From the 2006 NIDRR SCI Measures Meeting

Functional Recovery Measures for Spinal Cord Injury: An Evidence-Based Review for Clinical Practice and Research

Report of the National Institute on Disability and Rehabilitation Research Spinal Cord Injury Measures Meeting

Functional Recovery Outcome Measures Work Group:

Kim Anderson, PhD¹; Sergio Aito, MD²; Michal Atkins, MA, OTR/L³; Fin Biering-Sørensen, MD⁴; Susan Charlifue, PhD⁵; Armin Curt, MD⁶; John Ditunno, MD⁷; Clive Glass, PhD⁸; Ralph Marino, MD, MSCE⁷; Ruth Marshall, MBBS, DPRM, FAFRM⁹; Mary Jane Mulcahey, PhD, OTR/L¹⁰; Marcel Post, PhD¹¹; Gordana Savic, MD, MSc¹²; Giorgio Scivoletto, MD¹³; Amiram Catz, MD, PhD¹⁴

Received September 17, 2007; accepted October 11, 2007

Abstract

Background/Objective: The end goal of clinical care and clinical research involving spinal cord injury (SCI) is to improve the overall ability of persons living with SCI to function on a daily basis. Neurologic recovery does not always translate into functional recovery. Thus, sensitive outcome measures designed to assess functional status relevant to SCI are important to develop.

Method: Evaluation of currently available SCI functional outcome measures by a multinational work group.

Results: The 4 measures that fit the prespecified inclusion criteria were the Modified Barthel Index (MBI), the Functional Independence Measure (FIM), the Quadriplegia Index of Function (QIF), and the Spinal Cord Independence Measure (SCIM). The MBI and the QIF were found to have minimal evidence for validity, whereas the FIM and the SCIM were found to be reliable and valid. The MBI has little clinical utility for use in the SCI population. Likewise, the FIM applies mainly when measuring burden of care, which is not necessarily a reflection of functional recovery. The QIF is useful for measuring functional recovery but only in a subpopulation of people with SCI, and substantial validity data are still required. The SCIM is the only functional recovery outcome measure designed specifically for SCI.

Conclusions: The multinational work group recommends that the latest version of the SCIM (SCIM III) continue to be refined and validated and subsequently implemented worldwide as the primary functional recovery outcome measure for SCI. The QIF may continue to be developed and validated for use as a supplemental tool for the nonambulatory tetraplegic population.

J Spinal Cord Med. 2008;31:133–144

Key Words: Recovery of function; Spinal cord injuries; Outcome measures; Functional Independence Measure; Spinal Cord Independence Measure; Quadriplegia Index of Function; Modified Barthel Index; ASIA impairment scale; Paraplegia; Tetraplegia; Validity; Reliability

INTRODUCTION

The clinical care and management of spinal cord injury (SCI) has progressed significantly, resulting in an increasing population worldwide of individuals living with SCI. Paralleling these advances, basic science and clinical research have made great strides in understanding the underlying biological mechanisms altered by SCI, and potential therapies to promote repair and recovery of function are nearly “at our doorstep.”

Over the last few years, however, there has arisen an increasing demand among scientists for sensitive outcome measures, specific to SCI, that have been rigorously tested and are appropriate for use in clinical trials. The default outcome measure is typically the ASIA (American

Spinal Injury Association) Impairment Scale (AIS), which was originally designed as a clinical tool to assess motor and sensory neurologic intactness (1). However, the AIS does not assess autonomic neurologic intactness, nor does it assess functional ability. Furthermore, to date, there have been no drugs approved by the Food and Drug Administration (FDA) for SCI treatment based on AIS outcomes and no accepted degree of improvement based on the AIS to define a clinically meaningful effect of treatment required to justify FDA approval (2). The ultimate goal of SCI clinical care and clinical research is straightforward: to improve the overall functional ability of persons living with SCI. Because neurologic recovery and functional recovery do not always parallel each other

in SCI, it is imperative that tools for clinical assessment of functional recovery be designed specifically for the SCI population rather than being applicable to any neurologic disorder or global disability.

In an effort to address this problem, on June 24, 2006, the National Institute on Disability and Rehabilitation Research (NIDRR) sponsored a premeeting course on SCI outcome measures at the American Spinal Injury Association and International Spinal Cord Society (ASIA-ISCoS) joint meeting held in Boston, Massachusetts. Several work groups were formed prior to the premeeting course and given the task of reviewing published outcome measures for various SCI components (eg, pain, sensory/motor strength, gait/ambulation, quality of life, functional recovery, spasticity, neuroimaging criteria). The proceedings presented here are that of the multinational work group that reviewed the existing literature regarding *functional recovery outcome measures*. The charge of this work group was to identify and evaluate outcome measures that assess overall daily activities (ie, functional status) for persons with SCI and, more specifically, to identify the utility of the measures for use in the clinic setting or in clinical trials when assessing functional recovery. To qualify for review, outcome measures under any of the categories had to meet 3 preset criteria:

1. History of prior use with SCI
2. Cited in the literature during 2000 to 2005
3. Cited by at least 2 different researchers or research groups during 2000 to 2005

The views expressed in this article are the result of independent research and do not necessarily represent the views of the US Department of Education. The SCI Measures Meeting was supported by grant funds from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitation Services in the US Department of Education.

¹University of California at Irvine, CA; ²Careggi Hospital, Florence, Italy; ³Rancho Los Amigos National Rehabilitation Center, Downey, California; ⁴Rigshospitalet, Copenhagen, Denmark; ⁵Craig Hospital, Englewood, Colorado; ⁶British Columbia University, Vancouver, BC, Canada; ⁷Thomas Jefferson University, Philadelphia, Pennsylvania; ⁸Northwest Regional Spinal Injuries Centre, Southport, UK; ⁹Hampstead Rehabilitation Centre, Northfield, Australia; ¹⁰Shriners Hospitals for Children, Philadelphia, Pennsylvania; ¹¹Institute for Rehabilitation Research, Groningen, The Netherlands; ¹²Stoke-Mandeville Hospital, Aylesbury, UK; ¹³IRCCS Fondazione, Santa Lucia, Florence, Italy; ¹⁴Lowenstein Hospital, Tel Aviv, Israel

Please address correspondence to Kim Anderson, PhD, Reeve-Irvine Research Center, Department of Neurological Surgery, University of California at Irvine, 843 Medical Science Court, Hewitt Hall Rm 1113, Irvine, CA 92697-1385; phone: 949.824.0056; fax: 949.824.0057 (e-mail: kanderso@uci.edu).

