Outcome Measures in Spinal Cord Injury: Recent Assessments and Recommendations for Future Directions


1University of Alabama, Birmingham, AL, USA 2University of California, Irvine, CA, USA 3Rigshospitalet and University of Copenhagen, Denmark 4Acorda Therapeutics, Hawthorne, NY, USA 5National Institute on Disability and Rehabilitation Research (NIDRR), Washington, DC, USA 6Mount Sinai School of Medicine, New York, NY, USA 7Stanford University, Palo Alto, CA, USA 8Tel-Aviv University, Tel-Aviv, Israel 9International Collaboration On Repair Discoveries (ICORD), University of British Columbia and Vancouver Coastal Health, Vancouver, BC, Canada 10Memorial Hermann/TIRR, Houston, TX, USA 11Thomas Jefferson University, Philadelphia, PA, USA 12Imperial College, London, UK 13Danish Pain Research Center, Aarhus, Denmark 14Baylor College of Medicine, Houston, TX, USA 15Alseres Pharmaceuticals, Hopkinton, MA, USA 16Northwestern University Medical School, Chicago, IL, USA 17University of Wisconsin, Milwaukee, WI, USA 18University of Michigan, Ann Arbor, MI, USA 19National Institutes of Health/National Institute of Neurological Disorders and Stroke (NIH/NINDS), Bethesda, MD, USA 20Aarhus University Hospital, Aarhus, Denmark 21Craig Hospital, Englewood, CO, USA 22Center for Outcomes, Research, and Education, Evanston, IL, USA 23Shriners Hospital, Philadelphia, PA, USA 24Balgrist Hospital, Zurich, Switzerland 25Veterans Affairs (VA) Puget Sound Health Care, Seattle, WA, USA 26Kessler Medical Rehabilitation Research and Education Center, West Orange, NJ, USA

Abstract

Study Design—review by the Spinal Cord Outcomes Partnership Endeavor (SCOPE), which is a broad-based international consortium of scientists and clinical researchers representing academic institutions, industry, government agencies, not-for-profit organizations and foundations.

Objectives—assessment of current and evolving tools for evaluating human spinal cord injury (SCI) outcomes for both clinical diagnosis and clinical research studies.

Methods—a framework for the appraisal of evidence of metric properties was used to examine outcome tools or tests for accuracy, sensitivity, reliability and validity for human SCI.

Results—imaging, neurological, functional, autonomic, sexual health, bladder/bowel, pain, and psycho-social tools were evaluated. Several specific tools for human SCI studies have or are being developed to allow the more accurate determination for a clinically meaningful benefit

*The views expressed here are those of the authors and do not represent those of the National Institutes of Health (NIH), the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute on Disability and Rehabilitation Research (NIDRR), the Department of Veterans Affairs (VA) or the US Government.

*Correspondence: John Steeves, ICORD, 818 West 10th Avenue, Vancouver, BC, V5Z 1M9, Canada (steeves@icord.org)
(improvement in functional outcome or quality of life) being achieved as a result of a therapeutic intervention.

**Conclusion**—significant progress has been made, but further validation studies are required to identify the most appropriate tools for specific targets in a human SCI study or clinical trial.

**Keywords**
-outcome tools; upper and lower limb; diagnostic tools; spinal cord clinical trial targets; functional assessment; quality of life; community participation

**Introduction**

**Development of “Spinal Cord Outcomes Partnership Endeavor” (SCOPE)**

Today, there are more human spinal cord injury (SCI) studies in progress, or planned, than ever before. In light of this fact, it is important to the international research community and to people living with SCI that both clinical and scientific organizations, along with private sector industry partners, take an active leadership role to ensure the objective and valid conduct of these human studies. This can be accomplished, in part, through the identification and development of appropriate clinical tools and valid measures that are specific to relevant therapeutic targets.

Over the past few years, multiple complementary ventures have been initiated, including those by the American Spinal Injury Association (ASIA) with the National Institute on Disability and Rehabilitation Research (NIDRR), the International Spinal Cord Society (ISCoS), and the International Campaign for Cures of spinal cord injury Paralysis (ICCP) with the International Collaboration On Repair Discoveries (ICORD). Many of these activities have now coalesced into the Spinal Cord Outcomes Partnership Endeavor (SCOPE, www.scopesci.org). SCOPE is a broad-based consortium of scientists and clinical researchers representing academic institutions, industry, government agencies, not-for-profit organizations and foundations. SCOPE’s mission is to enhance the development of human study protocols (e.g. clinical trials) to accurately assess therapeutic interventions for SCI, which lead to the adoption of improved best practices. Of the major SCI clinical trials that have been undertaken and completed, none of the tested pharmaceutical therapeutic interventions has become a universally accepted standard of clinical care. These trials have highlighted some of the difficulties that must be adequately addressed for the successful completion of future clinical trials, including: the choice of an appropriate and valid primary clinical endpoint, selection of trial participants, stratification of subjects, effective timing, length and route of administration for a therapeutic intervention after SCI, and the coordination and standardization of trial protocols across multiple participating centers.

