. (2000). Neuropsychological effects of valproate in traumatic brain injury: a randomized trial. *Neurology* 54, 895–902.
 57. U.S. Food and Drug Administration. Section 314.125 - Adequate and Well-Controlled Studies. <https://www.gpo.gov/fdsys/granule/CFR-2010-title21-vol5/CFR-2010-title21-vol5-sec314-126>.
 58. Cordoba, G., Schwartz, L., Woloshin, S., Bae, H., and Gøtzsche, P.C. (2010). Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ* 341, c3920.
 59. Temkin, N.R., Anderson, G.D., Winn, H.R., Ellenbogen, R.G., Britz, G.W., Schuster, J., Lucas, T., Newell, D.W., Mansfield, P.N., Machamer, J.E., Barber, J., and Dikmen, S.S. (2007). Magnesium sulfate for neuroprotection after traumatic brain injury: a randomised controlled trial. *Lancet Neurol.* 6, 29–38.
 60. Miller, L.S., and Rohling, M.L. (2001). A statistical interpretive method for neuropsychological test data. *Neuropsychol. Rev.* 11, 143–169.
 61. Reeve, B.B., Hays, R.D., Bjorner, J.B., Cook, K.F., Crane, P.K., Teresi, J.A., Thissen, D., Revicki, D.A., Weiss, D.J., Hambleton, R.K., Liu, H., Gershon, R., Reise, S.P., Lai, J.S., and Cella, D.; PROMIS Cooperative Group. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med. Care* 45, Suppl 1, S22–S31.
 62. Reise, S.P., Morizot, J., and Hays, R.D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Qual. Life Res.* 16, Suppl 1, 19–31.