The 4 measures that fulfilled these criteria were:

1. Modified Barthel Index (MBI)
2. Functional Independence Measure (FIM)
3. Quadriplegia Index of Function (QIF)
4. Spinal Cord Independence Measure (SCIM)

METHODS

Parameters to analyze outcome measures were established prior to the review for all work groups to follow. For details on the development of those parameters, see Johnston and Graves, "Toward Guidelines for Evaluation of Measures: An Introduction with Application to Spinal Cord Injury" (3). The following properties of each measure were analyzed, if published data were available:

1. Underlying construct: general property/topic/aspect being measured, subscales/parameters measured, and any relevant statistics currently available
2. Administration: type/mode/equipment needed, performance, language/multicultural issues, and extent of burden/risk
3. Population applicability: populations in which the measure had been utilized, extent of use in SCI, and availability of normative data
4. Reliability/reproducibility and bias
5. Sensitivity to change and measurement range (ceiling/floor)
6. Validity (eg, construct, criterion-related, predictive) and clinical utility

The overall rating of validity, incorporating reliability, was then established using the scale below:

- = No formal validity/reliability data published, content inappropriate, not reviewed
- * = Questionable or insufficient
- ** = Minimal validity
- *** = Reliability and validity shown
- **** = Extensively validated and widely used

The review process for each functional outcome measure included 4 steps:

1. Primary literature review. Each outcome measure was reviewed independently by 2 members of the work group. The primary review consisted of critical evaluation of the published literature regarding the outcome measure and then completion of the "Schema for Rating Measures of Physical and Biological Constructs" (created by the organizational committee) (3).
2. Secondary review. For each outcome measure, the results of the primary reviews described above underwent a secondary review by a separate referee. The secondary referee summarized the primary reviews and recorded his/her own comments, scoring, and recommendations.
3. Refinement. The secondary reviews were submitted to all of the work group members for additional

comments, which were incorporated into the final documents.

4. Integration of all reviews for presentation at the SCI Outcome Measures NIDRR-sponsored preconference.

The findings of each primary review were reported in a table according to the instructions in the “Schema for Rating Measures of Physical and Biological Constructs.” In addition, reviewers were requested to assign to the reviewed outcome measure an overall score from 0 to 100, summing 7 subscales/items (Table 1).

Based on the primary and secondary reviews, 2 members prepared a draft of the results, which was then available for comment by all work group members in preparation of the final presentation at the NIDRR-sponsored preconference during the 2006 joint ASIA-ISCoS meeting.

RESULTS

For each individual measure, a detailed review of the 6 properties described above is provided.

Modified Barthel Index

Construct. The MBI measures the individual’s performance on 10 activities of daily living (ADL) functions in the area of self-care, continence, and locomotion; it measures the individual’s performance of daily functions (activities in terms of the International Classification of Functioning, disability and health) (4). The scale has been modified from the original BI (5) to increase the sensitivity to change of each item, most of all as to the quality and quantity of assistance. The MBI does not assess the easiest and most difficult functions for someone with SCI, such as respiratory management (ie, using a ventilator or a tracheal tube, taking care of tube hygiene, performing suction, assisting cough) and wheelchair-ground transfers.

There are 10 items that are scored based on the amount of physical assistance required to perform the task. Each item has 5 categories (instead of 3–4 in the original BI), which should increase the sensitivity and facilitate accuracy of coding, relative to the BI. The maximal score of each item represents the weighting of that item for its relative contribution to the total score. For example, personal hygiene constitutes a maximum 5 of 100 points, dressing 10 of 100 points, and chair/bed transfers 15 of 100 points. There are no subtotal scores because there are no subscales.

Administration. The MBI is administered by observation of the individual’s performance of ADL skills. It may require effort by the subject to perform each task the best s/he can, but administering the MBI does not require special equipment or particular training. Little time and little/no expense is required to administer the MBI. Cultural issues could be relevant. Kucukedvici (6), in translating the scale into Turkish, reported the need of conceptual translation for some items, meaning

that certain tasks may be interpreted differently in different cultures. There are no significant burdens or risks associated with administering the MBI.

Population Applicability. The MBI is somewhat of a generic measure, having been used mostly for cerebrovascular diseases, such as stroke. It has also been used in other disease populations, including geriatric-orthopedic injury, traumatic brain injury, and multiple sclerosis (7–10). It has not been used as extensively in SCI. There were 3 SCI references in the 2000 to 2005 time frame (6,11,12) and 3 references prior to 2000 (13–15). There are no SCI-specific normative data available for the MBI.

Reliability, Reproducibility, and Bias. The internal consistency of the MBI in SCI patients is good (alpha 0.88–0.90). The inter-rater reliability is sufficient at the item level (kappa 0.50–0.78) and good for the overall inter-rater agreement (intraclass correlation coefficient [ICC] 0.77) (4,6). However, there is no test-retest information available and no signal:noise ratio information. In addition, environmental changes can bias/influence the results (4,6).

Sensitivity to Change. The MBI has been utilized in different papers to evaluate functional outcomes after SCI rehabilitation. It has also been utilized to evaluate SCI subjects with and without pain (12) and to evaluate the relationships between functional ability and fitness (15). The MBI may be more sensitive to functional changes than the original BI in stroke (16,17), but its responsiveness has not been compared with that of other measures after SCI. The MBI is sensitive to changes after rehabilitation, and the scale does not detect cognitive changes, which may contribute to its sensitivity to changes in function, in SCI patients (18). The bowel and bladder items, however, are mostly focused on the presence/absence of incontinence and, therefore, cannot not be sensitive enough to changes in management for SCI individuals. There may be a floor effect at admission to rehabilitation, but there do not appear to be any ceiling/floor effects at discharge, at least for stroke and traumatic brain injury.

Validity. Rasch analysis generally supported the construct validity of the scale but also showed that bladder and bowel items had a considerable level of misfit (ie, they did not measure the same construct: disability) (6). Convergent criterion-related validity for SCI has been demonstrated by relating the MBI to the ASIA impairment scale and to ASIA motor (Spearman’s correlation coefficient was 0.55 at admission and 0.76 at discharge) and sensory score (Spearman’s correlation coefficient 0.43 at admission and 0.51 at discharge) (6).

The MBI at admission to rehabilitation has demonstrated validity in the prediction of the function status at discharge, rehabilitation efficacy, and length of rehabilitation stay of stroke patients (19,20); in the prediction of length of stay after traumatic brain injury (21); in the prediction of survival in cancer patients (22);

Table 1. Items Composing Overall Score for Reviewed Outcome Measures

Points Range	Aspect Evaluated
0–20	Relevance to recovery after spinal cord injury of the properties measured by the reviewed instrument. To what extent does the measure represent important functional achievements of the spinal cord injury individual? To what extent does the measure score higher functional achievements that do not involve medical compromise or that depend less than other functional achievements on technological aids? (For example, does it score an individual on intermittent catheterization higher than one on an indwelling catheter, or an individual who walks with aids higher than one who is ambulating independently in a wheelchair?)
0–30	Level of evidence of validity and reliability/reproducibility.
0–10	Sensitivity to functional changes in tasks of specific relevance after spinal cord injury (eg, sphincter management, mobility).
0–10	Clinical utility. (Are decisions in clinical practice altered depending on the measure?)
0–10	Cross-cultural comparability.
0–10	Significance of the total score. Is the total score meaningful? Are the item scores weighted for their relative contribution to the overall primary activities of daily living? Is their summation a reasonable representation of the overall performance in activities of daily living?
0–10	Convenience in use. How much time is required to obtain the scoring? Is it easy to perform it? Is scoring accurate without a manual at hand? Is scoring sufficiently accurate when performed by staff members who have not undergone training in using the instrument?

and in the prediction of discharge timing after hip arthroplasty (23). However, no predictive information exists for SCI.

