

REVIEW

A systematic review of functional ambulation outcome measures in spinal cord injury

T Lam^{1,2}, VK Noonan^{2,3,4}, JJ Eng^{2,3,5} and the SCIRE Research Team

¹School of Human Kinetics, University of British Columbia, Vancouver, British Columbia, Canada; ²International Collaboration on Repair Discoveries (ICORD), University of British Columbia, Vancouver, British Columbia, Canada; ³School of Rehabilitation Sciences, University of British Columbia, Vancouver, British Columbia, Canada; ⁴Department of Health Care and Epidemiology, University of British Columbia, Vancouver, British Columbia, Canada and ⁵GF Strong Rehabilitation Centre, Vancouver, British Columbia, Canada

Study design: Systematic review.

Objectives: To systematically review the psychometric properties of outcome measures used to assess ambulation in people with spinal cord injury (SCI).

Setting: Vancouver, BC, Canada.

Methods: A keyword literature search of original articles that evaluated the psychometric properties of ambulation outcome measures in the SCI population was conducted using multiple databases. Multidimensional scales of function were included if specific data were available on ambulation-related subscales. Reliability, validity and responsiveness values were extracted and conclusions drawn about the psychometric quality of each measure.

Results: Seven outcome measures were identified and were broadly categorized into timed and categorical measures of ambulation. Timed measures included timed walking tests that showed excellent reliability, construct validity and responsiveness to change. The psychometric properties of the categorical scales were more variable, but those that were developed specifically for the SCI population had excellent reliability and validity. Categorical scales also exhibited some floor or ceiling effects.

Conclusion: Excellent tools are available for measuring functional ambulation capacity. Further work is required to develop and evaluate outcome measures to include environmental factors that contribute to the ability to achieve safe, functional ambulation in everyday settings.

Sponsorship: Rick Hansen Man-in-Motion Foundation and Ontario Neurotrauma Fund.

Spinal Cord (2008) 46, 246–254; doi:10.1038/sj.sc.3102134; published online 9 October 2007

Keywords: spinal cord injury; walking; measurement; validity; reliability; responsiveness

Introduction

The question of how the spinal cord can be repaired after damage has been pursued for several years. Potential interventions for the recovery of function following spinal cord injury (SCI) likely lie in a combination of pharmacological, surgical and rehabilitation approaches. In order to test the efficacy of any such approaches, reliable, valid and responsive measures of neurological and functional outcome are essential.¹ From the patient's perspective, improvements in the ability to function in everyday activities will be the most meaningful determinant of treatment efficacy.

Almost half of all spinal cord injuries are functionally incomplete,² meaning that there is some sparing of function

below the level of the lesion (American Spinal Injury Association Classification C or D).³ The probability of functional recovery is more substantial in incomplete SCI, and the majority of people with incomplete SCI are able to recover walking.^{2,4,5} Thus given that a significant proportion of people who sustain an SCI may be expected to recover some walking, ambulation outcomes will be an important measurement of the efficacy of new medical or rehabilitation interventions.

Walking not only involves the ability to move the legs, but also requires the intricate coordination of neural commands to regulate upright balance and posture and the ability to adapt gait to environmental constraints. Functional ambulation may therefore be defined as 'the ability to walk, with or without the aid of appropriate assistive devices (such as prostheses, orthoses, canes or walkers), safely and sufficiently to carry out mobility-related activities of daily living.'⁶ When assessing ambulation, we considered two

Correspondence: Dr T Lam, School of Human Kinetics, University of British Columbia, 6081 University Blvd, Vancouver, British Columbia, Canada V6T 1Z1.

E-mail: tania.lam@ubc.ca

Received 2 July 2007; revised 23 August 2007; accepted 7 September 2007; published online 9 October 2007

constructs: capacity and performance.⁷ Using this framework, we define ambulatory capacity as the highest level of walking function achieved within a standardized environment. Ambulatory performance is defined by what an individual actually achieves in his or her environment. Performance therefore is dependent not only on the ability of an individual to execute a given task, but also on the constraints posed by the surrounding environment. The measurement of such outcomes will be important for assessing the efficacy and impact of clinical interventions for enhancing function in people with SCI. Therefore, the objective of this review was to examine the evidence for the validity, reliability and sensitivity of current outcome measures used to measure ambulation in the SCI population.

Materials and methods

Outcome measures were identified using a keyword search of electronic databases (MEDLINE/PubMed, CINAHL, EMBASE, PsychInfo) from 1980 to 2007. The following keywords were used in the search: spinal cord injury/paraplegia/tetraplegia/quadruplegia, ambulation/gait/walking, measure/scale. References of studies were also hand searched for additional studies. Outcome measures were included if information was published on its psychometric properties (that is, reliability, validity and/or responsiveness) using individuals with SCI. Multidimensional outcome measures were included only if psychometric data were available for specific gait/ambulation subscales.

The following psychometric properties were assessed for each of the measures: reliability, validity and responsiveness. Reliability includes reproducibility and internal consistency. Reproducibility examines the degree to which the score is free from random error and includes test–retest reliability and interobserver reliability. Reliability coefficients have been reported using Pearson's product-moment correlation coefficient (r), Spearman's rank correlation (ρ), kappa-statistic (κ) or intraclass correlation coefficients. Internal consistency assesses the homogeneity of the items and is measured using Cronbach's α (α) or split-half reliability. The minimum standard for reproducibility and internal consistency coefficients for group comparisons is 0.70.⁸ We considered reliability coefficients ≥ 0.75 excellent, 0.40–0.74 moderate and ≤ 0.39 poor.⁹ For internal consistency, α scores ≥ 0.80 were considered excellent, 0.70–0.79 considered adequate and ≤ 0.69 poor.⁹

Validity assesses if the instrument actually measures what it intends to measure. Since there are no 'gold standards' for measuring ambulation in SCI, we assessed construct validity. Validity was considered poor if correlation coefficients ≤ 0.49 , moderate if between 0.50 and 0.69 and excellent if ≥ 0.70 .¹⁰ The minimum standard for validity is a correlation coefficient of 0.60.¹¹

