SPECIAL COMMUNICATION

Recommendations for the Use of Common Outcome Measures in Traumatic Brain Injury Research

Elisabeth A. Wilde, PhD, Gale G. Whiteneck, PhD, Jennifer Bogner, PhD, Tamara Bushnik, PhD, David X. Cifu, MD, Sureyya Dikmen, PhD, Louis French, PsyD, Joseph T. Giacino, PhD, Tessa Hart, PhD, James F. Malec, PhD, Scott R. Millis, PhD, Thomas A. Novack, PhD, Mark Sherer, PhD, David S. Tulsky, PhD, Rodney D. Vanderploeg, PhD, Nicole von Steinbuechel, PhD

ABSTRACT. Wilde EA, Whiteneck GG, Bogner J, Bushnik T, Cifu DX, Dikmen S, French L, Giacino JT, Hart T, Malec JF, Millis SR, Novack TA, Sherer M, Tulsky DS, Vanderploeg RD, von Steinbuechel N. Recommendations for the use of common outcome measures in traumatic brain injury research. Arch Phys Med Rehabil 2010;91:1650-60.

This article summarizes the selection of outcome measures by the interagency Traumatic Brain Injury (TBI) Outcomes Workgroup to address primary clinical research objectives, including documentation of the natural course of recovery from TBI, prediction of later outcome, measurement of treatment effects, and comparison of outcomes across studies. Consistent with other Common Data Elements Workgroups, the TBI Outcomes Workgroup adopted the standard 3-tier system in its selection of measures. In the first tier, core measures included valid, robust, and widely applicable outcome measures with proven utility in TBI from each identified domain, including global level of function, neuropsychological impairment, psychological status, TBI-related symptoms, executive functions, cognitive and physical activity limitations, social role participation, and perceived health-related quality of life. In the second tier, supplemental measures were recommended for consideration in TBI research focusing on specific topics or populations. In the third tier, emerging measures included important instruments currently under development, in the process of validation, or nearing the point of published findings that have significant potential to be superior to some older ("legacy") measures in the core and supplemental lists and may eventually replace them as evidence for their utility emerges.

Key Words: Outcome assessment; health care; Brain injuries; Neurobehavioral manifestations; Research; Rehabilitation.

THE PURPOSE OF THE common data elements traumatic brain injury Outcomes Workgroup was to address the need for a common set of outcome measures for TBI research across agencies and populations, as outlined in Thurmond et al[1] (see p. 1633-6, this issue). The work group was composed of physicians, psychologists, neuropsychologists, and others with expertise in TBI outcomes research. Many work group members also had previous collaborative experience in large multicenter TBI research projects, such as the NIH Clinical Trials Network and the National Institute on Disability and Rehabilitation Research TBI Model Systems program, as well as specific TBI-related clinical trials and multicenter studies.

SELECTION OF TBI OUTCOME DOMAINS AND MEASURES

The work group considered several factors in selecting outcome domains that should be assessed after TBI. First, we wanted to cover outcomes at multiple levels of the International Classification of Functioning, Disability, and Health; in other words, function, activity, and participation.[2] Second, we targeted outcome domains previously shown to be affected by TBI and of importance to consumers, scientists, and practitioners. Third, we sought a set of measures that collectively would cover the continua from acute to long-term outcomes and from mild to severe TBI. Thus, the work group examined measures of global outcome; recovery of consciousness; neuropsychological impairment; psychological status; TBI-related symptoms; performance of activities loading on behavioral, cognitive, and physical demands; social role participation; and perceived health-related quality of life, as well as health economic measures. Additionally, a multidimensional domain of patient-reported outcomes was identified as a promising

From the Departments of Physical Medicine and Rehabilitation (Wilde, Sherer), Neurology (Wilde), and Radiology (Wilde), Baylor College of Medicine, Houston and TIRR Memorial Hermann, Houston (Sherer); Michael E. DeBakey Veterans' Administration Medical Center, Houston (Wilde); University of Texas Medical School at Houston, Houston, TX (Sherer); Craig Hospital, Englewood, CO (Whiteneck); Department of Physical Medicine and Rehabilitation, Ohio State University, Columbus, OH (Bogner); Department of Rehabilitation Medicine, Rusk Institute for Rehabilitation, New York, NY (Bushnik); Department of Physical Medicine and Rehabilitation, Virginia Commonwealth University, PM&R Service, Hunter Holmes McGuire Veterans Administration Medical Center, Richmond, VA (Cifu); Department of Rehabilitation Medicine, University of Washington, Seattle, WA (Dikmen); Department of Orthopaedics and Rehabilitation, Walter Reed Army Medical Center, Washington DC (French); JFK Johnson Rehabilitation Institute, Edison, NJ (Giacino); Spaulding Rehabilitation Hospital/Harvard Medical School, Boston, MA (Giacino); Moss Rehabilitation Research Institute, Elkins Park, PA (Hart); Department of Physical Medicine and Rehabilitation, Indiana University School of Medicine, Indianapolis (Malec); Rehabilitation Hospital of Indiana, Indianapolis, IN (Malec); Departments of Physical Medicine and Rehabilitation and Emergency Medicine, Wayne State University School of Medicine, Detroit, MI (Millis); Department of Physical Medicine and Rehabilitation, University of Alabama at Birmingham, AL (Novack); Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI (Tulsky); Psychology Service, James A. Haley Veterans' Hospital, Tampa (Vanderploeg); Departments of Psychology and Psychiatry, University of South Florida, Tampa, FL (Vanderploeg); and Department of Medical Psychology and Medical Sociology, University Medical Center Göttingen Georg-August-University, Göttingen, Germany (von Steinbuechel).

Supported by the National Institutes of Health (NIH; National Institute of Neurological Disorders and Stroke), U.S. Department of Veterans Affairs (VA), U.S. Department of Defense (DoD), and U.S. Department of Education/National Institute on Disability and Rehabilitation Research.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Gale G. Whiteneck, PhD, is Chair of the Traumatic Brain Injury Outcomes Workgroup.

Views expressed are those of the authors and do not necessarily reflect those of the agencies or institutions with which they are affiliated, including the U.S. Department of VA, the U.S. DoD, the U.S. Department of Health and Human Services, the NIH, the National Institute of Mental Health, the U.S. Department of Education, or the Uniformed Services University of the Health Sciences. This work is not an official document, guidance, or policy of the U.S. Government, nor should any official endorsement be inferred.

Correspondence to Elisabeth A. Wilde, PhD, Baylor College of Medicine, 1709 Dryden Rd, Ste 1200, Houston, TX 77030, e-mail: *ewilde@bcm.edu.* Reprints are not available from the author.

0003-9993/10/9111-00369$36.00/0
doi:10.1016/j.apmr.2010.06.033

List of Abbreviations

ADL	activity of daily living
ASSIST	Alcohol, Smoking, and Substance Use Involvement Screening Test
AUDIT	Alcohol Use Disorders Identification Test
BSI-18	Brief Symptom Inventory 18
BVMT-R	Brief Visuospatial Memory Test–Revised
CDE	common data element
CHART	Craig Handicap Assessment and Reporting Technique
CHART-SF	Craig Handicap Assessment and Reporting Technique Short Form
Cog-FIM	FIM Cognition Subscale
COWAT	Controlled Oral Word Association Test
CRS-R	JFK Coma Recovery Scale–Revised
CWIT	Color-Word Interference Test
DRS	Disability Rating Scale
FAD	Family Assessment Device
FrSBe	Frontal Systems Behavior Scale
GOS	Glasgow Outcome Scale
GOS-E	Glasgow Outcome Scale (Extended)
GPT	Grooved Pegboard Test
MCS	minimally conscious state
MPAI-4	Mayo-Portland Adaptability Inventory
MMPI-2	Minnesota Multiphasic Personality Inventory 2
MMPI-2-RF	Minnesota Multiphasic Personality Inventory 2, Restructured Form
NIH	National Institutes of Health
NINDS	National Institute on Neurological Disorders and Stroke
NOS-TBI	Neurological Outcome Scale for Traumatic Brain Injury
NSI	Neurobehavioral Symptom Inventory
PART	Participation Assessment With Recombined Tools
PCL-C	Posttraumatic Stress Disorder Check List–Civilian Version
PCL-M	Posttraumatic Stress Disorder Check List–Military Version
PCL-S	Posttraumatic Stress Disorder Check List–Stressor Specific Version
PI	Participation Index
PROMIS	Patient-Reported Outcomes Measurement Information System
PTSD	posttraumatic stress disorder
QOLIBRI	Quality of Life After Brain Injury
RAVLT	Rey Auditory Verbal Learning Test
RPQ	Rivermead Post Concussion Symptom Questionnaire
SF-12	Medical Outcomes Study 12-Item Short Form Health Survey
SF-36	Medical Outcomes Study 36-Item Short Form Health Survey
SWLS	Satisfaction With Life Scale
TBI	traumatic brain injury
TBI-QOL	Traumatic Brain Injury–Quality of Life
TMT	Trail Making Test
VA	Veterans Affairs
WAIS-III	Wechsler Adult Intelligence Scale, Third Edition
WAIS-IV	Wechsler Adult Intelligence Scale, Fourth Edition
WHO	World Health Organization
WRAT-4	Wide Range Achievement Test, Fourth Edition

area represented by outcome measures currently in development. These domains are described further in table 1.

Factors of Importance in Selecting Outcome Measures Within the Domains

Within each domain, measures were selected to maximize the ability of clinical researchers to (1) document the natural course of recovery after TBI, (2) enhance the prediction of later outcome, (3) measure the effects of treatment, and (4) facilitate comparisons across studies.

The work group divided into smaller subgroups based on interests and expertise to develop lists of names and detailed characteristics of potential measures for each domain. Measures were identified using the following criteria: (1) sufficient representation in the scientific literature and/or widespread use in the TBI clinical and research community in diagnosis, outcome measurement and prediction, or treatment effectiveness; (2) evidence of sound psychometric properties, including (when applicable) construct validity, internal consistency, sensitivity to change, test-retest reliability, intra-/interrater agreement (including subject/proxy and telephone/in-person administration); (3) well-established normative data; (4) applicability across a range of injury severity and functional levels; (5) availability in the public domain; (6) ease of administration; and (7) brevity. The panel also considered factors that would render the measures appropriate for international use, such as availability in different languages and validation in different ethnic groups. For measures of health-related quality of life, activity/participation, and psychological function, consideration also was given to flexibility of formats; for example, telephone interview versus in-person administration or self versus proxy respondent. Finally, for objective neuropsychological measures, the availability of alternate forms to moderate the potential impact of practice effects was considered. The work group considered measures that could be applied in both adult and pediatric populations, but recommended that a separate work group be convened to address pediatric outcome recommendations.

Distinguishing Core, Supplemental, and Emerging Outcome Measure Recommendations

In accordance with other CDE Workgroups, 3 tiers of CDE were recommended: core, supplemental, and emerging (see Thurmond et al,[1] p. 1633-6, this issue). First, well-established core measures covering outcome domains relevant to most TBI studies were included. A listing of 9 core measures was selected, with the idea that most could be applied across large TBI studies either as a comprehensive battery or in addition to other outcome measures selected by the investigator. Use of these measures should be tempered by the objectives, study design, and target population. In the second tier, additional supplemental measures were recommended for consideration in TBI research focusing on more specific topics or populations. For example, a study in which neuropsychological outcome is of particular interest may draw on measures from the supplemental list that target cognitive functions not tapped by the core. In the third tier, emerging measures include important instruments currently under development, in the process of validation, or nearing the point of published findings that have significant potential to be superior to some older ("legacy") measures currently in the core and supplemental lists.

General Process for Selecting CDEs

Each member of the panel selected 1 or 2 outcome domains based on his/her interests and expertise or was assigned a

Table 1: Outcome Domains and Descriptions

Domain Name	Domain Description and Relevance in TBI
Global outcome	Global outcome measures summarize the overall impact of TBI, incorporating functional status, independence, and role participation
Recovery of consciousness	Duration of coma, level of consciousness, and rate of recovery contribute significantly to functional outcome and have a key role in treatment and disposition planning
Neuropsychological impairment	Objective measures of neuropsychological functions, such as attention, memory, and executive function, are very sensitive to effects of TBI and often affect everyday activities and social role participation
Psychological status	Psychological issues associated with TBI that affect outcomes include adjustment problems, personality changes (eg, impulsivity), or mood disturbances. In addition, substance use disorders are prevalent in persons with TBI and can have a substantial impact on long-term outcomes
TBI-related symptoms	TBI-related symptoms include somatic (eg, headaches, visual disturbances), cognitive (eg, attention and memory difficulties), and emotional (eg, irritability) symptoms. They commonly are reported after TBI or concussion and may persist in some cases at all levels of TBI severity
Behavioral function	Behavioral dysfunction commonly is reported after TBI and may contribute to difficulties in return to work/school, personal relationships, and social functioning. Common examples are aggression and childlike behavior
Cognitive activity limitations	Cognitive activity measures describe the impact of neuropsychological impairments on cognitively loaded real-world tasks, such as instrumental ADLs, functional communication, and health and safety-related behaviors
Physical function	People with TBI (particularly severe TBI) may manifest difficulties in physical or neurologic functioning, including cranial or peripheral nerve damage; impairment in motor functioning, strength, and/or coordination; or impairment in sensation. These impairments may contribute to difficulties performing day-to-day activities safely and independently
Social role participation	Participation is defined by the WHO as "involvement in life situations"[3] and commonly includes engagement in endeavors within one's community. TBI affects many areas of participation, including work/productive activity, recreation and leisure pursuits, and social/family role function
Perceived generic and disease-specific health-related quality of life	TBI may create significant limitations in multiple areas of functioning and well-being, often reducing perceived quality of life with regard to multiple generic and disease-specific dimensions
Health economic measures	Health economic measures assess the magnitude of benefit in relation to costs spent; eg, they identify the most cost-effective therapeutic procedure in terms of cost per QALY
Patient-reported outcomes (future multidimensional tools)	No single measure to date can adequately capture the multiplicity of difficulties that people with TBI may face. This domain includes emerging large-scale measurement tools for patient-reported outcomes across several domains for generic medical populations, neurologic compromise, and TBI-related symptoms

Abbreviation: QALY, quality-adjusted life-year.

domain. Subgroups of panel members developed initial lists of potential measures within each domain and provided information about the criteria detailed. Potential measures were discussed among the entire panel through a series of conference calls, and a more limited set of measures for each outcome domain was selected for further discussion in the panel at a face-to-face meeting in March 2009. In preparation for the meeting, all panel members assisted in composing a series of detailed tables with relevant information about general administration characteristics, psychometric properties, and advantages and limitations of each potential measure.

The primary objective of the meeting was to further examine, refine, and limit the list of potential outcome measures by using the information collected and reviewed. In accordance with other CDE work groups, a final set of measures was selected and organized into the 3 tiers described after further discussion of the relative advantages and limitations of each measure. Selection of the final measures for each CDE level was done by work group consensus.

Description and Selection of Core, Supplemental, and Emerging CDEs

The rationale behind the core measures was to create a primary set of well-established measures that cover outcome domains important to many studies. Primary emphasis was given to selecting a single measure (or limited set of measures) that best covered each domain. For the core measures, established use in subjects with TBI and favorable psychometric characteristics were considered primary criteria. Brevity and ease of administration also influenced the selection of core measures because the intent was to recommend measures that would be feasible to administer in a reasonable time (ie, $<$90min). Availability of different validated formats, such as self and other proxy response formats, also was considered given that self-report is impossible or unreliable for some people with TBI. Finally, applicability of each measure across a range of postinjury functional levels also was considered highly important because one of the primary objectives of the CDEs was to foster comparability of outcome measurement across different studies.