Clinical Utility. There is no information regarding clinical utility in SCI. Overall, the MBI has construct validity, with the exception of the bladder and bowel items, and it has been validated and widely used in non-SCI populations. There is little validation in SCI. The MBI was rated ** (minimal validity) and assigned a score of 60.

Functional Independence Measure

Construct. The FIM, the most widely used functional measure, was developed to evaluate functional ability in daily activities, as represented mainly by the burden of care of a disability (24). It was not designed specifically for SCI. It evaluates 6 areas of function (self-care, sphincter management, transfers, locomotion, communication, and social cognition) based on 18 tasks. Scores of each task range from 1 to 7, with 1 reflecting full assistance and 7 being complete independence. The scale reflects the time, energy, effort, and equipment that are used to achieve the task. Bladder and bowel care items include an additional scoring system assessing frequency of accidents (1 being 5 or more accidents in the past 7 days and 7 being no accidents). Locomotion is measured in terms of distance. Cognition is measured in percentage of prompting or cuing.

Multiple correlational analyses have been performed comparing the FIM with a wide range of other functional measures that are relevant for SCI patients, including the Canadian Occupational Performance Measure (25), the

Capabilities of Upper Extremity (26), the QIF (27,28), the SCIM (29,30), and the Walking Index for Spinal Cord Injury (31), showing a fair level of correlation in some, but not all, areas and indicating that FIM and the other measures share common components in their underlying construct.

Administration. The FIM score is typically collected within 72 hours after admission to a rehabilitation unit, within 72 hours before discharge, and from 80 to 180 days after discharge. To administer the FIM initially, the clinician observes the client performing the task. Rating the FIM is quite complex and requires training (including study of a 200-page manual). The FIM is a trademarked instrument administered by Uniform Data System for Medical Rehabilitation (UDSmr). UDSmr certifies institutions, and the individual institutions certify their clinicians. Certification (or credentialing) is a process of ensuring the reliability of the FIM data submitted to the UDSmr. It includes a written test of FIM scoring skills and a technical review of the facility's data report. UDSmr collects data quarterly, for certified institutions only. Some detailed institution-specific rating decisions are tailored to the institution, and, therefore, a clinician who is certified must get recertified if s/he moves to a different facility.

The 18 tasks are designed to be evaluated by any certified clinician in the facility, not by a specific discipline (eg, nursing, occupational therapy, physical therapy, speech, and social work), but most rehabilitation facilities assign each functional area to a discipline. Equipment can be used by the patient to complete each task as needed according to the clinician's professional judgment.

Subjects are required to perform each task the best they can and perform as many of the 18 tasks as they are able according to the clinician's professional judgment. Subjects with C1-C4 ASIA A tetraplegia and often those with C1-C4 ASIA B cannot perform any task in the motor section.

The FIM has been used internationally; however, institutional certification (UDSmr) is limited to the USA and recently Australia through a special licensing arrangement. The FIM has been utilized across North America and worldwide and translated into several languages [eg, Turkish (32) Japanese (33)]. Rasch analysis of multicountry European data suggests that the present scoring system lacks cross-cultural validity (34).

There is fundamentally no burden or risk associated with administering the FIM, but when the FIM is administered during the first 72 hours of admission, individuals with SCI may not be up to performing the task (either being tired or sick or in pain). Because rating depends on observing the subject performing the task, clinicians must request that the subject participate, which may put undue pressure on both clinicians and clients. This is not, however, a criticism of the scale per se, but rather its method of use in some places, and it is not specific to the FIM. Self-ratings (35–37) and questioning (38) including telephone interview have been used as modes of administration (36,39). Shorter versions and self-ratings can shorten the time for data collection, which, together with the possibility of telephone interview, make it possible to use the FIM in a variety of settings and for different purposes. However, self-ratings do not provide data that are comparable to observer-administered ratings.

Population Applicability. The FIM has been used extensively across a wide range of disability groups [eg, multiple sclerosis (40), stroke (41), traumatic brain injury (42), SCI] and across the adult age band (16+ years). A specific instrument has been developed for children: WeeFIM (43) (outside the scope of the current review). As mentioned previously, however, the FIM is not sensitive to change in subjects with high tetraplegia (C1–C4): most patients with those injury levels score the minimum on the motor items at admission and discharge, except for (electric) wheelchair locomotion.

In the USA, most Model SCI System Centers use the FIM. Additionally, it is used worldwide in SCI centers both for research and for clinical purposes. Standardized forms are used for data collection; however, normative values for various combinations of SCI level and completeness have not been published.

Reliability, Reproducibility, and Bias. There is an extensive publication history regarding reliability and reproducibility of the FIM in large ($n > 1,000$) patient groups. Cronbach's alpha is high for all FIM total scores and subscales (>0.70). Coefficients are nearly all high ($\alpha > 0.90$) for admission and discharge total scores,

with only the locomotion subscale having a low internal consistency (44). Social cognitive items showed low internal consistency (45), however, and these items are not sensitive to change in SCI subjects. The inter- and intra-rater reliability is high ($r = 0.81$ – 0.96). In addition, the ICC for total FIM scores is 0.97, with individual subscales ranging from 0.93 to 0.96. The item kappa mean is 0.71 (range 0.61–0.76) (46–48), and the weighted kappa exceeds 0.70 for all subscales except memory (49). A subset selection of 7 FIM items has an ICC of 0.99 (39).

Potential bias can be introduced if the treating therapists are completing the scale rather than nontreating FIM evaluators (44). Comparisons between clinician ratings and self-ratings (37) and between questioning and observation (38) have shown strong correlations ($r = 0.828$ and 0.94 , respectively). The self-report version of the FIM (FIM-SR) showed high test-retest reliability in persons with SCI (individual item kappa coefficients near or above 0.6 and most higher than 0.7 and total score coefficient 0.89) (35,36).

Sensitivity to Change. FIM items lack adequate sensitivity to detect changes associated with return of function after SCI. The locomotion items are the least sensitive to change, in contrast to the transfer items of the FIM (29,30,44,50,51). Further, the FIM is less sensitive to small gains compared with the QIF (27) in persons with tetraplegia, less sensitive to walking recovery than the walking index for spinal cord injury (31), and less sensitive to functional changes in SCI than the SCIM (29,30).

The ceiling effect for cognitive items limits the utility of the FIM for detecting change over time (34,36,52,53). If the FIM were to be used in clinical research settings to assess motor function, it may be more useful to only use the motor subscale portion of the FIM (eg, "motor FIM") and disregard the cognitive subscale portion. Overall, the floor and ceiling effects for most items are small but not in all populations: caution must be taken with FIM interpretation in populations that are either severely disabled or very mildly disabled (54). There is a ceiling effect for bed transfer in paraplegia and a floor effect for tetraplegia (50). In addition, total scores are influenced by floor and ceiling effects for differing lesion levels.

Validity. Rasch analysis has been applied frequently to the FIM (34,55–58) to assess validity. A 4-category scale containing 2 dimensions rather than 1 (Motor and Cognition) (59) measures dependence more precisely in the clinical setting (56). There is low cross-cultural validity unless adjustments for differential item functioning are made (57), including omission of a few country-specific items, to allow for the valid pooling of data (34). However, low cross-cultural validity does not affect the scale utility in clinical practice.