Responsiveness determines the ability of a measure to detect clinically important change over time. When possible, effect size (Cohen's d effect size estimate with Hedges adjustment for sample size)¹² was calculated from available

data to measure responsiveness. When only median and range values were reported, the mean and standard deviation of the data were estimated according to the method of Hoza *et al.*¹³ in order to calculate effect size. An effect size of 0.20 is considered small, 0.50 medium and >0.80 large. Floor or ceiling effects, which result when scale items are inappropriately scaled at either extreme, were considered problematic when $>20\%$ of subjects received either minimum (floor) or maximum (ceiling) scores.¹⁰

Reproducibility and responsiveness are measures of within-subject variability that together provide information about how well an outcome measure can detect change, while taking into account changes due to measurement error or random variability. The standard error of measurement (s.e.m.) indicates how many units of change in the measure are necessary until a change beyond expected error is detectable.¹⁴ When data were available, we calculated the s.e.m. by the equation, $s.d. \times \sqrt{1-r}$, where s.d. is the standard deviation of a set of scores and r is the test–retest reliability coefficient of the measurement set.¹⁴ In addition, we calculated the smallest real difference (SRD), calculated by $1.96 \times s.e.m. \times \sqrt{2}$, which has been defined as the smallest change that represents a real (clinical) change beyond 0, with 95% confidence.¹⁵ It provides an indication of how well an outcome measure would be able to detect a clinically relevant change.¹⁵

Results

The search yielded thirteen outcome measures: 10-m walk test (10MWT),^{16–18} 6-min walk test (6MWT),^{16,17} Barthel index (BI),^{19–21} Clinical Outcome Variables Scale (COVS),^{22–24} Functional Independence Measure (FIM),^{25–35} Functional Standing Test (FST),³⁶ Motor Capacities Scale (MCS),³⁷ Needs Assessment Checklist (NAC),^{38,39} Rivermead Mobility Index (RMI), Spinal Cord Injury-Functional Ambulation Index (SCI-FAI),⁴⁰ Spinal Cord Independence Measure (SCIM-I,⁴¹ SCIM-II,⁴² and SCIM-III⁴³), Timed-Up and Go Test (TUG),¹⁷ and the Walking Index for Spinal Cord Injury (WISCI-I⁴⁴ and WISCI-II⁴⁵). Five of these measures were omitted for not directly measuring gait/ambulation (FST, MCS) or lacking psychometric data on gait/ambulation subscales in the SCI population (BI, COVS, RMI). One multidimensional measure (NAC) was omitted because its mobility subscale broadly measured transfers, wheelchair skills, as well as ambulation. The remaining seven measures were divided into two categories: timed measures of ambulation (10MWT, 6MWT, TUG and the Temporal/distance Score of the SCI-FAI) and categorical measures of ambulation (FIM, SCIM, WISCI (I, II) and the Gait Score and Assistive Device component of the SCI-FAI). Timed measures reflect the ability to transport the body from one place to another in a timely manner. Two of the categorical measures (FIM, SCIM) are multidimensional scales of function. The psychometric properties of their ambulation-related subscales (FIM locomotor, FIM_L; SCIM mobility (indoor/outdoor), SCIM_{IOMob}) will be reviewed here.

Table 1 Instrument properties

Instrument	Scale items	Scoring	Characteristics of tested subjects
<i>Timed measures of ambulation</i>			
10MWT	1	Gait speed (m s^{-1}) calculated by measuring the total time required to walk 10 m	Incomplete SCI; C2-L1; <12 months postinjury; all ambulatory within first month postinjury or able to complete the walking tests ^{17,18} or able to walk 10 m within 3 months postinjury ¹⁶
6MWT	1	Total distance (m) walked in 6 min	Incomplete SCI; C2-L1; <12 months postinjury; all ambulatory within first month postinjury ¹⁸ or able to walk 10 m within 3 months postinjury ¹⁶
TUG	1	Total time (s) required to stand up from chair, walk 3 m, turn around, and return to chair	ASIA A, B, C, D; C2-L1; only subjects who were able to complete the walking test were included ¹⁷
SCI-FAI ^a	1 (Temporal/Distance)	Distance walked in 2 min	Incomplete SCI; C3-L1; all able to weight-bear and take at least eight steps ⁴⁰
<i>Categorical measures of ambulation</i>			
SCI-FAI ^b	2 (Gait score and Assistive Devices)	Therapist-rated quality of gait (weight shift, step width, step rhythm, step height, foot contact and step length) and use of assistive devices	Incomplete SCI; C3-L1; all able to weight-bear and take at least eight steps ⁴⁰
FIM _L (locomotor items of motor subscale)	2 (walking or wheelchair propulsion, stair climbing)	The ability to perform the scale items is rated on a seven-point ordinal scale ranging from 1 (total dependence/maximum assistance) to 7 (total independence)	ASIA A, B, C; para- and tetraplegia; <43 years postinjury; distribution of mobility status not reported
SCIM _{10Mob}	9 (includes level walking indoors and outdoors, stair climbing)	Therapist- or patient-rated scale of level of independence/device-reliance (including wheelchair use) for each item	Frankel/ASIA A, B, C, D; para- and tetraplegia; <7 months postinjury; distribution of mobility status not reported
WISCI-II	1	Therapist-rated scale of level of dependence on assistive devices (on a 21-point scale)	ASIA A, B, C, D; para- and tetraplegia; <12 months postinjury; 53% nonwalkers, 5% walking with physical assistance, 42% walking without physical assistance (with or without aids) ⁴⁶

^aIncludes both timed and categorical assessment of ambulation.

^bIncludes both timed and categorical assessment of ambulation.

Instrument properties

Table 1 describes the properties (number of items, scoring method) of each instrument as well as the ambulatory status of the subjects who were tested on these measures.

Timed measures of ambulation

Three of the measures (10MWT, 6MWT, SCI-FAI) include timed measures of overground gait speed or distance, while one measure (TUG) measures the time required to rise from a chair, walk 3 m and return to the chair. Note that only subjects who are able to walk (for example, take at least eight steps⁴⁰) were used for the psychometric testing of these measures (Table 1).

