

Outcome Oriented

March 2010
Issue 13

The Online Newsletter of the
Center for Outcome Measurement in Brain Injury (COMBI)

Measuring Up!

The COMBI continues to add more important scales to its resource center. As of March 2010 there are currently 31 measures featured and detailed in the COMBI.

Agitated Behavior Scale (ABS)
Alcohol and Substance use items
Apathy Evaluation Scale (AES)
Awareness Questionnaire (AQ)
The Cognitive Log (Cog-Log)
Coma/Near Coma Scale (CNC)
Coma Recovery Scale-Revised (CRS-R)
Community Integration Questionnaire (CIQ)
Confusion Assessment Protocol (CAP)
The Craig Handicap Assessment and Reporting Technique (CHART)
The CHART Short Form (CHART-SF)
The Craig Hospital Inventory of Environmental Factors (CHIEF)
Disability Rating Scale (DRS)
Employment variables after TBI
The Family Needs Questionnaire (FNQ)
Functional Assessment Measure (FAM)
Functional Independence Measure (FIM)
Glasgow Outcome Scale (GOS)
Extended Glasgow Outcome Scale (GOS-E)
High Level Mobility Assessment Tool (HiMat)
Levels of Cognitive Functioning Scale (LCFS)
Mayo Portland Adaptability Inventory (MPAI)
Moss Attention Rating Scale (MARS)
Mississippi Aphasia Screening Test (MAST)
Neurobehavioral Functioning Inventory (NFI)
The Orientation Log (O-Log)
Participation Objective, Participation Subjective (POPS)
The Patient Competency Rating Scale (PCRS)
Satisfaction With Life Scale (SWLS)
Service Obstacle Scale (SOS)
Supervision Rating Scale (SRS)

Choosing the Best Outcome Measure

Stephanie Kolakowsky-Hayner, Ph.D.
Director of Rehabilitation Research
Rehabilitation Research Center
Santa Clara Valley Medical Center

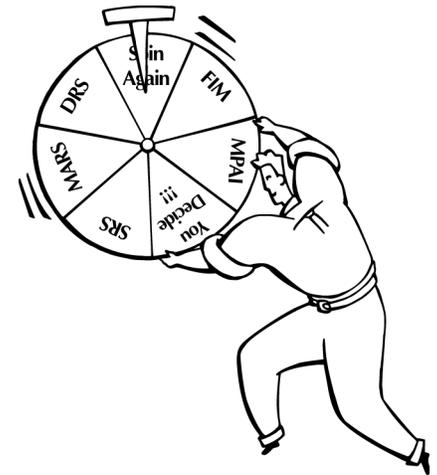
When trying to determine which outcome measure to use for a research study or to measure the impact of clinical treatment, one must consider issues including determining the purpose of the test and its technical characteristics. When determining the purpose of the test, the clinician or researcher should ask themselves, "What am I trying to measure?" With regard to the technical characteristics, s/he should investigate the reliability, validity, types of norms available and how those norms were obtained. This is often a lengthy process including a thorough review of the test materials, the norms, and a comprehensive literature review.

"the COMBI has minimized the amount of time needed to conduct a comprehensive evaluation by providing detailed information on a variety of outcome measures"

However, the COMBI (Center for Outcome Measurement in Brain Injury) has minimized the amount of time needed to conduct a comprehensive evaluation by providing detailed information on a variety of outcome measures for use with persons with brain injury. The COMBI provides detailed summaries of the measures including syllabi, training information, forms, bibliographies of published materials on the measures, frequently asked questions about the measures, and most importantly technical data regarding reliability and validity.

What is reliability?

Reliability refers to the stability or consistency of test scores, regardless of what the test is actually measuring. A reliable measure is psychometrically consistent and dependable, yielding the same or similar score over multiple replications, within a certain margin of error, for the same person. For example, if the subject or patient takes the test today and again tomorrow, will s/he get approximately the same score under the same testing conditions? If the answer is yes, the measure is reliable. A reliable test should measure systematic change or variance of a particular trait being measured. For example, if a



Don't let chance decide your outcome measure!

test is given to measure depression, the participant or subject then undergoes therapy to improve the depression, the change or variability in the test score should be due to an actual change in the trait of depression.