**Previous Initiatives**

The ICCP is an affiliation of ’not for profit’ organizations (www.campaignforcure.org/iccp), which aims to facilitate the translation of valid treatment strategies for SCI paralysis. The 24 international members of ICCP’s SCI Clinical Guidelines Panel developed an initial set of guidelines1·2·3·4 regarding the design of clinical trials to protect or repair the injured spinal cord. These four papers focused on experimental cell-based and pharmaceutical drug treatments, and addressed both acute and chronic stages of SCI. This focus was selected because of the substantial risks and potential benefits of these types of treatments, and because some treatments of this type have been offered without clinical trial data on safety and efficacy.

*Spinal Cord. Author manuscript; available in PMC 2010 August 1.*
ASIA is a professional organization of physicians from multiple disciplines, as well as allied health professionals and researchers with special expertise in the care of persons with SCI. During active discussions at annual ASIA meetings, it was noted that comprehensive measurement tools to accurately document the effects of treatments for conditions arising from SCI, are in the early stages of development. Moreover, many previous studies have been performed using outcome measures with unknown or limited sensitivity or did not address functional relevance. In 2005, based on these discussions, NIDRR and ASIA convened multiple panels of experts to undertake a systematic review of the published literature regarding specific diagnostic and research outcome measures after SCI. These groups evaluated the strengths and limitations of methods to sensitively, accurately, and reliably measure either a clinically meaningful change in a functional outcome, a significant change in Activities of Daily Living (ADL) or an improvement in Quality of Life (QoL).

ISCoS is a worldwide international professional society of physicians and surgeons, as well as members of allied professions (e.g. scientists and therapists) with activity in the treatment of patients with spinal cord afflictions or in research into a subject relating to SCI. A major ISCoS initiative in collaboration with ASIA has been the development of International SCI Data Sets to standardize the collection and reporting of information necessary to evaluate and compare results of published studies (www.iscos.org.uk & www.asia-spinalinjury.org). Additional modules of International SCI Data Sets are being developed by panels of experts to identify critical variables for specific topics of research and provide recommended standards for collecting and reporting of that information.

ICORD coordinated the ICCP clinical guidelines initiative, including the development of a document for the general public on “Experimental treatments for SCI: What you should know...”, which is directed to people living with a spinal cord injury (SCI), their families and friends, as well as health care professionals and scientists when discussing experimental treatments for SCI. ICORD researchers also coordinated a meta-analysis of SCI Rehabilitation Evidence (SCIRE), which objectively reviewed the strength of support for a large number of SCI rehabilitation practices and strategies. Both documents are available as free downloads (www.icord.org).

All these international efforts led to an inclusive coalition for the development of improved outcome measurement activities by ASIA, ISCoS, ICCP, ICORD, representatives of the National Institute of Health (National Institute of Neurological Disorders and Stroke and the National Center for Medical Rehabilitation Research), the US Food and Drug Administration, the Veterans Administration Rehabilitation Research and Development Service, as well as corporate partners Acorda Therapeutics, Alseres Pharmaceuticals, Clinical Assistance Programs, and Cyberkinetics.

The goal of this paper is to provide a synopsis of the current status of SCI outcome measurements and to identify the unmet needs and challenges in providing improved objective outcomes that can be used for upcoming therapeutic intervention trials.

**Methods**

Specific and objective processes have been developed by ASIA, ICCP, ICORD, ISCoS and SCOPE to evaluate SCI outcome measures. For example, as part of the ASIA-NIDRR initiative a framework for the appraisal of evidence of metric properties was developed and designed to be useful both for reviewing past studies and for planning future research. Key features of this framework included:

- Pre-established criteria in an evidence table for grading of reliability and validity indicators on a multi-level scale.
Reliance on a foundation of the principles of Classical Test Theory and established works on measurement standards in rehabilitation and other fields.\(^6\)\(^7\).

Incorporation of evidence from Modern Test Theory (Item Response Theory and Rasch analysis).

Introduction of “internal validity” as a key consideration, based on analysis of items within the measurement domain.

Grading of evidence of “external validity,” including predictive and consequential validity and emphasizing evidence of utility in practice.

Emphasis on defining the construct to be measured, including both its content and key external characteristics.