The rationale behind creating a set of supplemental measures was to recommend additional measures in each domain that could be considered for more in-depth outcome assessment within a certain domain or for patients at a specific functional level. For example, in studies in which neuropsychological outcome is of particular interest, investigators may draw on additional outcome measures from the supplemental list that target additional aspects of cognitive functioning not covered by the core measures (eg, visual memory, verbal fluency, fine motor control). Similarly, in studies focusing on patients with

Table 2: Core, Supplemental, and Emerging Measures for Each Domain

Domain Name	Core Measure(s)	Supplemental Measures	Emerging Measure(s)
Global outcome	GOS-E	MPAI-4	
		DRS	
		SF-36, version 2	
Recovery of consciousness		CRS-R	
Neuropsychological impairment	RAVLT	BVMT-R	NIH Toolbox cognitive battery
	TMT	Letter-Number Sequencing subtest of the WAIS-III/WAIS-IV	
	Processing Speed Index from WAIS-III/WAIS-IV	COWAT	
		CWIT	
		Digit Span subtest of the WAIS-III/ WAIS-IV	
		Word Reading subtest of the WRAT-4	
		GPT	
Psychological status	BSI-18	MMPI-2-RF	NIH Toolbox emotional battery
		AUDIT	
		Substance use questions from the TBI Model Systems data set	
		ASSIST	
		PCL-C/M/S	
		FAD	
TBI-related symptoms	RPQ	NSI	
Behavioral function		FrSBe	
Cognitive activity limitations	Cog-FIM		
Physical function	FIM motor subscale		NIH Toolbox motor and sensory batteries
			NOS-TBI
Social role participation	CHART-SF		PART
Perceived generic and disease-specific health-related quality of life	SWLS		QOLIBRI
Health economic measures		EuroQOL	
Patient-reported outcomes (future multidimensional tools)			PROMIS
			Neuro-QOL
			TBI-QOL

disorders of consciousness, the CRS-R was recommended because it was designed specially for assessment of this target population. Finally, additional measures of psychological and/or family functioning or substance abuse may be of importance, depending on the study design, functional level, or target population.

The third tier consists of emerging measures, meaning those currently under development, in validation, or nearing publication and that have significant potential to be superior to some older (legacy) measures in the core or supplemental sets. These emerging measures were selected because they fill existing gaps in the measurement of TBI-related sequelae or use more sophisticated validation techniques than older measures. Additionally, some of these measures may better facilitate comparison across patient groups (eg, different disease populations, broader age range, more comprehensive sampling of domains of function). Because established use in subjects with TBI and psychometric characteristics were considered primary criteria for core and supplemental measures, the emerging measures will require further consideration as CDEs as evidence accumulates about their psychometric characteristics, normative data, and utility in TBI research.

In this vein, the work group acknowledges that the selection of recommended outcome measures is a flexible and dynamic process that will undergo further evolution as additional evidence emerges and as testing of these measures as CDEs is undertaken. For example, some subtests from the WAIS-III were selected as core measures, although version IV has been released and the WAIS-III may not be available for purchase from the publisher in the future. Although version III currently is recommended because of its better fit with the selection criteria, particularly its use in TBI research, we acknowledge that the WAIS-IV versions of the subtests are likely to replace the WAIS-III versions in the core set pending further research, and that either version is acceptable because both measures will yield reliable and valid results. With any effort such as this attempt to create a set of CDEs, there is a dynamic tension between the desire to maintain consistency among a stable set of measures and the desire to adopt new improved measures as they become available. All core and supplemental measures listed here have been selected as recommended measures at the time of this publication; nevertheless, the work group advises the reader to consult the CDE web site for any updates to this listing. It is particularly important to track the progress of emerging measures because these are believed to have the potential to replace items in the existing core and/or supplemental set.

RECOMMENDATIONS FOR TBI OUTCOME MEASURES

Recommended CDEs (all 3 tiers) are listed in table 2 and described briefly next. The reader also is referred to www.

CommonDataElements.ninds.nih.gov for detailed supplemental information about each measure.

Core Data Elements

Glasgow Outcome Scale (Extended). The GOS[4] is a single-item scale that summarizes patient status in 1 of 5 categories: dead, vegetative state, severe disability, moderate disability, and good recovery. The GOS-E[5] is a revision of the GOS that provides 8 categories of outcome: dead, vegetative state, lower severe disability, upper severe disability, lower moderate disability, upper moderate disability, lower good recovery, and upper good recovery. GOS-E ratings are based on a structured interview and are easily recoded to GOS ratings. Together, these scales are the most commonly used TBI global outcome measure, and their use permits comparison to much of the world literature on TBI outcome.

Rey Auditory Verbal Learning Test. This measure of word list learning is brief, available in the public domain, covers a wide age range, and has alternate forms. The RAVLT is one of the most widely studied measures of cognition, has extensive normative data,[6-8] and has been used in different languages, cultures, and ethnic groups. It has good psychometric properties and is sensitive to neurologic conditions. The RAVLT will be used for validating the episodic memory measure of the NIH Toolbox or will be included in the Toolbox itself.

Trail Making Test. The TMT[9] is a measure of attention, speed, and mental flexibility. It is brief, widely used by neuropsychologists,[10] sensitive to TBI-associated cognitive impairment, and reliable.[11] Demographically adjusted normative data are available for a wide age range,[9] and there are adult and child versions. Arabic, Chinese, and Hebrew versions are available. Practice effects are found over short retest intervals, but disappear after several administrations; at longer intervals, scores show only modest change in healthy adults.

WAIS-III/WAIS-IV Processing Speed Index. This index is based on the Digit Symbol Coding and Symbol Search subtests of the WAIS-III (or WAIS-IV),[12,13] which has extensive normative data and excellent psychometric properties. As a measure of information processing rate, it is highly sensitive to the effects of TBI and its severity. It has been used in different languages, cultures, and ethnic groups and is usable across literacy levels. This measure is being used as a legacy measure to validate NIH Toolbox processing speed measures.[12,14,15]

Brief Symptom Inventory 18. The BSI-18[16] is a short form of the Symptom Checklist-90-Revised.[17] It is a brief self-report measure of psychological distress with 3 subscales (Depression, Anxiety, and Somatization) and a Global Severity Index. The BSI-18 was selected as a core measure because of its brevity, global assessment of common psychological issues in people with TBI, and sound psychometric characteristics. It can be used to monitor change in response to treatment and can be completed using paper-and-pencil or computerized administration formats.

Rivermead Post Concussion Symptom Questionnaire. The RPQ is a measure of postconcussion symptom presence and severity after TBI. It contains 16 items, which the participant rates in relation to premorbid functioning by means of written self-report or in-person or telephone interview. The RPQ has been used most often in assessing postconcussion symptoms in persons with mild to moderate TBI, but also has been used in patients with severe TBI,[18] and the measurement of symptoms contained in this measure may be applicable at all levels of severity. The RPQ was selected as a core measure based on its sound psychometric characteristics and capacity to detect clinical changes in patients with mild TBI. The scale has been used to investigate the relationship between behavioral and neurophysiologic markers of injury[19-21] and outcome prediction.[22]

FIM Cognition Subscale. The FIM was selected as a core measure of both physical and cognitive activity limitations because of its widespread clinical use in TBI populations, multiple validated response formats (observational ratings, self- or proxy-report in person or by telephone),[23,24] and extensive use in studies of diagnostic accuracy, outcome prediction, and treatment effectiveness. Item response analysis of the FIM has confirmed a motor domain consisting of 13 items and a cognitive domain consisting of 5 items.[25] There is low correlation between the Cog-FIM and mental and physical health measures, suggesting discriminant validity.[24] Ceiling effects may limit the FIM's utility for longitudinal studies of TBI, although ceiling effects are less extreme for the Cog-FIM versus the Motor FIM in those with moderate/severe TBI.[26,27]

Craig Handicap Assessment and Reporting Technique Short Form. The CHART-SF was designed to provide a simple objective measure of the degree to which impairments and disabilities result in handicaps (participation restriction) in the years after initial rehabilitation.[28] It contains the 6 Rasch-validated subscales of Physical Independence, Cognitive Independence, Mobility, Occupation, Social Integration, and Economic Self-Sufficiency. It shows good interrater, test-retest, and subject-proxy reliability. It has been shown to discriminate between people with TBI and stroke who report lower scores than those with other disabilities.[29,30] CHART Cognitive Independence scores correlate more highly with Cog-FIM scores than FIM motor subscale scores.[29]

Satisfaction With Life Scale. The SWLS is a global measure of life satisfaction.[25] The SWLS consists of 5 items that are completed by the subject. It has shown consistent differences between populations that would be expected to have different quality of life (eg, psychiatric patients or male prison inmates). The SWLS also has been found to change in the expected directions in response to major life events[31] and in patients receiving psychotherapy.[32]

Supplemental Data Elements

JFK Coma Recovery Scale–Revised. The CRS-R is a standardized behavioral assessment instrument designed to measure neurobehavioral function in patients with disorders of consciousness.[33] It is composed of 6 subscales designed to assess auditory, visual, motor, oromotor/verbal, communication, and arousal functions. The CRS-R is the only standardized assessment measure that directly incorporates diagnostic criteria for coma, vegetative state, MCS, and emergence from MCS and thus is strongly recommended for all studies of disorders of consciousness. The CRS-R shows adequate sensitivity and specificity,[34-36] correlates well with functional outcome,[37] is useful for monitoring treatment effectiveness,[3,38] and is available in 11 languages.

Mayo-Portland Adaptability Inventory. The MPAI-4 (and the PI, a component of the MPAI-4)[39-41] was designed for outcome measurement after acquired brain injury in the post-acute stage of recovery. The MPAI-4 is the product of 15 years of development using item response and classic psychometric theory and has established concurrent, construct, and predictive validity. There is a total score and subscale scores for Ability, Adjustment, and Participation. Strengths include ease of administration, flexibility (by telephone or in person), and normative values based on ratings by people with brain injury, significant others, and clinical staff. The PI (independent of the entire MPAI-4) has not been used extensively in research applications, although studies support the use of the entire MPAI-4 with adult and pediatric samples.[42-44]

Disability Rating Scale. The DRS is intended to measure general functioning during the course of recovery.[45] The DRS provides a single score based on level of arousal; cognitive ability to perform basic ADLs, including eating, grooming, and toileting; independence in the home; and employability. The DRS was selected because it is applicable across a wide range of injury severity and recovery intervals. The DRS may be useful in studies of subjects with moderate to severe TBI with serial measurement,[46] particularly when initial measurement occurs in the acute postinjury interval.

Medical Outcomes Study 36-Item Short Form Health Survey. The SF-36 is the most widely used subjective health status measure, developed by using classic psychometric test theory methods.[47] It contains 11 items and generates subscale scores in physical functioning, physical role function, emotional role function, bodily pain, vitality, mental well-being, social functioning, and general health perception. An additional item assesses changes in health status during the last year. Two summary scores can be computed, a physical component score and a mental component score. In TBI research, more than 30 studies using the SF-36 showed high internal consistencies of all scales,[48] as well as sound values for construct, discriminant, and content validity[49,50] and sensitivity to treatment-related changes. The work group is mindful that the SF-12, has been selected as a core element by the PTSD Workgroup.[51] However, the SF-12 items are contained within and may be scored from the SF-36.[52] Thus, use of the latter scale, which we recommend for TBI studies for comprehensive evaluation, will readily permit comparison with studies using the SF-12.

Brief Visuospatial Memory Test–Revised. The BVMT-R is a multitrial measure of visual-spatial memory/learning requiring reproduction of geometric forms. It was normed for a wide adult age range (age range, 18 to ≥79y) and has no significant sex- or education-related effects. It has good test-retest reliability for total score and interrater reliability. The BVMT-R has multiple alternate forms, correlates well with other measures of memory, and shows sensitivity to neurologic conditions. It is being used as a legacy measure for validating the NIH Toolbox measure of episodic memory.[53]

WAIS-III/WAIS-IV Letter-Number Sequencing subtest. This is a measure of auditory working memory that appears in both the WAIS-III and Wechsler Memory Scale-Third Edition (or WAIS-IV).[12,13] The subtest has extensive normative data and good psychometric properties, as well as clinical sensitivity. This measure is being used as a legacy measure to validate the NIH Toolbox working memory measure.[14,54,55]

Controlled Oral Word Association Test. The COWAT[56] measures attentional control, working memory, and other components of executive function. There is a strong association between focal frontal lobe injuries after TBI and impaired performance on the COWAT.[57] Several alternate forms and a Spanish version are available. It does not have a low ceiling in people without neurologic disorders.[15] Demographically adjusted normative data are available for ages 20 to 85 years.[9]

Color-Word Interference Test. The CWIT,[58] a variant of the Stroop procedure, measures cognitive flexibility, selective attention, and the capacity to inhibit an overlearned response. The CWIT version is a subtest from the Delis-Kaplan Executive Function System. Normative data are available for people aged 8 to 89 years.

WAIS-III or WAIS-IV Digit Span subtest. This test[12,13] provides a brief assessment of auditory attention. The digits backward component is particularly informative as a simple measure of working memory. The test is widely available, easy and quick to administer, and well normed. Digit Span has been used as a marker of cognitive deficit and recovery.[59,60] Its potential use as a symptom validity measure is an added benefit.[61,62]

Word Reading subtest of the WRAT-4. The WRAT-4[63] provides a quick measure of academic achievement based on well-established norms. Studies have shown stability in people with TBI, allowing results to be used as an estimate of premorbid cognitive ability.[64,65] Results can be affected adversely by visual difficulty, severe language disorder, and preexisting learning disability.

Grooved Pegboard Test. The GPT has proved to be a sensitive indicator of brain functioning, with diminished performance noted after even milder injury. It is readily available, easy and quick to administer, and well normed. The GPT can be used to document existing deficits and predict outcome.[66,67] One drawback is that performance can be influenced by peripheral injury, such as arm or hand fracture or problems with visual acuity.

Minnesota Multiphasic Personality Inventory 2, Restructured Form. The MMPI-2-RF is a revised 338-item version of the MMPI-2. There are 50 scales: Restructured Clinical Scales, Validity Scales, Specific Problem Scales, Interest Scales, and Personality Psychopathology Five Scales.[68] The MMPI-2 is the most extensively used and researched of the comprehensive personality assessment tools. The MMPI-2-RF provides a more time-efficient approach to using the MMPI-2. It is psychometrically up to date and is linked to current models of psychopathology and personality.

Alcohol Use Disorders Identification Test. The AUDIT[69] and the substance use questions from the TBI Model Systems data set indicate the extent of "problematic" substance use. The AUDIT was developed by the WHO and has been used extensively with a range of populations, including persons with TBI.[70] An abbreviated 3-item version (AUDIT-C)[71] screens for extent of alcohol consumption. The substance use questions from the TBI Model Systems data set query the use of alcohol and other drugs and are based on questions from population-based surveys, thus allowing comparisons with statistics from the general population.[72] A dichotomous variable indicating the presence of problem substance use (unhealthy use) can be derived from these questions.

Alcohol, Smoking, and Substance Use Involvement Screening Test. For a more comprehensive assessment of substance use, the ASSIST[73] also was developed by the WHO, has been validated in 9 countries, and is easily administered, reliable, and valid. Recently completed work indicates that the ASSIST is sensitive to change and specifically to the effects of a brief intervention.[73]

PTSD Check List–Civilian, –Military, and –Stressor Specific versions. The PTSD Checklist is a 17-item self-report measure composed of the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* symptoms of PTSD.[74] Respondents rate on a 5-point scale how much they were bothered by each symptom "in the past month." The PTSD Checklist-Military Version asks about problems in response to "stressful military experiences." The PTSD Checklist-Civilian Version is not focused on one traumatic event. The PTSD Checklist-Stressor Specific Version requires the respondent to provide responses in relation to a specific event. The PCL measures also can be used for evaluation of PTSD severity and to monitor change in response to treatment. These are public domain measures and are widely used.