Reliability can be impacted by a variety of sources. First, test scoring, or lack of agreement between multiple raters can cause unsystematic variation in a person's test scores. Variability may also be a result of test content. For example, if there are multiple versions of the same test given, one may have slightly easier or slightly harder questions. The third area of unsystematic variance in reliability may be due to testing conditions. All reliable tests should have a standard method of administration including directions for administration, time limits (or specified lack thereof), and suggested physical arrangements. Additionally, the test administrator should do all s/he can to ensure consistent external interference. For example, if the test is administered in a room next to the railway station and trains commonly pass on the half hour, making significant noise, it is the administrator's responsibility to look for an alternate testing room that will improve noise pollution or ensure all subjects/patients receive the test under the same noisy conditions. Lastly, personal variance can play a role in reliability. If a person is sick or stressed, their score on a particular test could fluctuate unsystematically.

Most information on test reliability found in journals, manuals, and scoring forms, is formulated based on classical test theory. However, item

Continued on Page 2

Choosing the Best Scale (continued)

response theory and generalizability theory are becoming more popular in the literature. Within classical test theory, reliability can be determined using test-retest reliability, inter-rater reliability, alternate form reliability, and internal consistency. Test-retest reliability compares the consistency of scores across multiple testing of the same individual. The reliability is the Pearson correlation of the first test scores with the second test scores, typically referred to as the stability coefficient. Inter-rater reliability is a comparison of multiple raters of the same test. Inter-rater reliability is often expressed in terms of a Pearson correlation as well. Alternate or parallel-form reliability is the Pearson correlation of the scores from multiple versions of the same test to the same patient/subject. Alternate form reliability is not very common as there are not likely two versions of the same test available for examination. Internal consistency is calculated using either split-half or odd-even splits of the test items. One half of the scores are correlated with the other half of the scores, often expressed as a correlation corrected with the Spearman-Brown correction, Kuder-Richardson formula, or Cronbach's alpha. In addition to the correlation, the standard error of measurement describes the practical interpretability of the test scores.

“since a test cannot be valid unless it is reliable, it is more important to have a highly valid test than a highly reliable one”

One pressing question still remains, how high should a reliability coefficient be? Unfortunately there is no hard and fast answer for this question. It truly depends on what the test scores are being used for. For example, if the test scores are being used to determine eligibility for certification, or prescription of a particular medication, the test scores should be highly reliable (.90+). If the test is among a large battery of other tests being used to form an overall picture of a patient/subject, moderately high reliability may be acceptable (.80-.90). Lastly, if the scores will be aggregated across a group and be used for something such as quality improvement, there is not a lesser degree of reliability may be acceptable (.70-.80). Typically, reliability below .60 is unacceptable.

While reliability is very important when determining the type of outcome measure to use for a research study or clinical practice, it is not the only technical aspect of the test to consider. One must be certain to evaluate the validity of the measure. A measure can be reliable without being valid. However, since a test cannot be valid unless it is reliable, it is more important to have a highly valid test than a highly reliable one.

What is validity?

Validity is the most important technical aspect of a test. Regardless of high reliability and great norms, if a test isn't moderately to considerably valid, it is not worth using. Validity is the measurement of whether the test scores are actually measuring what the test has set out to measure. If a test was designed to measure cognitive functioning, is it actually measuring cognitive functioning or is it measuring intelligence, or ADLs, or something completely unrelated? Validity is the degree to which the construct or trait is being accurately measured by the test scores. If the construct or trait of interest exceeds the capacity of the test measuring such construct or trait, the amount of information not included within the test is considered the construct underrepresentation. Further, if the test measures beyond the scope of the construct or trait one wishes to meas-

ure, the excess test items are considered construct irrelevant variance.

Validity has been measured in a variety of ways. Traditionally, journals, manuals, and score reports describe content validity, criterion-related validity, concurrent validity, predictive validity, construct validity, convergent or discriminant validity, and others. All of these measures of validity take an empirical approach to describing the test scores as interpretable for a specific purpose. An initial step toward determining validity may be examining face validity – looking at the test questions and determining that they look like they are measuring the construct – the what you see is what you get approach. Unfortunately this is not always the case and it is best to rely on empirically based validity.