This process was used for each of the areas addressed in this summary paper with the exception of upper extremity (UE) measures and spasticity. Given the evolving nature for many of the outcome tools, many of the above psychometric criteria have yet to be satisfied. To the extent possible, individual groups evaluated measures that have been used by at least two independent SCI research groups since 2000. Many of these findings and reviews are available online at the following websites (www.asia-spinalinjury.org; www.iscos.org.uk; www.icord.org). Furthermore, additional international academic and corporate experts have been recruited to participate in summarizing these reviews and have included any new, relevant outcome measures to ensure a concise yet comprehensive review.

Results

Neuroimaging

The group assessing neuroimaging included representatives from the fields of SCI medicine, neurosurgery, and neuroradiology. Ninety-nine clinical and pre-clinical articles published between 1984 and early 2006 have been reviewed in this rapidly expanding field.\(^8\) Magnetic Resonance Imaging (MRI) was judged to be the neuroimaging modality of choice for assessment of SCI because of its ability to define the location of injury, degree of cord compression, as well as the presence of hemorrhage/contusion, and edema. MRI studies have been shown to contribute to the understanding of injury severity and prognosis. MRI-Diffusion Weighted Imaging may be useful in diagnostically quantifying the extent of axon loss after SCI, but remains an evolving research tool due to resolution limitations imposed by the small cross-sectional size of the cord and the technical challenges posed by motion artifact (e.g. respiratory and cardiac gating). Functional MRI (fMRI) was found to be useful for assessing the correlation between sensorimotor activities of persons with chronic SCI with imaging of metabolic activity of the brain or spinal cord; however, it is not likely to be used as an acute clinical outcome tool due to the lengthy time constraints required for adequate data collection.

Magnetic Resonance Spectroscopy can be utilized in research studies for the assessment of biochemical characteristics of the spinal cord after injury. Intraoperative Spinal Sonography was judged to be useful in assessing spine and spinal cord anatomy and gross pathology during surgical procedures. Grading of the clinical neuroimaging articles showed a paucity of the highest level of evidence, suggesting that more rigorous development is needed for all imaging modalities before MRI can even be considered as a surrogate outcome measure.

Motor and Sensory Function

The group summarizing motor and sensory function included neurologists, physiatrists, and scientists. Clinical and laboratory-based measures of motor and sensory function have been...
evaluated for their utility in tracing preserved residual and/or recovered function after SCI1-2,9. The International Standards for Neurological Classification of SCI (ISNCSCI) including the ASIA Impairment Scale (AIS) has become a standardized and routinely applied neurological assessment and classification scale for patients suspected of suffering a SCI10. However, certain aspects of the ISNCSCI (e.g. AIS grades) may be insensitive or highly variable as an outcome measure for assessing the possible benefits of an intervention and currently there is no method for measuring upper cervical, thoracic, or sacral motor function1,2,10. New measures of motor function are being considered to address these gaps (e.g. by specifically examining trunk motor function). For SCI research, the ASIA Motor Score (AMS) is composed of upper and lower extremity motor scores, which should be tracked separately2,11. As a guide to establishing more accurate therapeutic thresholds for determining whether a treatment as a functional clinical benefit, the ICCP Clinical Guidelines Panel is currently calculating the degree of spontaneous change in the AMS and the AIS motor level in ‘untreated’ SCI populations, from a number of previous datasets. An alternative strategy that classifies only the presence or absence of activity in a larger number of muscle groups is also being examined.

Manual Muscle Testing (MMT) is an easily accessible and reliable method of determining the strength of individual muscles and may be more reliable than myometry. MMT is accurate within a functional range, though not sensitive to changes in the upper range of strength. Electrophysiological measurements such as electromyography (EMG), and motor-evoked potential (MEP) recordings provide objective data (latencies and amplitudes) for assessing spinal conductivity that can be quantitatively analyzed by a blinded investigator. Surface EMG recordings provide a sensitive measure for trace muscle function; however, they are not widely used. Abnormal activity such as spasms may confound data so such interpretation is best undertaken with simultaneous, multi-muscle recordings. With further development, a combination of somatosensory evoked potential (SEP), MEP and/or EMG measurements could provide information about spinal cord function that is not retrievable by other clinical means and may have additional value in predicting functional clinical benefit12. A new objective measure of motor control called the voluntary response index (VRI) has been developed from EMG recordings, but needs further validation13.

Clinical sensory testing using the Light Touch and Pin Prick tests defined in the ASIA standards has been shown to be a reliable diagnostic method, especially preservation of pin-prick sensation. The sensory score is less predictive for incomplete motor deficits than motor complete SCI. Quantitative Sensory Testing (QST), employing thermal, mechanical, vibratory and electrical stimuli is developing14. These methods may assist in differentiating the contributions from small and large diameter peripheral sensory afferent projections or distinguish the contributions of ascending spinal sensory pathways (spinothalamic and dorsal columns), but further development is ongoing. The sensitivity of QST, including the emerging electrical perception threshold test15, to detect abnormality or preserved innervation may be superior to SEP recordings and ASIA sensory scores.