Family Assessment Device. The FAD[75] is a 60-item self-report instrument based on the McMaster Model of Family Functioning.[76] Patients and/or family members read and respond to the items. The FAD assesses structural and organizational properties of the family group and the patterns of trans-

actions among family members that have distinguished between healthy and unhealthy families. It can be used to evaluate the family unit in a broadly defined manner to include any type of committed or enduring relationship structure. The FAD has 7 subscales: 1 General Functioning scale, which assesses overall health and pathology of the family, and 6 dimensional subscales: Problem Solving, Communication, Roles, Affective Responsiveness, Affective Involvement, and Behavior Control. Subscale reliabilities and test-retest reliability are adequate, and it has low correlations with social desirability and moderate correlations with other self-report measures of family functioning.[77]

Neurobehavioral Symptom Inventory. The NSI, also known as the Post Mild TBI Symptom Checklist,[78] was designed to measure 22 common postconcussion symptoms after TBI without regard to the existence of preinjury symptoms. The severity of each symptom is measured using a 5-item scale (0 indicates none to 4 indicates very severe) that asks participants to indicate the extent to which each symptom has disturbed them in the previous 2 weeks. NSI total score is the sum of severity ratings of the 22 symptoms. Cluster scores (physical, cognitive, affective, and sensory domains) were derived.[78] The NSI was selected as a supplemental measure because it currently is being used by both the U.S. Department of Defense and the VA as part of their Operation Iraqi Freedom/Operation Enduring Freedom TBI evaluation process of postconcussional symptoms.

Frontal Systems Behavior Scale. The FrSBe[79] assesses the severity of behavior problems associated with frontal lobe function as rated by the injured person and/or a caregiver. There is a total score and 3 subscales (apathy, disinhibition, and executive dysfunction) confirmed by using factor analysis.[80] Norms are available based on sex, age, and education for persons with TBI and caregivers. The sample of non-neurologically impaired people used to derive the norms was small and constrained in terms of education and race, but the FrSBe has been used effectively in a few studies.[81,82]

EuroQoL. The EuroQoL is a generic self-rating instrument to assess health-related quality of life and health status. It generates an index of health for use in economic evaluation, has good psychometric properties, is available in many languages, and consists of self-rating of a set of health states and background information about the respondent's health. The 5 dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression result in a health state profile. Combined with clinical data (eg, survival) it gives quality-adjusted life-years. In TBI, the instrument has been used in some outcome studies with good success.[83-87]

Please see Supplementary Table 1 for information related to the psychometric properties of all core and supplemental measures.

Emerging Data Elements

NIH Toolbox. The NIH Toolbox (Cognitive, Emotional, Motor, and Sensory components) is part of the NIH Blueprint initiative. It seeks to assemble brief comprehensive assessment tools that will be useful in a variety of settings with particular emphasis on measuring outcomes in epidemiologic studies and clinical trials across the life span. The ultimate goal is to help improve communication within and between fields of biomedical research to advance knowledge by using CDEs. The battery will examine various cognitive (episodic memory, language, processing speed, working memory, executive functions, attention), emotional (negative affect, positive affect, stress and coping, social relationships), sensory (vestibular, audition, olfaction, taste, vision), and motor functions (dexterity, strength, locomotion, endurance, balance). The battery is designed to measure these domains in subjects aged 3 through 85 years, is normed for both English and Spanish speakers, and will be available at a nominal cost and will take no more than 2 hours to administer. The battery has gone through extensive work to identify and pretest the constructs to be measured. Validation of the NIH Toolbox batteries have been completed, with norming planned in about 4500 subjects (please see http://www.nihtoolbox.org for additional information).

Neurological Outcome Scale for TBI. The NOS-TBI is a brief measure of neurologic functioning (including level of consciousness, cranial nerve functioning, limb strength, language, ataxia) modeled after the NIH Stroke Scale, but containing items specific to and validated in a TBI population. The NOS-TBI contains 15 items, some of which have subparts (eg, for lateralization). Administration and scoring guidelines are provided for patients who are comatose, obtunded, or aphasic, rendering this a measure that can be used across a wide range of injury severity and chronicity. Available preliminary psychometric results indicate excellent reliability and validity.[88-90]

Participation Assessment with Recombined Tools. The PART is a measure of community participation developed by the TBI Model Systems by combining the primary measures found in the TBI literature (Community Integration Questionnaire, original and revised[91,92]; Participation Objective; Participation Subjective[93]; and the CHART).[28] The measure has been administered to persons with TBI and other sources of disability and to a population-based sample. Psychometric data have not yet been published, but available results indicate that the PART is reliable and valid, maintaining the strengths and overcoming some of the weaknesses of its component measures.

Quality of Life After Brain Injury. The QOLIBRI is the first TBI disease-specific quality-of-life cross-culturally and consensually developed patient-reported outcome tool for clinical trials and individual use (www.qolibrinet.com).[94-96] It has been validated in 2 large multinational TBI populations (N>1500, N>900) with different grades of disease, showing good psychometric properties. Based on classic and modern test theory, it yields 37 items in 6 Likert-formatted scales, 4 assessing satisfaction (Cognition, Self, Daily Life and Autonomy, Social Relationships) and 2 assessing the feeling of botheredness (Emotions and Physical Problems). A total score and a 6-item screener also are available. It is brief, will be in the public domain from February 2010 onward, and exists in more than 10 languages.

Finally, there are 3 interrelated measurement systems (the PROMIS, Neuro-QOL, TBI-QOL) of patient-reported outcomes measures being developed to measure emotional functioning, social participation, and physical and medical functioning across a wide array of domain areas.

Patient-Reported Outcomes Measurement Information System. The PROMIS[97] is a new measurement system that is part of the NIH Roadmap to improve the clinical research enterprise. The PROMIS Network has developed and tested a large bank of items measuring patient-reported outcomes over several domains, including physical functioning, sleep disturbance, fatigue, anxiety, depression, anger, social roles, and social activities. Item banks have been calibrated, allowing the test to be administered as a computerized adaptive test or as short forms to ensure brevity. Researchers can select domains of functioning relevant to their specific research question. The PROMIS is designed as a generic measure that is to be used across all medical populations.

Neuro-QOL. The Neuro-QOL also is a patient-reported outcome measurement system funded through a contract

method by the NINDS.[98,99] The Neuro-QOL team has developed separate item banks covering the domains of Mobility/Ambulation, ADLs/Upper Extremity, Depression, Anxiety, Positive Psychological Functioning, Stigma, Perceived and Applied Cognition (includes communication), Social Role Performance, Social Role Satisfaction, Fatigue, Personality and Behavioral Change, and Sleep Disturbances. Embedded in several of the Neuro-QOL domains are a significant number of PROMIS items. The Neuro-QOL is designed to be a common outcome variable across all clinical trials research sponsored by the NINDS.

Traumatic Brain Injury–Quality of Life. The TBI-QOL is a new multifaceted patient-reported outcomes measure that is in development.[100,101] It will embed Neuro-QOL and PROMIS items and will cover the domains of functioning of Perceived and Applied Cognition (includes communication), Personality and Behavioral Change (includes impulsivity), Depression, Anxiety, Positive Psychological Functioning, Stigma, Social Role Performance, Social Role Satisfaction, Fatigue, and Sexuality. Five TBI Model System centers and 4 VA Polytrauma/Defense and Veterans Brain Injury Centers are collaborating to develop this instrument.

Because the PROMIS, Neuro-QOL, and TBI-QOL contain common items and have been developed as calibrated item banks using item response theory, the researcher will not administer common item banks from these instruments (eg, both the TBI-QOL and PROMIS depression item banks), but instead select one or the other. Linking tables or cross-walks between the PROMIS, Neuro-QOL, and TBI-QOL will be developed, allowing researchers to compute a PROMIS and Neuro-QOL equivalency score from TBI-QOL item banks. Similarly, PROMIS equivalency scores will be derived for people who complete the relevant Neuro-QOL item banks.

For additional information about all core, supplemental, and emerging CDEs, please consult www.CommonDataElements.ninds.nih.gov. This site contains a table with descriptions of the measures; a listing of variables and permissible values; information about administration length, training requirements, and appropriate populations for use; and references or contact information.

FUTURE ISSUES AND RESEARCH NEEDS

The work group identified several areas in which additional research would enhance outcome measurement in TBI. First, as indicated in the discussion of emerging measures, there is a need for further validation and testing of measures, such as the NIH Toolbox, to specifically evaluate their utility in TBI. Second, TBI causes characteristic cognitive and communication impairments that can compromise the validity of self-report. Despite the great promise of the new patient-reported outcome measures under development, the work group also identified the need for more research about the applicability and validity of proxy report. In addition, as has been done for other populations, such as those with serious mental illness, we see the need for further development of standardized measures that directly test performance on cognitively demanding activities, such as using transportation, adhering to medication schedules, and exercising judgment in the home and community. Third, we recognize that measurement of vocational outcomes is minimally represented in the CDE recommended here, a circumstance that partly reflects the state of the science. More work is needed to develop standard measures of employment post-TBI, taking into account the diversity of important outcomes (return to work vs new employment, long-term job maintenance, pay, satisfaction, and so on). Finally, the work group acknowledged the need for additional measures of executive functioning that keep pace with theoretical developments in clinical neuroscience.

SUMMARY

In accordance with other CDE work groups, the following 3 tiers of CDE were recommended: (1) 9 core measures covering outcome domains relevant to most TBI studies that could be applied as either a comprehensive battery or in addition to other outcome measures selected by the investigator, (2) supplemental measures for consideration in TBI research focusing on more specific topics or subpopulations, and (3) emerging measures, which include promising instruments currently under development, in the process of validation, or nearing the point of published findings that have significant potential to be superior to some measures currently in the core and supplemental lists. Selection of the CDE measures is intended to facilitate comparison of findings from large-scale research efforts designed to document the natural course of recovery from TBI, enhance the prediction of outcome, and/or measure the effects of treatment. The work group acknowledges that although these measures were chosen after substantial review of available evidence and discussion within the group, any selection of CDE is a dynamic process that must accommodate some shift and evolution in the measures within each category as new evidence emerges and selected measures continue to be tested.

References

1. Thurmond VA, Hicks R, Gleason T, et al. Advancing integrated research in psychological health and traumatic brain injury: common data elements. Arch Phys Med Rehabil 2010;91:1633-6.
2. World Health Organization. International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001.
3. Schnakers C, Hustinx R, Vandewalle G, et al. Measuring the effect of amantadine in chronic anoxic minimally conscious state. J Neurol Neurosurg Psychiatry 2008;79:225-7.
4. Jennett B, Bond M. Assessment of outcome after severe brain damage. Lancet 1975;1:480-4.
5. Wilson JT, Pettigrew LE, Teasdale GM. Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. J Neurotrauma 1998;15:573-85.
6. Schmidt M. Rey Auditory Verbal Learning Test. Los Angeles: Western Psychological Services; 1996.
7. Mitrushina M, Boone KB, Razani J, D'Elia LF. Handbook of normative data for neuropsychological assessment. 2nd ed. New York: Oxford University Pr; 2005.
8. Ivnik RJ, Malec JE, Tangalos EG, Peterson RC, Kokmen E, Kurland LT. Mayo's older American's normative studies: updated AVLT norms for ages 56 to 97. Clin Neuropsychol 1992;6:83-104.
9. Heaton RK, Miller SW, Taylor MJ, Grant I. Revised comprehensive norms for an expanded Halstead-Reitan Battery: demographical adjusted neuropsychological norms for African American and Caucasian adults professional manual. Lutz: Psychological Assessment Resources Inc; 2004.
10. Rabin LA, Barr WB, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: a survey of INS, NAN, and APA Division 40 members. Arch Clin Neuropsychol 2005;20:33-65.
11. Strauss E, Sherman EMS, Spreen O. Rey-Osterrieth Auditory Verbal Learning Test. Compendium of neuropsychological tests. New York: Oxford University Pr; 2006. p 776–807.
12. Wechsler D. Wechsler Adult Intelligence Scale III. San Antonio: Harcourt Assessment Inc; 1997.

13. Wechsler D. Wechsler Adult Intelligence Scale IV. San Antonio: Harcourt Assessment Inc; 2008.
14. WAIS-III-WMS-III technical manual. San Antonio: Psychological Corp; 2008.
15. Strauss E, Sherman EMS, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. 3rd ed. New York: Oxford University Pr; 2006.
16. Derogatis LR. Brief Symptom Inventory 18 (BSI 18): administration, scoring and procedures manual. Minneapolis: NCS Pearson Inc; 2001.
17. Derogatis LR. Symptom Checklist-90-R (SCL-90-R): administration, scoring and procedures manual. 3rd ed. Minneapolis: NCS Pearson, Inc; 1994.
18. King NS, Crawford S, Wenden FJ, Moss NE, Wade DT. The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. J Neurol 1995;242:587-92.
19. Ingebrigtsen T, Romner B, Marup-Jensen S, et al. The clinical value of serum S-100 protein measurements in minor head injury: a Scandinavian multicentre study. Brain Inj 2000;14: 1047-55.
20. Wilde EA, McCauley SR, Hunter JV, et al. Diffusion tensor imaging of acute mild traumatic brain injury in adolescents. Neurology 2008;70:948-55.
21. Smits M, Dippel DW, Houston GC, et al. Postconcussion syndrome after minor head injury: brain activation of working memory and attention. Hum Brain Mapp 2009;30:2789-803.
22. Crawford S, Wenden FJ, Wade DT. The Rivermead Head Injury Follow up Questionnaire: a study of a new rating scale and other measures to evaluate outcome after head injury. J Neurol Neurosurg Psychiatry 1996;60:510-4.
23. Tepper S, Beatty P, DeJong G. Outcomes in traumatic brain injury: self-report versus report of significant others. Brain Inj 1996;10:575-81.
24. Hobart JC, Lamping DL, Freeman JA, et al. Evidence-based measurement: which disability scale for neurologic rehabilitation? Neurology 2001;57:639-44.
25. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985;49:71-5.
26. Hall KM, Bushnik T, Lakisic-Kazazic B, Wright J, Cantagallo A. Assessing traumatic brain injury outcome measures for long-term follow-up of community-based individuals. Arch Phys Med Rehabil 2001;82:367-74.
27. Hammond FM, Hart T, Bushnik T, Corrigan JD, Sasser H. Change and predictors of change in communication, cognition, and social function between 1 and 5 years after traumatic brain injury. J Head Trauma Rehabil 2004;19:314-28.
28. Whiteneck GG, Charlifue SW, Gerhart KA, Overholser JD, Richardson GN. Quantifying handicap: a new measure of long-term rehabilitation outcomes. Arch Phys Med Rehabil 1992;73: 519-26.
29. Mellick D, Walker N, Brooks CA, Whiteneck GG. Incorporating the cognitive independence domain into CHART. J Rehabil Outcomes Meas 1999;3:12-21.
30. Walker N, Mellick D, Brooks CA, Whiteneck GG. Measuring participation across impairment groups using the Craig Handicap Assessment Reporting Technique. Am J Phys Med Rehabil 2003;82:936-41.
31. Vitaliano PP, Russo J, Young HM, Teri L, Maiuro RD. Predictors of burden in spouse caregivers of individuals with Alzheimer's disease. Psychol Aging 1991;6:392-402.
32. Pavot W, Diener E. Review of the Satisfaction With Life Scale. Psychol Assess 1993;5:164-72.
33. Giacino JT, Kalmar K, Whyte J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabil 2004;85:2020-9.
34. Schnakers C, Giacino J, Kalmar K, et al. Does the FOUR score correctly diagnose the vegetative and minimally conscious states? Ann Neurol 2006;60:744-5; author reply 745.
35. Schnakers C, Majerus S, Giacino J, et al. A French validation study of the Coma Recovery Scale-Revised (CRS-R). Brain Inj 2008;22:786-92.
36. Schnakers C, Vanhaudenhuyse A, Giacino J, et al. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. BMC Neurol 2009;9:35.
37. Giacino JT, Kalmar K. The vegetative and minimally conscious states: a comparison of clinical features and functional outcome. J Head Trauma Rehabil 1997;12:36-51.
38. Schiff ND, Giacino JT, Kalmar K, et al. Behavioural improvements with thalamic stimulation after severe traumatic brain injury. Nature 2007;448:600-3.
39. Malec JF. Comparability of Mayo-Portland Adaptability Inventory ratings by staff, significant others and people with acquired brain injury. Brain Inj 2004;18:563-75.
40. Malec JF. The Mayo-Portland Participation Index: a brief and psychometrically sound measure of brain injury outcome. Arch Phys Med Rehabil 2004;85:1989-96.
41. Malec JF, Kragness M, Evans RW, Finlay KL, Kent A, Lezak MD. Further psychometric evaluation and revision of the Mayo-Portland Adaptability Inventory in a national sample. J Head Trauma Rehabil 2003;18:479-92.
42. Winstanley J, Simpson G, Tate R, Myles B. Early indicators and contributors to psychological distress in relatives during rehabilitation following severe traumatic brain injury: findings from the Brain Injury Outcomes Study. J Head Trauma Rehabil 2006;21: 453-66.
43. Murrey GJ, Hale FM, Williams JD. Assessment of anosognosia in persons with frontal lobe damage: clinical utility of the Mayo-Portland Adaptability Inventory (MPAI). Brain Inj 2005;19:599-603.
44. Oddson B, Rumney P, Johnson P, Thomas-Stonell N. Clinical use of the Mayo-Portland Adaptability Inventory in rehabilitation after paediatric acquired brain injury. Dev Med Child Neurol 2006;48:918-22.
45. Granger CV. The emerging science of functional assessment: our tool for outcomes analysis. Arch Phys Med Rehabil 1998;79: 235-40.
46. Hall K, Cope DN, Rappaport M. Glasgow Outcome Scale and Disability Rating Scale: comparative usefulness in following recovery in traumatic head injury. Arch Phys Med Rehabil 1985;66:35-7.
47. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
48. Findler M, Cantor J, Haddad L, Gordon W, Ashman T. The reliability and validity of the SF-36 Health Survey questionnaire for use with individuals with traumatic brain injury. Brain Inj 2001;15:715-23.
49. McNaughton HK, Weatherall M, McPherson KM. Functional measures across neurologic disease states: analysis of factors in common. Arch Phys Med Rehabil 2005;86:2184-8.
50. Emanuelson I, Andersson Holmkvist E, Bjorklund R, Stalhammar D. Quality of life and post-concussion symptoms in adults after mild traumatic brain injury: a population-based study in western Sweden. Acta Neurol Scand 2003;108:332-8.
51. Kaloupek D. Common data elements for posttraumatic stress disorder research. Arch Phys Med Rehabil 2010;91:1684-91.
52. Resnick B, Parker R. Simplified scoring and psychometrics of the revised 12-item Short-Form Health Survey. Outcomes Manag Nurs Pract 2001;5:161-6.