The 10MWT has been widely used as a measure of gait training studies for people with incomplete SCI. There have been several variations to this measure. Some investigators calculate the time required to walk over the full 10-m walkway,⁴⁷ while others account for acceleration and deceleration and measure the time required to walk over the middle 6 m of a 10-m walkway.⁴⁸ van Hedel *et al.*^{16,18} have tested the responsiveness of the 10MWT whereby a 'flying start' is permitted (the patient walks 14 m and the time required to cover the middle 10 m is measured).

The 6MWT measures the distance a patient is able to walk over 6 min (including rests as needed).⁴⁹ It was originally developed as test of aerobic capacity in patients with cardiopulmonary disease.⁵⁰

The SCI-FAI consists of three domains: a Gait Score indicating the quality of gait, an Assistive Devices score indicating the use of assistive devices, and a Temporal/Distance score. The first two domains are therapist rated, while the Temporal/Distance score consists of a timed walking test (distance walked in 2 min) and a self-report ambulation classification score.⁴⁰

The TUG is a timed walking test designed to measure mobility and balance. It was originally developed as a clinical measure of balance in elderly individuals.⁵¹ The individual is instructed to stand up from an arm chair, walk 3 m, return to the chair and sit down at their preferred walking speed. Assistive devices can be used.

Categorical measures of ambulation

Four of the measures describe ambulation with respect to the extent of external support required, in addition to certain domains of the SCI-FAI. Two of these measures are devoted entirely to ambulation (SCI-FAI, WISCI), while the others are subscales of larger multidimensional measures (FIM_L, SCIM_{10Mob}). Note that the categorical measures can capture both nonwalkers as well as walkers or wheelchair users (Table 1).

In addition to the timed domains described above, the SCI-FAI also includes domains that describe the quality of gait (Gait Score) as well as the use of assistive devices. The Gait Score domain was developed in collaboration with physical

Table 2 Reliability

Instrument	Internal consistency	Test-retest	Interobserver	s.e.m.	SRD	SRD (%)
10MWT		$r = 0.983$ ($n = 22$) ¹⁷	$r = 0.974$ ($n = 20$) ¹⁷	0.05 m s ⁻¹	0.13 m s ⁻¹	39
6MWT		$r = 0.981$ ($n = 22$) ¹⁷	$r = 0.971$ ($n = 20$) ¹⁷	16.5 m	45.8 m	22
TUG		$r = 0.979$ ($n = 20$) ¹⁷	$r = 0.973$ ($n = 22$) ¹⁷	3.9 s	10.8 s	30
SCI-FAI		Gait score ICC = 0.850–0.956 ($n = 22$) ⁴⁰	100% agreement for Assistive Devices and Temporal/distance Gait score ICC = 0.703–0.840 ($n = 22$) ⁴⁰	0.7 points ^a	1.9 points	13
FIM _L	Cronbach's ($\alpha < 0.41$) ^{26,33}	Self-report FIM _L Cronbach's $\alpha = 0.64$ ($n = 84$) ³³	Walk/wheelchair item: $r = 0.62$ ($n = 57$) ³⁵ $\kappa = 0.59$ ($n = 1018$), ²⁹ 0.44 ($n = 57$) ³⁵ $\kappa = 0.65$ ($n = 50$) ³⁰ (Turkish version) Stairs item: $r = 0.32$ ($n = 57$) ³⁵ $\kappa = 0.66$ ($n = 1018$), ²⁹ 0.95 ($n = 57$) ³⁵ $\kappa = 0.82$ ($n = 50$) ³⁰ (Turkish version)	1.6 points ^b	4.4 points	49
SCIM _{IO} Mob			$\kappa = 0.840$ – 0.983 ($n = 30$) ⁴¹ (original SCIM) $\kappa = 0.690$ – 0.791 ($n = 28$) ⁴² (SCIM-II) Rasch reliability index = 0.91 ($n = 425$) ⁴³ (SCIM-III)			
WISCI			$\rho = 1.0$ ($n = 40$) ⁴⁴ (original WISCI) $\rho = 1.0$ ($n = 284$) ⁴⁶ (WISCI-II)			

Abbreviations: s.e.m., standard error of the measure; SRD, smallest real difference.

^aMaximum possible score of 20.

^bTwo items that are each rated on a 1–7 scale, for a maximum possible score of 14.

therapists that helped to identify and rank its six parameters: weight shift, step width, step rhythm, step height, foot contact and step length.⁴⁰

In the WISCI, walking is evaluated by determining the patient's dependence on physical assistance, braces or walking aids to cover 10 m. The original WISCI consisted of 19 levels ranging from 'ambulates in parallel bars, with braces and physical assistance of two persons, less than 10 m' (Level 1) to 'ambulates with no devices, no braces and no physical assistance, 10 m (Level 19).⁴⁴ The WISCI has since been modified (WISCI-II) to include two additional scale items for a total of 21 levels: Level 0 to indicate inability to stand or walk with assistance and a new Level 18 to indicate no devices, use of braces and no assistance.⁴⁵

The FIM is a multidimensional scale that assesses the burden of care and functional impairment across a range of domains.⁵² The motor subscale includes 2 locomotor-related items (FIM_L): walking or wheelchair propulsion and stair climbing. Each item is scored on a seven-point scale ranging from 1 (total dependence/maximum assistance) to 7 (total independence). The FIM_L does not consider the use of devices or braces to enable independence. The tool is completed by trained health professionals who observe patient performance.

The SCIM is a new disability scale developed by Catz and colleagues,⁴¹ that specifically addresses patients with spinal cord lesions in order to describe their ability to accomplish activities of daily living. It has since undergone two revisions,^{42,43} the most recent one resulting in the SCIM-III.⁴³ The Mobility (indoors/outdoors) subscale of the SCIM (SCIM-III_{IO}Mob) consists of six items (mobility indoors, mobility for moderate distances (10–100 m), mobility outdoors (>100 m), stair management (up and down three steps), transfers: wheelchair–car, transfers: ground–wheelchair) that are each scored on a 2- to 9-level

categorical scale. Higher scores reflect a higher level of independence.