**Functional Potential**

The group assessing functional potential included physiatrists, physical and occupational therapists, spinal cord medicine physicians, clinical researchers and scientists. This group had expertise in evaluating outcome measures that assess overall ADL (i.e. functional capacity) in persons with SCI for either clinical evaluation, or functional recovery assessment. Four measures were studied in depth, including: the Modified Barthel Index (MBI), the Functional Independence Measure (FIM), the Quadriplegia Index of Function (QIF), and the Spinal Cord Independence Measure (SCIM)16. The FIM and SCIM were found to be reliable and valid, while validity of the MBI and QIF has not been sufficiently investigated. Unlike the MBI and FIM, the SCIM and QIF were specifically designed for the
SCI population. While the SCIM comprehensively assesses functional recovery, the QIF is focused on persons with tetraplegia. The FIM has some limitations, as it was designed to assess a broad range of disabling medical conditions (e.g., it generally assesses burden of care requirements) and might not specifically reflect functional recovery after SCI.

The work group recommends optimizing the SCIM and QIF by institutions throughout the world, rather than spending time and resources on the development of a new functional recovery measure for SCI. The latest version of the SCIM (SCIM III) should continue the refinements and psychometric validation so it might subsequently be implemented worldwide as the primary functional recovery outcome measure for SCI (e.g., as a primary outcome measure for pivotal phase 3 clinical trials). Given the important health care and societal costs of tetraplegia, the accurate assessment of upper extremity function is viewed as priority. Thus, the QIF and other upper extremity functional outcome tools should undergo continued development and validation as a tool for cervical level SCI.

**Upper Extremity Function**

The UE is often evaluated using performance-based outcomes measures; however, it is also important to evaluate impairment and capacity of the UE independent of performance. There is a general consensus that generic tests of hand function are ill-suited for use with persons with SCI. The Grasp and Release Test (GRT) developed to evaluate opening and closing of the hand by a person with SCI with a neuroprosthesis, met the general criteria for UE SCI measures and good reliability was documented. The Capacity of the Arm and Hand Test (CAHT) is being developed to measure actual performance in arm and hand function; however, it also needs reliability and validity testing. The Graded and Redefined Assessment of Strength, Sensibility and Prehension (GRASSP) is also being developed as a clinical research tool that is responsive and would track the extent of spontaneous recovery or possible outcomes of a surgical or pharmacological intervention in a clinical trial. GRASSP evaluates changes within the motor and sensory systems, but also has a prehension component to relate impairment level changes to complex hand function tasks. GRASSP is currently undergoing international reliability and validity testing.

**Ambulation**

The group assessing ambulation included physiatrists, physical therapists and clinical research scientists. Six measures were reviewed: the Walking Index for Spinal Cord Injury II (WISCI II), 50 Foot Walk Test (50FTWT), 6 Minute Walk Test (6MWT), 10 Meter Walk Test (10MWT), Spinal Cord Injury-Functional Ambulation Inventory (SCI-FAI) and Functional Independence Measure-Locomotor (FIM-L). Findings suggested that the WISCI II and 10MWT were the most valid and clinically useful tests as primary outcome measures for gait and ambulation for incomplete SCI, as they demonstrated criterion-oriented validity, reliability, and sensitivity to change. Conversely, the FIM-L was found to have the least validity and utility for human studies, as it had poor sensitivity to change and limited clinical utility in certain populations. Both the 50FTWT and the 6MWT were rated as acceptable, but will need further validation and improvements to be considered as primary outcome measures. The SCI-FAI measured gait quality, but validity has only been shown among trained physical therapists. Ideally, the most comprehensive assessment of ambulation would include evaluations of speed, endurance, and functional capacity and would require the use of a combination of tests, such as the 10MWT and WISCI II.

**General Autonomic Function**

The assessment of general autonomic function was performed by a group of basic scientists, pulmonologists, cardiologists and physiatrists. Uniform operational definitions for autonomic dysfunctions related to SCI and 25 autonomic tests were selected for appraisal.
The group assessed the potential usefulness and applicability of these tests to SCI individuals and five tests were selected for detailed analysis: sympathetic skin responses, blood pressure and heart rate variability analyses, sit-up and tilt-table orthostatic challenge tests, and mental stress testing. These tests were evaluated for validity, reliability and reproducibility in determining autonomic function following SCI26. The review of studies using these tests showed that 3 tests have content validity and metric reliability (blood pressure and heart rate variability analyses, sit-up and tilt-table orthostatic challenge tests), one test had minimal validity (sympathetic skin response), and no formal validation had been performed for the mental stress test. The group was not able to identify validated tests for sweating abnormalities and other temperature deregulation. The group is in the process to examining possible additions for evaluation of the autonomic control of respiratory functions. The addition of autonomic measures to the International Standards for the Neurologic Classification is discussed below.