53. Benedict RH. Brief Visuospatial Memory Test-Revised. Odessa: Psychological Assessment Resources Inc; 1997.
54. Donders J, Tulsky DS, Zhu J. Criterion validity of new WAIS-II subtest scores after traumatic brain injury. J Int Neuropsychol Soc 2001;7:892-8.
55. DeLuca J, Chelune GJ, Tulsky DS, Lengenfelder J, Chiaravalloti ND. Is speed of processing or working memory the primary information processing deficit in multiple sclerosis? J Clin Exp Neuropsychol 2004;26:550-62.
56. Spreen O, Benton AL. Neurosensory Center comprehensive examination for aphasia. Victoria: Univ of Victoria Neuropsychology Laboratory; 1977.
57. Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. Neuropsychology 2004;18:621-8.
58. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan Executive Function System. San Antonio: The Psychological Corp; 2001.
59. Kersel DA, Marsh NV, Havill JH, Sleigh JW. Neuropsychological functioning during the year following severe traumatic brain injury. Brain Inj 2001;15:283-96.
60. Millis SR, Rosenthal M, Novack TA, et al. Long-term neuropsychological outcome after traumatic brain injury. J Head Trauma Rehabil 2001;16:343-55.
61. Greiffenstein MF, Baker WJ, Gola T. Validation of malingered amnesia meaures with a large clinical sample. Psychol Assess 1994;6:218-24.
62. Meyers JE, Volbrecht M. Validation of reliable digits for detection of malingering. Assessment 1998;5:303-7.
63. Wilkinson GS, Robertson GJ. Wide Range Achievement Test-4 professional manual. Lutz: Psychological Assessment Resources Inc; 2006.
64. Orme DR, Johnstone B, Hanks R, Novack TA. The WRAT-3 Reading Subtest is a measure of premorbid intelligence among persons with brain injury. Rehabil Psychol 2004;49:250-3.
65. Ball JD, Hart RP, Stutts ML, Turf E, Barth JT. Comparative utility of Barona formulae, Wtar demographic algorithms, and WRAT-3 reading for estimating premorbid ability in a diverse research sample. Clin Neuropsychol 2007;21:422-33.
66. Boake C, Millis SR, High WM Jr, et al. Using early neuropsychologic testing to predict long-term productivity outcome from traumatic brain injury. Arch Phys Med Rehabil 2001;82:761-8.
67. Sherer M, Sander AM, Nick TG, High WM Jr, Malec JF, Rosenthal M. Early cognitive status and productivity outcome after traumatic brain injury: findings from the TBI Model Systems. Arch Phys Med Rehabil 2002;83:183-92.
68. Tellegen A, Ben-Porath YS. MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2): technical manual. Minneapolis: University of Minneapolis Pr; 2008.
69. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption–II. Addiction 1993;88:791-804.
70. Ponsford J, Whelan-Goodinson R, Bahar-Fuchs A. Alcohol and drug use following traumatic brain injury, a prospective study. Brain Inj 2007;21:1385-92.
71. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998;158:1789-95.
72. Corrigan JD, Bogner J, Lamb-Hart G, Sivak-Sears N. Technical report on problematic substance use variables 2003. Available at: http://www.tbims.org/combi/subst. Accessed September 4, 2009.
73. Humeniuk R, Dennington V, Ali R. The effectiveness of a brief intervention for illicit drugs linked to the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in primary health care settings: a technical report of phase III findings of the WHO ASSIST randomized controlled trial 2008. Available at: http://www.who.int/substance_abuse/activities/assist/en/index.html. Accessed August 17, 2009.
74. Weathers F, Litz B, Herman D, Huska J, Keane T. The PTSD Checklist (PCL): reliability, validity, and diagnostic utility. Proceedings of the 9th Annual Convention of the International Society of Traumatic Stress Studies, October 25, San Antonio, Texas; 1993.
75. Epstein NB, Baldwin LM, Bishop DS. The McMaster Family Assessment Device. J Marital Fam Ther 1983;9:171-80.
76. Epstein NB, Bishop DS, Ryan C, Miller IW, Keitner GI. The McMaster model view of healthy family functioning. In: Walsh F, editor. Normal family processes. New York: Guilford Pr; 1993. p 138-60.
77. Miller IW, Epstein NB, Bishop DS, Keitner GI. The McMaster Family Assessment Device: reliability and validity. J Marital Fam Ther 1985;11:345-56.
78. Cicerone KD, Kalmar K. Persistent postconcussion syndrome: the structure of subjective complaints after mild traumatic brain injury. J Head Trauma Rehabil 1995;10:1-17.
79. Grace J, Malloy PF. Frontal Systems Behavior Scale professional manual. Lutz: Psychological Assessment Resources Inc; 2001.
80. Stout JC, Ready RE, Grace J, Malloy PF, Paulsen JS. Factor analysis of the Frontal Systems Behavior Scale (FrSBe). Assessment 2003;10:79-85.
81. Norton LE, Malloy PF, Salloway S. The impact of behavioral symptoms on activities of daily living in patients with dementia. Am J Geriatr Psychiatry 2001;9:41-8.
82. Reid-Arndt SA, Nehl C, Hinkebein J. The Frontal Systems Behaviour Scale (FrSBe) as a predictor of community integration following a traumatic brain injury. Brain Inj 2007;21:1361-9.
83. EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16:199-208.
84. Bell KR, Temkin NR, Esselman PC, et al. The effect of a scheduled telephone intervention on outcome after moderate to severe traumatic brain injury: a randomized trial. Arch Phys Med Rehabil 2005;86:851-6.
85. Klose M, Watt T, Brennum J, Feldt-Rasmussen U. Posttraumatic hypopituitarism is associated with an unfavorable body composition and lipid profile, and decreased quality of life 12 months after injury. J Clin Endocrinol Metab 2007;92:3861-8.
86. Jakola AS, Muller K, Larsen M, Waterloo K, Romner B, Ingebrigtsen T. Five-year outcome after mild head injury: a prospective controlled study. Acta Neurol Scand 2007;115:398-402.
87. EQ-5D value sets: inventory, comparative review and user guide. Dordrecht: Springer; 2008.
88. McCauley SR, Wilde EA, Kelly TM, et al. The Neurological Outcome Scale for Traumatic Brain Injury (NOS-TBI): II. Reliability and convergent validity. J Neurotrauma 2010;27:991-7.
89. Wilde EA, McCauley SR, Kelly TM, et al. Feasibility of the Neurological Outcome Scale for Traumatic Brain Injury (NOS-TBI) in adults. J Neurotrauma 2010;27:975-81.
90. Wilde EA, McCauley SR, Kelly TM, et al. The Neurological Outcome Scale for Traumatic Brain Injury (NOS-TBI): I. Construct validity. J Neurotrauma 2010;27:983-9.
91. Willer B, Ottenbacher KJ, Coad ML. The Community Integration Questionnaire. A comparative examination. Am J Phys Med Rehabil 1994;73:103-11.
92. Johnston MV, Goverover Y, Dijkers M. Community activities and individuals' satisfaction with them: quality of life in the first year after traumatic brain injury. Arch Phys Med Rehabil 2005;86:735-45.
93. Brown M, Dijkers MP, Gordon WA, Ashman T, Charatz H, Cheng Z. Participation Objective, Participation Subjective: a

measure of participation combining outsider and insider perspectives. J Head Trauma Rehabil 2004;19:459-81.
94. von Steinbuechel N, Petersen C, Bullinger M. Assessment of health-related quality of life in persons after traumatic brain injury—development of the QOLIBRI, a specific measure. Acta Neurochir Suppl 2005;93:S43-9.
95. von Steinbuechel N, Wilson L, Gibbons H, et al. Quality of Life After Brain Injury (QOLIBRI)—scale validity and correlates of quality of life. J Neurotrauma 2010;27:1167-85.
96. von Steinbuechel N, Wilson L, Gibbons H, et al. Quality of Life After Brain Injury (QOLIBRI)—scale development and metric properties. J Neurotrauma 2010;27:1157-65.
97. Ader D. Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care 2007;45(Suppl):S1-2.
98. Perez L, Huang J, Jansky L, et al. Using focus groups to inform the Neuro-QOL measurement tool: exploring patient-centered, health-related quality of life concepts across neurological conditions. J Neurosci Nurs 2007;39:342-53.
99. Miller D, Nowinski C, Victorson D, Peterman A, Perez L. The Neuro-QOL project: establishing research priorities through qualitative research and consensus development. Qual Life Res 2005;14:2031.
100. Tulsky DS, Carlozzi NE, Kisala P, Victorson D, Cella D. New initiatives to develop patient reported outcomes measures for research with neurological populations. Proceedings of the International Neuropsychological Society, February 11–14, 2009, Atlanta, GA; 2009.
101. Tulsky DS, Kisala P, Victorson D, Carlozzi NE, Cella D. Development of tailored outcomes measures for traumatic brain injury and spinal cord injury. Proceedings of the American Academy of Neurology, April 25–May 2, 2009, Seattle, Washington; 2009.

Supplementary Table 1: Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
Global outcome				
GOS/GOS-E	IRR of the GOS is very good, with 92%–95% agreement.[1,2] Agreement for standard in-person administration for the GOS-E is less at ~78%.[2] TRT weighted κ is .92 for in-person vs telephone GOS and GOS-E, which is excellent. Absolute agreement for in-person vs telephone was 71%–77% for GOS-E and 86%–90% for GOS. Interrater weighted κ was .85 and .84 for GOS and GOS-E, respectively.[3] Weighted κ was .94 and .98 for TRT reliability using postal questionnaires for GOS and GOS-E, respectively. Overall agreement is ~85% for both. Agreement between the telephone and mail-administered questionnaire is not as strong, at 68.6% for GOS-E and 86% for GOS.[4] Despite excellent reliability data in many reports, others have reported misclassification rates of 17%–40% for GOS outcomes in clinical trials, with resulting decreases in power.[5]	The GOS is sensitive to recovery from 3–6mo postinjury, but less sensitive for recovery from 6mo–1y postinjury.[1] GOS scores at hospital discharge are not valid predictors of return to work at 6mo and predicted only 6-mo GOS scores for those who did not reach good recovery.[6] The 3-mo GOS score predicted 12mo and 56% showed improvement from 3–12mo, partially countering concerns about sensitivity to change.[7] However, in another study, 3-mo GOS score predicted 15-mo GOS score for patients with good early outcome, but not those with poorer early outcome.[8] Patients in this study had milder injuries. GOS-E scores are associated with neuropsychological test findings and disability measures. indicating validity as an index of TBI outcome.[9]		
MPAI-4	Good ICC[10] by Rasch (person reliability=.88; item reliability=.99) and classic metrics (Cronbach α=.89); good interrater agreement on individual items among staff, patients, and significant others, with 58%–88% agreement within $\pm$1.[11]	Concurrent/construct validity established by correlation with DRS (ρ=.81).[12] Bohac et al[13] reported that MPAI factors correlated with associated neuropsychological measures. Predictive validity shown through correlations of admission MPAI ratings with outpatient rehabilitation outcomes, ie, Goal Attainment Scaling (ρ=−.47), Independent Living Scale (ρ=−.26), Vocational Independence Scale (ρ=−.32).[14] Using logistic regression, Malec et al[15] showed that staff MPAI (χ^2=8.30; P<.01) and time since injury (χ^2=9.70; P<.01) were the best predictors (69% correct classification) of job placement after participation in vocational rehabilitation. Malec[16] found that staff MPAI was the best predictor of long-term vocational (correct classification=67%; χ^2=5.33; P<.05) and independent living outcome (correct classification=70%; χ^2=6.85; P<.01) 1y after completion of comprehensive day rehabilitation in a logistic model that included age, education, severity of injury,	Sensitive to change in studies of rehabilitation interventions[14,16,18] and to frontal lobe damage.[19]	