“Once a clinician or researcher determines which test is most appropriate for their needs and whether it is a reliable and valid measure, there are a number of other practical issues to consider”

Content validity evaluates the content of the test in relation to a particularly defined domain. Does the test cover all aspects of the topic at hand? For example, if the domain is neurobehavioral recovery, does the test cover all aspects of the domain or only one or two? With regard to criterion-related validity, test scores are evaluated based on their relationship with performance on some other criterion already established as interrelated to the construct being measured. Two subcategories of criterion related validity include predictive validity and concurrent validity. Predictive validity uses test scores to forecast some future event or criterion. Concurrent validity, on the other hand, measures agreement between test scores and some other current event or criterion. Two other concepts related to criterion-related validity include convergent or discriminant validity. Convergent validity measures the correlation between test scores and some other criterion already proven to measure the same construct. Conversely, when trying to show that the test scores are unrelated to another already established measure of a particular construct, that is called discriminant validity. For example, if trying to measure the discriminant validity of a measure of anxiety, one may show there is little or no correlation with a previously proven valid measure of depression or psychosis. Further, construct validity, often the most difficult to measure, examines the tests scores' ability to measure an unclearly defined construct or one with no obvious reference points. Construct validity is often measured through factor analysis and examination of patient/subject responses processes. Lastly, differential validity, a hot topic of late due to issues with cultural competency of measures, examines whether the test scores measure inequitably for different groups of test takers – e.g., test bias.

Practical Considerations

Once a clinician or researcher determines which test is most appropriate for their needs and whether it is a reliable and valid measure, there are a number of other practical issues to consider. For example, how much does the test cost and is it easy to obtain? Some of the COMBI measures are available for immediate download and include training information as well. Additionally, the clinician or researcher may be interested in whether or not their particular measure of interest is available in languages other than English. The COMBI includes alternate language forms for some of the measures provided. ☑

Development of the MPAI as a Rating Tool

Vicki Eicher, MSW; Director, Quality Management & Training; ReMed; Paoli, PA

James F. Malec, PhD, ABPP-Cn, Rp; Research Director, Rehabilitation Hospital of Indiana; Adjunct Professor of PM & R, Indiana University School of Medicine; Emeritus Professor of Psychology, Mayo Clinic

Thomas Murphy, CEO, Inventive Software Solutions, Philadelphia, PA

The Mayo Portland Adaptability Index 4 (MPAI-4) is a robust patient/person evaluation tool that is the product of over 15 years of research. The MPAI-4, was developed with collaboration from Dr. James Malec, Dr. Murial Lezak, Dr. Randy Evans, Dr. Karen Finlay, Dr. Miriam Kragness, Anne Moessner, RN, MSN and Ann Kent.

The MPAI – 4 has a total of 29 items in the 3 subscales. It also has 6 additional items to record pre injury and post injury information that is not added into the total score, but allows one to capture data regarding alcohol use, drug use, psychotic symptoms, law violations and other conditions causing physical impairment or cognitive impairment. The MPAI is normed on a national sample of individuals with ABI (Acquired Brain Injury). Based on this, the total score and each of the subscale scores can be converted to standardized T scores.

The MPAI – 4 forms, manual and other language versions (Spanish, Danish, French, German, Swedish and Portugese) can be downloaded from the COMBI (Center for Outcomes Measurement in Brain Injury) website. The MPAI-4 can be administered by a Single Professional, Professional Consensus, Person with Brain Injury or Significant Other.

Several studies have shown that the Staff MP AI correlates moderately well with the Disability Rating Scale, the Rancho Scale, neuropsychological measures and the Significant Other MP AI. Each version of the MP AI has showed satisfactory to excellent reliability. With the MP AI-4, it has shown Person Reliability of .88 and Item Reliability of .99. Studies have also shown that it has very acceptable concurrent and predictive validity. Overall, it is recognized that the MP AI-4 is a highly useful research tool because of its ease of administration and scoring, the ability of different groups to use the tool to rate and compare, and that it allows for staff consensus. Finally, the T score conversions demonstrate a range of scores that reflects severity of impairments typical for persons who have an ABI.