**Colon and Rectal Function**

The group assessing colon and rectal function included physiatrists and gastroenterologists. Impairment measures reviewed include anal manometry, rectal electromyography, rectal impedance planometry, and colonic transit time. Anorectal manometry, determining anal resting and squeeze pressure, as well as anorectal sensibility testing with standardized rectal distension or electrical stimulation of the anal canal are standard procedures in anorectal physiology laboratories worldwide. Those methods provide valuable information about anorectal physiology27, but use is limited by extensive equipment needs and a lack of clinical utility for the information obtained. Total or segmental colorectal transit times determined by oral intake of radio-opaque markers and subsequent abdominal x-rays have been extensively used28, however, the reproducibility and the association between colorectal transit times and bowel symptoms remain to be described.

Colorectal scintigraphy, rectal impedance planimetry, anorectal electromyography, the activity of defecation or the modified activity of bowel care for stool elimination using the Events and Intervals of Bowel Care along with stool weights is useful to measure the effectiveness and efficiency of defecation29. Recently, a Neurogenic Bowel Dysfunction Score has been formulated and used in populations of individuals with SCI, but its validity and reliability need to be proven30. Patient-centred Faecal Incontinence Scales have been written that include Quality of Life (QoL) measures (subject response questionnaires) attempting to quantify participation, but none have been designed for SCI31. A Cochrane review concluded that treatment of bowel dysfunction in central neurological diseases must remain empirical, until large well-designed trials have been performed32.

**Lower Urinary Tract Function**

The assessment of lower urinary tract function was performed by a group of urologists and physiatrists. Standardization of urodynamic terminology and technique has been proposed by the International Continence Society. Outcome measures including voiding and continence diaries, post-void residual volume measurement, urodynamic studies and bladder-related QoL measures have been reviewed34-35. Findings suggest that diary-based measures of continence and voiding are not well standardized and have limited sensitivity, accuracy and reliability. Measurement of post-void residual volumes by ultrasound is sufficiently reliable for clinical purposes, but measurement by catheterization is more accurate for research studies. Urodynamic measurements of filling and voided volumes, bladder and sphincter pressures, urine flow rates and electromyography of the pelvic floor are accurate and important for evaluating clinical management of the neuropathic bladder, but their sensitivity and reliability for evaluating spinal cord treatments have not been well established. Electrophysiological measurement of sacral nerve function is accurate, sensitive.
and reliable and has potential for evaluating conus medullaris and cauda equina lesions.
Objective measurement of bladder sensation is in its infancy. Quality of life in relation to
bladder function after SCI can be measured with good sensitivity, accuracy and reliability by
the Qualiveen questionnaire.

Sexual Function

The assessment of sexual function was performed by a group of urologists, physiatrists, a
sexologist, and a basic scientist with expertise in sexual functioning and SCI. Sexual
function was divided into male and female sexuality, male fertility and female fertility and,
within categories, measures were chosen for detailed review based on expert consensus.
Vaginal pulse amplitude was found to be the most reliable measure to evaluate vaginal
blood flow and it has been used in SCI; however, its use is limited to laboratory testing and
it is not practical for clinical trials as there is limited equipment availability and the testing is
somewhat invasive. The Female Sexual Function Index (FSFI) was found to have good
discriminant and divergent validity and has been used successfully in clinical trials;
however, there are no published results yet in SCI females.

The International Index of Erectile Function (IIEF) has documented internal consistency,
divergent and convergent validity and discriminant validity. It has been used successfully in
multiple clinical trials involving men with SCI. With regards to male fertility, measurement
of ejaculatory potential through penile vibratory stimulation or electroejaculation, and
standard semen analysis were considered the only options available. No measures were
available to document female reproductive capability. It is the consensus of the committee
that the IIEF and FSFI are appropriate measures to use in clinical trials; however, further
documentation of their validity is needed.