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
		traumatic vs nontraumatic injury, time since injury, and Rasch-converted staff MPAI score. Malec and Degiorgio[17] reported that logistic regression of the MPAI and time since injury could be used to estimate the probability of community-based employment as a result of outpatient rehabiliitation.		
DRS	IRR of DRS was established among 3 raters on a sample of 88 TBI rehabilitation inpatients.[20] Pearson correlations were .97–.98. In a separate study by Gouvier et al,[21] Spearman ρ correlation coefficient was .98 in 3 raters on a sample of 37–45 subjects. Novack et al[22] reported IRR in a study of 27 severely brain-injured persons. A comparison of DRS ratings by family members vs rehabilitation professionals yielded significant correlations for both rehabilitation admission (r=.95) and discharge (r=.93) ratings. TRT reliability was shown by Gouvier,[21] reporting Spearman ρ correlation of .95.	Concurrent validity was established in the initial publication for the DRS,[20] in which abnormality ratings of auditory, visual, and somatosensory brain evoked potentials significantly correlated with DRS ratings (r=.35–.78). Additional validation of the scale is documented in a published article by Hall et al.[23] Correlation of DRS with simultaneously obtained GOS scores at 2 times was shown in a sample of 70 TBI inpatients (r=.50 at admission, r=.67 at discharge).[24] Gouvier[21] found Spearman ρ correlation was .92 between the rehabilitation admission DRS and Stover Zeiger Scale (1976). Rehabilitation discharge DRS correlated, with .81 for the discharge Stover Zeiger Scale, .80 for the GOS,[1,251] and .85 for the GOS-E.[26]	Rasch analysis was completed on the 8 DRS item scores at rehabilitation admission for 266 cases. Composite scores of 1–29 were obtained (0=normal, 30=dead: clients rated "normal" were omitted). Findings were as follows: relative level of difficulty between admission and discharge ratings of DRS items for 256 cases was consistent; range of difficulty reflected in the scale is excellent, from items measuring very simple functioning to those measuring complex functions.[21] Each of the following domains are scored: Eye Opening, Communication Ability, Motor Response, Feeding, Toileting, Grooming, Level of Functioning, and Employability. Difficulty levels of the 3 items Cognitive Ability for Feeding, Toileting, and Grooming were very similar. There is a gap between Cognitive Ability for Feeding, Toileting and Grooming and Level of Functioning (ie, ability to live independently) and between the latter and Employability. The functional difficulty of each item is substantially different, with no intervening items to reflect intermediate abilities, consistent with the observation of less sensitivity to change in the DRS in people at high functional levels.	A limitation of the DRS is its relative insensitivity at the low end of the scale (mild TBI) and inability to reflect more subtle but sometimes significant changes in a person within a specific limited window of recovery. Average DRS scores at rehabilitation admission, discharge, 1y, and 2y postinjury for all cases with data in the TBIMS database were analyzed for ceiling and floor effects. Ceiling is defined as mean score of 0, 1, or 2 on the DRS (top 10% of scale). These ceiling scores define independent or modified independent status. The DRS has virtually no ceiling effect at discharge, 1y, and 2y postinjury on a consistent sample over time. Results including all cases with data available at any time were similar, with sample sizes ranging from 598–206. The DRS was developed with the continuum of recovery in mind, consistently shows good scale properties, and predicts employment well. At 1y postinjury, 28% of FIM and FIM+FAM scale reflects independence/modified independence (scores=6 and 7 on 7-point scale) and only 10% of DRS summed score represents this level of independence (scores=0, 1 and 2 on 30-point scale). This difference gives the DRS an advantage in regard to ceiling effect.

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
			In summary, Rasch analysis provided transformed scores for use in interval scale data analyses and validated observations about the DRS: a scale that measures a wide range of disability with less sensitivity at the high end (mild TBI). Items discriminate varying levels of disability well and relative difficulty of items remains constant between admission and discharge. In analysis of TBIMS national database data, average DRS untransformed score at rehabilitation admission was 12 (rounded); at discharge, 5; and at 1y postinjury follow-up, 3, in a sample of 70 cases with complete data for all times.	
SF-36	There are >200 studies of ICC and >30 with data for TRT reliability. Reliability estimates of single scales generally are >.80. Sum score reliabilities (physical, mental) generally are >.90. In TBI specifically: MacKenzie et al[27] reported in patients with multiple injuries ((N=1197; 45% with head injuries): α coefficient=.77 (GH) to .93 (Physical Functioning). Findler et al[28] reported scale α coefficients of .79–.92 in moderate/severe TBI patients (N=228), .83–.91 in mild TBI (N=98), and .68–.87 in 271 healthy controls.	Validity has been established in numerous studies.[29,30] In TBI specifically: Findler et al[28] reported convergent validity in 326 patients; correlations of physical SF-36 scales with Physical Symptoms scale of SCL were −.50 to −.63, and with the HPL were −.60 to −.75. There were robust correlations between BDI-II scores and SF-36 scales, with the largest value for Mental Health (−.77). McNaughton et al[31] examined construct validity of the mental and physical CS scores shown in joint factor analysis with several functional measures in 89 patients. In a study examining discriminant validity by Paniak et al,[32] significant differences between 120 MTBI patients and 120 healthy controls in all SF-36 scales (except GH), mental CS, and physical CS were found. In another study of discriminant validity by Emanuelson et al,[33] reduced values were found on all SF-36 subscale, mental CS, and physical CS scores in a study of 173 MTBI patients and age-/sex-matched healthy controls.		

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
Recovery of consciousness				
CRS-R	ICC, TRT reliability, and IRR were shown to be good to excellent by original investigators[34] (IRR, r=.84; TRT, r=.94). Schnakers et al[35] reported IRR to be good (κ=.80) in a French CRS-R validation study.[36] A recently completed Norwegian study found IRR (r=.65–.82) and TRT (r=.77–.83) to be acceptable to very good and noted that both IRR and TRT reliability were influenced by level of experience with the CRS-R. Reliability data also are available for specific CRS-R subscales, with most values in the moderate to good range.[34,36,37] Reliability data for the original version of the CRS were published by Giacino et al[38] and O'Dell et al.[39]	Criterion validity has been shown in comparative analyses with the GCS,[34,40] DRS,[34,37] WHIM,[36] and FOUR.[36] Total CRS-R scores significantly correlated with WHIM, FOUR, and GCS scores in both acute and long-term patient samples.[36] However, CRS-R performance is related most closely to scores on the WHIM (r=.76), a scale designed primarily for use in rehabilitation settings. Lower correlations have been reported with the FOUR (r=.63) and GCS (r=.59), which are intended for use in intensive care and trauma settings, respectively.	Diagnostic validity has been established in 4 separate studies investigating the sensitivity and specificity of the CRS-R for detection of MCS.[34,36,40] Giacino et al[34] found that the CRS-R detected behavioral signs of consciousness in 10 of 80 patients misdiagnosed with VS on the DRS. Similarly, Schnakers et al[40] reported that the CRS-R identified 7 cases of MCS (n=25) in which VS was misdiagnosed by using the FOUR. A more recent study by Schnakers et al[36] found that MCS was diagnosed in 45 of 77 patients with disorders of consciousness after examination with the CRS-R compared with 36, 32, and 24 for the WHIM, FOUR and GCS.	The CRS-R has been used in a range of studies exploring the relationship between behavioral and neurophysiologic markers of consciousness.[41-43] Evidence of cognitive processing after exposure to linguistic stimuli has been reported in 3 fMRI studies involving patients who failed to show behavioral signs of conscious awareness.[41-43] The scale also has been used to characterize the course of recovery from VS, MCS,[44,45] and locked-in syndrome[46] and has sufficient sensitivity to capture salient functional changes associated with pharmacologic interventions[35] and deep brain stimulation.[47]
Neuropsychological impairment				
RAVLT	TRT reliability is good for total recall over 5 trials, .60–.70 over 1y.[48] Internal reliability of total score is high (α coefficients$>$.90).[49]	Extensive literature regarding good validity, including construct, criterion, and predictive. For specific information, refer to Strauss et al.[50]	Sensitive to a variety of diseases of the brain and their severity. Has been used in TBI. Sensitive to change. Has good floor and ceiling. Has extensive norms. Refer to Strauss.[50]	Familiar and widely-used and accepted measure of memory and learning. For many years, it was part of the TBIMS data set. It is a legacy measure of episodic memory of the NIH Toolbox.
TMT	TRT reliability varies with age range and population studied, but is adequate, especially for Part B. Dikmen et al[51] tested 384 healthy adults who were retested 11mo after the initial session. Coefficients were adequate for Part A (.79) and high for Part B (.89). Similar findings were reported by Levine et al.[52]	Parts A and B are moderately intercorrelated (r=.31–.36), suggesting they measure similar, but somewhat different, functions. TMT is sensitive to a wide range of neurologic disorders, including TBI. TMT completion time shows a dose-response relationship with TBI severity: time increases with increasing TBI severity.[55]	Practice effects are found over short retest intervals, but disappear after several administrations. After longer intervals, TMT scores show only modest change in healthy adults. Performance on TMT is affected by age, with performance declining as age increases.	

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
	Reliabilities in clinical groups are not as high. Goldstein & Watson[53] found similar reliability coefficients (.69–.94 for Part A; .66–.86) for various neurologic groups. IRR has been reported as .94 for Part A and .90 for Part B.[54]	Longitudinal studies have reported marked heterogeneity of TMT outcome after moderate to severe TBI; 5y after injury, a substantial proportion of persons with moderate to severe TBI continued to show deficits on TMT. In Millis et al,[56] 43% showed marked impairment on Part A and 33% had impaired performance on Part B. TMT may be less useful in mild TBI. Poor discrimination was reported by Iverson[57] in differentiating mild TBI from substance abuse. Several studies have shown that psychosocial outcome after TBI can be predicted by TMT.	IQ has a moderate relationship with TMT. Sex has little impact on performance. Cultural and linguistic variables may affect test scores.	
WAIS-III Processing Speed Index	Internal consistency is high at .80–.89, as well as TRT reliability at 80–.89 (WAIS/WMS technical manual[58])[50]	Good construct and criterion validity. Very good sensitivity to acquired brain damage. For more specific information, refer to WAISIII/WMS III manual[58] and Strauss.[50]	Extensive normative data through the Wechsler norming and additional studies.[50]	This is a widely known index from WAIS III. It is a legacy measure of Processing Speed for the NIH Toolbox.
BVMT-R	TRT is .80 for total score, which is very good for a memory measure. IRR is high (.90.)[50,59]	Highly correlated with HVLT, VR WMS, and Rey Figure (r=.65–.80); probably measures verbal and nonverbal memory; seems to show reasonable convergent and divergent validities.[50]	A variety of studies support its sensitivity to neurologic condition of the brain.[50]	One of the measures chosen by Matrics for studies in schizophrenia based on extensive review of the literature. It also is 1 of the 2 legacy measures for the Memory Measure of the NIH Toolbox.
WAIS-III Number-Letter Sequencing subtest	ICC is .80–.89: TRT reliability is .70–.79.[50]	Good criterion, construct, discriminant validities. Good clinical sensitivity. Refer to WAIS/WMS III manual[58] and Strauss.[50]	Extensive normative data through the Wechsler plus additional studies	It is a legacy measure for the Working Memory measure for the NIH Toolbox.
COWAT	ICC is high: coefficient α is .83 using total number of words generated for each letter as individual items.[60] In healthy adults, TRT reliability typically is $>$.70.[51] IRR is high (.99[61]).	Correlations among phonemic fluency tasks (eg, FAS, CFL) are high, ranging from .85–.94. Phonemic fluency shows a stronger relationship to Verbal IQ (r=.42–.48) than Performance IQ (r=.29–.36). COWAT is sensitive to severity of TBI.[62] A meta-analysis[63] found that as with patients with focal frontal (but not temporal) lobe injuries, TBI patients were impaired similarly on tests of phonemic and semantic fluency. Phonemic fluency also was significantly more sensitive to the presence of TBI than the Wisconsin Card Sorting Test. In a mixed neurologic sample, Burgess et al[64] found that poor performance on COWAT was moderately associated with caregiver ratings of patients' problems in everyday life and patients' lack of insight into their problems (r=.29–.35).	COWAT has been used in treatment studies (eg, Sarno et al[65]). Higher education level is associated with better performance on COWAT. There is little evidence of sex differences on COWAT.	

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
CWIT	ICC is .70–.79. TRT reliability is .70–.79.	Stroop-like tests frequently have been used in a wide variety of patient groups thought to have executive function deficits. TBI patients typically are slower is responding to all conditions, although they do not consistently show disproportionate impairment on the interference condition (eg, Batchelor et al[66]). Stroop-like tests may have limited diagnostic sensitivity in mild TBI.[67] Baseline interference scores on the Golden-version Stroop were predictive of functional status at 1-y follow-up in patients with vascular dementia.[68]	Although women tend to have superior color-naming skills, sex differences on the color-word interference condition are not consistently present.[69] Education is modestly associated with interference score ($r<.30$[69]).	
WAIS-III Digit Span subtest	TRT stability coefficient is .83. Average reliability coefficient across age is .90.[58]	The Digit Span test, particularly the Digits Backward component, has been identified as a marker of cerebral disorder after TBI.[56,70]		Can be used as a measure of symptom validity.[71,72]
WRAT-4 Word Reading subtest	ICC reliability coefficient[73] is .90–.96 (by age). Alternate form reliability (by age) is .85–.95.	External validity[73]: correlations with the following measures: WIAT-II Word Reading, .71; Woodcock Johnson-III Basic Reading .66; KTEA-II Comprehensive Letter/Word Rec, .76; WAIS-III FSIQ, .79. Studies have focused on earlier versions of the Word Reading subtest, but administration has not changed. With learning disability screened out, WRAT-3 oral reading was considered a reasonable predictor of premorbid ability.[74] Similar results were obtained, with the WRAT-R reading subtest predicting verbal intellectual ability.[75] Stability of WRAT-3 Word Reading across 1y in people with TBI was shown by Orme et al,[76] although slight nonsignificant increases were evident in the most severely injured.		
GPT	TRT reliability[50] is .67–.86 (at 4–24mo).	External validity[50]: correlations with the following measures: Tapping Speed, −.35; Near Visual Acuity, −.62; Reaction Time, .31; TMT-B, .46; Digit Symbol, −.60; Block Design, −.34; Object Assembly, −.45. More than 70% of those experiencing moderate-severe TBI show impairment on the GPT using established cutoff values.[77,78] The GPT is among tests in this population that predict outcome in terms of productivity.[78-80]		

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
Psychological status				
BSI-18	TRT reliability in TBI patients (median retest interval, 1y): GSI of .66; somatization of .67; depression of .63; anxiety of .57.[81] In TBI outpatients (n=176), ICC estimates (Cronbach α) were GSI of .91, somatization of .75, depression of .84, anxiety of .83. For TBI inpatients (n=81), ICC estimates were lower: GSI of .84, somatization of .61, depression of .64, anxiety of .74.[81] In community populations, the BSI-18 has good ICC (coefficient α for global severity=.89, somatization=.74, depressive=.84; anxiety=.89).[82]	Validity analyses using TBI outpatients (n=176) found BSI-18 GSI correlated significantly with psychosocial and functional outcomes: NFI depression (r=.68), NFI aggression (r=.55), PANAS negative affectivity (r=.49).[81] In community populations, BSI-18 scales have shown excellent correlation with the SCL-90-R (Pearson correlation coefficient for global severity=.93, somatization=.91, depressive=.93, anxiety=.96), which in turn has shown acceptable convergent validity with other measures of somatization, depression, and anxiety.[82]	After controlling for demographic and TBI injury characteristics, the BSI-18 accounted for 4% of total variance in FIM scores, 3% of total variance in DRS scores, 3% of total variance in CIM scores, and 8% of total variance in SWLS scores.[81]	
MMPI-2-RF	Extensive psychometric information for the 50 scales of the MMPI-2-RF is presented in chapt 3 of the technical manual.[83] Estimates in the form of Cronbach α coefficient are reported for men and women of the normative sample, an outpatient community mental health sample, a psychiatric inpatient sample tested at a general community hospital, and male psychiatric inpatients tested at a VA hospital. TRT reliability estimates are reported for a combined-sex subset of the MMPI-2 normative sample. Members of the sample completed the MMPI-2 twice, with 1wk between test administrations. Compared with the original MMPI-2, the MMPI-2-RF had similar or improved reliability. TRT correlations and ICC values of the Higher-Order, Restructured Clinical, and PSY-5 scales of the MMPI-2-RF mostly were $>$.80. The α values derived from the normative sample are somewhat lower because of truncated distributions. SEMs generally are $\leq$8 T score points, and most scales have SEMs $\leq$6 points.	The MMPI-2-RF has several validity scales that provide information regarding threats to the validity of a test protocol that must be considered before scores on the clinical scales can be interpreted: inconsistent responding, overreporting, and underreporting indexes. Extensive correlate data are listed in Appendix A of the technical manual.[83] External validity data were gathered from a wide range of settings that document the convergent and discriminant validity and corroborate the construct validity of the substantive scales. Empirical correlates are reported for clinical, forensic, medical, and nonclinical samples. Correlates include a broad range of criteria, including therapist ratings, clinical diagnoses, intake staff ratings, admissions records, biographical information, and other self-report measures.	The 338 items of the MMPI-2-RF are embedded within the MMPI-2 item pool. Hence, MMPI-2-RF profiles can be generated from original MMPI-2 profiles.	
AUDIT	Mean ICC was .83 across 18 studies. TRT reliability κ range was .70–.89 using a cutoff of 8; intraclass correlations range was .87–.95.[84]	Factor analyses indicate a 2 factor structure (consumption and adverse consequences), supporting the use of the abbreviated AUDIT-C as a measure of consumption.	As with other measures, AUDIT does not perform well with the elderly. AUDIT and AUDIT-C have been used successfully	

