

# Assessment of Disability in Patients with Acute Traumatic Spinal Cord Injury: A Systematic Review of the Literature

Julio C. Furlan,<sup>1,3,4</sup> Vanessa Noonan,<sup>2</sup> Anoushka Singh,<sup>3</sup> and Michael G. Fehlings<sup>1,3,4</sup>

## Abstract

Given the importance of accurately and reliably assessing disability in future clinical trials, which will test therapeutic strategies in acute spinal cord injury (SCI), we sought to appraise comprehensively studies that focused on the psychometric properties (i.e., reliability, validity, and responsiveness) of all previously used outcome measures in the SCI population. The search strategy included Medline, CINAHL, EMBASE, and Cochrane databases. Two reviewers independently assessed each study regarding eligibility, level of evidence (using Sackett's criteria), and quality. Of 363 abstracts captured in our search, 36 full articles fulfilled the inclusion and exclusion criteria. Eight different outcome measures were used to assess disability in the SCI population, including Functional Independence Measure (FIM), Spinal cord Injury Measure (SCIM), Walking Index for Spinal Cord Injury (WISCI), Quadriplegia Index of Function (QIF), Modified Barthel Index (MBI), Timed Up & Go (TUG), 6-min walk test (6MWT), and 10-m walk test (10MWT). While 19 of 36 studies provided level-4 evidence, the remaining 17 articles were classified as level-2b evidence. Most of the instruments showed convergent construct validity in the SCI population, but criterion validity was not examined due to the lack of a gold standard for assessment of disability. All instruments were tested in the rehabilitation and/or community setting, but only FIM was examined in the acute care setting. Based on our results of quality assessment, the SCIM has the most appropriate performance regarding the instrument's psychometric properties. Nonetheless, further investigations are required to confirm the adequate performance of the SCIM as a comprehensive measure of functional recovery in patients with SCI in rehabilitative care. The expert panel of the Spinal Cord Injury Solutions Network (SCISN) that participated in the modified Delphi process endorsed these conclusions.

**Key words:** clinical research; disability; spinal cord injury; systematic review

## Introduction

**S**PINAL CORD INJURY (SCI) HAS A SIGNIFICANT worldwide health and social impact with an incidence of between 10.4 and 59 spinal cord injured individuals per million inhabitants per year (Wyndaele and Wyndaele, 2006). Currently, only limited treatments to enhance spinal cord function are available in the clinical setting (Hawryluk et al., 2008). However, there are a number of novel emerging neuroprotective and neuroregenerative strategies with the potential to reduce neuronal death after CNS injury, enhance the intrinsic growth capacity of postmitotic neurons, or modify of the CNS extracellular milieu that is hostile to neuronal growth (Baptiste and Fehlings, 2008; Jacobs and Fehlings, 2003). Given that those strategies are in the early phases of clinical translation, reliable

and responsive adequate outcome measures for assessment of the degrees of impairment and disability are required.

While measures of impairment are commonly used as the primary outcome measure in clinical trials of SCI, assessment of disability after SCI is an important secondary endpoint. As defined by the International Classification of Functioning, Disability, and Health from the World Health Organization (WHO), disability is related to the level of "activity" (WHO, 2001). Ideally, the instrument of choice for assessment of disability in the clinical practice and research areas of SCI should be appropriate for descriptive and evaluative purposes in accordance with the framework of Kirschner and Guyatt (Kirschner, 1985).

Given the significance of assessment of disability in anticipated future clinical trials, which will test therapeutic

<sup>1</sup>Department of Genetics and Development, Toronto Western Research Institute, University Health Network, Toronto, Ontario, Canada.

<sup>2</sup>University of British Columbia, Vancouver, British Columbia, Canada.

<sup>3</sup>Krembil Neuroscience Centre, Spinal Program, Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada.

<sup>4</sup>Department of Surgery, Division of Neurosurgery and Spinal Program University of Toronto, Toronto, Ontario, Canada.

strategies in patients with acute SCI, we sought to review comprehensively the studies that focused on the psychometric properties (i.e., reliability, validity, and responsiveness) of all previously used outcome measures in the SCI population.

## Methods

This systematic review included all outcome measures of disability after traumatic SCI that were published at least twice in the literature. Based on the examination of their psychometric properties including reliability, validity, and responsiveness, we sought to answer the following key question: What is the most reliable, validated, and responsive outcome measure of disability for patients with acute traumatic SCI?

### *Inclusion and exclusion criteria*

For this purpose, we selected all original articles that examined at least one the psychometric properties of an outcome measure of disability in the setting of traumatic SCI. We included only those outcome measures of disability of which psychometric properties were examined in at least two publications. We excluded case reports, editorial articles, and meeting abstracts.

### *Literature search strategy*

The primary literature search was performed using Medline, CINAHL, EMBASE, and Cochrane databases. A secondary search strategy included articles referred in meta-analysis, and systematic and non-systematic review articles that were captured in the primary search strategy.

The literature searches addressed publications from 1966 to April 2008. The search strategy included the following key words: "disability," "activity," "activity of daily living," "functional outcome," and "functional recovery." Those specific key words were paired with the following Medical Subject Headings (MeSHs): "spinal cord injury," "SCI," "tetraplegia," "quadriplegia," and "paraplegia." The literature search was limited to papers written in English only. Subsequently, the terms of each outcome measure captured using the above search tactic were paired again with the generic MeSHs including "spinal cord injury," "SCI," "tetraplegia," "quadriplegia," and "paraplegia."

### *Data abstraction and synthesis*

For the culling process, two reviewers (JCF and VN) independently selected the articles that fulfilled the inclusion and exclusion for each topic. Disagreements were solved by a debate and consensus between both reviewers.

A research assistant extracted the relevant data from each selected article. Subsequently, both reviewers examined all clinical studies with regard to the extracted data and, hereafter, determined the level of evidence according to Sackett and associates (2000). Using the quality criteria by Terwee and associates, every instrument of assessment of disability was examined with regard to its psychometric properties (Higginson, 2007; Terwee et al., 2007). Divergences during those steps were solved by consensus between both reviewers. The main results from each article and the reviewers' assessments were included in summary tables.

### *Definitions of the psychometric properties*

For the purpose of this systematic review, the psychometric properties were classified according to Terwee and associates (2007) (Table 1). *Content validity* refers to the extent to which the items in the instrument comprehensively represent the concepts of interest (Guyatt et al., 1993). *Internal consistency* refers to the extent to which items in the instrument (sub)scale are homogeneously correlated and, hence, measure the same concept (Terwee et al., 2007). *Criterion validity* refers to the degree to which the instrument measures in comparison with the criterion or "gold standard" (Furlan et al., 2008). Given that there is no well-established gold standard for assessing disability in the SCI population, criterion validity was not assessed. *Construct validity* is commonly divided into convergent or divergent. While *convergent construct validity* indicates the degree of similarity between two constructs that theoretically should be related to each other, *divergent construct validity* reveals how dissimilar two constructs are that in theory should not be related to each other (Furlan et al., 2008). *Reproducibility* refers to the degree to which repeated measurements in steady patients provide similar results (Terwee et al., 2007). Reproducibility is generally divided into agreement and reliability. While *agreement* reflects the absolute measurement error, *reliability* refers to the degree to which patients can be distinguished from each other, regardless of measurement error (Terwee et al., 2007). *Responsiveness* concerns the ability of a measurement instrument to detect change accurately when it has occurred (de Bruin et al., 1992). *Floor or ceiling effects* occur when more than 15% of examined patients reach the lowest or highest possible score respectively (McHorney and Tarlov, 1995). Finally, *interpretability* concerns the degree to which one can assign qualitative meaning to quantitative scores (Lohr et al., 1996).

### *Establishment of recommendations*

Based on the review information summarized in the tables, the authors answered the specific question formerly elaborated. Using a modified Delphi method, an expert panel comprised of scientists and clinicians in the Spinal Cord Injury Solutions Network (SCISN) examined the summary tables and answer to the focused question, and eventually determined the evidence-based recommendations (Reid, 1993).