Pain

The group performing the assessment of pain included physiatrists, basic scientists,
psychologists and clinical researchers with expertise in SCI pain. Recommendations were
made within the different domains for which outcome measures were available that met
review criteria. A 0-10 Point Numerical Rating Scale (NRS) is recommended to be used
to measure pain intensity after SCI, while the 7-Point Guy/Farrar Patient Global Impression
of Change (PGIC) scale is recommended to measure global changes in pain. The SF-36
single pain interference question and the Multidimensional Pain Inventory (MPI) or Brief
Pain Inventory (BPI) pain interference items are recommended as the measures for pain
interference after SCI. Brush or cotton wool and at least one high-threshold von Frey
filament are recommended for testing of mechanical allodynia/hyperalgesia, while a Peltier-
type thermistor is recommended to test thermal allodynia/hyperalgesia. The International
Association for the Study of Pain (IASP) or Bryce-Ragnarsson pain taxonomies are
recommended for classification of pain after SCI, while the Neuropathic Pain Scale (NPS)
is recommended for measuring neuropathic pain symptoms and any subsequent changes.
The Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) should be used for
discriminating between neuropathic and nociceptive pain. It was the consensus of the
committee that for each of these domains, further evaluation of reliability and validity in SCI
populations should occur.

Spasticity

Incomplete SCI often leaves the individual with altered motor control or spasticity,
manifest as a variety of clinical signs and symptoms including a diminution in intensity, and
diminished or increased motor output. The common definition of spasticity, “increased
resistance to passive stretch,” and scales such as the Modified Ashworth Scale (MAS) that
describe this aspect, capture only a small portion of what is really a multidimensional
phenomena. Other psychometrically evaluated scales include the Penn spasm frequency scale, the spinal cord assessment tool for spasticity, the visual analog scale, and the Wartenburg pendulum test. Objective alternatives include the use of surface electromyography recordings that characterize motor control in detail, and isokinetic dynamometry to quantify the force of spastic contraction. Recently, a self-assessment scale, designed to capture the patient’s experience of spasticity, has been introduced. To best characterize the multidimensional nature of spasticity, a battery of tests subject to additional validation testing and structured along the International Classification of Functioning, Disability and Health (ICF), would provide improved resolution of mechanisms and intervention targets.

Depression

A panel of experts including clinical and rehabilitation psychologists with expertise in depression identified seven depression measures in 24 studies reporting psychometric data in the peer-reviewed English literature since 1980, including: Beck Depression Inventory (BDI), Zung Self-Rating Depression Scale, Center for Epidemiological Studies Depression Scale (CES-D), Older Adult Health and Mood Questionnaire, the Structured Clinical Interview for the DSM-IV (SCID), Inventory to Diagnose Depression (IDD), and the Patient Health Questionnaire Depression Scale (PHQ-9). These measures require few modifications for administration to SCI respondents and are generally brief (<10 minutes) with the exception of semi-structured interviews (i.e., SCID).

The overall paucity of psychometric data on depression measures used among people with SCI is surprising given the focus on depression in this population. However, from the available evidence, it appears that the different measures perform equally well. Thus selection of a particular depression measure used in SCI research cannot be made on the grounds of psychometric superiority, but instead on feasibility, acceptability to patients, ease of administration and scoring, and the purpose of evaluation. For measuring symptom severity, the CES-D has been widely used in SCI research, second only to the BDI. For screening measures (i.e., criterion-referenced to diagnostic criteria) the PHQ-9 is widely used with its inclusion in the SCI Model Systems National Database and the IDD shows some promise as well. Nevertheless, more research is clearly needed to facilitate our ability to target interventions on the most problematic symptoms, endorse one or more measurement tools, evaluate the implementation of depression screening programs, which will ultimately determine the effectiveness of an intervention in clinical practice. Finally, it is important to validate any uniform measure of depression so that outcomes of clinical interventions can validly be compared across studies.

Quality of Life

The review of health-related QoL for an individual’s life was performed by a group of clinical and rehabilitation psychologists. QoL was defined as a multi-dimensional construct that includes physical functioning, functional ability, emotional functioning and satisfaction with life. Four QoL scales met the above criteria, including the SF-36/SF-12, the Sickness Impact Profile (SIP-68), the Life Satisfaction Questionnaire (LISAT-9, LISAT-11) and the Satisfaction with Life Scales (SWLS). The SF-36/SF-12 measures were the most widely used and both reflect health status. The original SIP was developed as an assessment of general health-related functioning and the behavioral impact of “sickness” for physical, emotional, and social functioning in everyday life. The shortened SIP-68 has been re-conceptualized as a measure of individualized levels of disability. The LISAT-9, LISAT-11 and SWS are measures of life satisfaction and tap into only one domain within a HRQOL framework.
Several instruments did not meet review criteria and are currently in development but deserve mention. The Patient Reported Outcomes Measurement Information System (PROMIS), the Neuro-QoL, and the related SCI-QoL and SCI-CAT instruments are in development, using a grounded theory approach to guide item development and large scale field testing to calibrate the item difficulties using Item Response Theory. Plans are to develop these measures as computerized adaptive tests. These scales are being designed for use as patient reported outcome measures in clinical trials and the SCI-QoL/SCI-CAT scales will cover issues targeted to individuals with SCI.