Face, content, construct, discriminant, and criterion-related concurrent validity for the FIM are high (42,44,47,60,61). The FIM correlates well with other

function scales [eg, QIF (27,28) and Capabilities of Upper Extremity (26)]. The predictive validity of the FIM is supported by score decrease with ascending level of SCI and significant improvements between admission and discharge detectable for all impairment groups except C1 to C4 AIS A/B SCI (44). Functional status, as measured by the FIM, has been used to successfully predict life satisfaction after SCI and traumatic brain injury (62).

Clinical Utility. The clinical utility of the FIM for measuring functional recovery in SCI is questionable. Some studies have failed to detect changes in function in subjects with SCI over time, and changes in function are dependent on age/lesion level (63). The FIM is unable to detect distinct differences in the continuum from tetraplegia to paraplegia (44), and there are no differences when comparing scores for C8, T1 to T5, and T6 or below (64). There is, however, a better delineation between paraplegia and tetraplegia when 5 mobility and locomotor skills are added (50).

The FIM was developed by a multiprofessional team in the 1980s and has been used across a number of countries and patient populations, including SCI. It is a valid measure of the burden of care (42) and has excellent reliability and validity, although caution is required with interpretation of cognitive items (37,52,53). The high correlations among several of the motor items suggest redundancy (39,52).

Overall, the validity and reliability of the FIM for measuring the burden of care is high, but it is not specifically designed for the SCI population. As such, it lacks specificity, it underscores in the area of sphincter management, and it does not evaluate respiratory management. The FIM was rated *** (reliability and validity shown) and assigned a score of 64.

Quadriplegia Index of Function

Construct. The QIF was specifically developed for assessment of people living with tetraplegia. The aim was to provide a sensitive scale for measuring functional gains during rehabilitation in this subpopulation of SCI because the existing instruments were deemed insensitive to detect minor but significant functional improvements in ADL in people with cervical injuries (65). The QIF (65) measures the level of independence in 10 tasks of ADL (called “categories” in publications describing the QIF): (a) transfers, (b) grooming, (c) bathing, (d) feeding, (e) dressing, (f) wheelchair mobility, (g) bed activities, (h) bladder program, (i) bowel program, and (j) understanding of personal care. The tasks consist of up to 10 items each, and all items are weighted for their relative contribution to the total score. Each item represents a different activity. For the first 7 tasks, each item is scored separately 0 to 4 (independent, independent with devices, supervision, assistance needed, dependent); the last 3 tasks have separate sets of scoring criteria (65,66). The total QIF score ranges from 0 to 100. Subscores for each task can also be used

separately, as can the scores for individual items within each task. The items are weighted in terms of their assumed clinical relevance. Some authors have reported using a portion of the QIF in their studies (67–70). The QIF does not assess walking, stair management, respiration management, or bladder management.

A short-form version of the QIF (66) was developed and consists of 6 tasks that were found to be the best predictors of the total score through regression analysis (wash/dry hair, turn supine to side in bed, put on lower body clothing, open carton/jar, transfer from bed to chair, and lock wheelchair). The total short-form QIF score [0–24] is a simple sum of the 6 tasks, each scored 0 to 4 (no task weighting was used).

Administration. Administration of the QIF is simple if done by an experienced assessor. It requires up to 30 minutes for the full QIF and only a few minutes for the short-form version (65,66). The QIF can be administered by interview and/or observation. Self-report through interview appears to be the most frequently used mode of administration, but tasks 1 to 9 can be evaluated through observation (which would make the process much longer than the stated 30 minutes). No specific equipment is required for administration, and the instructions for scoring are on the form.

The QIF seems straightforward for international use. It has been used in various countries [Turkey (27), Serbia (69), USA (28,66–68,70), and New Zealand (71)]. It is not clear whether the QIF actually was translated into Turkish and Serbian languages or whether the observers used the English version. There have been no cross-cultural comparisons, however.

Population Applicability. The QIF was designed specifically for use with people with cervical SCI, in particular nonambulatory individuals, because it only focuses on ADLs requiring hand function. It could be applicable for people with other neurological disorders affecting the upper limbs, but there are no reports in the literature regarding use outside the SCI population. The QIF has been used many times in SCI but not as extensively as some other measures. There were only 8 citations (28,66–71) after the original paper (65), and only 2 of those were since 2000 (70,71). It has been used as an outcome measure in assessing the relationship between level of injury and functional independence (67,68), as well as for evaluation of efficacy of treatment [Bionic glove (69), upper limb tendon transfer surgery (71), and implanted Freehand system (70)]. There are no SCI-specific normative data available for the QIF.

Reliability, Reproducibility, and Bias. In the original QIF study (65), good inter-rater reliability was shown among 3 raters studying the same 20 subjects, all with complete tetraplegia. Pearson’s correlation coefficient ranged from 0.62 to 0.95 ($P < 0.001$) for individual tasks (65), but no subsequent studies have reported inter-rater reliability.

Sensitivity to Change. The QIF is more sensitive to change than the BI, Kenny Self-Care Evaluation (65), and

FIM (27,28). There is also a stronger correlation between the QIF and the ASIA Upper Extremity Motor Score than between the FIM and the ASIA Upper Extremity Motor Score (27,28). The QIF has not been compared with the SCIM. Use of part of the QIF allows assessment of individual tasks. A total score based on tasks 1 to 3 shows fairly even distribution, with clustering only at the maximum score. There is a possible ceiling effect in subjects with low cervical SCI (good hand function). A problem with the short-form QIF involves poor discrimination between C7 and C8 lesions (66).

Validity. The QIF has face validity in that it was originally developed by an experienced multidisciplinary SCI team (65). It includes relevant variables for measuring performance after cervical SCI. The QIF uses weighting of items for their contribution to the total score, but parts of the total QIF can be used with a simple scoring system if subsections/subscores are analyzed separately. The short-form QIF relies on this latter method. In regard to criterion-related validity, the QIF has a good correlation with the FIM [$r = 0.97$ (27) and 0.93 (28)] and strong correlations with ASIA motor scores [$r = 0.91$ (27), 0.91 (28)] and is also correlated with ASIA sensory scores ($r = 0.64$ for light touch, 0.65 for pinprick) (27). Correlation is particularly strong between QIF grooming, bathing, and feeding scores and ASIA Upper Extremity Motor Score. No predictive validity data have been established.

Clinical Utility. QIF appears to be useful for individuals with cervical SCI (a) for monitoring functional progress, (b) for evaluating rehabilitation efficacy in neurologically stable patients, and (c) as an outcome measure in clinical trials.

Overall, the QIF appears to be specifically useful for evaluating ADL reflecting hand function in non-ambulatory tetraplegia. Only 3 comparative validity studies exist and only 1 study of reliability. There is no justification presented for the original item weighting and no rules for scoring if there are missing items. The QIF was rated ** (minimal validity) and assigned a score of 77.