Note that the FIM_L and SCIM-III_{IO}Mob are not pure ambulation measures since wheelchair use is one of the scoring options. The FIM_L requires one to indicate whether a wheelchair or walking was used as the mode of locomotion. The SCIM_{IO}Mob is scored along a continuum, which extends from wheelchair use to walking with aids to walking without aids. These measures, therefore, may be more applicable to a broader range of individuals with SCI.

Instrument reliability

Table 2 data supporting the reliability of each of the instruments. Most of the timed measures of ambulation showed excellent test-retest and interobserver reliability, with the 10MWT, 6MWT and TUG having the strongest scores. The reliability of most of the measures that assess ambulatory dependence was relatively poorer, with only the WISCI and SCIM-III_{IO}Mob providing superior reliability.

The s.e.m. values show that the amount of change necessary to detect differences beyond expected error were 0.05 m s⁻¹ for gait speed over 10 m, 16.5 m for the distance covered over 6 min, 3.9 s for the TUG, 0.7 points on the Gait Score of the SCI-FAI and 1.6 points on the FIM_L. The SRD tells us the minimum difference required for each of the outcome measures to detect real (clinical) change. This means that a change of 0.13 m s⁻¹ is required for the 10MWT to detect real change, 45.8 m for the 6MWT, 10.8 s for the TUG, 1.9 points for the SCI-FAI Gait Score and 4.4 points on the FIM_L in order for real (clinical) changes to be detected.

Instrument validity

Table 3 provides details on the available data supporting the validity of each of the instruments. Since there is no 'gold

standard' of an ambulation outcome measure for SCI, construct validity was assessed. Construct validity of the timed measures of ambulation (10MWT, 6MWT, TUG) between each other and with the WISCI-II is very strong.¹⁷ Rasch analysis showed that the SCIM-III_{IOMob} has excellent construct validity.⁴³ The SCIM-I_{IOMob} also correlated strongly with the WISCI-II ($\rho = 0.97$, 284 patients).⁴⁶ However, some limitations in the SCIM-I_{IOMob}'s validity for assessing ambulation is suggested by the finding that individuals with a WISCI-II score of 13 (ambulates with walker, no braces, no physical assistance) had no comparable SCIM-I_{IOMob} score.⁴⁶ The validity of the FIM_L appears poor, according to misfits of its items with the Rasch model.^{30,31} Construct validity of the FIM_L with other measures of neural impairment ranged from poor (for example, admission ASIA motor score) to excellent (discharge ASIA motor score).³⁰

Instrument responsiveness and floor/ceiling effects

Available data on the responsiveness of each measure is presented in Table 4. The timed measures of ambulation

(10MWT, 6MWT) were able to detect changes in ambulation between 1 and 3 months and 3 and 6 months postinjury in patients who were able to stand or walk within the first 3 months after injury.¹⁶ Among the categorical measures of ambulatory dependence, there was also a large effect size in WISCI-II scores in changes between 1 and 3 months postinjury.¹⁶

The measures of ambulatory dependence exhibit floor (SCIM-III_{IOMob}, WISCI-II) and ceiling (FIM_L, WISCI-II) effects. In particular, the WISCI-II demonstrated severe ceiling among ambulatory SCI individuals⁵⁵ and severe floor effects among a representative group of SCI patients upon discharge from rehabilitation.⁴⁶ Floor/ceiling effects have not been evaluated for the timed walking tests or the SCI-FAI (Table 5).

Discussion

Ongoing efforts in basic science and applied clinical research promise to bring new strategies for enhancing functional

Table 3 Validity

Instrument	Construct
10MWT	WISCI-II $\rho = -0.68$ ($n = 67$) ¹⁷ TUG $r = 0.89$ ($n = 70$) ^{a17} 6MWT $r = -0.95$ ($n = 62$) ¹⁷
6MWT	10MWT $r = -0.95$ ($n = 62$) ¹⁷ TUG $r = -0.88$ ($n = 62$) ¹⁷ WISCI $\rho = 0.60$ ($n = 60$) ¹⁷
TUG	10MWT $r = 0.89$ ($n = 70$) ¹⁷ 6MWT $r = -0.88$ ($n = 62$) ¹⁷
SCI-FAI	Gait speed $r = -0.74$ ($n = 22$) ⁴⁰ Self-report walking mobility ⁵³ $r = 0.697$ ($n = 22$) ⁴⁰
FIML	Craig Handicap Assessment and Reporting Technique (CHART) $r = 0.13$ (CHART physical subscale) and 0.26 (CHART mobility subscale) ³³ Admission ASIA motor score $\rho = 0.38$ ($n = 61$) ³⁰ Discharge ASIA motor score $\rho = 0.81$ ($n = 58$) ³⁰ Admission ASIA sensory score $\rho = 0.30$ ($n = 61$) ³⁰ Discharge ASIA sensory score $\rho = 0.53$ ($n = 58$) ³⁰ Admission Rasch infit = 1.24 (stairs subscore), 0.94 (walk/wheelchair subscore) ($n = 62$) ³⁰ Admission Rasch outfit = 0.68 (stairs subscore), 0.70 (walk/wheelchair subscore) ($n = 62$) ³⁰ Discharge Rasch infit = 1.42 (stairs subscore), 0.54 (walk/wheelchair subscore) ($n = 59$) ³⁰ Discharge Rasch outfit = 1.01 (stairs subscore), 0.64 (walk/wheelchair subscore) ($n = 59$) ³⁰ Rasch infit = 3.56 (stairs subscore), 1.29 (walk/wheelchair subscore) ($n = 358$) ³¹ Rasch outfit = 4.70 (stairs subscore), 1.53 (walk/wheelchair subscore) ($n = 358$) ³¹
SCIM-III (IOMob)	Rasch infit = 0.94 ($n = 425$) ⁴³ Rasch outfit = 0.89 ($n = 425$) ⁴³ Unidimensionality = 98.8% of variability ⁴³
WISCI-II	SCIM _{IOMob} $\rho = 0.97$ ($n = 76$) ⁴⁶ FIML $\rho = 0.765$ ⁴⁴ (original WISCI); $\rho = 0.7$ ⁴⁶ ($n = 76$); $\rho = 0.92, 0.89, 0.88^a$ ($n = 104-121$) ⁵⁴ 6MWT $\rho = 0.76, 0.68, 0.69^b$ ($n = 80-88$) ⁵⁴ Rivermead Mobility Index (ambulation items) $\rho = 0.67$ ($n = 76$) ⁴⁶ Barthel Index (ambulation item) $\rho = 0.67$ ($n = 76$) ⁴⁶ LEMS $\rho = 0.58$ ($n = 200$) ⁴⁶ ; LEMS $\rho = 0.85, 0.85, 0.88^c$ ($n = 107-121$) Berg Balance Scale $\rho = 0.91, 0.89, 0.92^b$ ($n = 107-120$) ⁵⁴