**Participation**

The review of participation was performed by a group that included experts with backgrounds in rehabilitation psychology, speech communication, and occupational therapy. People with SCI experience barriers to participation within their society and/or resident community, including: reduced mobility and employment, limited social and family role functioning, and decreased access to recreational and leisure activities. High quality instruments would help describe participation needs and monitor efforts to ameliorate restrictions. Three instruments met the review criteria: The Craig Handicap Assessment and Reporting Technique (CHART), Assessment of Life Habits (LIFE-H), and the Impact on Participation and Autonomy (IPA). They reflect different perspectives in participation measurement. The LIFE-H uses a qualitative approach while the CHART adopts a quantitative approach; both are based on societal norms of participation. The IPA integrates individual choice and control in its definition. CHART is the most widely used instrument, though its development predates the more recent International Classification of Functioning, Disability and Health (ICF). The IPA is a relatively new instrument and its psychometric properties have only recently been published.

Several instruments did not meet inclusion criteria, but deserve monitoring, including the Participation Measure for Post-Acute Care (PM-PAC), the Participation Survey/Mobility (PARTS/M), the PRO-PAR, and Community Participation Indicators (CPI). The PM-PAC reflects participation as conceptualized by the ICF. The PRO-PAR complements activity assessments with items designed to cover more complex life experiences in the ICF participation domain. PARTS/M addresses participation by people with mobility impairments. The CPI used a grounded theory approach to guide the development of an instrument for people with disabilities, especially those who are disenfranchised through the experience of disability and are also at an economic or social disadvantage.

**Discussion**

**Conclusions and Future Trends**

While perfect and complete evidence is not possible, use of an objective and systematic framework to evaluate measures will encourage sound development and application of these measures, both in research and clinical practice. Reviewers and researchers are encouraged to use objective and standardized frameworks, adapting it if necessary, to identify and validate the most critical issues.

From these recent reviews we can see that a significant amount of development has been accomplished, but further work is needed for adequately establishing reliable and sensitive outcome measures after SCI. There is also little consensus of what size of change (threshold) in any of these measures should be considered to reflect a clinical meaningful benefit that is statistically different from spontaneous functional recovery. In several domains a combination of measures would likely be optimal. Such combinations would also need to be...
carefully evaluated, weighted and validated, as the burden of multiple assessments on subjects must also be considered.

Selection of measures within many domains for clinical trials will depend on their initial (baseline) value as diagnostic screening or tools for monitoring symptoms. We recommend that when planning a trial, consideration should be given to those measures identified here for specific clinical trial targets.

In addition to the specific outcome measures discussed here, standard clinical information is needed about subjects in clinical trials. The field of SCI is fortunate that the ISNCSCI (includes the ASIA impairment scale or AIS) is available so there is a standardized terminology to clinically describe the neurological level and severity of a person’s spinal injury. With regards to clinically diagnosing more details about an individual’s SCI, there are two ongoing initiatives that will improve standardization of clinical care and it is possible that these new modifications may also improve therapeutic assessments in future clinical trials.

The first modification is the development of an adjunct to the ISNCSCI that describes the impact of SCI on autonomic function. An international committee has been working on this addition to the assessment protocol since 2005 and will publish a recommended format for accurately assessing the impact of SCI on bladder, bowel, sexual, cardiovascular, pulmonary, thermoregulatory and sudomotor function in 2008. There is also an online electronic training program being developed for this new version of the ISNCSCI called INSTeP (International Standards Training e Program). Evolving versions are available for review at www.asia-spinalinjury.org/eLearning.

The other significant development in the field is the ISCoS/ASIA led International SCI Data Sets initiative. It has been recommended that common data be collected internationally on individuals with SCI to facilitate comparisons regarding injuries, treatments, and outcomes between patient groups, study centers and countries. To facilitate this, data sets are being developed that are simple and relevant to specific aspects of SCI. The data sets are available free for use without any restrictions (www.iscos.org.uk and www.asia-spinalinjury.org).

A structure and terminology has been developed following the format of the ICF58 (Figure 1). It is recommended that the Core Data Set data be included as a descriptive table in publications describing individuals with SCI. A Basic SCI Data Set is the minimal number of data elements, which should be collected in daily clinical practice. The various Basic Data Sets may be the basis for a structured record in SCI centers worldwide. Extended SCI Data Sets are more detailed modules, which may be valuable for human research studies. For each data set a syllabus is being developed, including definitions, instructions on how to collect each data item, and coding schemes.