Spinal Cord Independence Measure

Construct. The SCIM is a scale developed specifically for people with SCI to evaluate their performance of ADL and to make functional assessments of this population sensitive to change (29). SCIM scores a task higher in patients who accomplish it with less assistance, aids, or medical compromise than other patients (72). Three versions of the SCIM (I, II, III) (29,72–75) have been developed since its first publication in 1997. The most recent version of the SCIM [SCIM III (72,75)] is composed of 19 items in 3 subscales: (a) self-care (6 items, subscore 0–20), (b) respiration and sphincter management (4 items, subscore 0–40), and (c) mobility (9 items, subscore 0–40). The total score ranges from 0 to 100. Mobility is subdivided into “room and toilet” and “indoors and outdoors” (29). The items are weighted in terms of their assumed clinical relevance.

Administration. The SCIM can be administered by observation. Subjects are asked to show their best performance, and no measurement tool is necessary other than a form that is filled out by the observer. An interview method comparable to the observation method has also been reported (76). There is no burden or risk associated with SCIM administration. The assessment is performed by therapists and/or nurses and requires 30 to 45 minutes. The interview requires less time. The SCIM has been evaluated in multiple countries in North America, Europe, and the Middle East, and, so far, it appears to be resistant to cross-cultural differences (72,75).

Population Applicability. The SCIM is a condition-specific measure for traumatic and nontraumatic SCI (30,73). It provides information complementary to the neurological deficit in SCI and can be applied in acute and chronic conditions. The use of the SCIM is steadily increasing. Information on the validation of the latest version of the SCIM, SCIM III, has been recently published (72,75), and earlier versions of the SCIM have been used by different research groups involved in clinical trials (77,78) or descriptive studies (79,80) or as a benchmark for another instrument (31,81). The SCIM is being used increasingly in Europe. A few centers in North America have started to add the SCIM to the FIM, although in North America the FIM is an administrative requirement to obtain Medicare reimbursement (UDSmr participation is voluntary).

The SCI Ability Realization Measurement Index is a method to set norms for SCIM scores that could be predicted, based on the motor impairment (ASIA total motor index) of the subject (82). The developers, however, have refrained from setting SCIM norms for the level-ASIA grade combination because of the difficulty to determine the hierarchy of certain level-ASIA grade combinations (which impairment is more severe, that of a C5D SCI or that of L3A?). Currently, therefore, there are no SCIM values from large samples related to level and completeness of lesion, although they could potentially be developed in the future. It is important to note that the SCIM was developed on a mixed population of traumatic and nontraumatic spinal cord lesions. The outcome of subjects with nontraumatic spinal cord lesions may vary from those with traumatic lesions because of diverse causes and clinical course (79,83,84), but a study of differential item functioning has shown that the SCIM III measurement is similar for traumatic and nontraumatic SCI (72).

Reliability, Reproducibility, and Bias. Extensive reports, by the SCIM developers, of reliability and reproducibility are available. Inter-rater reliability examination showed a kappa coefficient between 0.66 and 0.98 for the items in the original SCIM scale (29). Rasch analysis of SCIM II showed no differential item functioning between first and second examinations: task difficulty was similar in both examinations (74). The reliability of SCIM III was reflected

in kappa values of 0.64 to 0.84, ICC > 0.94, Cronbach's alpha > 0.7, and lack of significant differential item functioning in Rasch analysis ($P < 0.05$) across contrasted subgroups of age, gender, lesion etiology, and lesion level (tetraplegia vs paraplegia) (72,75). SCIM III tasks seem to provide a proper hierarchical ordering of their categories (72,74).

The one potential bias is that the multidisciplinary team, who participated in the development of the original scale, also participated in the assessment of the original scale (29), but validation of SCIM III has been performed since then in a multicenter study by individuals who did not participate in the development of the scale (72).

Sensitivity to Change. The SCIM is more sensitive to change than the FIM. When compared, the SCIM detected all functional changes that the FIM detected but FIM missed 26% of the changes detected by the SCIM (29). Similar findings were subsequently reported based on SCIM versions II and III (72–74). The change in SCIM scores was associated with neurological recovery and progression through rehabilitation (82). In 2 small randomized clinical trials (fewer than 30 subjects each), in which SCIM II was utilized for evaluating the success of interventions in SCI, there was no difference noted between control and intervention groups (77,78). This could be related to insufficient sensitivity to change of SCIM II, but it may well reflect failure of the intervention to introduce a functional change. However, there were within group changes over time, indicating that the SCIM is sensitive to change. Overall, the SCIM is sensitive to change in function in persons with SCI and demonstrates the ability to capture changes in function not captured by the FIM. The content advantages of the SCIM over the FIM in the areas of bladder, bowel, respiration, and mobility have clinical relevance and are important measures in the disability domain for clinical trials.

The SCIM is insensitive to changes or differences in walking speed because it does not measure this component of function. In the SCIM III study (72), some ceiling and floor effects were found, but it remains unclear if these effects are due to the nature of the subjects involved (cohort of subjects that does not represent the whole range of impairment) or due to the SCIM itself. It may be that the SCIM has too few very easy or very difficult items.

Validity. The initial face validity of the SCIM was determined by the multidisciplinary team that developed and tested it at one center (29). Construct validity was reported for SCIM II, in a study of a sample of more than 200 subjects from 1 country (74), and for SCIM III, in another study of 425 subjects from 13 centers in 6 countries by Rasch analysis (72). The findings revealed a good fit (0.8–1.4; $P < 0.05$) for all tasks (items) and most subscales (72,74). Criterion-related validity was tested by comparison with the FIM [correlation between SCIM I and FIM $r = 0.85$ (29), correlation between SCIM III and

FIM $r = 0.79$ (75)], because there is currently no “gold standard” SCI ability measure. Strong correlations ($r = 0.68$ – 0.85) were also found between SCIM subscales and corresponding scales of the needs assessment checklist (81).

There are no studies assessing the accuracy of prediction of outcome, such as neurological or functional status at the end of rehabilitation, based on initial SCIM scores. However, in one study, initial highest SCIM II scores in patients with similar ASIA motor scores accurately predicted the initial ASIA motor scores of these patients ($r = 0.96$) (82). In another study, initial SCIM II scores and the SCIM II gain predicted the length of stay in patients who remained in the hospital for 34 to 70 days after admission ($r = -0.73$ and 0.81 , respectively) (80).

Clinical Utility. The developers of the SCIM indicate that it is valid for use in clinical practice to evaluate function and plan goals of rehabilitation. As yet, there are no published reports on the utilization of SCIM data for treatment decision making by clinical SCI programs. Published information regarding clinical utility would be beneficial because the refinement of the SCIM III appears to indicate that it is a valuable measure of ability to perform ADLs in the SCI population.

Overall, the SCIM III represents the most sensitive, reliable, and valid measure of global disability that exists for individuals with SCI. These properties were demonstrated in a multinational study. The SCIM was rated *** (reliability and validity shown) and assigned a score of 75.

DISCUSSION

A summary of the star rating and overall score of each outcome measure is provided in Table 2. Both the SCIM and FIM were rated *** for having validity and reliability. The SCIM is useful for measuring global disability specific to the SCI population, whereas the FIM is mainly useful in measuring burden of care. The SCIM was given an overall score of 75, and the FIM was given a score of 64.