PARF Outcomes Benchmarking Project

In 2004, PARF (Pennsylvania Association of Rehabilitation Facilities) launched a collaborative outcomes benchmarking project with 5 post acute brain injury providers in Pennsylvania. These providers agreed to collect outcomes data using the MP AI-4 scale along with identified demographic data. The group selected the MP AI-4 not only because of its robust validity, but because this scale focuses on function rather than impairment. For each item, it is not if the individual has the impairment, but how much that impairment is impacting their everyday functioning that drives the rating. Dr. Malec helped the group clarify issues and items in question on the MP AI. The PARF Outcomes Benchmarking group worked closely with Inventive Software Solutions and their software developer who customized a secure web-based database application that is accessible on the Internet to the providers for this group outcomes project. Each provider always has access to their own data and can run their own reports from that data. Annually, a collaborative report is produced that reflects the de-identified data by various demographic indicators (e.g. years post injury, funder type), by program groupings, by the MP AI-4 Subscales, etc. The PARF group now has 7 post acute providers participating from both PA and NJ. The group has been able to provide State funders and other stakeholders with meaningful data regarding the numbers and needs of persons with brain injuries. For the individual provider, the data is used to provide information regarding the progress/impairment level of an individual person or to provide comparisons between program groupings (e.g. behavioral programs vs. community re-entry programs) within their own organization. This data also helps provide feedback to the provider regarding the effectiveness of their programs.

Malec, Inventive Software Solutions & Oregon Research Institute Secure Federal Grant

In 2007, Dr. Jim Malec, Inventive Software Solutions, creator of the software and database, and the Oregon Research Institute (ORI) secured a Small Business Technology Transfer grant for an "Internet-based Evaluation System for Postacute Acquired Brain Injury". Given that living with the functional impact of a brain injury continues for a lifetime, they supported the need for a national data-driven system for quality monitoring in post-acute brain injury rehabilitation. This post-acute system would

expand upon the current National Institute for Disability and Rehabilitation Research (NDIRR) sponsored TBI Model System data base, which focuses primarily on inpatient rehabilitation. Also, collection of the MP AI-4 and other outcomes measurements in a post-acute setting would improve upon the only widely and consistently used system of patient progress and program evaluation, the Functional Independence Measure (FIM); designed for use in inpatient rehabilitation. Focus on the FIM has not kept pace with the rapid evolution of brain injury rehabilitation practice and the rapidly shifting emphasis from inpatient to post-acute rehabilitation.

The purpose of Phase I of the grant was to develop a software tool and database that could support a national MP AI-4 database. In Phase I, revisions were made to improve the ease of completion of the MP AI-4 tool and input of the demographics, along with the scoring of the MP AI. The system is designed to provide individual scores and T-scores as well as a demographic profile at time of scoring, along with pre-defined composite reports to the organization. Twice per year, comparative analysis reports to other like providers and client/patient populations can be provided.

In 2009, Phase II of this grant was approved with goals to enhance the data analysis capabilities of the software, expand the number of providers participating in the national project, and provide the structures needed for the ongoing financial viability of the product. See "Collaborative Outcomes Research Project" for a summary of the grant.

The purpose of a national MP AI-4 database is to enhance research opportunities to support the efficacy and improvement of brain injury rehabilitation treatment. The national database will provide for a statistically sound and representative measure of long-term outcome after acquired brain injury. Such an outcome measure will allow rehabilitation providers to assess the effectiveness of their services and programs relative to other providers working with similar patients and methods in addition to specifying the effects of ABI on people's lives for rehabilitation planning. Acquiring such information can also play a role in advocacy and policy efforts directed at optimizing funding for programs and services designed to maximize quality of life after ABI. ✓

Future Directions

We are looking to add more training and testing materials for COMBI measures, and to make the existing materials more interactive (automatic email of results from testing exercises).

Reporting of web use statistics from web log files will continue in the next issue. Currently the COMBI receives roughly 1500 visitors per day!