## Results

### *Literature search*

Of 363 abstracts captured in our search, 36 full articles fulfilled the inclusion and exclusion criteria and were reviewed by the two reviewers. There were eight different outcome measures that were used to assess disability in the SCI population as follows:

- i. Functional Independence Measure (FIM) (Beninato et al., 2004; Davidoff et al., 1990; Dijkers and Yavuzer, 1999; Grey and Kennedy, 1993; Hall et al., 1999; Kucukdeveci et al., 2001; Lawton et al., 2006; Lundgren-Nilsson et al., 2006; Marino et al., 1993; Masedo et al., 2005; Nilsson et al., 2005; Roth et al., 1990; Segal et al., 1993; Yavuz et al., 1998);
- ii. Spinal Cord Injury Measure (SCIM) (Catz et al., 1997, 2002, 2007; Catz, Itzkovich, Agranov et al., 2001; Catz,

TABLE 1. QUALITY CRITERIA FOR PSYCHOMETRIC PROPERTIES OF INSTRUMENTS ACCORDING TO TERWEE AND ASSOCIATES (2007)

<i>Psychometric property</i>	<i>Quality criteria<sup>a,b</sup></i>
1. Content validity	+A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection ? A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method –No target population involvement
2. Internal consistency	0 No information found on target population involvement +Factor analyses performed on adequate sample size (7 * # items and $\geq 100$ ) AND Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95 ? No factor analysis OR doubtful design or method –Cronbach's alpha(s) $< 0.70$ or $> 0.95$ , despite adequate design and method
3. Criterion validity	0 No information found on internal consistency +Convincing arguments that gold standard is "gold" AND correlation with gold standard $\geq 0.70$ ? No convincing arguments that gold standard is "gold" OR doubtful design or method –Correlation with gold standard $< 0.70$ , despite adequate design and method
4. Construct validity	0 No information found on criterion validity +Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses ? Doubtful design or method (e.g., no hypotheses) –Less than 75% of hypotheses were confirmed, despite adequate design and methods
5. Reproducibility	0 No information found on construct validity +MIC $< SDC$ OR MIC outside the LOA OR convincing arguments that agreement is acceptable
5.1. Agreement	? Doubtful design or method OR (MIC not defined AND no convincing arguments that agreement is acceptable) –MIC $\geq SDC$ OR MIC equals or inside LOA, despite adequate design and method
5.2. Reliability	0 No information found on agreement +ICC or weighted Kappa $\geq 0.70$ ? Doubtful design or method (e.g., time interval not mentioned) –ICC or weighted Kappa $< 0.70$ , despite adequate design and method
6. Responsiveness	0 No information found on reliability +SDC or SDC $< MIC$ OR MIC outside the LOA OR RRO1.96 OR AUC $\geq 0.70$ ? Doubtful design or method –SDC or SDC $\geq MIC$ OR MIC equals or inside LOA OR RR $\leq 1.96$ OR AUC $< 0.70$ , despite adequate design and methods
7. Floor and ceiling effects	0 No information found on responsiveness + $\leq 15\%$ of the respondents achieved the highest or lowest possible scores ? Doubtful design or method $\geq 15\%$ of the respondents achieved the highest or lowest possible scores, despite adequate design and methods
8. Interpretability	0 No information found on interpretation +Mean and SD scores presented of at least four relevant subgroups of patients and MIC defined ? Doubtful design or method OR less than four subgroups OR no MIC defined 0 No information found on interpretation

MIC, minimal important change; SDC, smallest detectable change; LOA, limits of agreement; ICC, Intraclass correlation; SD, standard deviation.

<sup>a</sup>+ = positive rating; ? = indeterminate rating; – = negative rating; 0 = no information available.

<sup>b</sup>Doubtful design or method: lacking of a clear description of the design or methods of the study, sample size smaller than 50 subjects (should be at least 50 in every [subgroup] analysis), or any important methodological weakness in the design or execution of the study.

- Itzkovich, Steinburg et al., 2001; Itzkovich et al., 2002, 2003, 2007);
- iii. Walking Index for Spinal Cord Injury (WISCI) (Dittuno and Dittuno, 2001; Dittuno et al., 2000, 2007, 2008; Kim et al., 2007; Morganti et al., 2005; van Hedel et al., 2005, 2006);
  - iv. Quadriplegia Index of Function (QIF) (Gresham et al., 1986; Marino and Goin, 1999; Marino et al., 1993, 1995; Yavuz et al., 1998);
  - v. Modified Barthel Index (MBI) (Roth et al., 1990; Kucukdeveci et al., 2000);
  - vi. Timed Up & Go (TUG) (van Hedel et al., 2005, 2008);
  - vii. 6-min walk test (6MWT) (Olmos et al., 2008; van Hedel et al., 2005, 2006); and
  - viii. 10-m walk test (10MWT) (Olmos et al., 2008; van Hedel et al., 2005, 2006, 2008).

Our search also captured the Spinal Cord Injury Functional Ambulation Inventory (SCI FAI) as another instrument for the assessment of disability in the SCI population (Field-Fote et al., 2001). However, there was only the original publication of this outcome measure in the literature and, hence, this instrument was excluded from our systematic review.

While 19 of 36 studies provided level-4 evidence, the remaining 17 articles were classified as level-2b evidence (Table 2). Convergent construct validity ( $n=23$ ) and reproducibility ( $n=14$ ) were the most commonly studied psychometric properties. In addition, other psychometric properties were also examined, including content validity ( $n=6$ ), internal consistency ( $n=3$ ), item generation/reduction ( $n=2$ ), and responsiveness ( $n=2$ ).

Using the criteria of Terwee and associates (2007), each instrument was assessed with regard to its quality based on the literature (Table 3). Generally speaking, most of the instruments showed convergent construct validity in the SCI population, but criterion validity was not examined due to the lack a gold standard for the assessment of disability. Only QIF and WISCI had the content validity examined. The most reliable instruments included FIM and SCIM, whereas many other instruments were not tested for reproducibility. While appropriate responsiveness was reported for SCIM, 6MWT, and 10 WMT, it was uncertain for QIF and WISCI. Negative ceiling/floor effects were documented for WISCI and FIM, whereas TUG, 6MWT, and 10WMT showed adequate results in the evaluations of floor/ceiling effects. All instruments were tested in the rehabilitation and/or community setting, but they were not examined in the acute care setting, except for FIM. Of note, inadequate construct validity and negative floor/ceiling effects were observed in that study, where FIM was tested in the acute care setting (Davidoff et al., 1990).

## Discussion

Our systematic review identified 36 clinical studies that examined psychometric properties of eight instruments of disability assessment. Those included FIM, SCIM, WISCI, QIF, MBI, TUG, 6MWT, and 10WMT, which were reported in original articles of level-4 or level-2b evidence. Although criterion validity was not examined due to the lack of gold standard, several other psychometric properties were studied, including item generation/reduction, reproducibility, internal consistency, convergent construct validity, content valid-

ity, and responsiveness. While all instruments were tested in the rehabilitation and/or community setting, only FIM was examined in the acute care setting, where inadequate construct validity and floor/ceiling effects were found.

### Modified Barthel Index (MBI)

The Barthel Index, which was published in 1965, was originally developed for use in rehabilitation patients with stroke and other neuromuscular or musculoskeletal disorders (Mahoney and Barthel, 1965). In 1979, following a number of revisions to the Barthel Index, the MBI was derived. The MBI is 10-item ordinal scale (range: 0 to 100) with ratings for feeding, moving from wheelchair to bed and return, grooming, transferring to and from a toilet, bathing, walking on a level surface, going up and down stairs, dressing, and continence of the bowels and bladder (Granger et al., 1979).

In our systematic review, there were only two prior studies that examined psychometric properties of the MBI in patients with SCI. A previous study evaluated the agreement among raters, but its quality varied from adequate to inadequate. Internal consistency was adequate in one prior study that used a Turkish version of the MBI. While one previous study indicated an adequate construct validity of the MBI, another study showed inconsistencies in the MBI with regard to its construct validity.

### Quadriplegia Index of Function (QIF)

The QIF was originally developed to overcome the limitations of the Barthel Index in the assessment of disability of patients with tetraplegia (Gresham et al., 1986). The QIF is a 10-item ordinal scale (range: 0 to 100) that includes assessments of transfers, grooming, bathing, feeding, dressing, wheelchair mobility, bed activities, bladder management, bowel management, and understanding of personal care (Gresham et al., 1986).

Our systematic review identified five original articles where psychometric properties of the QIF were studied. While internal consistency was assessed as adequate in one prior study, reliability was examined in another study that had an indeterminate rating of quality due to a lack of information or poor study design/method. Convergent construct validity was consistently assessed as adequate in all five prior studies. However, a prior study on content validity and another study on responsiveness of QIF had an indeterminate rating of quality due to a lack of information or poor study design/method.