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
	AUDIT-C TRT was .65–.85 for 3-mo interval,[85] .98 for 1-mo interval.[86]	Sensitivity and specificity have been studied extensively, with satisfactory identification of both hazardous drinking and harmful use. Lower cutoff scores are recommended for women and for the identification of hazardous drinking vs harmful/dependence.[84]	with adolescents, psychiatric populations, and across various countries and cultures, with studies indicating adequate reliability and validity.[87,88] AUDIT has been used with persons with TBI.[89]	
TBIMS questions (based on BRFSS and NSDUH)	BRFSS TRT κ for any alcohol use was .82, and for binge consumption was .64; correlation for number of drinks/mo was .72, with lower values for blacks and Hispanics.[90]	Population estimates derived from BRFSS questions correlate with similarly worded questions from the NSDUH (r=.82). Higher estimates are obtained from the latter version, which has been attributed to computer-assisted administration.[91] The questions have been included in both national surveys and the TBIMS national data set for >10y and have been used in multiple studies for monitoring and analyzing national trends.	Comparisons have been made between the general population and a population of persons hospitalized 1y earlier due to TBI.[92]	
ASSIST	Average TRT κ for question stems ranged from .58–.90; for substance class, from .61 for sedatives to .78 for opioids.[93] Cronbach α was >.80 for most domains.[94]	Concurrent validity: significantly associated with the MINI Plus (r=.93 for lifetime use; r=.76 with MINI-derived score for severity of abuse and dependence), AUDIT (r=.82), and ASI frequency of use (r=.84). Construct validity: significant and positive correlations between ASSIST scores reflecting abuse and dependence and MINI-derived scores of severity of abuse or dependence (r=.76 and r=.75, respectively). Discriminant validity: discriminates between groups classified based on use, abuse, or dependence; better with discriminating between use and abuse (ROC=.84–.97) than between abuse and dependence (ROC=.62–.84, except for sedatives, ROC=.45). Predictive validity: there are no significant differences between ASSIST scores obtained at baseline and 3-mo follow-up.[95]	In a cross-cultural RCT, ASSIST was sensitive to change associated with an ASSIST-linked brief intervention (WHO ASSIST Phase III Study Group).[95]	
PCL-C/M/S	TRT stability coefficient over 2–3d was .96 for Vietnam veterans.[96] ICC α coefficients in Vietnam and Persian Gulf veterans,[96] victims of MVCs, and sexual assault survivors were .97 and .94, respectively, with internal consistencies of .92 to .93 for each subscale).[96]	Combat veterans with PTSD score significantly higher (63.58±14.14 [SD]) than those without PTSD (34.40±14.09).[96] The same pattern is true with MVC-related and sexual assault PTSD.[97] In Vietnam veterans, PCL-M significantly correlated with other measures of PTSD (Spearman ρ range, .77–.93).[96]	Factor analysis for data derived from Persian Gulf war veterans suggested that items are best accounted for by a single factor.[96] Diagnostic sensitivity and specificity of the PCL: cutoff of 50 on the PCL-M resulted in	There are several versions, including the PCL-C, PCL-S, and PCL-M. The PCL-C is available in Spanish. The PCL was developed by the National Center for PTSD and is in the public domain. It maps directly onto DSM criteria.

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
		In Persian Gulf veterans, PCL-M score was associated significantly with another measure of PTSD (.85).[96] The PCL-M highly correlated with the Mississippi Scale for Combat-Related PTSD (.93), the PK scale of the MMPI (.77), and the Impact of Event Scale (.90).[98]	sensitivity of .82 and specificity of .84 in 1 study[96] and sensitivity of .78 and specificity of .86 in another.[97]	Published cutoff values should be used with caution because they were derived from samples with high prevalence rates of current PTSD. Reviews of the PCL can be found in Orsillo[99] and Norris & Hamblen.[100]
FAD	TRT stability coefficients across scales during 1wk ranged from .66–.76.[101] ICC: In a large white sample (n=1302), Cronbach α ranged from .73 for the Roles scale to .87 for the General Function scale; they were slightly lower in a similar Hispanic sample (n=323), ranging from .59 for the Roles scale to .82 for the General Function scale.[102] These were slightly higher than in another study, in which α ranged from .57–.86; only the Roles scale had $\alpha<.70$. In the latter study, α values were marginally higher in clinical than nonclinical samples.[103]	Validity: correlations between FAD scales and clinician ratings of family functioning based on a semistructured clinical interview (McMaster Structured Interview For Families) ranged from .38–.62.[104] Moderate correlation also was found between FAD scales and the FES and scales on the SCL-90-R.[105] Concurrent validity was shown between children's FAD ratings and mother's ratings of family functioning.[106]	One confirmatory factor analytic study supported the factor structure of the FAD, finding similar factor structure in nonclinical, psychiatric, and medical samples,[103] whereas a second found low goodness-of-fit indexes, but good residual error fit indexes.[102]	Reviews of the FAD can be found in Epstein et al.[101] The FAD has shown good reliability and validity across cultural groups, including in China, The Netherlands, Great Britain, Italy, and Canada and in the United States with different racial groups.
TBI-related Symptoms				
RPQ	The measure's developers[107] presented scatterplots (no reliability coefficients reported) that suggest good TRT reliability during a 24-h period at a mean 8d postinjury for 41 adult patients with mild to moderate TBI. A second scatterplot was presented that included 46 adults with mild to severe TBI at ~6mo postinjury that suggested good TRT reliability during a mean 10-d TRT interval.	Significant correlations were reported between a head injury follow-up questionnaire regarding common problems after TBI (eg, problems conversing, problems with facets of community reentry, fatigue at work, getting along with others) and the patient's RPQ total score (Spearman $\rho=.67$ at 3mo postinjury and $\rho=.56$ at 6mo postinjury).[108] No differences were found between patient-completed and interview-format responses. Modest predictive validity ($r=.37$; $P<.05$) was reported[109] between 1-wk and 6-mo RPQ scores. At 3mo after mild TBI, the RPQ distinguished between patients with and without PCS, and those who were vs were not "on sick leave" from work. Ingebrigtsen et al.[110] reported a trend between RPQ total score and serum S-100B protein level in patients with mild TBI 24h postinjury. However, Savola & Hillbom[111] found that S-100B on hospital admission was a significant predictor of RPQ total score at 1mo postinjury.	RPQ total score does not appear to be associated with age, sex, cause of injury, severity of injury (GCS score), or duration of PTA in patients with mild TBI.[114] Chan[115] found no sex effect on RPQ total score. Eyres et al[116] reported that all RPQ items functioned well across age and sex. No differences in RPQ total score were noted between patients with chronic pain and mild TBI.[117] A significant difference was found on RPQ total scores between adolescents with mild TBI and an uninjured control group assessed an average of 3d postinjury.[113]	Rasch analysis suggested the RPQ was not unidimensional and the authors suggested splitting 3 items (headache, dizziness, nausea) into a separate scale. The resulting RPQ-13 and RPQ-3 performed well in terms of external construct validity with a head injury follow-up questionnaire (RPQ-13, .83; RPQ-3, .62) and 2-wk TRT reliability (RPQ-13, .89; RPQ-3, .72).

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
		Higher RPQ total scores related to greater activations in fMRI tasks of working memory and selective attention in patients with mild TBI.[112] FA of the corpus callosum measured by using DTI significantly related (ρ =.76) to RPQ total score in a sample of adolescents with mild TBI assessed an average of 3d postinjury.[113]		
NSI	Schwab et al[118] reported that Afghanistan and Iraq war veterans reporting a probable TBI had a higher prevalence of having ≥3 problematic PCS (using the NSI) symptoms than that of veterans who did not report TBI (64% vs 41%; $P<.001$).	Schwab[118] reported a moderate association between number of TBI-related problems reported on a TBI screening interview and number of moderate/severe PCS symptoms reported on the NSI ($r=.477$; $P<.001$). Schwab[118] also showed that in participants reporting a TBI on a brief TBI screening interview, the prevalence of ≥3 moderate/severe PCS symptoms was higher in those with than in those without self-reported TBI-related problems (74% vs 40%; $P=.003$).		
Behavioral function				
FrSBe	Intrascale reliability (normative sample)[119]: Family (total), .92; Apathy, .78; Disinhibition, .80; Ex Dysfunction, .87. Self-rating (total), .88; Apathy, .72; Disinhibition, .75; Ex Dysfunction, .79. Intrascale reliability (neurologic sample): Family (total), .94; Apathy, .87; Disinhibition, .84 Ex Dysfunction: .91. Self-rating (total), .92; Apathy, .83; Disinhibition, .78; Ex Dysfunction, .84.	Correlation of FrSBe to Neuropsychiatric Inventory[120]: total, $r=.64$; $P<.001$; Apathy, $r=.37$; $P=.04$; Disinhibition, $r=.62$; $P<.001$. Construct validity: the FrSBe has shown significant differences in before and after ratings for people with frontal system lesions[121] and also has differentiated frontal lesion populations from controls.[122] The FrSBe has shown stronger correlation with a measure of community reentry than tests of executive functioning.[123]	Principal factor analysis yielded 3 factors corresponding to a priori domains of apathy, disinhibition, and executive dysfunction that accounted for >46% of variance.[124]	
Cognitive activity limitation				
Cog-FIM Please see next section for general information about the FIM instrument. This section provides information specific to Cog-FIM items.	Interrater reproducibility was .95.[125] Both Cog-FIM and Motor FIM have excellent ICC[126] (α=.86–.97; α=.89 for Cog-FIM).[125]	Correlates (.51) with WAIS VIQ.[125] Low correlations (discriminant validity) with physical and mental health status measures.[125] Predicts amount of supervision (vs physical assistance) received in the home setting.[127] Predicts falls more robustly than Motor FIM in rehabilitation inpatients.[128]	Ceiling issues: in the TBIMS, maximum score (all items 7) was attained at 1y postinjury by 16% using the Cog-FIM total; 20% for Memory, 56% for Social Interaction, 45% for Comprehension/Expression.[129] Between 1 & 5y postinjury, 26% improved on Cog-FIM, 61% stayed the same, 14% worsened.[129]	Routinely collected at most rehabilitation facilities, so it may be cost-effective to obtain in that setting. Extensive use in studies of TBI and a wide range of other patient populations. Allows comparison across patient populations. Sensitive to cognitive rehabilitation vs functional skills training approach in RCT of subacute TBI.[130]

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
Physical function				
FIM	Total FIM has outstanding reliability, with TRT reliability, IRR, and ICC much >0.90.[131]	Extensive demonstration of construct and predictive validity in TBI and a wide range of other patient populations. The FIM has clinically appropriate validity and interrater agreement.[131]	Sensitive to improvements in function for up to 1y post-TBI. Evaluation of the metric properties of the FIM has been reported extensively.[132-136] Precision (ability of the instrument to detect meaningful change in level of function during rehabilitation) has been high.[137] In a Rasch analysis of the FIM, 2 separate domains of items were defined: the motor domain consisting of 13 items and the cognitive domain consisting of 5 items.[133,136] Previous analyses of FIM data from the SCI Model Systems suggested that the cognitive domain may be inappropriate for people with SCI.[138] Ceiling effects of the FIM at rehabilitation discharge and particularly at 1y postinjury were observed in the moderate and severely injured TBI population.[139] Of the sample, 49% and 84% had attained independence (average score, 7 or 6) by discharge and 1y postinjury, respectively; ie, the FIM is not sensitive to more subtle changes expected after acute inpatient rehabilitation discharge.	Possible ceiling effect after 1y. Because FIM ratings affect facility reimbursement and many facilities use FIM change as a quality indicator, there may be pressure for the lowest possible admission FIM and highest possible discharge FIM scores to be obtained if rated by clinicians.
Social role participation				
CHART-SF	CHART-SF subscales closely approximate scores of subscales gathered by the original CHART. Reliability data are based on original CHART without Cognitive Domain.[140] IRR (1-wk interval) was .80–.95 (n=135 with SCI). Subject-proxy was .69 (economic self-sufficiency), .28 (social integration), 0.80–.83 for remaining scales (n=135 persons with SCI and proxies).	CHART-SF subscales closely approximate scores of subscales gathered by the original CHART. Validity data are based on original CHART without Cognitive Domain.[140] Rehabilitation professionals rated 135 persons with SCI as either high or low levels of handicap. CHART scores and subscales (except for economic self-sufficiency) were significantly different in the expected direction.		

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
	TRT (2-wk interval) for Cognitive Independence was .87.[141] Subject-proxy for Cognitive Independence was .81.	Rasch analysis indicated satisfactory separation of items along the handicap dimension. Items within each subscale fit well. A sample of 2259 participants weighted to represent the population of Colorado in 1999 showed that those who reported no activity limitations scored significantly higher than those who reported activity limitations on the BRFSS. A sample of 1110 participants was administered the CHART (including Cognitive Independence), persons with TBI or stroke had lower scores than those with multiple sclerosis, SCI, amputation, or burn.[141,142] Correlation coefficients were higher between CHART Cognitive Independence and Cog-FIM subscale compared with CHART Cognitive Independence and FIM motor subscale.[141]		
Perceived generic and disease-specific HRQOL				
SWLS[143]	ICC was $>.80$.[144] TRT (2-mo interval) was .82 in 76 students.[143] With 2-wk interval, it was .89.[144] Factor and Rasch analysis support a single factor; however, item 5 (If I could live my life over, I would change almost nothing) is the least well associated.	Content validity: initially 48 items were included; factor analysis showed that 10 items loaded highly ($>.60$) on a factor reflecting cognitive-judgmental evaluative processes; 5 items were redundant, resulting in the current 5-item scale. Criterion-validity: original validation studies compared SWLS scores with 10 measures of subjective well-being; all correlated at $r \geq .50$. Construct validity: TRT stability decreases as interval increases. Consistent differences between populations in the expected directions have been found. Scores change in expected directions when major life events occur.	Normative data are available for people with TBI (TBIMS national database).[145-147] Norms are available for other populations.[144]	
Health economic measures				
EuroQOL[148]	The MVH Group[149] reported that a large valuation study of a general British population of 221 respondents had very reliable mean ICCs of .78 for questions and .73 for the VAS. Van Agt et al[150] assessed TRT reliability in a Dutch population (N=208) after TBI by using several specific method approaches that indicated good TRT reliability.	Brazier et al[151] found evidence for construct validity of the EuroQol comparing it with the SF-36 in a large British sample (N=1582). Sintonen[152] reported correlations of the 15-D HRQOL with EuroQol, among others, giving evidence for construct validity. In an RA study, Hurst et al[153] showed clinically relevant correlations with other condition-specific instruments, indicating EuroQol construct validity in RA.	Sensitive to change-intervention effects.	