^aConcurrent validity between the 10MWT and the TUG or 6MWT was excellent whether patients required assistance or not during walking (Van Hedel *et al.*, 2005).

^bAs reported by the authors.

^cAt 3, 6 and 12 months, respectively, following entry into a 3-month gait training intervention.

Table 4 Responsiveness

Instrument	Method	Responsiveness
10MWT	Mean change b/w 1 and 3 months postinjury ($P < 0.001$) ¹⁶ Mean change b/w 3 and 6 months postinjury ($P < 0.001$) ¹⁶	Effect size: 0.92 (0.51, 1.32) ^a Effect size: 0.47 (0.07, 0.86)
6MWT	Mean change b/w 1 and 3 months postinjury ($P < 0.001$) ¹⁶ Mean change b/w 3 and 6 months postinjury ($P < 0.001$) ¹⁶	Effect size: 0.92 (0.51, 1.33) Effect size: 0.50 (0.10, 0.89)
TUG	NA	NA
SCI-FAI (Gait score)	Mean change after an experimental gait intervention ⁴⁰	Change by 44.7% ($P < 0.001$)
FIM _L (self-report)	Discrimination between paraplegia and tetraplegia ($P > 0.001$) ³³ Detection of change between admission and discharge from neurorehabilitation. ³⁴	Effect size: 0.38 (-0.06, 0.81) Effect size: 0.90 ^b
SCIM _{IOMob} (original SCIM)	Responsiveness compared with FIM _L .	Detected 48% (incomplete tetraplegia), 55% (incomplete paraplegia) and 33% (complete paraplegia) more of the functional changes compared to FIM. ⁴²
WISCI-II	Discrimination between 1 and 3 months postinjury ($P < 0.001$) ¹⁶ Discrimination between 3 and 6 months postinjury ($P = 0.004$) ¹⁶	Effect size: 2.05 (1.57, 2.53) Effect size: 0.73 (0.33, 1.13)

Abbreviation: NA, not available.

^aEffect size = Hedges g effect size estimate calculated from data (lower bound, upper bound 95% confidence interval).

^bAs reported by the authors.

Table 5 Floor/ceiling effects

Instrument	Floor/ ceiling	
10MWT	NA	
6MWT	NA	
TUG	NA	
SCI-FAI	NA	
FIM _L	Ceiling	Observed in >20% of paraplegics upon discharge from rehabilitation ²⁸
SCIM-III _{IOMob}	Floor	Observed in ~37% of subjects scored within 1 week of admission to rehabilitation program ⁴³
WISCI-II	Ceiling	17% (42/249) of subjects upon discharge from rehabilitation had WISCI = 20 ⁴⁶ 48% (24/50) ambulatory subjects >1 year postinjury had WISCI = 20 ⁵⁵
	Floor	53% (131/249) of subjects with WISCI = 0 upon discharge from rehabilitation ⁴⁶

Abbreviation: NA, not available.

recovery following SCI. Reliable, valid and responsive outcome measures will be necessary for accurately assessing the efficacy of any new intervention.¹ In this review, we evaluated the psychometric properties of available measures for assessing functional ambulation in the SCI population.

The timed measures of ambulation (10MWT, 6MWT, TUG) consistently fared the best in terms of test-retest and interobserver reliability, construct validity and responsiveness to change (particularly over the first 3 months postinjury) in a subgroup of ambulatory SCI patients. Of the categorical measures, the reliability of the SCI-FAI, SCIM_{IOMob} and WISCI was moderate to excellent while the FIM_L was the worst, with poor to moderate reliability and poor internal consistency. The s.e.m. provides valuable insight into how much change is necessary to detect differences beyond random measurement error and the SRD indicates the minimum difference between two scores that represents a real change beyond 0.¹⁵ Beckerman *et al.*¹⁵ emphasizes that there must be a distinction between the SRD and what clinicians define as 'clinically relevant'. The SRD is

a psychometric property of an instrument that indicates how well it can detect a given size of change (that is, the 10MWT can detect a change of at least 0.13 m s⁻¹). For an instrument to be sensitive to a clinically relevant level of change, its SRD should be less than what is deemed clinically relevant.¹⁵ We recommend that future work focus on establishing the relationship of clinically deemed relevant differences with the SRD of these measures in order to improve our ability to accurately measure functional change in SCI.

Construct validity of these scales with respect to other measures of ambulation or neurological impairment ranged from poor (for example, FIM_L Rasch analysis^{30,31}) to excellent (SCIM-III_{IOMob}⁴³). The responsiveness of the WISCI-II to change over the first 3 months postinjury was particularly strong in a subgroup of ambulatory SCI patients;¹⁶ effect sizes for other comparisons and other scales were somewhat more modest. Many of the categorical measures of ambulation also exhibited floor or ceiling effects, especially the FIM_L upon discharge (ceiling), the SCIM-III_{IOMob} upon admission to rehabilitation (floor) and the WISCI-II (floor and ceiling) upon discharge from rehabilitation. One of the advantages of the timed walking tests is that there is no conceivable ceiling effect. However, these tests are valid only for those individuals who are able to walk, whereas a more representative SCI population can be evaluated with the categorical tests. In addition to their utility in tracking both walkers and nonwalkers, categorical tests can also capture the transition from nonambulatory to ambulatory status.