### Functional Independence Measure (FIM)

Although FIM was developed for assessment of disability in patients with stroke and to assess the requirements for burden of care, this instrument has been widely used in the assessment of disability in spinal cord injured patients (Kirshblum et al., 2004). The FIM is an 18-item ordinal scale (range: 0 to 126) with seven levels per item (from complete independence to total assist) that includes assessment of disability in the areas of self-care, sphincter control, mobility, locomotion, communication, psychosocial adjustment, and cognitive function (Keith et al., 1987). The physical FIM subscore refers to the summed subscores for self-care, sphincter control, mobility, and locomotion items, whereas cognitive

TABLE 2. SUMMARY OF THE ARTICLES WITH LEVEL OF EVIDENCE (LOE) INCLUDED IN OUR SYSTEMATIC REVIEW

Reference	Psychometric properties	Sample size	Inclusion and exclusion criteria	Study population	Injury features	Instrument	LOE
Davidoff et al. (1990)	Convergent construct validity	N (total) = 63 n (case) = 41 n (control) = 22	Patients admitted to The university medical center for acute care and rehabilitation within 45 days of injury, aged 18 to 55 years were included. A group of paid neurologically intact volunteers was included as a control group	Age range (total) = 18–55 years Mean age: • Case = 30.4 years • Control = 27.7 years Percentage of males: • Case = 80% • Control = 55%	Level of SCI: • Cervical = 41% • Thoracic = 39% • Lumbosacral = 20% Cause of SCI: • MVA = 70% • Gunshot = 10% • Assault = 10% • Fall = 5% • Sports = 2.5% • Others = 2.5% Associated injuries (TBI) = 57% Severity of SCI: • Complete = 18 • Incomplete = 11 Causes of SCI: • MVA = 51.7% • Fall = 13.8% Levels of SCI: • Cervical = 28 • Thoracic = 1 Mean time since SCI = 20 weeks (Range: 2–72)	FIM	2b
Yavuz et al. (1998)	Convergent construct validity	N = 29	Consecutive tetraplegic patients admitted between May 1994 and January 1996	Mean age = 37 years Age range = 14–66 yrs Males/females = 20/9	Severity of SCI: • Complete = 18 • Incomplete = 11 Causes of SCI: • MVA = 51.7% • Fall = 13.8% Levels of SCI: • Cervical = 28 • Thoracic = 1 Mean time since SCI = 20 weeks (Range: 2–72)	• FIM • QIF	2b
Dijkers and Yavuzer (1999)	Inter-rater reliability	N = 4128	Patients in the National Spinal Cord Injury Database were included. Patients with missing Form I discharge neurological category were excluded.	Mean age = 37.5 years Median age = 36 years Age range = 6–99 years Percentage of males = 80.5%	Level of injury: • Complete tetraplegia = 23.7% • Incomplete tetraplegia = 28.1% • Complete paraplegia = 31.1% • Incomplete paraplegia = 17.0% Mean time since SCI = 6 years (Range: 1–25 years).	FIM	2b
Marino and Goin (1999)	• Item reduction • Convergent construct Validity	N = 95	Patients admitted from December 1987 to August 1992 with tetraplegia on initial evaluation, who were non-ambulatory at 6 months post-injury were included. Patients with concomitant upper extremity fractures, TBI, or psychiatric disturbances were excluded.	Mean age = 31.2 years Age range = 16–68 years Males/females = 85/10	Severity of SCI: • Frankel A = 60 • Frankel B = 23 • Frankel C = 11 • Frankel D = 1 (none was community ambulatory)	• QIF • Short-form QIF	2b

(continued)

TABLE 2. (CONTINUED)

Reference	Psychometric properties	Sample size	Inclusion and exclusion criteria	Study population	Injury features	Instrument	LOE
Hall et al. (1999)	Responsiveness	n (admission) = 3971 n (discharge) = 4033 n (year 1) = 903 n (year 2) = 714 n (year 5) = 570	Patients admitted to inpatient rehabilitation within 60 days after injury were included. Only SCI patients aged 16 years or older with functionally complete SCI at inpatient rehabilitation admission were included.	Not reported	Severity of SCI: AIS A = 2447 AIS B = 679 AIS C = 845 Level of SCI: Cervical = 1796 Thoracic = 1718 Lumbar/sacral = 457 Mean time since injury = 8 days Median time since SCI = 1 day Maximum time since SCI = 60 days Level of SCI: • Tetraplegia = 8 • Paraplegia = 14 Severity of SCI: • Frankel A/B = 6 • Frankel C/D = 16 Etiology of SCI: • Traumatic = 11 • Non-traumatic injury = 11 Mean follow-up time = 3 months (Range: 1–6 months)	FIM	2b
Catz, Itzkovich, Agranov et al. (2001)	Convergent construct validity	N = 22	Only those patients with complete data were included (that is SCIM and FIM scores obtained in the first week after admission and every month until discharge).	Mean age = 48.1 yrs Age range = 23–76 yrs Males/females = 17/5	Level of SCI: • Tetraplegia = 8 • Paraplegia = 14 Severity of SCI: • Frankel A/B = 6 • Frankel C/D = 16 Etiology of SCI: • Traumatic = 11 • Non-traumatic injury = 11 Mean follow-up time = 3 months (Range: 1–6 months)	SCIM I	2b
Kucukdeveci et al. (2001)	• Inter-rater reliability • Convergent construct validity • Content validity	N (SCI) = 62	Patients with stroke or SCI admitted to Department of Physical Medicine and Rehabilitation from 1994 to 1998 were included. (Only data of SCI patients are analyzed herein.)	Mean age (SCI) = 32.7 years Males = 44%	Level of SCI: • Cervical = 21% • Thoracic = 42% • Lumbar = 37% Mean time since SCI = 16.4 years (Range: 1–210 months)	FIM (Turkish version)	2b
Masedo et al. (2005)	Intra-rater reliability	N (SCI) = 84	Patients with amputation or SCI were included. Only participants included were those with: either an amputation or SCI > 6 months before enrollment; pain > 3 months; average pain rating in the last month > 2 for amputation and > 3 for SCI. Patients younger than 18 years, older than 65 years, with cardiovascular disease or abnormalities in a screening ECG, with seizures, pregnant, received any type of antidepressant medication or consumed > 2 alcoholic drinks per day were excluded.	Mean age = 41.43 years Males = 80%	Level of SCI: • Cervical = 53.6% • Thoracic = 38.1% • Lumbar/sacral = 7.1% Mean time since injury = 13.96 years (Range: 8 months to 43 years)	Self-report FIM	2b

Morganti et al. (2005)	Convergent construct validity	N = 284	Subjects included patients with a recent onset of traumatic and non-traumatic SCI admitted between 1997 and 2001. Patients without the cognitive ability to participate in the rehabilitation program were excluded.	Mean age = 50.4 years Age range = 12-86 years Males/females = 184/100	Lesion of SCI: • Cervical = 81 • Thoracic = 148 • Lumbosacral = 55 Severity of SCI: • AIS A = 84 • AIS B = 19 • AIS C = 129 • AIS D = 52 Etiology of SCI: • Traumatic = 107 • Non-traumatic injury = 177 Mean time since SCI = 57 days Not reported	WISCI II	2b
Nilsson et al. (2005)	Content validity	N = 1661 n (SCI) = 358	Patients admitted to one of the Scandinavian rehabilitation units with stroke, SCI, and TBI were included.	Median age (total) = 48 years Age range (total) = 16-90 years Males (total) = 64%		FIM	2b
van Hedel et al. (2005)	• Convergent construct validity (WISCI II, TUG, 6MWT, 10MWT) • Intra and Inter-rater reliability (TUG, 6MWT, 10MWT)	n (validity) = 75 n (reliability) = 22	Patients with traumatic or ischemic SCI were included. Only patients admitted for rehabilitation or ambulant check-up and were seen by physical therapist in the hospital from May 2001 to August 2003. Patients with additional gait impairment deficits and those who scored zero of WISCI II were excluded.	• Validity: Mean age = 54 years Age range = 17-84 years Males/females = 45/30 • Reliability: Mean age = 52 yrs Age range = 21-77 yrs Males/Females = 14/8	Validity Level of SCI: • Cervical = 25 • Thoracic = 21 • Lumbosacral = 29 Severity of SCI: • AIS A/B = 7 • AIS C = 7 • AIS D = 61 Reliability Level of SCI: • Cervical = 7 • Thoracic = 7 • Lumbosacral = 8 Severity of SCI: • AIS A = 1 • AIS C = 3 • AIS D = 18 Not reported	• WISCI II • TUG • 6MWT • 10MWT	2b
Lawton et al. (2006)	Content validity	N = 647	Patients recently admitted with a diagnosis of SCI in one of the 19 rehabilitation facilities were included.	Median age = 46 years Age range = 11-93 years Males = 69%		FIM	2b
Lundgren- Nilsson et al. (2006)	Content validity	N = 471 n (SCI) = 157	Patients admitted to one of the nine Scandinavian rehabilitation units with SCI.	Age range = 11-90 years Males = 70%	Not reported	FIM	2b
van Hedel et al. (2006)	• Convergent construct validity (WISCI II, 6MWT, 10MWT) • Responsiveness (6MWT, 10MWT)	N = 22	Only incomplete patients who were able to stand or walk within the first month after SCI were selected for this study (WISCI II score $\geq$ 1). All patients should have performed the 6MWT and 10MWT at all four times.	Mean age = 45.5 years Age range = 17-78 years Males/females = 18/4	Level of SCI: • Cervical = 13 • Thoracic = 1 • Lumbosacral = 8 Etiology of SCI: • Traumatic = 19 • Non-traumatic = 3	• WISCI II • 6MWT • 10MWT	2b