Supplementary Table 1 (Cont'd): Psychometric Properties of Recommended TBI Outcomes CDEs in Core and Supplemental Tiers

Measure	Reliability	Validity	Additional Psychometric	Other
		In TBI, Klose et al[154] found decreased scores on the EuroQoL VAS in patients with posttraumatic hypopituitarism 12mo after injury. Bell et al[155] reported, among other measures, significantly increased EuroQoL scores as an effect of a scheduled telephone intervention in patients with moderate to severe TBI.		

NOTE. Supplemental measures are as follows: Global outcome: MPAI-4, DRS, and SF-36; Recovery of consciousness: CRS-R; Neuropsychological impairment: BVMT-R, WAIS-III Number-Letter Sequencing subtest, COWAT, CWIT, WAIS-III Digit Span subtest, WRAT-4 Word Reading subtest, and GPT; Psychological status: MMPI-2-RF, AUDIT, TBIMS questions (based on BRFSS and NSDUH), ASSIST, PCL-C/M/S, and FAD; Postconcussive/TBI symptoms: NSI; and Behavioral function: FrSBe. shaded in gray.
Abbreviations: ASI, Addiction Severity Index; AUDIT-C, Alcohol Use Disorders Identification Test Alcohol Consumption Questions; BDI-II, Beck Depression Inventory-II; BRFSS, Behavioral Risk Factor Surveillance Survey; CFL, CFL verbal fluency test; CIM, Community Integration Measure; CS, component summary; DSM, *Diagnostic and Statistical Manual of Mental Disorders;* DTI, diffusion tensor imaging; FA, fractional anisotropy; FAM, Functional Assessment Measure; FAS, FAS verbal fluency test; FES, Family Environment Scale; fMRI, functional magnetic resonance imaging; FOUR, Full Outline of Unresponsiveness; FSIQ, full scale intelligence quotient; GCS, Glasgow Coma Scale; GH, General Health; GSI, ; HPL, Health Problems List; HRQOL, health-related quality of life; HVLT, Hopkins Verbal Learning Test; ICC, internal consistency; IRR, interrater reliability; KTEA-II, Kaufman Test of Educational Achievement, Second Edition; MINI Plus, Mini International Neuropsychiatric Interview Plus; MTBI, mild TBI; MVC, motor vehicle collision; MVH, measurement of valuation of health; NFI, Neurobehavioral Functioning Inventory; NSDUH, National Survey on Drug Use and Health; PANAS, Positive and Negative Affect Schedule; PCS, Postconcussion Syndrome; PK, Keane PTSD scales of the MMPI; PTA, posttraumatic amnesia; PSY-5, Personality Psychopathology Five; RA, rheumatoid arthritis; RCT, randomized controlled trial; ROC, receiver operating characteristic; SCI, spinal cord injury; SCL, Symptoms Check List; SCL-90-R, Symptom Checklist-90-Revised; TBIMS, TBI Model Systems; TRT, test-retest; VAS, visual analog scale; VIQ, verbal intelligence quotient; VS, vegetative state; VR WMS, Visual Reproduction subscale of the Wechsler Memory Scale; WHIM, Wessex Head Injury Matrix; WIAT-II, Wechsler Individual Achievement Test, Second Edition; WMS III, Wechsler Memory Scale, Third edition.

References

1. Jennett B, Snoek J, Bond MR, Brooks N. Disability after severe head injury: observations on the use of the Glasgow Outcome Scale. J Neurol Neurosurg Psychiatry 1981;44:285-93.
2. Wilson JT, Pettigrew LE, Teasdale GM. Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. J Neurotrauma 1998; 15:573-85.
3. Pettigrew LE, Wilson JT, Teasdale GM. Reliability of ratings on the Glasgow Outcome Scales from in-person and telephone structured interviews. J Head Trauma Rehabil 2003;18:252-8.
4. Wilson JT, Edwards P, Fiddes H, Stewart E, Teasdale GM. Reliability of postal questionnaires for the Glasgow Outcome Scale. J Neurotrauma 2002;19:999-1005.
5. Lu J, Murray GD, Steyerberg EW, et al. Effects of Glasgow Outcome Scale misclassification on traumatic brain injury clinical trials. J Neurotrauma 2008;25:641-51.
6. Gabbe BJ, Simpson PM, Sutherland AM, et al. Functional measures at discharge: are they useful predictors of longer term outcomes for trauma registries? Ann Surg 2008;247:854-9.
7. King JT Jr, Carlier PM, Marion DW. Early Glasgow Outcome Scale scores predict long-term functional outcome in patients with severe traumatic brain injury. J Neurotrauma 2005;22: 947-54.
8. Miller KJ, Schwab KA, Warden DL. Predictive value of an early Glasgow Outcome Scale score: 15-month score changes. J Neurosurg 2005;103:239-45.
9. Levin HS, Boake C, Song J, et al. Validity and sensitivity to change of the extended Glasgow Outcome Scale in mild to moderate traumatic brain injury. J Neurotrauma 2001;18:575-84.
10. Malec JF, Kragness M, Evans RW, Finlay KL, Kent A, Lezak MD. Further psychometric evaluation and revision of the Mayo-Portland Adaptability Inventory in a national sample. J Head Trauma Rehabil 2003;18:479-92.
11. Malec JF. Comparability of Mayo-Portland Adaptability Inventory ratings by staff, significant others and people with acquired brain injury. Brain Inj 2004;18:563-75.
12. Malec JF, Thompson JM. Relationship of the Mayo-Portland Adaptability Inventory to functional outcome and cognitive performance measures. J Head Trauma Rehabil 1994;9:1-15.
13. Bohac DL, Malec JF, Moessner AM. Factor analysis of the Mayo-Portland Adaptability Inventory: structure and validity. Brain Inj 1997;11:469-82.
14. Malec JF, Moessner AM, Kragness M, Lezak MD. Refining a measure of brain injury sequelae to predict postacute rehabilitation outcome: rating scale analysis of the Mayo-Portland Adaptability Inventory. J Head Trauma Rehabil 2000;15:670-82.
15. Malec JF, Buffington AL, Moessner AM, Degiorgio L. A medical/vocational case coordination system for persons with brain injury: an evaluation of employment outcomes. Arch Phys Med Rehabil 2000;81:1007-15.
16. Malec JF. Impact of comprehensive day treatment on societal participation for persons with acquired brain injury. Arch Phys Med Rehabil 2001;82:885-95.
17. Malec JF, Degiorgio L. Characteristics of successful and unsuccessful completers of 3 postacute brain injury rehabilitation pathways. Arch Phys Med Rehabil 2002;83:1759-64.
18. Constantinidou F, Thomas RD, Scharp VL, Laske KM, Hammerly MD, Guitonde S. Effects of categorization training in patients with TBI during postacute rehabilitation: preliminary findings. J Head Trauma Rehabil 2005;20:143-57.
19. Murrey GJ, Hale FM, Williams JD. Assessment of anosognosia in persons with frontal lobe damage: clinical utility of the Mayo-Portland Adaptability Inventory (MPAI). Brain Inj 2005;19:599-603.

20. Rappaport M, Hall KM, Hopkins K, Belleza T, Cope DN. Disability rating scale for severe head trauma: coma to community. Arch Phys Med Rehabil 1982;63:118-23.
21. Gouvier WD, Blanton PD, LaPorte KK, Nepomuceno C. Reliability and validity of the Disability Rating Scale and the Levels of Cognitive Functioning Scale in monitoring recovery from severe head injury. Arch Phys Med Rehabil 1987;68:94-7.
22. Novack TA, Bergquist TF, Bennett G, Gouvier WD. Primary caregiver distress following severe head injury. J Head Trauma Rehabil 1991;6:69-77.
23. Hall KM, Hamilton BB, Gordon WA, Zasler ND. Characteristics and comparisons of functional assessment indices: Disability Rating Scale, Functional Independence Measure and Functional Assessment Measure. J Head Trauma Rehabil 1993;8:60-74.
24. Hall K, Cope DN, Rappaport M. Glasgow Outcome Scale and Disability Rating Scale: comparative usefulness in following recovery in traumatic head injury. Arch Phys Med Rehabil 1985;66:35-7.
25. Jennett B, Bond M. Assessment of outcome after severe brain damage. Lancet 1975;1:480-4.
26. Smith RM, Fields FRJ, Lenox JL, Morris HO, Nolan JJ. A functional scale of recovery from severe head trauma. Clin Neuropsychol 1979;1:48-50.
27. MacKenzie EJ, McCarthy ML, Ditunno JF, et al. Using the SF-36 for characterizing outcome after multiple trauma involving head injury. J Trauma 2002;52:527-34.
28. Findler M, Cantor J, Haddad L, Gordon W, Ashman T. The reliability and validity of the SF-36 health survey questionnaire for use with individuals with traumatic brain injury. Brain Inj 2001;15:715-23.
29. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
30. Ware JE, Gandek B, Aaronson NK, et al. The SF-36 Health Survey: development and use in mental health research and the IQOLA project. Int J Mental Health 1994;23:49-74.
31. McNaughton HK, Weatherall M, McPherson KM. Functional measures across neurologic disease states: analysis of factors in common. Arch Phys Med Rehabil 2005;86:2184-8.
32. Paniak C, Phillips K, Toller-Lobe G, Durand A, Nagy J. Sensitivity of three recent questionnaires to mild traumatic brain injury-related effects. J Head Trauma Rehabil 1999;14:211-9.
33. Emanuelson I, Andersson Holmkvist E, Bjorklund R, Stalhammar D. Quality of life and post-concussion symptoms in adults after mild traumatic brain injury: a population-based study in western Sweden. Acta Neurol Scand 2003;108:332-8.
34. Giacino JT, Kalmar K, Whyte J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabil 2004;85:2020-9.
35. Schnakers C, Hustinx R, Vandewalle G, et al. Measuring the effect of amantadine in chronic anoxic minimally conscious state. J Neurol Neurosurg Psychiatry 2008;79:225-7.
36. Schnakers C, Majerus S, Giacino J, et al. A French validation study of the Coma Recovery Scale-Revised (CRS-R). Brain Inj 2008;22:786-92.
37. Lovstad M, Froslie K, Giacino JT, Anke A, Schanke A. Reliability and diagnostic characteristics of the JFK Coma Recovery Scale-Revised: exploring the influence of rater's level of experience. J Head Trauma Rehabil 2010;25:349-56.
38. Giacino JT, Kezmarsky MA, DeLuca J, Cicerone KD. Monitoring rate of recovery to predict outcome in minimally responsive patients. Arch Phys Med Rehabil 1991;72:897-901.
39. O'Dell MW, Jasin P, Stivers M, Lyons N, Schmidt S, Moore DE. Interrater reliability of the Coma Recovery Scale. J Head Trauma Rehabil 1996;11:61-6.
40. Schnakers C, Giacino J, Kalmar K, et al. Does the FOUR score correctly diagnose the vegetative and minimally conscious states? Ann Neurol 2006;60:744-5; author reply 745.
41. Coleman MR, Rodd JM, Davis MH, et al. Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. Brain 2007;130:2494-507.
42. Coleman MR, Davis MH, Rodd JM, et al. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. Brain 2009;132:2541-52.
43. Giacino JT, Schnakers C, Rodriguez-Moreno D, Kalmar K, Schiff N, Hirsch J. Behavioral assessment of patients with disorders of consciousness: gold standard or fool's gold? Prog Brain Res 2009;177:33-48.
44. Giacino JT, Kalmar K. The vegetative and minimally conscious states: a comparison of clinical features and functional outcome. J Head Trauma Rehabil 1997;12:36-51.
45. Giacino JT, Trott CT. Rehabilitative management of patients with disorders of consciousness: grand rounds. J Head Trauma Rehabil 2004;19:254-65.
46. Smart CM, Giacino JT, Cullen T, et al. A case of locked-in syndrome complicated by central deafness. Nat Clin Pract Neurol 2008;4:448-53.
47. Schiff ND, Giacino JT, Kalmar K, et al. Behavioural improvements with thalamic stimulation after severe traumatic brain injury. Nature 2007;448:600-3.
48. Mitrushina M, Satz P. Effect of repeated administration of a neuropsychological battery in the elderly. J Clin Psychol 1991; 47:790-801.
49. van den Burg W, Kingma A. Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): parallel test-retest reliabilities with an interval of 3 months and normative data. Arch Clin Neuropsychol 1999;14:545-59.
50. Strauss E, Sherman EMS, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. New York: Oxford University Pr; 2006.
51. Dikmen SS, Heaton RK, Grant I, Temkin NR. Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. J Int Neuropsychol Soc 1999;5:346-56.
52. Levine AJ, Miller EN, Becker JT, Selnes OA, Cohen BA. Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. Clin Neuropsychol 2004;18:373-84.
53. Goldstein G, Watson J. Test-retest reliability of the Halstead-Reitan battery and the WAIS in a neuropsychiatric population. Clin Neuropsychol 1989;3:265-73.
54. Fals-Stewart W. An interrater reliability study of the Trail Making Test (Parts A and B). Percept Mot Skills 1992;74:39-42.
55. Dikmen S, Machamer JE, Winn HR, Temkin NR. Neuropsychological outcome at 1 year post head injury. Neuropsychology 1995;9:80-90.
56. Millis SR, Rosenthal M, Novack TA, et al. Long-term neuropsychological outcome after traumatic brain injury. J Head Trauma Rehabil 2001;16:343-55.
57. Iverson GL. Outcome from mild traumatic brain injury. Curr Opin Psychiatry 2005;18:301-17.
58. Wechsler D. Wechsler Adult Intelligence Scale III. San Antonio: Harcourt Assessment Inc; 1997.
59. Benedict RH. Brief Visuospatial Memory Test-Revised. Odessa: Psychological Assessment Resources Inc; 1997.
60. Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol 1999;14:167-77.
61. Ross TP. The reliability of cluster and switch scores for the Controlled Oral Word Association Test. Arch Clin Neuropsychol 2003;18:153-64.