The audience was queried during the functional recovery session of the NIDRR-sponsored preconference on SCI outcome measures. For all data below, N refers to the number of voters in the audience who voted on a particular question. The voters/audience consisted of SCI clinicians and scientists attending the ASIA-ISCoS preconference session on Functional Assessment. The majority (85% and 83%, respectively) voted that assessment of functional recovery for clinical and research purposes, in persons with SCI, should take into account “patient preference, achievement's medical value, specific assistive devices, + personal assistance required” ($N = 40$ and $N = 48$, respectively). When asked which measure (MBI, FIM, QIF, or SCIM) best represented the SCI individual's functional status for clinical purposes, 94% voted for the SCIM (4% for FIM, 2% for MBI, 0% for QIF; $N = 52$). When asked which measure was most useful for assessing functional improvement after therapeutic

Table 2. Summary of Star Ratings and Overall Scores

Outcome Measure	Star Rating	Overall Score
Modified Barthel Index	**	60
Functional Independence Measure	***	64
Quadriplegia Index of Function	**	77
Spinal Cord Independence Measure	***	75

interventions, such as in SCI clinical trials, 86% voted for the SCIM (10% for QIF, 4% for MBI, 0% for FIM; N = 49).

Overall, based on the findings of this systematic review, the SCIM has been found to be the most valid and useful measure of functional recovery in the SCI population for clinical and research purposes. The latest version of the SCIM (SCIM III) appears to represent the most sensitive, reliable, and valid measure of SCI-specific global disability to date, and its properties are resistant, to the extent that they have been tested, to potential multinational differences. Ongoing study of the clinical and research utility and validity of the SCIM III promises to yield additional information needed to further refine the measure in order to improve the ability to detect, measure, and predict functional recovery in the SCI population.

The QIF is considered useful for evaluating several aspects of ADL function in nonambulatory tetraplegia, particularly fine hand functions, but requires further validation or changes/improvement. It is less useful, however, for incomplete tetraplegia because walking is not assessed. Ceiling effects limit its use in mixed groups of tetraplegia and paraplegia. When the audience at the functional recovery session of the NIDRR course was queried as to whether the QIF should be used in addition to the MBI, FIM, or SCIM for subjects with complete tetraplegia, 32% voted “yes, always,” 38% voted “yes, but only for research purposes,” 13% voted “yes, but only for clinical purposes,” and 17% voted “no” (N = 47).

There is a noticeable discrepancy between the 2 global ratings for the QIF, with the QIF receiving the highest overall score, 77 (because it is good at what it measures: small improvements mainly in hand function), but only ** and a mixed reaction from the audience (because it is only applicable in nonambulatory patients with tetraplegia). The problem is that the scoring system (0–100), used by all of the work groups for the NIDRR-sponsored preconference on SCI outcome measures, to rate the quality of the measures was designed for the *entire* SCI population. Thus, it is likely that this scoring system is inappropriate for use in assessing an outcome measure designed for a very specific *subpopulation* of SCI. Studies comparing the sensitivity to change of the SCIM and QIF may be useful: QIF seems to be most sensitive for changes in people with complete tetraplegia but has never been compared with SCIM. Regardless, further validation and reliability studies need to be performed for

the QIF. In addition, it is recommended that modifications addressing the weighting of categories and rules for scoring when items are missing be pursued.

The MBI appears to have applicable content with good validity and reliability in SCI but has little clinical utility in SCI. It may be useful only when moderate to large differences in groups are important. Only 4% of the audience at the functional recovery session voted that the MBI was useful for assessing functional improvement after therapeutic interventions in SCI.

The FIM is currently probably the most widely used clinical evaluation tool for functional ability in the SCI population, and there is certainly a large body of validity and reliability data for multiple disabilities. However, the biggest limitation of the FIM in relation to SCI is in regard to insensitivity to change. The FIM is less sensitive to small gains of function than the QIF in the tetraplegic population, a population in which recovery of a small amount of function can translate into a large degree of independence. The FIM is less sensitive to walking recovery than the Walking Index for Spinal Cord Injury (85) and less sensitive to functional changes than the SCIM. Further, there is no relationship between the FIM social cognition or communication subscale and neuropsychological tests of well being in the SCI population. At the functional recovery session, 0% of the audience voted that the FIM was useful for assessing functional improvement after therapeutic interventions in SCI (N = 49). Hence, in view of being able to detect small but functionally significant amounts of change in the SCI population in future clinical trials, the FIM is not a suitable outcome tool.

CONCLUSIONS

The multinational work group recommends continued development and validation of the SCIM III as the primary outcome measure to assess functional recovery. It would be a misuse of precious time and resources to develop brand new outcome measures. Further, it is recommended that the leading institutions and organizations specializing in SCI care and research throughout the world work together in testing, utilizing, and validating the SCIM III so that the most sensitive and appropriate outcome measure to assess functional recovery in the SCI population can be further developed in a timely manner. Continued research is needed with the SCIM III with regard to assessing recovery of the ability to perform tasks in the acute/subacute phase of SCI, and metric research is needed to establish norms for the improvement in ability allowed by the degree of natural neurological recovery. Such norms may be important when differentiating between natural recovery and the effect of clinical trial interventions and can be achieved through further development of norms for SCIM scores (ie, using the Spinal Cord Injury Ability Realization Measurement Index).

When evaluating subpopulations of SCI, it may be appropriate to supplement the SCIM III with a more selective outcome measure. As such, it is recommended that the QIF continue to be developed and validated for use as a supplemental tool when evaluating the nonambulatory tetraplegic population.