(continued)

TABLE 2. (CONTINUED)

Reference	Psychometric properties	Sample size	Inclusion and exclusion criteria	Study population	Injury features	Instrument	LOE
Catz et al. (2007)	Internal consistency	N = 425	All patients 18 years or older admitted to the 13 spinal units with spinal cord lesions were included. Patients with concomitant impairments such as brain injury or mental disease, or any other condition that might influence their everyday function were excluded.	Mean age = 46.93 years Males/females = 309/116	Level of SCI: • Tetraplegia = 188 • Paraplegia = 237 Severity of SCI: • AIS A = 151 • AIS B = 59 • AIS C = 92 • AIS D = 119 Etiology of SCI: • Traumatic = 261 • Non-traumatic = 164	SCIM III	2b
Ditunno et al. (2007)	Convergent construct validity	N = 146	Patients with incomplete upper motor neuron (UMN) lesion (C4-T11) or lower motor neuron lesion (T11 to L3 with no UMN signs) were included. Patients with complete SCI or time from injury > 8 weeks were excluded.	Mean age = 32 years Age range = 16-69 years Males = 78%	Level of SCI: • Cervical = 58% • Thoracic = 18% • Lumbar = 24% Severity of SCI: • AIS B = 36 • AIS C = 90 • AIS D = 20	WISCI II	2b
Ditunno et al. (2008)	Convergent construct validity	N = 150	Patients with acute traumatic SCI at C2 to L3 admitted within 1 month of SCI and followed in 6-12 months were included. Patients with significant TBI, upper/lower extremity fracture, an immobilization device or those with WISCI 20 on initial evaluation were excluded.	SCI	Level of SCI: • Paraplegia = 66 • Tetraplegia = 84 Severity of SCI: • AIS A = 59 • AIS B = 19 • AIS C = 32 • D = 40 Range of time since SCI = 3-12 months All complete SCI at C4 to C8.	WISCI II	2b
Gresham et al. (1986)	• Convergent construct validity • Inter-rater reliability	n = 30 (validity) n = 20 (reliability)	Patients with complete tetraplegia (at admission to and discharge from inpatient medical rehabilitation) were included.	Not reported		QIF	4
Roth et al. (1990)	Convergent construct validity	N = 41	Patients admitted within 45 days of traumatic SCI to a regional model system SCI care center for acute management and comprehensive rehabilitation were included.	Mean age = 30.4 years Males = 85%	Level of SCI: • Cervical = 17 • Thoracic = 16 • Lumbosacral = 8 Cause of SCI: • MVA = 29 • Gunshot = 4 • Assault = 4 • Fall = 2 • Other = 2	• "Adapted" FIM (Only the self-care and mobility FIM subscales) • MBI	4
Grey and Kennedy (1993)	Inter-rater reliability	N = 40	Only patients aged 16-65 years with traumatic SCI on admission were included.	Mean age = 29.6 years Age range = 17-54 years Males = 80%	Cause of SCI: • MVA = 45% • Falls = 25% • Sports = 22.5% • Other = 7.5% Level of SCI: • Tetraplegia = 32.5% • Paraplegia = 67.5%	FIM/self-report FIM	4



Marino et al. (1993)	Convergent construct validity	N = 22	Only patients with traumatic quadriplegia admitted within 72h of SCI from November 1989 to March 1991 were included. Patients were excluded if medically unstable, acute medical conditions, amputation, stroke, TBI, upper extremity fractures, knife or gunshot wounds at the time of evaluation, or variability in functional performance.	Males = 100% Mean age = 33 years Age range = 18-63 years	Severity of SCI • Frankel A = 4 • Frankel B = 11 • Frankel C = 2 • Frankel D = 5 Level of SCI: • C4 = 6 • C5 = 7 • C6 = 4 • C7 = 5 Time since SCI: • 3 months = 7 • 6 months = 10 • 12 months = 5 Not reported	• FIM • Partial QJF (only grooming, bathing and feeding)	4
Segal et al. (1993)	Inter-rater reliability	N = 57	Patients admitted from January 1989 to May 1990, transferred without interruption of care directly from the acute care rehabilitation setting to the ongoing rehabilitation setting were included.	Not reported		FIM	4
Marino et al. (1995)	Convergent construct validity	N = 50	Patients with motor complete traumatic cervical SCI at a neurological level of C4-8 and Frankel grades of A or B on admission (from 1988 to 1992) were included. Individuals who had upper extremity fractures or were without data at 12 months post injury were excluded. Only patients with documented neurological examination and QJF scores.	Age range = 16-68 years Males/females = 47/3	Severity of SCI: • Frankel A = 36 • Frankel B = 14 Etiology of SCI: • Diving = 16 • MVA = 11 • Fall = 6 • Gunshot = 7 • Hit by a car = 2 • Farm = 2 • Fight = 3 • Sports = 2 Level of SCI: • C4 = 29 • C5 = 12 • C6 = 6 • C7 = 2 • C8 = 1	Feeding QJF	4
Catz et al. (1997)	• Inter-rater reliability • Convergent construct validity	N = 30	Patients admitted from December 1994 to October 1995 were included.	Mean age = 45 years Age range = 17-76 years Males/females = 22/8	Level of SCI: • Tetraplegia = 9 • Paraplegia = 21 Severity of SCI: • Frankel A/B = 9 • Frankel C/D = 21 Etiology of SCI: • Traumatic = 15 • Non-traumatic injury = 14 Mean follow-up time = 3 months (Range: 1-6 months)	SCIM I	4

(continued)

TABLE 2. (CONTINUED)

Reference	Psychometric properties	Sample size	Inclusion and exclusion criteria	Study population	Injury features	Instrument	LOE
Ditunno et al. (2000)	<ul style="list-style-type: none"> <li>• Inter-rater reliability</li> <li>• Item generation and reduction</li> <li>• Content validity</li> <li>• Convergent construct validity</li> </ul>	N = 40	"Representative of patients walking"	SCI	Not reported	WISCI I	4
Kucukdeveci et al. (2000)	<ul style="list-style-type: none"> <li>• Internal consistency</li> <li>• Inter-rater reliability</li> <li>• Convergent construct validity</li> </ul>	N = 100; n (stroke) = 50 n (SCI) = 50	Patients with stroke or SCI admitted from 1993 to 1997 were included.	Mean age (SCI) = 31.5 years Males = 44% Mean follow-up time = 3.6 months (Range: 1–24 months)	<ul style="list-style-type: none"> <li>• Level of SCI:</li> <li>• Cervical = 22%</li> <li>• Thoracic = 46%</li> <li>• Lumbar = 32%</li> </ul>	Modified Barthel Index	4
Catz, Itzkovich, Steinburg et al. (2001)	<ul style="list-style-type: none"> <li>• Inter-rater reliability</li> <li>• Convergent construct validity</li> </ul>	N = 28	Patients with traumatic and non-traumatic SCI were included. Patients with concurrent medical or cognitive disabilities likely to impair daily function were excluded.	Mean age = 46 years Age range = 20–79 years Males/females = 18/10	<ul style="list-style-type: none"> <li>• Level of SCI:</li> <li>• Tetraplegia = 6</li> <li>• Paraplegia = 22</li> </ul> Severity of SCI: <ul style="list-style-type: none"> <li>• Frankel A/B = 7</li> <li>• Frankel C/D = 21</li> <li>• Traumatic = 11</li> <li>• Non-traumatic injury = 17</li> </ul> Range of follow-up time = 2–7 months Severity of SCI: <ul style="list-style-type: none"> <li>• AIS A = 14</li> <li>• AIS B = 18</li> <li>• AIS C = 52</li> <li>• AIS D = 19</li> </ul>	SCIM II	4
Ditunno and Dittuno (2001)	Convergent construct validity	N = 103	Not stated	SCI	Severity of SCI: <ul style="list-style-type: none"> <li>• AIS A = 14</li> <li>• AIS B = 18</li> <li>• AIS C = 52</li> <li>• AIS D = 19</li> </ul>	WISCI II	4
Catz et al. (2002)	Inter-rater reliability	N = 28	Patients with traumatic and non-traumatic SCI were included. Patients with concomitant medical problems that might have influenced everyday function, such as malignancy, brain injury or mental disease were excluded.	Mean age = 46 years Age range = 20–79 years Males/females = 18/10	Level of SCI: <ul style="list-style-type: none"> <li>• Tetraplegia = 6</li> <li>• Paraplegia = 22</li> </ul> Severity of SCI: <ul style="list-style-type: none"> <li>• Frankel A/B = 7</li> <li>• Frankel C/D = 21</li> <li>• Traumatic = 11</li> <li>• Non-traumatic injury = 17</li> </ul>	SCIM II	4
Itzkovich et al. (2002)	Convergent construct validity	N = 202	Patients with spinal cord lesion.	Mean age = 46.78 years Age range = 14–94 years Males/females = 138/64	Level of SCI: <ul style="list-style-type: none"> <li>• Tetraplegia = 78</li> <li>• Paraplegia = 124</li> </ul> Severity of SCI: <ul style="list-style-type: none"> <li>• AIS A/B = 65</li> <li>• AIS C/D = 137</li> </ul> Etiology of SCI: <ul style="list-style-type: none"> <li>• Traumatic = 97</li> <li>• Non-traumatic injury = 105</li> </ul>	SCIM II	4