62. Iverson GL, Franzen MD, Lovell MR. Normative comparisons for the controlled oral word association test following acute traumatic brain injury. Clin Neuropsychol 1999;13:437-41.
63. Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. Neuropsychology 2004;18:621-8.
64. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. J Int Neuropsychol Soc 1998;4:547-58.
65. Sarno MT, Postman WA, Cho YS, Norman RG. Evolution of phonemic word fluency performance in post-stroke aphasia. J Commun Disord 2005;38:83-107.
66. Batchelor J, Harvey AG, Bryant RA. Stroop colour word test as a measure of attentional deficit following mild head injury. Clin Neuropsychol 1995;9:180-6.
67. Cicerone KD, Azulay J. Diagnostic utility of attention measures in postconcussion syndrome. Clin Neuropsychol 2002;16:280-9.
68. Boyle PA, Paul RH, Moser DJ, Cohen RA. Executive impairments predict functional declines in vascular dementia. Clin Neuropsychol 2004;18:75-82.
69. Ivnik RJ, Malec J, Smith GE, Tangalos EG. Neuropscyhological test norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, Stroop, TMT, and JLO. Clin Neuropsychol 1996;10:262-78.
70. Kersel DA, Marsh NV, Havill JH, Sleigh JW. Neuropsychological functioning during the year following severe traumatic brain injury. Brain Inj 2001;15:283-96.
71. Greiffenstein MF, Baker WJ, Gola T. Validation of malingered amnesia meaures with a large clinical sample. Psychol Assess 1994;6:218-24.
72. Myers JE, Volbrecht M. Validation of reliable digits for detection of malingering. Assessment 1999;5:303-7.
73. Wilkinson GS, Robertson GJ. Wide range achievement test-4 professional manual. Lutz: Psychological Assessment Resources Inc; 2006.
74. Ball JD, Hart RP, Stutts ML, Turf E, Barth JT. Comparative utility of Barona Formulae, Wtar demographic algorithms, and WRAT-3 reading for estimating premorbid ability in a diverse research sample. Clin Neuropsychol 2007;21:422-33.
75. Johnstone B, Callahan CD, Kapila CJ, Bouman DE. The comparability of the WRAT-R reading test and NAART as estimates of premorbid intelligence in neurologically impaired patients. Arch Clin Neuropsychol 1996;11:513-9.
76. Orme DR, Johnstone B, Hanks R, Novack TA. The WRAT-3 Reading Subtest is a measure of premorbid intelligence among persons with brain injury. Rehabil Psychol 2004;49:250-3.
77. Kreutzer JS, Gordon WA, Rosenthal M, Marwitz J. Neuropsychological characteristics of patients with brain injury: preliminary findings from a multicenter investigation. J Head Trauma Rehabil 1993;8:47-59.
78. Boake C, Millis SR, High WM Jr, et al. Using early neuropsychologic testing to predict long-term productivity outcome from traumatic brain injury. Arch Phys Med Rehabil 2001;82:761-8.
79. Ip RY, Dornan J, Schentag C. Traumatic brain injury: factors predicting return to work or school. Brain Inj 1995;9:517-32.
80. Sherer M, Sander AM, Nick TG, High WM Jr, Malec JF, Rosenthal M. Early cognitive status and productivity outcome after traumatic brain injury: findings from the TBI model systems. Arch Phys Med Rehabil 2002;83:183-92.
81. Meachen SJ, Hanks RA, Millis SR, Rapport LJ. The reliability and validity of the Brief Symptom Inventory-18 in persons with traumatic brain injury. Arch Phys Med Rehabil 2008;89:958-65.
82. Derogatis LR. Brief Symptom Inventory 18 (BSI 18): administration, scoring and procedures manual. Minneapolis: NCS Pearson Inc; 2001.
83. Tellegen A, Ben-Porath YS. MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2): technical manual. Minneapolis: Univ of Minneapolis Pr; 2008.
84. Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test: an update of research findings. Alcohol Clin Exp Res 2007;31:185-99.
85. Bradley KA, Bush KR, McDonell MB, Malone T, Fihn SD. Screening for problem drinking: comparison of CAGE and AUDIT. J Gen Intern Med 1998;13:379-88.
86. Bergman H, Kallmen H. Alcohol use among Swedes and a psychometric evaluation of the Alcohol Use Disorders Identification Test. Alcohol Alcohol 2002;37:245-51.
87. Moussas G, Dadouti G, Douzenis A, et al. The Alcohol Use Disorders Identification Test (AUDIT): reliability and validity of the Greek version. Ann Gen Psychiatry 2009;8:11.
88. Hides L, Cotton SM, Berger G, et al. The reliability and validity of the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in first-episode psychosis. Addict Behav 2009;34:821-5.
89. Ponsford J, Whelan-Goodinson R, Bahar-Fuchs A. Alcohol and drug use following traumatic brain injury, a prospective study. Brain Inj 2007;21:1385-92.
90. Stein AD, Lederman RI, Shea S. The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample. Am J Public Health 1993;83:1768-72.
91. Miller JW, Gfroerer JC, Brewer RD, Naimi TS, Mokdad A, Giles WH. Prevalence of adult binge drinking: a comparison of two national surveys. Am J Prev Med 2004;27:197-204.
92. Horner MD, Ferguson PL, Selassie AW, Labbate LA, Kniele K, Corrigan JD. Patterns of alcohol use 1 year after traumatic brain injury: a population-based, epidemiological study. J Int Neuropsychol Soc 2005;11:322-30.
93. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction 2002;97:1183–94.
94. Humeniuk R, Ali R, Babor TF, et al. Validation of the Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST). Addiction 2008;103:1039-47.
95. Humeniuk R, Dennington V, Ali R. The effectiveness of a brief intervention for illicit drugs linked to the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in primary health care settings: a technical report of phase III findings of the WHO ASSIST randomized controlled trial. Geneva: World Health Organization; 2008.
96. Weathers F, Litz B, Herman D, Huska J, Keane T. The PTSD Checklist (PCL): reliability, validity, and diagnostic utility. 9th Annual Convention of the International Society for Traumatic Stress Studies, October 1993, San Antonio, TX. Available at: http://www.pdhealth.mil/library/downloads/PCL_sychometrics.doc. Accessed October 15, 2010.
97. Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. Psychometric properties of the PTSD Checklist (PCL). Behav Res Ther 1996;34:669-73.
98. Keane TM, Caddell JM, Taylor KL. Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: three studies in reliability and validity. J Consult Clin Psychol 1988;56:85-90.
99. Orsillo SM. Measures for acute stress disorder and posttraumatic stress disorder. In: Antony MM, Orsillo SM, editors. Practitioner's guide to empirically based measures of anxiety. New York: Kluwer Academic/Plenum; 2001. p 255-307.
100. Norris FH, Hamblen JS. Standardized self-report measures of civilian trauma and PTSD. In: Wilson JP, Keane TM, Martin T, editors. Assessing psychological trauma and PTSD. New York: Guilford Pr; 2004. p 63-102.

101. Epstein NB, Baldwin LM, Bishop DS. Family Assessment Device (FAD). Handbook of psychiatric measures. Washington (DC): American Psychiatric Association; 2000.
102. Aarons GA, McDonald EJ, Connelly CD, Newton RR. Assessment of family functioning in Caucasian and Hispanic Americans: reliability, validity, and factor structure of the Family Assessment Device. Fam Process 2007;46:557-69.
103. Kabacoff RI, Miller IW, Bishop DS, Epstein NB, Keitner GI. A psychometric study of the McMaster Family Assessment Device in psychiatric, medical, and nonclinical samples. J Fam Psychol 1990;3:431-9.
104. Barney MC, Max JE. The McMaster Family Assessment Device and clinical rating scale: questionnaire vs interview in childhood traumatic brain injury. Brain Inj 2005;19:801-9.
105. Wenniger WFDB, Hageman WJJM, Arrindell WA. Cross-national validity of dimensions of family functioning: first experiences with the Dutch version of the McMaster Family Assessment Device (FAD). Pers Individual Differences 1993;14:769-81.
106. Bihun JT, Wamboldt MZ, Gavin LA, Wamboldt FS. Can the Family Assessment Device (FAD) be used with school aged children? Fam Process 2002;41:723-31.
107. King NS, Crawford S, Wenden FJ, Moss NE, Wade DT. The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. J Neurol 1995;242:587-92.
108. Crawford S, Wenden FJ, Wade DT. The Rivermead head injury follow up questionnaire: a study of a new rating scale and other measures to evaluate outcome after head injury. J Neurol Neurosurg Psychiatry 1996;60:510-4.
109. King NS, Crawford S, Wenden FJ, Caldwell FE, Wade DT. Early prediction of persisting post-concussion symptoms following mild and moderate head injuries. Br J Clin Psychol 1999; 38(Pt 1):15-25.
110. Ingebrigtsen T, Romner B, Marup-Jensen S, et al. The clinical value of serum S-100 protein measurements in minor head injury: a Scandinavian multicentre study. Brain Inj 2000;14:1047-55.
111. Savola O, Hillbom M. Early predictors of post-concussion symptoms in patients with mild head injury. Eur J Neurol 2003;10: 175-81.
112. Smits M, Dippel DW, Houston GC, et al. Postconcussion syndrome after minor head injury: brain activation of working memory and attention. Hum Brain Mapp 2009;30:2789-803.
113. Wilde EA, McCauley SR, Hunter JV, et al. Diffusion tensor imaging of acute mild traumatic brain injury in adolescents. Neurology 2008;70:948-55.
114. Ingebrigtsen T, Waterloo K, Marup-Jensen S, Attner E, Romner B. Quantification of post-concussion symptoms 3 months after minor head injury in 100 consecutive patients. J Neurol 1998; 245:609-12.
115. Chan RC. Base rate of post-concussion symptoms among normal people and its neuropsychological correlates. Clin Rehabil 2001; 15:266-73.
116. Eyres S, Carey A, Gilworth G, Neumann V, Tennant A. Construct validity and reliability of the Rivermead Post-Concussion Symptoms Questionnaire. Clin Rehabil 2005;19:878-87.
117. Smith-Seemiller L, Fow NR, Kant R, Franzen MD. Presence of post-concussion syndrome symptoms in patients with chronic pain vs mild traumatic brain injury. Brain Inj 2003;17:199-206.
118. Schwab KA, Ivins B, Cramer G, et al. Screening for traumatic brain injury in troops returning from deployment in Afghanistan and Iraq: initial investigation of the usefulness of a short screening tool for traumatic brain injury. J Head Trauma Rehabil 2007;22:377-89.
119. Grace J, Malloy PF. Frontal Systems Behavior Scale professional manual. Lutz: Psychological Assessment Resources Inc; 2001.
120. Caracuel A, Verdejo-Garcia A, Vilar-Lopez R, et al. Frontal behavioral and emotional symptoms in Spanish individuals with acquired brain injury and substance use disorders. Arch Clin Neuropsychol 2008;23:447-54.
121. Grace J, Stout JC, Malloy PF. Assessing frontal lobe behavioral syndromes with the frontal lobe personality scale. Assessment 1999;6:269-84.
122. Norton LE, Malloy PF, Salloway S. The impact of behavioral symptoms on activities of daily living in patients with dementia. Am J Geriatr Psychiatry 2001;9:41-8.
123. Reid-Arndt SA, Nehl C, Hinkebein J. The Frontal Systems Behaviour Scale (FrSBe) as a predictor of community integration following a traumatic brain injury. Brain Inj 2007;21:1361-9.
124. Stout JC, Ready RE, Grace J, Malloy PF, Paulsen JS. Factor analysis of the Frontal Systems Behavior Scale (FrSBe). Assessment 2003;10:79-85.
125. Hobart JC, Lamping DL, Freeman JA, et al. Evidence-based measurement: which disability scale for neurologic rehabilitation? Neurology 2001;57:639-44.
126. Stineman MG, Shea JA, Jette A, et al. The Functional Independence Measure: tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories. Arch Phys Med Rehabil 1996;77:1101-8.
127. Corrigan JD, Smith-Knapp K, Granger CV. Validity of the Functional Independence Measure for persons with traumatic brain injury. Arch Phys Med Rehabil 1997;78:828-34.
128. Lee JE, Stokic DS. Risk factors for falls during inpatient rehabilitation. Am J Phys Med Rehabil 2008;87:341-50; quiz, 51, 422.
129. Hammond FM, Hart T, Bushnik T, Corrigan JD, Sasser H. Change and predictors of change in communication, cognition, and social function between 1 and 5 years after traumatic brain injury. J Head Trauma Rehabil 2004;19:314-28.
130. Vanderploeg RD, Schwab K, Walker WC, et al. Rehabilitation of traumatic brain injury in active duty military personnel and veterans: Defense and Veterans Brain Injury Center randomized controlled trial of two rehabilitation approaches. Arch Phys Med Rehabil 2008;89:2227-38.
131. Hamilton BB, Laughlin JA, Granger CV, Kayton RM. Interrater agreement of the seven-level Functional Independence Measure (FIM). Arch Phys Med Rehabil 1991;72:790.
132. Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. Am J Phys Med Rehabil 1993;72:84-9.
133. Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 1994;75:127-32.
134. Dodds TA, Martin DP, Stolov WC, Deyo RA. A validation of the functional independence measurement and its performance among rehabilitation inpatients. Arch Phys Med Rehabil 1993; 74:531-6.
135. Heinemann AW, Kirk P, Hastie BA, et al. Relationships between disability measures and nursing effort during medical rehabilitation for patients with traumatic brain and spinal cord injury. Arch Phys Med Rehabil 1997;78:143-9.
136. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Relationships between impairment and physical disability as measured by the functional independence measure. Arch Phys Med Rehabil 1993;74:566-73.
137. Granger CV, Cotter AC, Hamilton BB, Fiedler RC, Hens MM. Functional assessment scales: a study of persons with multiple sclerosis. Arch Phys Med Rehabil 1990;71:870-5.
138. Ditunno JF, Cohen ME, Formal CS, Whiteneck GG. Functional outcomes in spinal cord injury. In: Stove SL, Whiteneck GG, DeLisa J, editors. Spinal cord injury: clinical outcomes from the model systems. Gaithersburg: Aspen Publications; 1995.

139. Hall K, Mann N, High W, Wright J, Kreutzer J. Functional measures after traumatic brain injury: ceiling effects of FIM, FIM+FAM, DRS, and CIQ. J Head Trauma Rehabil 1996;11: 27-39.
140. Whiteneck GG, Charlifue SW, Gerhart KA, Overholser JD, Richardson GN. Quantifying handicap: a new measure of long-term rehabilitation outcomes. Arch Phys Med Rehabil 1992;73: 519-26.
141. Mellick D, Walker N, Brooks CA, Whiteneck GG. Incorporating the cognitive independence domain into CHART. J Rehabil Outcomes Meas 1999;3:12-21.
142. Walker N, Mellick D, Brooks CA, Whiteneck GG. Measuring participation across impairment groups using the Craig Handicap Assessment Reporting Technique. Am J Phys Med Rehabil 2003;82:936-41.
143. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985;49:71-5.
144. Pavot W, Diener E. Review of the Satisfaction With Life Scale. Psychol Assess 1993;5:164-72.
145. Corrigan JD, Smith-Knapp K, Granger CV. Outcomes in the first 5 years after traumatic brain injury. Arch Phys Med Rehabil 1998;79:298-305.
146. Corrigan JD, Bogner JA, Mysiw WJ, Clinchot D, Fugate L. Life satisfaction following traumatic brain injury. J Head Trauma Rehabil 2001;16:543-55.
147. Sokol K, Heinemann A, Bode R, Shin J, Van de Venter L. Community participation after TBI: factors predicting return to work and life satisfaction. 127th Annual Meeting of the American Public Health Association, November 7-11, 1999, Chicago, IL.
148. EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16:199–208.
149. Dolan P, Gudex C, Kind P, Williams A. The MVH Group. The measurement and valuation of health. First report on the main survey. Centre for Health Economics. pp 1-48. York: Univ of York; 1994.
150. van Agt HME, Essink-Bot ML, Krabbe PFM, Bonsel GJ. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. Soc Sci Med 1994;39:1537-44.
151. Brazier J, Jones N, Kind P. Testing the validity of the EuroQOL and comparing it with the SF-36 health survey questionnaire. Qual Life Res 1993;2:169-80.
152. Sintonen H. The 15-D measure of health-related quality of life: reliability, validity and sensitivity of its health state descriptive system. Melbourne: National Centre for Health Program Evaluation; 1995.
153. Hurst NP, Jobanputra P, Hunter M, Lambert M, Lochhead A, Brown H. Validity of EuroQOL—a generic health status instrument—in patients with rheumatoid arthritis. Economic and Health Outcomes Research Group. Br J Rheumatol 1994;33:655-62.
154. Klose M, Watt T, Brennum J, Feldt-Rasmussen U. Posttraumatic hypopituitarism is associated with an unfavorable body composition and lipid profile, and decreased quality of life 12 months after injury. J Clin Endocrinol Metab 2007;92:3861-8.
155. Bell KR, Temkin NR, Esselman PC, et al. The effect of a scheduled telephone intervention on outcome after moderate to severe traumatic brain injury: a randomized trial. Arch Phys Med Rehabil 2005;86:851-6.