Although the FIM is often considered the gold standard for assessing activities of daily living, it is not SCI specific and the results of this review reveal its limited reliability, construct validity, responsiveness to different lesion levels and ceiling effects in people with paraplegia. The poor reliability of the FIM_L may be attributed to the fact that this subscale groups walking and wheelchair mobility together in one item when they are actually two separate modes of locomotion. One recommendation to improve internal consistency of the FIM_L is to split these items.²⁶ Rasch analysis has also suggested that the psychometric properties

of the FIM could be improved by reducing the seven-point scale rating to a four- or five-point scale.³¹

On the other hand, the SCIM, which was developed specifically for the SCI population,⁴¹ shows promise in becoming the gold standard for the comprehensive assessment of basic function in SCI.⁵⁶ The SCIM-III_{IOMob} has been shown to have excellent interobserver reliability and construct validity⁴³ and may be a useful indicator of functional ambulation as part of a larger multidimensional assessment of function. Further research should determine whether the revised version of the SCIM (SCIM-III) has improved precision in capturing all levels of ambulatory ability.⁴⁶

The WISCI-II, which was also developed specifically for the SCI population, is solely used for the assessment of ambulatory function rather than overground locomotion in general (that is, where both wheelchair use and walking are options).⁴⁶ It therefore provides a more comprehensive consideration of the use of braces and assistive devices to achieve overground ambulation not found in the other categorical measures. However, the WISCI exhibits ceiling effects, which could limit its use in assessing individuals with only minor impairments. The WISCI-II also does not consider gait speed or energy consumption and does not provide any indication of endurance since the distance covered is only 10 m. It has been suggested that the WISCI-II would benefit from additional information on walking speed to improve responsiveness and to decrease its ceiling effect.^{17,46,55}

One suggestion has been to use a combination of the WISCI-II and a timed test (for example, 10MWT) to assess functional ambulation in individuals with SCI.^{1,55} Walking velocity measured by either the 6MWT or the 10MWT is comparable in people with incomplete SCI who are able to complete both tests, although care must be taken in exactly how the tests are performed (for example, whether a flying start is permitted for the 10MWT or the dimensions of the track to be used for the 6MWT).^{16,18} Due to its shorter and easier implementation, the 10MWT has been recommended as the preferred timed ambulation test in people with SCI.¹⁶

Despite the appeal of using these quick and simple measures of functional ambulation (that is, WISCI-II and 10MWT), important information about the quality of walking may be missed. Only one of the ambulatory outcome measures we examined includes an assessment of the quality of the gait pattern (SCI-FAI). Further work is required to elucidate the relationship between common gait deviations with functional outcomes (for example, gait speed and dependence on ambulatory aids). In addition, the usefulness of such measures is obviously limited to only those individuals who are able to take at least some steps. Therefore, care must be taken in choosing appropriate outcome measures for specific subgroups of the SCI population. Categorical measures such as the SCIM-III may be valuable for capturing a broader range of locomotor abilities among SCI individuals, while scales such as the WISCI-II or the timed tests can provide a more precise and specific measure of ambulation.

An important safety aspect of ambulation is the ability to maintain upright balance while walking. None of the

measures we reviewed directly measure balance. The Berg Balance Scale is a commonly used clinical measure of static and dynamic balance in the elderly, stroke and Parkinson's populations,^{57,58} but has yet to be validated for SCI. There are otherwise limited tools available for assessing balance capacity during walking and none to our knowledge would be suitable for clinical settings. The development and validation of such a dynamic balance-walking tool would be useful as an adjunct to ambulation outcome measures in the SCI population.

Most of the measures we reviewed assess level gait, ranging from simple timed tests such as the 10MWT or 6MWT to the SCI-FAI, which describes specific qualities of leg movement during gait. Although level gait may be captured by a combination of the 10MWT and WISCI-II, these measures lack consideration of 'real-life' environmental constraints (mobility).⁵⁹ The only aspect of mobility, beyond level gait, covered by these measures is stair-climbing ability (rated in the FIM_L and SCIM_{IOMob}). Environmental surveys to help identify the various challenges faced in everyday walking situations are only beginning to emerge.^{60,61} Some common environmental challenges, such as stair-climbing and obstacles, have already been chosen for ambulatory scales for the stroke population. Measures, such as the Emory Functional Ambulation Profile,⁶² may provide a basis for developing an SCI-specific scale of functional ambulation. Measures that go beyond walking in a straight line may also prove useful in functional ambulation assessments in SCI. The L-test,⁶³ which requires different degrees of turning toward both sides of the body, is one such measure whose value should be assessed for the SCI population.

The International Classification of Functioning (ICF) classifies 'capacity' as the highest level of functioning that a person may reach in a given domain and is generally done in a 'standardized' environment.⁷ The measures described in this review fall within the construct of capacity as they are typically evaluated in laboratory or hospital settings and participants are trying to the best of their ability to obtain a better score. In contrast, the ICF classifies 'performance' as what an individual does in his or her current environment.⁷ The measurement of such outcomes will be important for assessing the efficacy and impact of clinical interventions for enhancing function in people with SCI. Direct measurement of such parameters are being implemented by some researchers with the emergence of improved technology for monitoring daily ambulation (for example, StepWatch3 Activity Monitor)⁶⁴ or remote tracking systems⁶⁰ that can monitor the quantity and quality of ambulation outside the standard, controlled environment of laboratories and clinics. Recently, a self-selected WISCI-II score corresponding to the level normally used to walk in the home or community (performance measure) was compared to the maximal WISCI-II achieved in a controlled setting (capacity measure) in a group of ambulatory, chronic SCI subjects.⁵⁵ It was found that subjects walked faster and with less energy expenditure at the self-selected WISCI-II level in the home or community. This highlights the importance of how changes in ambulation should be assessed in clinical trials (that is, whether the usual or minimal level of assistive

devices should be used).⁵⁵ The reliability of the SCIM-II when assessed by interview was also recently examined and found to be comparable with assessment by observation in a hospital setting.⁶⁵ Further research to validate these and other measures using interview or self-report in the home or community are encouraged to provide better performance measures of ambulation, which will provide complementary information to established capacity measures (for example, 10MWT).