Itzkovich et al. (2003)	Inter-rater reliability	N=28	<p>Patients with spinal cord lesions were included.</p> <p>Patients with concomitant medical problems such as malignancy, brain injury, or mental disease, which according to clinical judgment might influence everyday function, were excluded.</p>	<p>Mean age = 46 years Age range = 20–79 years Males/females = 18/10</p>	SCIM II	4
Beninato et al. (2004)	Convergent construct validity	N=20	<p>Patients with a lower cervical SCI (C5–C7), English speakers, and who were admitted to rehabilitation within 1 year Of initial SCI were included. Patients were excluded if they had TBI or orthopedic or medical diagnoses other than the neck injury that would affect function at discharge. Subjects with incomplete SCI with abdominal or lower extremity manual muscle testing scores greater than 2 of 5 were also excluded.</p>	<p>• Mean age = 36.8 years Median age = 36 years Age range = 18–62 years Males/females = 16/4 Ethnicity: • White = 19 • Hispanic = 1</p>	FIM	4
Kim et al. (2007)	• Content validity • Convergent construct validity	N=50	<p>Inclusion criteria: traumatic SCI &gt; 12 months, motor incomplete status and a motor level of C4 to L1 inclusive. Exclusion criteria: medication that could affect heart rate; history of heart disease, uncontrolled asthma, or other medical condition that could limit their ability to ambulate safely.</p>	<p>Mean age = 47.4 years Age range = 21–72 years Males = 86%</p>	WISCI II	4
Itzkovich et al. (2007)	Inter-rater reliability	N=425	<p>Patients with spinal cord lesions were included.</p> <p>Patients younger than 18 years, having concomitant impairments that might influence everyday function, such as brain injury or mental disease were excluded.</p>	<p>Mean age = 46.93 years Males/females = 309/116</p>	SCIM III	4

(continued)

TABLE 2. (CONTINUED)

Reference	Psychometric properties	Sample size	Inclusion and exclusion criteria	Study population	Injury features	Instrument	LOE
Olmos et al. (2008)	Internal consistency	N=18	<p>Patients between 16 and 75 yrs with an AIS-D SCI &gt; 6 months were included.</p> <p>Patients with active cardiac or pulmonary pathologic conditions, Parkinson's disease, or other active clinical conditions were excluded.</p>	<p>Mean age = 42.9 years</p> <p>Median age = 39.5 years</p> <p>Age range = 19-72 years</p> <p>Males/females = 12/6</p>	<p>Level of SCI:</p> <ul style="list-style-type: none"> <li>• C3-6 = 12</li> <li>• L3 = 2</li> <li>• T4-T11 = 4</li> </ul> <p>Severity of SCI:</p> <ul style="list-style-type: none"> <li>• AIS D = 22</li> </ul>	<ul style="list-style-type: none"> <li>• 6MWT</li> <li>• 10WMT</li> </ul>	4
van Hedel et al. (2008)	Convergent construct validity	N=917	<p>Inclusion criteria:</p> <ul style="list-style-type: none"> <li>• Patients with a single event (traumatic or ischemic) that resulted in para- or tetraplegia.</li> <li>• First assessment within the first 6 weeks after SCI</li> <li>• Patient capable and willing of giving written informed consent</li> </ul> <p>Exclusion criteria</p> <ul style="list-style-type: none"> <li>• Non-traumatic SCI except for a single ischemic event</li> <li>• Pre-injury dementia or severe reduction of intelligence</li> <li>• Peripheral nerve lesions above the level of SCI</li> <li>• Severe craniocerebral injury</li> </ul>	Not reported	Not reported	<ul style="list-style-type: none"> <li>• TUG</li> <li>• 10WMT</li> </ul>	4

SCI, spinal cord injury; MVA, motor vehicle accident; AIS, American Spinal Injury Association (ASIA) Impairment Scale; TBI, traumatic brain injury.

FIM subscore includes the subscores for communication, psychosocial adjustment, and cognitive function.

In our systematic review, FIM has the greatest number of publications among all instruments used for the assessment of disability in the SCI population. The reliability, internal consistency, and construct validity of FIM have been inconsistently found to be adequate. In addition, the responsiveness and interpretability of FIM were examined in a number of studies that were rated as indeterminate due to a lack of information or poor study design/method. Moreover, a negative ceiling effect, which refers to an effect whereby the instrument cannot detect changes in response on a value higher than some "ceiling," were consistently documented in four previous studies.

#### *Spinal Cord Injury Measure (SCIM)*

While the FIM is reportedly reliable and valid as a disability assessment instrument for various patient groups, including spinal cord injured individuals, the SCIM was specifically developed for patients with SCI, and its use is gradually increasing (Catz et al., 1997; Dodds et al., 1993; Ottenbacher et al., 1996). The SCIM is an attempt to minimize some of the shortcomings observed in the FIM when applied to patients with SCI, such as the ceiling effect (Catz et al., 1997; Hall et al., 1999). The SCIM is a 16-item ordinal scale (range: 0 to 100) that includes three levels of activity (i.e., self-care, respiratory and sphincter management, and mobility) that are weighted according to their clinical relevance (Catz et al., 1997). A second version (SCIM II) with improved phrasing of some of the components was reported to be reliable and valid among individuals with SCI (Catz, Itzkovich, Steinburg et al., 2001). Given the results of a Rasch analysis enriched by critiques of experts, the third version (SCIM III) incorporated several modifications including the addition of upper- and lower-body subitems for bathing and dressing, and the addition of another item on ground-wheelchair transfer (Itzkovich et al., 2002, 2007).

Based on the results of our systematic review, all three versions of the SCIM were examined in eight clinical studies that were multicenter in two occasions but led by the same group of investigators in Israel. Internal consistency was reportedly adequate in all four studies that evaluated this psychometric property in SCIM II and III. While the SCIM I and III consistently showed adequate reliability, there are concerns with regard to the reliability of SCIM II. Convergent construct validity was adequate for the SCIM II in one prior study, but it was assessed as inadequate for SCIM III in another previous study. The SCIM I and III were also reported to have adequate responsiveness in two previous studies.

#### *Walking Index for Spinal Cord Injury (WISCI)*

The WISCI was developed more specifically to evaluate patients with SCI with respect to their walking recovery because (i) this is of great interest for most individuals with SCI during rehabilitative care; (ii) it is poorly assessed by FIM; and (iii) there was supposedly a need for an instrument that assessed walking disability in humans matching with the commonly used locomotor scale developed by Basso, Beattie, and Bresnahan (BBB) for preclinical studies of treatment for SCI (Basso et al., 1996; Dittunno et al., 2000). The WISCI I is a 19-level hierarchical scale where the levels are scored from 1

(patients can ambulate less than 10 m using parallel bars, with braces, and with the physical assistance of two persons) to 19 (patient can ambulate at least 10 m with no devices, no braces, and no physical assistance) (Dittunno et al., 2000). Based on the experience of the use of WISCI in a randomized clinical trial of Body Weight Support Training, a revision of this scale was proposed, and therefore the WISCI II became a 21-level hierarchical scale where the levels are scored from 0 (patient is unable to walk) to 20 (patient can walk without braces and/or devices and without physical assistance for at least 10 m) (Dittunno and Dittunno, 2001).

The results of our systematic review suggest that WISCI I has an adequate agreement in one prior study, but the reliability of WISCI II was not confirmed in another study. While content validity and convergent construct validity were assessed as adequate for WISCI I in one previous study, there were considerable inconsistencies among the seven other studies that examined convergent construct validity of WISCI II. In addition to the uncertainty regarding the responsiveness of WISCI II in three prior studies, ceiling effects and inadequate interpretability of WISCI II were noticed in previous publications.