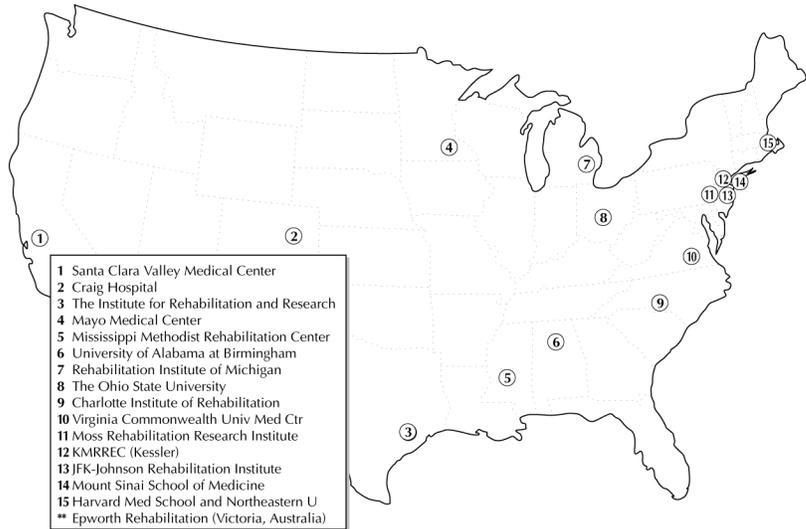
Please email us at <jerry.wright@hhs.sccgov.org> with your thoughts and suggestions. Let us know how we measure up! Thank you for allowing us to be your brain injury outcome measure resource! ☑

Outcome Oriented is a project of the Center for Outcome Measurement in Brain Injury (COMBI) which is funded by the U.S. Department of Education, Office of Special Education and Rehabilitative Services, National Institute on Disability and Rehabilitation Research. The contents of this newsletter were developed under a grant from the Department of Education. However those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal government.

Address inquiries to
Jerry Wright, Editor. Phone (408) 793-6430;
Email jerry.wright@hhs.sccgov.org

This document is available online at:
<www.tbims.org/combi/combinews.html>

CREDIT TO OUR COLLABORATORS



The COMBI is a collaborative project of sixteen brain injury centers. Without the expertise of these centers this project would not be possible. We would like to offer special recognition to the individuals at these facilities who have taken the time to prepare materials for the COMBI:

Tamara Bushnik, PhD, Jerry Wright, MS, Laura Jamison, Maurice Rappaport, MD, PhD, and Stephanie Kolakowsky-Hayner, PhD at Santa Clara Valley Medical Center (**Lead Center**)

Dave Mellick, MA and Cindy Harrison-Felix, PhD at Craig Hospital

Mark Sherer, PhD, ABPP-Cn, Corwin Boake, PhD and Angelle Sander, PhD at The Institute for Rehabilitation Research

James F. Malec, PhD, LP at the Mayo Medical Center

Risa Nakase-Richardson, PhD at the Mississippi Methodist Rehabilitation Ctr

Tom Novack, PhD at University of Alabama at Birmingham

Jennifer Bogner, PhD & John D. Corrigan, PhD at the Ohio State University

Flora Hammond, MD at Charlotte Institute of Rehabilitation

Jeffrey Kreutzer, PhD and Jenny Marwitz, MA at Virginia Commonwealth University Medical Center

Tessa Hart, PhD at Moss Rehabilitation Research Institute

Kessler Medical Rehabilitation Research and Education Corporation

Joseph Giacino, PhD and Kathleen Kalmar, PhD, at JFK-Johnson Rehabilitation Institute

Marcel Dijkers, PhD and Margaret Brown, PhD, at Mount Sinai School of Medicine

Therese O'Neil-Pirozzi, ScD, CCC-SLP and Mel Glenn, MD at Harvard Medical School and Northeastern Univ.

Gavin Williams, PhD, at Epworth Rehabilitation (Victoria, Australia) ☑



SANTA CLARA VALLEY MEDICAL CENTER
Rehabilitation Research Center
Santa Clara Valley Medical Center
751 South Bascom Avenue
San Jose, CA 95128

UPDATE

Center for Outcome Measurement
in Brain Injury (COMBI)
<www.tbims.org/combi>

March 2010