Organizations, societies etc. are invited to review the International SCI Data Set, and a process for approval and endorsement of the data sets has been established. Data sets are in development or have been published that include non-traumatic spinal injury, urinary tract function and imaging, pain, cardiovascular function, bowel function, vertebral injury and spinal surgery, male and female sexual function, as well as activity, participation and QoL. Once data sets are developed, it is recommended that relevant information be used in clinical trial outcomes analysis.

The primary concerns of the corporate sector, which are shared by all SCI researchers, is outcome measures in SCI trials should be practical for multi-center studies and should satisfy a regulatory agency’s requirement for approval and adoption, by allowing adequate demonstration of efficacy. In order to do this, outcome measures need to be 1) standardized,
so that clinicians know exactly how to perform them, 2) validated, so that their measurement characteristics are clear, and 3) capable of providing information about clinically meaningful change (i.e. benefit). This last requirement can be met either directly or by a process of mapping to other measures that represent meaningful benefit. For example, it may be possible to show that an improvement in an objective measurement of a discrete neurological dysfunction can be validated as clinically meaningful by reference to a softer, subjective measurement for an improvement of a “real-world” disability. This will require dedicated and carefully designed studies; it will not be sufficient for clinicians in the field to say that any improvement in neurological function is valuable. The regulatory goal is to have reliable information about real functional benefit to patients, which is balanced against information on the therapeutic risks.

Most of the treatments contemplated for direct treatment of SCI are expected to improve overall neurological function, through neuroprotection or neural repair, the details of which may be quite variable between individuals. There is no precedent from which regulatory bodies or the sponsors of clinical studies can draw on to derive a clear path for establishing this kind of efficacy. However, there is a tendency for even past failed trials to set a precedent for what a future trial should look like, in the minds of sponsors, experimentalists and regulators. This can be seen in the importance that has been placed on motor scores (e.g. due to their use in the NASCIS trials) in SCI trials, despite the poor measurement characteristics of such scores as outcome measures and our inability to map changes in these scores readily to a functional clinical benefit. There is no easy prescription for defining meaningful change, since subtle changes in strength can be reliably significant in one muscle group or behavioral activity, whereas larger changes may have little or no clinical impact for another functional behavior.

In the absence of a simple process for mapping from composite measures of neural function to global functional assessments of benefit, there is a real need for direct measures of clinically meaningful change. In this regard, the development of a SCI-specific measure of independence, the SCIM, is an improvement over the older and partly irrelevant FIM. However, such tools can be quite challenging as outcome measures in clinical trials, where significant effects may initially be quite small and variable between individuals, thus lending themselves to be documented first via an initial proof-of-principle study. Trials are also complicated by the need for accepted standards of rehabilitative care, and the economic challenges to their application, particularly in countries with patchwork health care. Without such standards, it is difficult to compare outcomes between trials or between different clinical centers.

There are a number of additional issues that deserve attention as we think about how to improve our tools and knowledge base in SCI. Beyond “clinical meaningfulness” there is limited ability to accurately address true QoL changes and societal health economics. Measurements of other important aspects of function, including the interplay between spasticity, pain and motor and sensory function are at an even earlier stage of development. There is also a concern about the potential for any treatment to produce heightened neuropathic pain (i.e. an adverse outcome), yet the tools we have to quantify dysesthesias and pain are not readily adapted as outcome measures, in part because of the multidimensional nature of these experiences.

Despite these concerns, the field of SCI treatment has benefitted from a rich history of coordinated clinical care efforts and, recently, a concerted effort to develop sensitive and accurate tools for therapeutic outcomes assessment. It is hoped that with the advent of SCOPE, the ICCP and other such initiatives, the coordination and iterative interplay...
between research and clinical practice in SCI will continue to evolve so that we can rapidly translate effective therapies into higher standards clinical care and treatment.

Acknowledgments

We are grateful for the support of ASIA, ICCP, ISCoS, and our corporate partners.

References


23. Sukhvinder Kalsi-Ryan. Quantification, sensitivity and reliability for the sensory module of the Graded and Redefined Assessment of Sensibility Strength and Prehension (GRASSP) hand measure; ASIA Annual Meeting; Tampa Bay FL. June 2007; poster presentation

24. Charles, T.; Carnel, MD.; Amie Jackson, MD.; John, F.; Ditunno, JR.; Mary Schmidt, MD.; Steve Williams, MD.; William Donovan, MD.; Mark Schmeler, MD.; Mike Boninger, MD. Outcome Measures for Gait and Ambulation in the Spinal Cord Injury Population Details. MD


Fig. 1. INTERNATIONAL SPINAL CORD INJURY DATA SETS