#### *Timed Up & Go (TUG)*

The TUG was originally developed as a measure of balance in elderly people (Mathias et al., 1986). The TUG is a timed walking test that measures the time (in seconds) for a patient to stand up from an armchair, walk 3 m, return to the chair, and sit down (Podsiadlo and Richardson, 1991).

In our systematic review, the two prior studies on the psychometric properties of TUG in patients with SCI indicate adequate performance of this test regarding floor/ceiling effects. In one previous study, agreement of the TUG was reported as inconsistent, but the TUG showed adequate convergent construct validity.

#### *6-min walk test (6MWT)*

The 6MWT was initially used to assess cardiovascular exercise capacity in elderly patients with congestive heart failure or chronic lung disease (Butland et al., 1982; Guyatt et al., 1985; Roomi et al., 1996). This is a straightforward measure of the distance (in meters) that a patient can walk within 6 min (Butland et al., 1982).

In the SCI population, the psychometric properties of the 6MWT were documented in three previous studies that were captured in our systematic review. For this timed walking test, reliability was uncertain, and agreement was found to be inadequate based on one previous study. While two previous studies assessed convergent construct validity of the 6MWT as adequate, a third study suggested its inadequacy. Responsiveness of the 6MWT was found to be adequate based on the results of one prior study. The adequate performance of the 6MWT in terms of floor/ceiling effects was consistently reported in all three previous studies.

#### *10-m walk test (10MWT)*

The 10MWT has been primarily used as a gait measure in patients with different neurologic movement disorders including stroke and Parkinson's disease (Rossier and Wade, 2001; Schenkman et al., 1997; Smith and Baer, 1999).

TABLE 3. SUMMARY OF THE QUALITY ASSESSMENT OF EACH INSTRUMENT ACCORDING TO THE CRITERIA OF TERWEE AND ASSOCIATES (2007) (SOME OF THE REFERENCES WERE CITED MORE THAN ONCE)

Reference	Instrument	Content validity	Internal consistency	Construct validity	Agreement	Reliability	Responsiveness	Floor/ceiling effect	Interpretability	Setting
Gresham et al. (1986)	QIF	(?)		(+)		(?)				Rehabilitation
Marino et al. (1993)	QIF			(+)						Rehabilitation
Marino et al. (1995)	QIF			(+)						Rehabilitation
Marino and Goin (1999)	Short-form QIF		(+)	(+)			(?)			Rehabilitation
Yavuz et al. (1998)	QIF			(+)	(+)		(?)			Rehabilitation
Yavuz et al. (2000)	WISCI I	(+)		(+)	(+)		(?)			NR
Dittunno et al. (2000)	WISCI II			(+ to -)						Rehabilitation & community
Dittunno et al. (2007)	WISCI II			(+)						Rehabilitation & community
Dittunno et al. (2008)	WISCI II			(+)			(?)		(-)	Rehabilitation & community
Kim et al. (2007)	WISCI II			(-)				(-)		Community
Morganti et al. (2005)	WISCI II			(+)				(-)		Community
van Hedel et al. (2005)	WISCI II			(+)		(?)		(-)		Rehabilitation
van Hedel et al. (2006)	WISCI II			(-)			(?)			Rehabilitation & community
Davidoff et al. (1990)	FIM			(-)				(-)		Acute care & rehabilitation
Roth et al. (1990)	FIM									Rehabilitation
Grey and Kennedy (1993)	FIM							(-)		Rehabilitation
Marino et al. (1993)	FIM									Rehabilitation
Segal et al. (1993)	FIM					(?)			(?)	Rehabilitation
Yavuz et al. (1998)	FIM			(+)			(?)			Rehabilitation
Dijkers and Yavuzer (1999)	Short-form motor FIM					(+)				Community
Hall et al. (1999)	FIM						(?)		(?)	Rehabilitation & community
Kucukdeveci et al. (2001)	FIM		(+ to -)	(+)		(+ to -)		(-)		Rehabilitation
Beninato et al. (2004)	FIM			(+)					(?)	Rehabilitation
Masedo et al. (2005)	FIM Self Report		(+)	(+)		(?)			(?)	Community
Nilsson et al. (2005)	FIM		(+ to -)	(+)						Rehabilitation
Lawton et al. (2006)	FIM		(+ to -)	(-)				(-)		Rehabilitation

Lundgren-Nilsson et al. (2006)	FIM	(-)	(-)				Rehabilitation
Catz et al. (1997)	SCIM I			(+)	(?)		Rehabilitation
Catz, Itzkovich, Agranov et al. (2001)	SCIM I			(+ to -)	(+)		Rehabilitation
Catz, Itzkovich, Steinberg et al. (2001)	SCIM II	(+)		(-)		(+)	Rehabilitation
Catz et al. (2002)	SCIM II	(+)	(+)	(-)			Rehabilitation
Itzkovich et al. (2002)	SCIM II	(+)	(+)	(+ to -)			Rehabilitation
Itzkovich et al. (2003)	SCIM II	(+)	(-)	(+)	(+)		Rehabilitation
Catz et al. (2007)	SCIM III	(+)	(-)	(+)	(+)		Rehabilitation
Itzkovich et al. (2007)	SCIM III	(+)	(+)	(+)	(+)		Rehabilitation
Roth et al. (1990)	Modified Barthel Index						Rehabilitation
Kucukdeveci et al. (2000)	Barthel Index (Turkish version)	(+)	(+ to -)	(+ to -)			Rehabilitation
van Hedel et al. (2005)	TUG	(+)	(+)	(+ to -)		(+)	Rehabilitation & community
van Hedel et al. (2008)	TUG			(+ to -)		(+)	Rehabilitation & community
van Hedel et al. (2005)	6MWT		(+)	(-)	(?)	(+)	Rehabilitation & community
van Hedel et al. (2006)	6MWT		(+)			(+)	Rehabilitation & community
Olmos et al. (2008)	6MWT		(-)			(+)	Rehabilitation & community
van Hedel et al. (2005)	10MWT		(+)	(+ to -)	(?)	(+)	Rehabilitation & community
van Hedel et al. (2006)	10MWT		(+)			(+)	Rehabilitation & community
Olmos et al. (2008)	10MWT		(+)			(+)	Rehabilitation & community
van Hedel et al. (2008)	10MWT		(+)			(+)	Rehabilitation & community

+, positive rating; -, negative rating; ?, indeterminate rating due to lack of information or poor study design /method; NA, not applicable; NR, not reported.

The 10MWT assesses the short-duration walking speed by measuring the time (in seconds) that a patient can walk a 10-m distance.

In our systematic review, reproducibility of the 10WMT was inconsistent in one previous study. However, all four previous studies that examined the 10WMT in the SCI population reported adequate convergent construct validity and appropriate performance with respect to floor/ceiling effect. In addition, this instrument was assessed as having adequate responsiveness in one previous study.

*What is the most reliable, validated, and responsive outcome measure of disability for patients with acute traumatic SCI?*

In our systematic review, there were eight instruments of disability assessment that have been examined in the SCI population. However, all those instruments were mostly tested in the rehabilitation and/or community setting. Only one study reported the use of FIM in the acute care and rehabilitation setting with negative evaluation in terms of its construct validity and ceiling effects (Davidoff et al., 1990). Of note, most of their spinal cord injured subjects were assessed using FIM between 2 and 4 months after SCI (or in the sub-acute stage of SCI) (Davidoff et al., 1990). Our systemic review provides a critique of the psychometric properties of the existing disability scales. Based on our results of quality assessment using the criteria of Terwee and associates, the SCIM has the most appropriate performance with regard to the instrument's psychometric properties. Nonetheless, the paucity of studies on validity, agreement, responsiveness, and interpretability in the setting of acute care suggests further investigations are required to confirm the adequate performance of the SCIM among patients with SCI in rehabilitative care.

*Recommendations*

In the Delphi process, a panel of clinical scientific experts in the field of acute SCI (including basic scientists, clinician-scientists, surgeons, rehabilitation specialists, nurses, and clinical epidemiologists) consensually endorsed the recommendation for use of SCIM III in the classification and evaluation of patients with acute SCI. However, the expert panel also recognized that the identification of a need to find a common objective means to assess the effect of surgery is not an isolated break in strategy, but reflects the general trend of modern science and medicine. Hence, there is the need for further investigations to confirm the performance of the SCIM in the acute care setting in a multi-centered trial.

**Author Disclosure Statement**

No competing financial interests exist. This work was supported by grants from the Rick Hansen Institute and Ontario Neurotrauma Foundation.