REFERENCES

- Maynard FM Jr, Bracken MB, Creasey G, et al. International Standards for Neurological and Functional Classification of Spinal Cord Injury. American Spinal Injury Association. *Spinal Cord*. 1997;35:266–274.
- Blight AR, Tuszynski MH. Clinical trials in spinal cord injury. *J Neurotrauma*. 2006;23:586–593.
- Johnston MV, Graves D. Towards guidelines for evaluation of measures: an introduction with application to spinal cord injury. *J Spinal Cord Med*. 2008;31:13–26.
- Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol*. 1989;42:703–709.
- Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J*. 1965;14:61–65.
- Kucukdeveci AA, Yavuzer G, Tennant A, Suldur N, Sonel B, Arasil T. Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey. *Scand J Rehabil Med*. 2000;32:87–92.
- Shah S, Muncer S. Stroke rehabilitation: Australian patient profile and functional outcome. *J Clin Epidemiol*. 1991;44:21–28.
- Crotty M, Whitehead C, Miller M, Gray S. Patient and caregiver outcomes 12 months after home-based therapy for hip fracture: a randomized controlled trial. *Arch Phys Med Rehabil*. 2003;84:1237–1239.
- Chua KS, Kong KH. Functional outcome in brain stem stroke patients after rehabilitation. *Arch Phys Med Rehabil*. 1996;77:194–197.
- Yap EC, Chua KS. Rehabilitation outcome after primary subarachnoid haemorrhage. *Brain Inj*. 2002;16:491–499.
- Nas K, Kemaloglu MS, Cevik R, et al. The results of rehabilitation on motor and functional improvement of the spinal tuberculosis. *Joint Bone Spine*. 2004;71:312–316.
- Yap EC, Tow A, Menon EB, Chan KF, Kong KH. Pain during in-patient rehabilitation after traumatic spinal cord injury. *Int J Rehabil Res*. 2003;26:137–140.
- Tow AM, Kong KH. Central cord syndrome: functional outcome after rehabilitation. *Spinal Cord*. 1998;36:156–160.
- Anderson PA, Bohlman HH. Anterior decompression and arthrodesis of the cervical spine: long-term motor improvement. Part II: improvement in complete traumatic quadriplegia. *J Bone Joint Surg Am*. 1992;74:683–692.
- Noreau L, Shephard RJ, Simard C, Pare G, Pomerleau P. Relationship of impairment and functional ability to habitual activity and fitness following spinal cord injury. *Int J Rehabil Res*. 1993;16:265–275.
- Shah S, Muncer S. Sensitivity of Shah, Vanclay and Cooper's modified Barthel Index. *Clin Rehabil*. 2000;14:551–552.
- Hocking C, Williams M, Broad J, Baskett J. Sensitivity of Shah, Vanclay and Cooper's modified Barthel Index. *Clin Rehabil*. 1999;13:141–147.
- Ng YS, Chua KS. States of severely altered consciousness: clinical characteristics, medical complications and functional outcome after rehabilitation. *NeuroRehabilitation*. 2005;20:97–105.
- Shah S, Vanclay F, Cooper B. Predicting discharge status at commencement of stroke rehabilitation. *Stroke*. 1989;20:766–769.
- Shah S, Vanclay F, Cooper B. Efficiency, effectiveness and duration of stroke rehabilitation. *Stroke*. 1990;21:241–246.
- Shah S, Muncer S. A comparison of rehabilitation outcome measures for traumatic brain injury. *OTJR*. 2003;23:2–9.
- Bennett M, Ryall N. Using the modified Barthel Index to estimate survival in cancer patients in hospice: observational study. *BMJ*. 2000;321:1381–1382.
- Wang T, Ackland T, Hall S, Gilbey H, Parsons R. Functional recovery and timing of hospital discharge after primary total hip arthroplasty. *Aust NZ J Surg*. 1998;68:580–583.
- Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. *Adv Clin Rehabil*. 1987;1:6–18.
- Donnelly C, Eng JJ, Hall J, et al. Client-centered assessment and the identification of meaningful treatment goals for individuals with a spinal cord injury. *Spinal Cord*. 2004;42:302–307.
- Marino J, Shea JA, Stineman MG. The Capabilities of Upper Extremity Instrument: reliability and validity of a measure of functional limitation in tetraplegia. *Arch Phys Med Rehabil*. 1998;79:1512–1521.
- Yavuz N, Tezyurek M, Akyuz M. A comparison of two functional tests in quadriplegia: the Quadriplegia Index of Function and the Functional Independence Measure. *Spinal Cord*. 1998;36:832–837.
- Marino RJ, Huang M, Knight P, Herbison GJ, Ditunno JF Jr, Segal M. Assessing self-care status in quadriplegia: comparison of the quadriplegia index of function (QIF) and the functional independence measure (FIM). *Paraplegia*. 1993;31:225–233.
- Catz A, Itzkovich M, Agranov E, Ring H, Tamir A. SCIM—spinal cord independence measure: a new disability scale for patients with spinal cord lesions. *Spinal Cord*. 1997;35:850–856.
- Catz A, Itzkovich M, Agranov E, Ring H, Tamir A. The spinal cord independence measure (SCIM): sensitivity to functional changes in subgroups of spinal cord lesion patients. *Spinal Cord*. 2001;39:97–100.
- Morganti B, Scivoletto G, Ditunno P, Ditunno JF, Molinari M. Walking index for spinal cord injury (WISCI): criterion validation. *Spinal Cord*. 2005;43:27–33.
- Kucukdeveci AA, Yavuzer G, Elhan AH, Sonel B, Tennant A. Adaptation of the Functional Independence Measure for use in Turkey. *Clin Rehabil*. 2001;15:311–319.
- Yamamoto M, Sonoda S, Hayashi T. New evaluation method of ADL—Functional Independence Measure (FIM) [Japanese]. *Japan J Nurs Techniques*. 1992;38:922–925.
- Lawton G, Lundgren-Nilsson A, Biering-Sorensen F, et al. Cross-cultural validity of FIM in spinal cord injury. *Spinal Cord*. 2006;44:746–752.
- Hoenig H, McIntyre L, Sloane R, Brancj LG, Truncali A, Horner RD. The reliability of a self-reported measure of disease, impairment, and function in persons with spinal cord dysfunction. *Arch Phys Med Rehabil*. 1998;79:378–387.