Future work should also be directed toward determining the contributing factors that enable independent community walking. For example, it has been suggested that self-paced walking speeds of 0.4 and 0.8 m s⁻¹ are the minimum criteria for limited and unlimited community ambulation in people with stroke.⁵³ The ability to manage curbs is also considered a critical task for independent community walking.⁵³ Individuals who achieve this level of walking also tend to score better on a quality of life scale.⁶⁶ Similar criteria have yet to be determined for the SCI population.

Acknowledgements

We thank the Rick Hansen Man-in-Motion Foundation and Ontario Neurotrauma Fund for their financial support. VKN is supported by a Canadian Institutes of Health Research fellowship. JJE is a senior scholar of Michael Smith Foundation for Health Research and Canadian Institutes of Health Research new investigator.

References

- 1 Steeves JD, Lammertse D, Curt A, Fawcett JW, Tuszynski MH, Ditunno JF *et al*. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP panel: clinical trial outcome measures. *Spinal Cord* 2007; **45**: 206–221.
- 2 Burns SP, Golding DG, Rolle Jr WA, Graziani V, Ditunno Jr JF. Recovery of ambulation in motor-incomplete tetraplegia. *Arch Phys Med Rehabil* 1997; **78**: 1169–1172.
- 3 Maynard Jr FM, Bracken MB, Creasey G, Ditunno Jr JF, Donovan WH, Ducker TB *et al*. International standards for neurological and functional classification of spinal cord injury. American spinal injury association. *Spinal Cord* 1997; **35**: 266–274.
- 4 Waters RL, Adkins RH, Yakura JS, Sie I. Motor and sensory recovery following incomplete tetraplegia. *Arch Phys Med Rehabil* 1994; **75**: 306–311.
- 5 Waters RL, Adkins RH, Yakura JS, Sie I. Motor and sensory recovery following incomplete paraplegia. *Arch Phys Med Rehabil* 1994; **75**: 67–72.
- 6 Stroke Engine. Glossary of terms. <http://www.medicine.mcgill.ca/strokeengine/definitions-en.html>. Accessed on August 23, 2007.
- 7 World Health Organization. *Towards a Common Language for Functioning, Disability and Health: The International Classification of Functioning, Disability and Health*. World Health Organization: Geneva, 2002.
- 8 Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002; **11**: 193–205.
- 9 Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000; **81**: S15–S20.
- 10 Andresen EM, Fouts BS, Romeis JC, Brownson CA. Performance of health-related quality-of-life instruments in a spinal cord injured population. *Arch Phys Med Rehabil* 1999; **80**: 877–884.
- 11 Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; **2**.
- 12 Devilly GJ. *The Effect Size Generator for Windows, version 2.3*. Centre for Neuropsychology, Swinburne University: Australia, 2004.
- 13 Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005; **5**: 13–22.
- 14 Domholdt E. *Physical Therapy Research. Principles and Applications*, 2nd edn. WB Saunders: Philadelphia, PA, 2000.
- 15 Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* 2001; **10**: 571–578.
- 16 van Hedel HJ, Dietz V, Curt A. Assessment of walking speed and distance in subjects with an incomplete spinal cord injury. *Neurorehabil Neural Repair* 2007; **21**: 295–301.
- 17 van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. *Arch Phys Med Rehabil* 2005; **86**: 190–196.
- 18 van Hedel HJ, Wirz M, Curt A. Improving walking assessment in subjects with an incomplete spinal cord injury: responsiveness. *Spinal Cord* 2006; **44**: 352–356.
- 19 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.
- 20 O'Connor RJ, Cano SJ, Thompson AJ, Hobart JC. Exploring rating scale responsiveness: does the total score reflect the sum of its parts? *Neurology* 2004; **62**: 1842–1844.
- 21 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.
- 22 Barker RN, Amsters DI, Kendall MD, Pershouse KJ, Haines TP. Reliability of the clinical outcome variables scale when administered via telephone to assess mobility in people with spinal cord injury. *Arch Phys Med Rehabil* 2007; **88**: 632–637.
- 23 Campbell J, Kendall M. Investigating the suitability of the clinical outcome variables scale (COVS) as a mobility outcome measure in spinal cord injury rehabilitation. *Physiotherapy Canada* 2003; **55**: 135–144.
- 24 Seaby L, Torrance G. Reliability of a physiotherapy functional assessment used in a rehabilitation setting. *Physiotherapy Canada* 1989; **41**: 264–271.
- 25 Dijkers MP, 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.
- 26 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.
- 27 Grimby G, Gudjonsson G, Rodhe M, Sunnerhagen KS, Sundh V, Ostensson ML. The functional independence measure in Sweden: experience for outcome measurement in rehabilitation medicine. *Scand J Rehabil Med* 1996; **28**: 51–62.
- 28 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.
- 29 Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level functional independence measure (FIM). *Scand J Rehabil Med* 1994; **26**: 115–119.
- 30 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.
- 31 Nilsson AL, Sunnerhagen KS, Grimby G. Scoring alternatives for FIM in neurological disorders applying Rasch analysis. *Acta Neurol Scand* 2005; **111**: 264–273.
- 32 Ota T, Akaboshi K, Nagata M, Sonoda S, Domen K, Seki 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.
- 33 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**: 167–176.