63. Crane, P.K., Carle, A., Gibbons, L.E., Insel, P., Mackin, R.S., Gross, A., Jones, R.N., Mukherjee, S., Curtis, S.M., Harvey, D., Weiner, M., and Mungas, D. (2012). Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging Behav.* 6, 502–516.
64. Gibbons, L.E., Carle, A.C., Mackin, R.S., Harvey, D., Mukherjee, S., Insel, P., Curtis, S.M., Mungas, D., and Crane, P.K. (2012). A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging Behav.* 6, 517–527.
65. Crane, P.K., Narasimhalu, K., Gibbons, L.E., Mungas, D.M., Haneuse, S., Larson, E.B., Kuller, L., Hall, K., and van Belle, G. (2008). Item response theory facilitated cocalibrating cognitive tests and reduced bias in estimated rates of decline. *J. Clin. Epidemiol.* 61, 1018–1027.
66. Crane, P.K., Narasimhalu, K., Gibbons, L.E., Pedraza, O., Mehta, K.M., Tang, Y., Manly, J.J., Reed, B.R., and Mungas, D.M. (2008). Composite scores for executive function items: demographic heterogeneity and relationships with quantitative magnetic resonance imaging. *J. Int. Neuropsychol. Soc.* 14, 746–759.
67. Bigler, E.D., Abildskov, T.J., Petrie, J., Farrer, T.J., Dennis, M., Simic, N., Taylor, H.G., Rubin, K.H., Vannatta, K., Gerhardt, C.A., Stancin, T., and Owen Yeates, K. (2013). Heterogeneity of brain lesions in pediatric traumatic brain injury. *Neuropsychology* 27, 438–451.
68. Rosenbaum, S.B., and Lipton, M.L. (2012). Embracing chaos: the scope and importance of clinical and pathological heterogeneity in mTBI. *Brain Imaging Behav.* 6, 255–282.
69. Carey, C.L., Woods, S.P., Gonzalez, R., Conover, E., Marcotte, T.D., Grant, I., and Heaton, R.K. (2004). Predictive validity of global deficit scores in detecting neuropsychological impairment in HIV infection. *J. Clin. Exp. Neuropsychol.* 26, 307–319.
70. Russell, E. (2011). *The Scientific Foundation of Neuropsychological Assessment*. Elsevier: London.
71. Binder, L.M., Iverson, G.L., and Brooks, B.L. (2009). To err is human: "abnormal" neuropsychological scores and variability are common in healthy adults. *Arch. Clin. Neuropsychol.* 24, 31–46.
72. Brooks, B.L., and Iverson, G.L. (2009). Comparing actual to estimated base rates of "abnormal" scores on neuropsychological test batteries: implications for interpretation. *Arch. Clin. Neuropsychol.* 25, 14–21.
73. Iverson, G.L., Brooks, B.L., and Young, A.H. (2009). Identifying neurocognitive impairment in depression using computerized testing. *Appl. Neuropsychol.* 16, 254–261.
74. Brooks, B.L., Holdnack, J.A., and Iverson, G.L. (2011). Advanced clinical interpretation of the WAIS-IV and WMS-IV: prevalence of low scores varies by level of intelligence and years of education. *Assessment* 18, 156–167.
75. Ivins, B.J., Lange, R.T., Cole, W.R., Kane, R., Schwab, K.A., and Iverson, G.L. (2014). Using base rates of low scores to interpret the ANAM4 TBI-MIL battery following mild traumatic brain injury. *Arch. Clin. Neuropsychol.* 30, 26–38.
76. Testa, S.M., Winicki, J.M., Pearlson, G.D., Gordon, B., and Schretlen, D.J. (2009). Accounting for estimated IQ in neuropsychological test performance with regression-based techniques. *J. Int. Neuropsychol. Soc.* 15, 1012–1022.
77. Novack, T.A., Bush, B.A., Meythaler, J.M., and Canupp, K. (2001). Outcome after traumatic brain injury: oathway analysis of contributions from premorbid, injury severity, and recovery variables. *Arch. Phys. Med. Rehabil.* 82, 300–305.
78. Revicki, D., Hays, R.D., Cella, D., and Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J. Clin. Epidemiol.* 61, 102–109.
79. International Initiative for Traumatic Brain Injury Research. Available at: http://ec.europa.eu/research/health/medical-research/brain-research/index_en.html. Accessed: October 6, 2015.
80. U.S. Food and Drug Administration. Clinical Outcome Assessment (COA): Glossary of Terms. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm#COU>. Accessed January 26, 2016.
81. Perel, P., Al-Shahi Salman, R., Kawahara, T., Morris, Z., Prieto-Merino, D., Roberts, I., Sandercock, P., Shakur, H., and Wardlaw, J. (2012). CRASH-2 (Clinical Randomisation of an Antifibrinolytic in Significant Haemorrhage) intracranial bleeding study: the effect of tranexamic acid in traumatic brain injury—a nested randomised, placebo-controlled trial. *Health Technol. Assess.* 16, 1–54.
82. Chiaravalloti, N.D., Sandry, J., Moore, N.B., and DeLuca, J. (2015). An RCT to treat learning impairment in traumatic brain injury: the TBI-MEM trial. *Neurorehabil. Neural. Repair* Epub ahead of print.
83. Carter, C.S., and Barch, D.M. (2007). Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: the CNTRICS initiative. *Schizophr. Bull.* 33, 1131–1137.
84. Yue, J.K., Pronger A.M., Ferguson, A.R., Temkin, N.R., Sharma, S., Rosand, J., Sorani, M. D., McAllister, T.W., Barber, J., Winkler, E.A., Burchard, E.G., Hu, D., Lingsma, H.F., Cooper, S.R., Puccio, A.M., Okonkwo, D.O., Diaz-Arrastia, R., and Manley, G.T. (2015). Association of a common genetic variant within ANKK1 with six-month cognitive performance after traumatic brain injury. *Neurogenetics* 16, 169–180.
85. Okonkwo, D.O., Yue, J.K., Puccio, A.M., Panczykowski, D.M., Inoue, T., McMahon, P.J., Sorani, M.D., Yuh, E.L., Lingsma, H.F., Maas, A.I., Valadka, A.B., Manley, G.T., and the Transforming Research and Clinical Knowledge in Traumatic Brain Injury Investigators. (2013). GFAP-BDP as an acute diagnostic marker in traumatic brain injury: results from the prospective transforming research and clinical knowledge in traumatic brain injury study. *J. Neurotrauma* 30, 1490–1497.
86. Yuh, E.L., Mukherjee, P., Lingsma, H.F., Yue, J.K., Ferguson, A.R., Gordon, W., Valadka, A.B., Schnyer, D.M., Okonkwo, D.O., Maas, A.I., and Manley, G.T. (2013). Magnetic resonance imaging improves 3-month outcome prediction in mild traumatic brain injury. *Ann. Neurol.* 73, 224–235.
87. Poon, W., Vos, P., Muresanu, D., Vester, J., von Wild, K., Hömberg, V., Wang, E., Lee, T.M., and Matula, C. (2015). Cerebrolysin Asian Pacific trial in acute brain injury and neurorecovery: design and methods. *J. Neurotrauma* 32, 571–580.
88. 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.
89. Manly, J.J. (2008). Critical issues in cultural neuropsychology: profit from diversity. *Neuropsychol. Rev.* 18, 179–183.
90. Larrabee, G.J. (2012). Performance validity and symptom validity in neuropsychological assessment. *J. Int. Neuropsychol. Soc.* 18, 625–630.
91. DeRight, J., and Jorgensen, R.S. (2015). I just want my research credit: frequency of suboptimal effort in a non-clinical healthy undergraduate sample. *Clin. Neuropsychol.* 29, 101–117.
92. Donders, J., and Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Inj.* 21, 319–326.
93. Duff, K., Westervelt, H.J., McCaffrey, R.J., and Haase R.F. (2001). Practice effects, test-retest stability, and dual baseline assessments with the California Verbal Learning Test in an HIV sample. *Arch. Clin. Neuropsychol.* 16, 461–476.
94. Kalmar, K., Novack, T.A., Nakase-Richardson, R., Sherer, M., Frol, A.B., Gordon, W.A., Hanks, R.A., Giacino, J.T., and Ricker, J.H. (2008). Feasibility of a brief neuropsychologic test battery during acute inpatient rehabilitation after traumatic brain injury. *Arch. Phys. Med. Rehabil.* 89, 942–949.
95. Giacino, J.T., Fins, J.J., Laureys, S., and Schiff, N.D. (2014). Disorders of consciousness after acquired brain injury: the state of the science. *Nat. Rev. Neurol.* 10, 99–114.
96. Larrabee, G.J., Binder, L.M., Rohling, M.L., and Ploetz, D.M. (2013). Meta-analytic methods and the importance of non-TBI factors related to outcome in mild traumatic brain injury: response to Bigler et al. (2013). *Clin. Neuropsychol.* 27, 215–237.
97. Iverson, G.L. (2010). Mild traumatic brain injury meta-analyses can obscure individual differences. *Brain Inj.* 24, 1246–1255.

Address correspondence to:

Noah D. Silverberg, PhD
 Rehabilitation Research Program
 4255 Laurel Street
 Vancouver
 British Columbia, V5Z 2G9
 Canada

E-mail: noah.silverberg@vch.ca