36. Masedo AI, Hanley M, Jensen MP, Ehde D, Cardenas DD. Reliability and validity of a self-report FIM™ (FIM-SR) in persons with amputation or spinal cord injury and chronic pain. *Am J Phys Med Rehabil.* 2005;84:167–178.
37. Grey N, Kennedy P. The Functional Independence Measure: a comparative study of clinician and self ratings. *Paraplegia.* 1993;31:457–461.
38. Karamahmetoglu SS, Karacan I, Elbasi N, Demirel G, Koyuncu H, Dosoglu M. The functional independence measure in spinal cord injured patients: comparison of questioning with observational rating. *Spinal Cord.* 1997; 35:22–25.
39. Dijkers MPJM, Yavuzer G. Short versions of the telephone motor functional independence measure for use with persons with spinal cord injury. *Arch Phys Med Rehabil.* 1999;80:1477–1484.
40. 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;71:870–875.
41. Granger CV, Cotter AC, Hamilton BB, Fiedler RC. Functional assessment scales: a study of persons after stroke. *Arch Phys Med Rehabil.* 1993;74:133–138.
42. Heinemann AW, Kirk P, Hastie BA, et al. Relationships between disability measures and nursing effort during medical rehabilitation for patients with traumatic brain injury and spinal cord injury. *Arch Phys Med Rehabil.* 1997; 78:143–149.
43. Garcia RA, Gaebler-Spira D, Sisung C, Heinemann AW. Functional improvement after pediatric spinal cord injury. *Am J Phys Med Rehabil.* 2002;81:458–463.
44. 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;74:531–536.
45. Segal ME, Ditunno JF, Staas MD. Inter-institutional agreement of individual functional independence measure (FIM) items measures at two sites on one sample of SCI patients. *Paraplegia.* 1993;31:622–631.
46. Ottenbacher KJ, Hsu Y, Granger CV, Fieldler RC. The reliability of the functional independence measure: a quantitative review. *Arch Phys Med Rehabil.* 1996;77: 1226–1231.
47. Kidd D, Stewart G, Baldry J, et al. The Functional Independence Measure: a comparative validity and reliability study. *Disabil Rehabil.* 1995;17:10–14.
48. Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Inter-rater reliability of the seven level Functional Independence Measure (FIM). *Arch Phys Med Rehabil.* 1994;26:115–119.
49. Chau N, Daler S, Andrew J, Patris A. Inter-rater agreement of two functional independence scales the Functional Independence Measure (FIM) and a subjective uniform continuous scale. *Disabil Rehabil.* 1994;16:63–71.
50. Middleton JW, Harvey LA, Batty J, Cameron I, Quirk R, Winstanley J. Five additional mobility and locomotor items to improve responsiveness of the FIM in wheelchair-dependent individuals with spinal cord injury. *Spinal Cord.* 2006;44:495–504.
51. Middleton JW, Truman G, Geraghty TJ. Neurological level effect on the discharge functional status of spinal cord injured persons after rehabilitation. *Arch Phys Med Rehabil.* 1998;79:1428–1432.
52. Hall KM, Cohen ME, Wright J, Call M, Werner P. Characteristics of the Functional Independence Measure in traumatic spinal cord injury. *Arch Phys Med Rehabil.* 1999;80:1471–1476.
53. Davidoff GN, Roth EJ, Haughton JS, Ardner MS. Cognitive dysfunction in spinal cord injury patients: sensitivity of the Functional Independence Measure subscales vs neuropsychologic assessment. *Arch Phys Med Rehabil.* 1990;71: 326–329.
54. Stineman MG, Marino RJ, Deutsch A, Granger CV, Maislin G. A functional strategy for classifying patients after traumatic spinal cord injury. *Spinal Cord.* 1999;37:717–725.
55. Linacre JM. FIM levels as ordinal categories. *J Outcome Meas.* 2000;42:616–633.
56. Nilsson AL, Sunnerhagen KS, Grimby G. Scoring alternatives for FIM in neurological disorders applying Rasch analysis. *Acta Neurol Scand.* 2005;111:264–273.
57. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation through differential item functioning within the framework of the Rasch model: the PRO-ESOR Project. *Med Care.* 2004;42(1 Suppl):I37–I48.
58. Heinemann AW, Linacre JM, Wright BP, Hamilton BB, Granger C. Relationships between impairment and physical disability as measured by the functional independence measure. *Arch Phys Med Rehabil.* 1993;74:566–573.
59. 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: 127–132.
60. Granger CV, Hamilton BB, Keith RA, Zielezky M, Sherwin ES. Advances in functional assessment for medical rehabilitation. *Geriatric Rehabil.* 1986;1:59–74.
61. 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;77:1101–1108.
62. Hicken BL, Putzke JD, Novack T, Sherer M, Richards JS. Life satisfaction following spinal cord and traumatic brain injury: a comparative study. *J Rehabil Res Dev.* 2002;39: 359–365.
63. Menter RR, Whiteneck GG, Charlifue SW, et al. Impairment, disability, handicap and medical expenses of persons ageing with spinal cord injury. *Paraplegia.* 1991;29:613–619.
64. Ota T, Akaboshi K, Nagata M, et al. Functional assessment of patients with spinal cord injury: measured by the motor score and the Functional Independence Measure. *Spinal Cord.* 1996;34:531–535.
65. Gresham GE, Labi ML, Dittmar SS, Hicks JT, Joyce SZ, Stehlik MA. The Quadriplegia Index of Function (QIF): sensitivity and reliability demonstrated in a study of thirty quadriplegic patients. *Paraplegia.* 1986;24:38–44.
66. Marino RJ, Goin JE. Development of a short-form Quadriplegia Index of Function scale. *Spinal Cord.* 1999;37:289–296.
67. Marciello MA, Herbison GJ, Ditunno JF Jr, Marino RJ, Cohen ME. Wrist strength measured by myometry as an indicator of functional independence. *J Neurotrauma.* 1995;12:99–106.

68. Marino RJ, Rider-Foster D, Maissel G, Ditunno JF. Superiority of motor level over single neurological level in categorizing tetraplegia. *Paraplegia*. 1995;33:510–513.
69. Popovic D, Stojanovic A, Pjanovic A, et al. Clinical evaluation of the bionic glove. *Arch Phys Med Rehabil*. 1999;80:299–304.
70. Mulcahey MJ, Betz RR, Kozin SH, Smith BT, Hutchinson D, Lutz C. Implantation of the Freehand System during initial rehabilitation using minimally invasive techniques. *Spinal Cord*. 2004;42:146–155.
71. Rothwell AG, Sinnott KA, Mohammed KD, Dunn JA, Sinclair SW. Upper limb surgery for tetraplegia: a 10-year re-review of hand function. *J Hand Surg (Am)*. 2003;28:489–497.
72. Catz A, Itzkovich M, Tesio L, et al. A multi-center International Study on the Spinal Cord Independence Measure, Version III: Rasch psychometric validation. *Spinal Cord*. 2007;45:275–291.
73. Catz A, Itzkovich M, Steinberg F, et al. The Catz-Itzkovich SCIM: a revised version of the Spinal Cord Independence Measure. *Disabil Rehabil*. 2001;23:263–268.
74. Itzkovich M, Tripolski M, Zeilig G, et al. Rasch analysis of the Catz-Itzkovich spinal cord independence measure. *Spinal Cord*. 2002;40:396–407.
75. Itzkovich M, Gelernter I, Biering-Sorensen F, et al. The Spinal Cord Independence Measure (SCIM) Version III: reliability and validity in a multi-center international study. *Disabil Rehabil*. 2007;29:1926–1933.
76. Itzkovich M, Tamir A, Philo O, et al. Reliability of the Catz-Itzkovich Spinal Cord Independence Measure assessment by interview and comparison with observation. *Am J Phys Med Rehabil*. 2003;82:267–272.
77. Grijalva I, Guizar-Sahagun G, Castaneda-Hernandez G, et al. Efficacy and safety of 4-aminopyridine in patients with long-term spinal cord injury: a randomized, double-blind, placebo-controlled trial. *Pharmacotherapy*. 2003;23:823–834.
78. Popovic MR, Thrasher TA, Adams ME, Takes V, Zivanovic V, Tonack MI. Functional electrical therapy: retraining grasping in spinal cord injury. *Spinal Cord*. 2006;44:143–151.
79. Catz A, Goldin D, Fishel B, Ronen J, Bluvshstein V, Gelernter I. Recovery of neurologic function following nontraumatic spinal cord lesions in Israel. *Spine*. 2004;29:2278–2282.
80. Ronen J, Itzkovich M, Bluvshstein V, et al. Length of stay in hospital following spinal cord lesions in Israel. *Spinal Cord*. 2004;42:353–358.
81. Berry C, Kennedy P. A psychometric analysis of the Needs Assessment Checklist (NAC). *Spinal Cord*. 2003;41:490–501.
82. Catz A, Greenberg E, Itzkovich M, Bluvshstein V, Ronen J, Gelernter I. A new instrument for outcome assessment in rehabilitation medicine: spinal cord injury ability realization measurement index. *Arch Phys Med Rehabil*. 2004;85:399–404.
83. McKinley WO, Seel RT, Gadi RK, Tewksbury MA. Non-traumatic vs. traumatic spinal cord injury: a rehabilitation outcome comparison. *Am J Phys Med Rehabil*. 2001;80:693–699.
84. Scivoletto G, Morganti B, Ditunno P, Ditunno JF, Molinari M. Effects on age on spinal cord lesion patients' rehabilitation. *Spinal Cord*. 2003;41:457–464.
85. Ditunno JF Jr, Ditunno PL, Graziani V, et al. Walking index for spinal cord injury (WISCI): an international multicenter validity and reliability study. *Spinal Cord*. 2000;38:234–243.