- 34 Cano SJ, O'Connor RJ, Thompson AJ, Hobart JC. Exploring disability rating scale responsiveness II: do more response options help? *Neurology* 2006; **67**: 2056–2059.
- 35 Segal ME, Ditunno JF, Staas WE. Interinstitutional agreement of individual functional independence measure (FIM) items measured at two sites on one sample of SCI patients. *Paraplegia* 1993; **31**: 622–631.
- 36 Triolo RJ, Bevelhimer T, Eisenhower G, Wormser D. Inter-rater reliability of a clinical test of standing function. *J Spinal Cord Med* 1995; **18**: 14–22.
- 37 Fattal C. Motor capacities of upper limbs in tetraplegics: a new scale for the assessment of the results of functional surgery on upper limbs. *Spinal Cord* 2004; **42**: 80–90.
- 38 Berry C, Kennedy P. A psychometric analysis of the needs assessment checklist (NAC). *Spinal Cord* 2002; **41**: 490–501.
- 39 Kennedy P, Hamilton LR. The needs assessment checklist: a clinical approach to measuring outcome. *Spinal Cord* 1999; **37**: 136–139.
- 40 Field-Fote EC, Fluet GG, Schafer SD, Schneider EM, Smith R, Downey PA *et al*. The spinal cord injury functional ambulation inventory (SCI-FAI). *J Rehabil Med* 2001; **33**: 177–181.
- 41 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.
- 42 Catz A, Itzkovich M, Steinberg F, Philo O, Ring H, Ronen J *et al*. The Catz–Itzkovich SCIM: a revised version of the spinal cord independence measure. *Disabil Rehabil* 2001; **23**: 263–268.
- 43 Catz A, Itzkovich M, Tesio L, Biering-Sorensen F, Weeks C, Laramie MT *et al*. A multicenter international study on the spinal cord independence measure, version III: Rasch psychometric validation. *Spinal Cord* 2007; **45**: 275–291.
- 44 Ditunno Jr JF, Ditunno PL, Graziani V, Scivoletto G, Bernardi M, Castellano V *et al*. Walking index for spinal cord injury (WISCI): an international multicenter validity and reliability study. *Spinal Cord* 2000; **38**: 234–243.
- 45 Dittuno PL, Dittuno Jr JF. Walking index for spinal cord injury (WISCI II): scale revision. *Spinal Cord* 2001; **39**: 654–656.
- 46 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.
- 47 Thomas SL, Gorassini MA. Increases in Corticospinal tract function by treadmill training after incomplete spinal cord injury. *J Neurophysiol* 2005; **94**: 2844–2855.
- 48 Wirz M, Zemon DH, Rupp R, Scheel A, Colombo G, Dietz V *et al*. Effectiveness of automated locomotor training in patients with chronic incomplete spinal cord injury: a multicenter trial. *Arch Phys Med Rehabil* 2005; **86**: 672–680.
- 49 American Thoracic Society. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; **166**: 111–117.
- 50 Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. *Br Med J (Clin Res Ed)* 1982; **284**: 1607–1608.
- 51 Podsiadlo D, Richardson S. The timed 'up & go': a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991; **39**: 142–148.
- 52 Granger CV, Hamilton BB, Keith RA, Zielesny M, Sherwin FS. Advances in functional assessment for medical rehabilitation. *Top Geriatr Rehabil* 1986; **1**: 59–74.
- 53 Perry J, Garrett M, Gronley JK, Mulroy SJ. Classification of walking handicap in the stroke population. *Stroke* 1995; **26**: 982–989.
- 54 Ditunno Jr JF, Barbeau H, Dobkin BH, Elashoff R, Harkema S, Marino RJ *et al*. Validity of the walking scale for spinal cord injury and other domains of function in a multicenter clinical trial. *Neurorehabil Neural Repair* 2007; **21**: 295–301.
- 55 Kim MO, Burns AS, Ditunno JF, Marino RJ. The assessment of walking capacity using the walking index for spinal cord injury: self-selected versus maximal levels. *Arch Phys Med Rehabil* 2007; **88**: 762–767.
- 56 Miller WC, Curt A, Elliott S, Hsieh J, Mortenson WB, Noonan V *et al*. Outcome measures. In: Eng JJ, Teasell RW, Miller WC, Wolfe DL, Townson AF, Aubut J, *et al*. (eds). *Spinal Cord Injury Rehabilitation Evidence*. ICORD: Vancouver, BC, 2006.
- 57 Berg K, Wood-Dauphinee S, Williams JL. The balance scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med* 1995; **27**: 27–36.
- 58 Franchignoni F, Martignoni E, Ferriero G, Pasetti C. Balance and fear of falling in Parkinson's disease. *Parkinsonism Relat Disord* 2005; **11**: 427–433.
- 59 Patla AE, Shumway-Cook A. Dimensions of mobility: defining the complexity and difficulty associated with community mobility. *J Aging Phys Act* 1999; **7**: 7–19.
- 60 Meng JT, Gage WH, Zabjek K, Brooks D, Maki BE, Mihailidis A *et al*. iWalker: a 'Real-World' Mobility Assessment Tool. *Festival of International Conferences on Caregiving, Disability, Aging and Technology (FICCDAT) Book of Abstracts*. Toronto, Canada, 2007, p 227.
- 61 Musselman KE, Yang JF. Walking tasks encountered by urban-dwelling adults and persons with incomplete spinal cord injuries. *J Rehabil Med* 2007; **39**: 567–574.
- 62 Wolf SL, Catlin PA, Gage K, Gurucharri K, Robertson R, Stephen K. Establishing the reliability and validity of measurements of walking time using the Emory functional ambulation profile. *Phys Ther* 1999; **79**: 1122–1133.
- 63 Deathe AB, Miller WC. The L test of functional mobility: measurement properties of a modified version of the timed 'up & go' test designed for people with lower-limb amputations. *Phys Ther* 2005; **85**: 626–635.
- 64 Behrman AL, Lawless-Dixon AR, Davis SB, Bowden MG, Nair P, Phadke C *et al*. Locomotor training progression and outcomes after incomplete spinal cord injury. *Phys Ther* 2005; **85**: 1356–1371.
- 65 Itzkovich M, Tamir A, Philo O, Steinberg F, Ronen J, Spasser R *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.
- 66 Schmid A, Duncan PW, Studenski S, Lai SM, Richards L, Perera S *et al*. Improvements in speed-based gait classifications are meaningful. *Stroke* 2007; **38**: 2096–2100.