**References**

Baptiste, D.C., and Fehlings, M.G. (2008). Emerging drugs for spinal cord injury. *Expert Opin. Emerg. Drugs* 13, 63–80.  
 Basso, D.M., Beattie, M.S., Bresnahan, J.C., Anderson, D.K., Faden, A.I., Gruner, J.A., Holford, T.R., Hsu, C.Y., Noble, L.J.,

Nockels, R., Perot, P.L., Salzman, S.K., and Young, W. (1996). MASCIS evaluation of open field locomotor scores: Effects of experience and teamwork on reliability. *Multicenter Animal Spinal Cord Injury Study. J. Neurotrauma* 13, 343–359.  
 Beninato, M., O'Kane, K.S., and Sullivan, P.E. (2004). Relationship between motor FIM and muscle strength in lower cervical-level spinal cord injuries. *Spinal Cord* 42, 533–540.  
 Butland, R.J., Pang, J., Gross, E.R., Woodcock, A.A., and Geddes, D.M. (1982). Two-, six-, and 12-minute walking tests in respiratory disease. *Br. Med. J. (Clin. Res. Ed.)* 284, 1607–1608.  
 Catz, A., Itzkovich, M., Agranov, E., Ring, H., and Tamir, A. (1997). SCIM – Spinal cord independence measure: A new disability scale for patients with spinal cord lesions. *Spinal Cord* 35, 850–856.  
 Catz, A., Itzkovich, M., Agranov, E., Ring, H., and Tamir, A. (2001). The spinal cord independence measure (SCIM): Sensitivity to functional changes in subgroups of spinal cord lesion patients. *Spinal Cord* 39, 97–100.  
 Catz, A., Itzkovich, M., Steinberg, F., Philo, O., Ring, H., Ronen, J., Spasser, R., Gepstein, R., Tamir, A. (2001). The Catz–Itzkovich SCIM: A revised version of the Spinal Cord Independence Measure. *Disabil. Rehabil.* 23, 263–268.  
 Catz, A., Itzkovich, M., Steinberg, F., Philo, O., Ring, H., Ronen, J., Spasser, R., Gepstein, R., Tamir, A. (2002). Disability assessment by a single rater or a team: A comparative study with the Catz–Itzkovich spinal cord independence measure. *J. Rehabil. Med.* 34, 226–230.  
 Catz, A., Itzkovich, M., Tesio, L., Biering-Sorensen, F., Weeks, C., Laramie, M.T., Craven, B.C., Tonack, M., Hitzig, S.L., Glaser, E., Zeilig, G., Aito, S., Scivoletto, G., Mecci, M., Chadwick, R.J., El Masry, W.S., Osman, A., Glass, C.A., Silva, P., Soni, B.M., Gardner, B.P., Savic, G., Bergstrom, E.M., Bluvshstein, V., and Ronen, J. (2007). A multicenter international study on the Spinal Cord Independence Measure, version III: Rasch psychometric validation. *Spinal Cord* 45, 275–291.  
 Davidoff, G.N., Roth, E.J., Haughton, J.S., and Ardner, M.S. (1990). Cognitive dysfunction in spinal cord injury patients: Sensitivity of the Functional Independence Measure subscales vs neuropsychologic assessment. *Arch. Phys. Med. Rehabil.* 71, 326–329.  
 de Bruin, A.F., de Witte, L.P., Stevens, F., and Diederiks, J.P. (1992). Sickness Impact Profile: The state of the art of a generic functional status measure. *Soc. Sci. Med.* 35, 1003–1014.  
 Dijkers, M.P., and Yavuzer, G. (1999). Short versions of the telephone motor Functional Independence Measure for use with persons with spinal cord injury. *Arch. Phys. Med. Rehabil.* 80, 1477–1484.  
 Dittuno, P.L., and Dittuno, J.F., Jr. (2001). Walking index for spinal cord injury (WISCI II): Scale revision. *Spinal Cord* 39, 654–656.  
 Dittuno, J.F., Jr., Barbeau, H., Dobkin, B.H., Elashoff, R., Harkema, S., Marino, R.J., Hauck, W.W., Apple, D., Basso, D.M., Behrman, A., Deforge, D., Fugate, L., Saulino, M., Scott, M., and Chung, J. (2007). Validity of the walking scale for spinal cord injury and other domains of function in a multicenter clinical trial. *Neurorehabil. Neural Repair* 21, 539–550.  
 Dittuno, J.F., Jr., Dittuno, P.L., Graziani, V., Scivoletto, G., Bernardi, M., Castellano, V., Marchetti, M., Barbeau, H., Frankel, H.L., D'Andrea Greve, J.M., Ko, H.Y., Marshall, R., Nance, P. (2000). Walking index for spinal cord injury (WISCI): An international multicenter validity and reliability study. *Spinal Cord* 38, 234–243.  
 Dittuno, J.F., Scivoletto, G., Patrick, M., Biering-Sorensen, F., Abel, R., and Marino, R. (2008). Validation of the walking



- index for spinal cord injury in a US and European clinical population. *Spinal Cord* 46, 181–188.
- Dodds, T.A., Martin, D.P., Stolov, W.C., and Deyo, R.A. (1993). A validation of the functional independence measurement and its performance among rehabilitation inpatients. *Arch. Phys. Med. Rehabil.* 74, 531–536.
- Field-Fote, E.C., Fluet, G.G., Schafer, S.D., Schneider, E.M., Smith, R., Downey, P.A., and Ruhl, C.D. (2001). The Spinal Cord Injury Functional Ambulation Inventory (SCI-FAI). *J. Rehabil. Med.* 33, 177–181.
- Furlan, J.C., Fehlings, M.G., Tator, C.H., and Davis, A.M. (2008). Motor and sensory assessment of patients in clinical trials for pharmacological therapy of acute spinal cord injury: Psychometric properties of the ASIA Standards. *J. Neurotrauma* 25, 1273–1301.
- Granger, C.V., Albrecht, G.L., and Hamilton, B.B. (1979). Outcome of comprehensive medical rehabilitation: Measurement by PULSES profile and the Barthel Index. *Arch. Phys. Med. Rehabil.* 60, 145–154.
- Gresham, G.E., Labi, M.L., Dittmar, S.S., Hicks, J.T., Joyce, S.Z., and Stehlik, M.A. (1986). The Quadriplegia Index of Function (QIF): Sensitivity and reliability demonstrated in a study of thirty quadriplegic patients. *Paraplegia* 24, 38–44.
- Grey, N., and Kennedy, P. (1993). The Functional Independence Measure: A comparative study of clinician and self ratings. *Paraplegia* 31, 457–461.
- Guyatt, G.H., Feeny, D.H., and Patrick, D.L. (1993). Measuring health-related quality of life. *Ann. Intern. Med.* 118, 622–629.
- Guyatt, G.H., Sullivan, M.J., Thompson, P.J., Fallen, E.L., Pugsley, S.O., Taylor, D.W., and Berman, L.B. (1985). The 6-minute walk: A new measure of exercise capacity in patients with chronic heart failure. *Can. Med. Assoc. J.* 132, 919–923.
- Hall, K.M., Cohen, M.E., Wright, J., Call, M., and Werner, P. (1999). Characteristics of the Functional Independence Measure in traumatic spinal cord injury. *Arch. Phys. Med. Rehabil.* 80, 1471–1476.
- Hawryluk, G.W., Rowland, J., Kwon, B.K., and Fehlings, M.G. (2008). Protection and repair of the injured spinal cord: A review of completed, ongoing, and planned clinical trials for acute spinal cord injury. *Neurosurg. Focus* 25, E14.
- Higginson, I.J. (2007). Quality criteria valuable with slight modification. *J. Clin. Epidemiol.* 60, 1315; author reply 1315–1316.
- Itzkovich, M., Gelernter, I., Biering-Sorensen, F., Weeks, C., Laramee, M.T., Craven, B.C., Tonack, M., Hitzig, S.L., Glaser, E., Zeilig, G., Aito, S., Scivoletto, G., Mecci, M., Chadwick, R.J., El Masry, W.S., Osman, A., Glass, C.A., Silva, P., Soni, B.M., Gardner, B.P., Savic, G., Bergstrom, E.M., Bluvshstein, V., Ronen, J., and Catz, A. (2007). The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study. *Disabil. Rehabil.* 29, 1926–1933.
- Itzkovich, M., Tamir, A., Philo, O., Steinberg, F., Ronen, J., Spasser, R., Gepstein, R., Ring, H., and Catz, A. (2003). Reliability of the Catz–Itzkovich Spinal Cord Independence Measure assessment by interview and comparison with observation. *Am. J. Phys. Med. Rehabil.* 82, 267–272.
- Itzkovich, M., Tripolski, M., Zeilig, G., Ring, H., Rosentul, N., Ronen, J., Spasser, R., Gepstein, R., and Catz, A. (2002). Rasch analysis of the Catz–Itzkovich Spinal Cord Independence Measure. *Spinal Cord* 40, 396–407.
- Jacobs, W.B., and Fehlings, M.G. (2003). The molecular basis of neural regeneration. *Neurosurgery* 53, 943–948; discussion 948–950.
- Keith, R.A., Granger, C.V., Hamilton, B.B., and Sherwin, F.S. (1987). The functional independence measure: A new tool for rehabilitation. *Adv. Clin. Rehabil.* 1, 6–18.
- Kim, M.O., Burns, A.S., Ditunno, J.F., Jr., and Marino, R.J. (2007). The assessment of walking capacity using the walking index for spinal cord injury: Self-selected versus maximal levels. *Arch. Phys. Med. Rehabil.* 88, 762–767.
- Kirshblum, S., Millis, S., McKinley, W., and Tulskey, D. (2004). Late neurologic recovery after traumatic spinal cord injury. *Arch. Phys. Med. Rehabil.* 85, 1811–1817.
- Krischner, B.G.G. (1985). A methodological framework for assessing health indices. *J. Chron. Dis.* 38, 27–36.
- Kucukdeveci, A.A., Yavuzer, G., Elhan, A.H., Sonel, B., and Tennant, A. (2001). Adaptation of the Functional Independence Measure for use in Turkey. *Clin. Rehabil.* 15, 311–319.
- Kucukdeveci, A.A., Yavuzer, G., Tennant, A., Suldur, N., Sonel, B., and Arasil, T. (2000). Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey. *Scand. J. Rehabil. Med.* 32, 87–92.
- Lawton, G., Lundgren-Nilsson, A., Biering-Sorensen, F., Tesio, L., Slade, A., Penta, M., Grimby, G., Ring, H., and Tennant, A. (2006). Cross-cultural validity of FIM in spinal cord injury. *Spinal Cord* 44, 746–752.
- Lohr, K.N., Aaronson, N.K., Alonso, J., Burnam, M.A., Patrick, D.L., Perrin, E.B., and Roberts, J.S. (1996). Evaluating quality-of-life and health status instruments: Development of scientific review criteria. *Clin. Ther.* 18, 979–992.
- Lundgren-Nilsson, A., Tennant, A., Grimby, G., and Sunnerhagen, K.S. (2006). Cross-diagnostic validity in a generic instrument: An example from the Functional Independence Measure in Scandinavia. *Health Qual. Life Outcomes* 4, 55.
- Mahoney, F.I., and Barthel, D.W. (1965). Functional evaluation: The Barthel index. *Md. Med. J.* 14, 61–65.
- Marino, R.J., and Goin, J.E. (1999). Development of a short-form Quadriplegia Index of Function scale. *Spinal Cord* 37, 289–296.
- Marino, R.J., Huang, M., Knight, P., Herbison, G.J., Ditunno, J.F., Jr., and Segal, M. (1993). Assessing self-care status in quadriplegia: Comparison of the quadriplegia index of function (QIF) and the functional independence measure (FIM). *Paraplegia* 31, 225–233.
- Marino, R.J., Rider-Foster, D., Maissel, G., and Ditunno, J.F. (1995). Superiority of motor level over single neurological level in categorizing tetraplegia. *Paraplegia* 33, 510–513.
- Masedo, A.I., Hanley, M., Jensen, M.P., Ehde, D., and Cardenas, D.D. (2005). 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.* 84, 167–176; quiz 177–169, 198.
- Mathias, S., Nayak, U.S., and Isaacs, B. (1986). Balance in elderly patients: The “get-up and go” test. *Arch. Phys. Med. Rehabil.* 67, 387–389.
- McHorney, C.A., and Tarlov, A.R. (1995). Individual-patient monitoring in clinical practice: Are available health status surveys adequate? *Qual. Life Res.* 4, 293–307.
- Morganti, B., Scivoletto, G., Ditunno, P., Ditunno, J.F., and Molinari, M. (2005). Walking index for spinal cord injury (WISCI): Criterion validation. *Spinal Cord* 43, 27–33.
- Nilsson, A.L., Sunnerhagen, K.S., and Grimby, G. (2005). Scoring alternatives for FIM in neurological disorders applying Rasch analysis. *Acta Neurol. Scand.* 111, 264–273.

- Olmos, L.E., Freixes, O., Gatti, M.A., Cozzo, D.A., Fernandez, S.A., Vila, C.J., Agrati, P.E., and Rubel, I.F. (2008). Comparison of gait performance on different environmental settings for patients with chronic spinal cord injury. *Spinal Cord* 46, 331–334.
- Ottenbacher, K.J., Hsu, Y., Granger, C.V., and Fiedler, R.C. (1996). The reliability of the functional independence measure: a quantitative review. *Arch. Phys. Med. Rehabil.* 77, 1226–1232.
- Podsiadlo, D., and Richardson, S. (1991). The timed "Up & Go": A test of basic functional mobility for frail elderly persons. *J. Am. Geriatr. Soc.* 39, 142–148.
- Reid, N. (1993). *Health Care Research by Degrees*. Blackwell: Oxford, England.
- Roomi, J., Johnson, M.M., Waters, K., Yohannes, A., Helm, A., and Connolly, M.J. (1996). Respiratory rehabilitation, exercise capacity and quality of life in chronic airways disease in old age. *Age Ageing* 25, 12–16.
- Rossier, P., and Wade, D.T. (2001). Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. *Arch. Phys. Med. Rehabil.* 82, 9–13.
- Roth, E., Davidoff, G., Haughton, J., and Ardner, M. (1990). Functional assessment in spinal cord injury: A comparison between the Modified Bathel Index and the "adapted" Functional Independence Measure. *Clinical Rehabilitation* 4, 277–285.
- Rowland, J.W., Hawryluk, G.W., Kwon, B., and Fehlings, M.G. (2008). Current status of acute spinal cord injury pathophysiology and emerging therapies: Promise on the horizon. *Neurosurg. Focus* 25, E2.
- Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W.M., and Haynes, R.B. (2000). *Evidence-Based Medicine: How To Practice and Teach EBM*. Churchill Livingstone: Toronto, Canada.
- Schenkman, M., Cutson, T.M., Kuchibhatla, M., Chandler, J., and Pieper, C. (1997). Reliability of impairment and physical performance measures for persons with Parkinson's disease. *Phys. Ther.* 77, 19–27.
- Segal, M.E., Ditunno, J.F., and Staas, W.E. (1993). Interinstitutional agreement of individual functional independence measure (FIM) items measured at two sites on one sample of SCI patients. *Paraplegia* 31, 622–631.
- Smith, M.T., and Baer, G.D. (1999). Achievement of simple mobility milestones after stroke. *Arch. Phys. Med. Rehabil.* 80, 442–447.
- Tator, C.H. (2006). Review of treatment trials in human spinal cord injury: Issues, difficulties, and recommendations. *Neurosurgery* 59, 957–982; discussion 982–957.
- Terwee, C.B., Bot, S.D., de Boer, M.R., van der Windt, D.A., Knol, D.L., Dekker, J., Bouter, L.M., and de Vet, H.C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* 60, 34–42.
- van Hedel, H.J., Wirz, M., and Curt, A. (2006). Improving walking assessment in subjects with an incomplete spinal cord injury: Responsiveness. *Spinal Cord* 44, 352–356.
- van Hedel, H.J., Wirz, M., and Dietz, V. (2005). Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. *Arch. Phys. Med. Rehabil.* 86, 190–196.
- van Hedel, H.J., Wirz, M., and Dietz, V. (2008). Standardized assessment of walking capacity after spinal cord injury: The European network approach. *Neurol. Res.* 30, 61–73.
- WHO (2001). *International Classification of Functioning, Disability, and Health*. World Health Organization: Geneva.
- Wyndaele, M., and Wyndaele, J.J. (2006). Incidence, prevalence and epidemiology of spinal cord injury: What learns a worldwide literature survey? *Spinal Cord* 44, 523–529.
- Yavuz, N., Tezyurek, M., and Akyuz, M. (1998). A comparison of two functional tests in quadriplegia: The quadriplegia index of function and the functional independence measure. *Spinal Cord* 36, 832–837.

Address correspondence to:

Michael G. Fehlings, M.D., Ph.D., FRCSC, FACS  
 Toronto Western Hospital  
 399 Bathurst Street, 4W449  
 Toronto, Ontario M5T 2S8  
 Canada

E-mail: Michael.Fehlings@uhn.on.ca