Original Article

Rasch analysis of the Catz-Itzkovich spinal cord independence measure

M Itzkovich^{1,2}, M Tripolski², G Zeilig^{2,3}, H Ring^{1,2}, N Rosentul¹, J Ronen^{1,2}, R Spasser¹, R Gepstein⁴ and A Catz^{*,1,2}

¹Loewenstein Rehabilitation Hospital, Raanana, Israel; ²Tel Aviv University, Tel Aviv, Israel; ³Sheba Medical Center, Tel Hashomer, Israel; ⁴Sapir Medical Center, Kfar Sava, Israel

Background: The Spinal Cord Independence Measure (SCIM) is a new disability scale developed specifically for patients with spinal cord lesions (SCL). Its original and second versions (SCIM and SCIM II) were found to be reliable and more sensitive than the Functional Independence Measure (FIM) to functional changes in SCL patients.

Objective: To further validate the SCIM II, examining its components on a larger population.

Design: Retrospective cohort study.

Setting: Two rehabilitation centers in Israel.

Subjects: Two hundred and two inpatients with SCL.

Interventions: Routine SCIM assessments by staff nurses. Rasch and accompanying analyses.

Main outcome measures: Unidimensionality of subscales (areas of function); goodness of fit of the tasks to the Rasch model; relationship of total-patient and single-task performance-ability; usability of task categories and the order of threshold locations between them; subscale discrimination of ability and difficulty and hierarchical nature; discrimination of task-categories ability, ie, distribution of thresholds along ability levels; and differential task behavior by age, gender and examination subgroups.

Results: Four unidimensional subscales were identified, and an acceptable goodness of fit to the Rasch model was demonstrated in most of their tasks (infit mean square = 0.8–1.2, outfit mean square = 0.6–1.4). However, some tasks showed overfit (bathing lower body) and some showed misfit (wheelchair-car transfer). Additional analyses performed to check for reasons for less than acceptable fit revealed flaws in a minority of the outcome measures.

Conclusions: The findings of this analysis confirm the validity and reliability of the SCIM II. To a large extent we can infer that the SCIM II construct allows for the detection of any level of disability in any patient with SCL. A few item categories, however, should be rephrased or removed.

Spinal Cord (2002) **40**, 396–407. doi:10.1038/sj.sc.3101315

Keywords: SCIM; Rasch analysis; spinal cord; disability assessment

Introduction

The Spinal Cord Independence Measure (SCIM) is a disability scale developed at Loewenstein Rehabilitation Hospital to assess the capacity of the individual patient with a spinal cord lesion (SCL) to perform daily tasks.¹ It consists of 94 categories defining 18 individual tasks (items) divided into three subscales (areas of function), as follows: Self-Care (score range 0–20), Respiration and Sphincter Management (0–40), and Mobility (0–40). Mobility is scored separately for room/toilet and indoors/outdoors. The total score

ranges between 0 and 100.¹ Patients are evaluated by observation. Individual capacity is defined by degree or by estimated value on the basis of the patient's comfort, the medical condition, and the economic burden. Thus, a functional ability or 'activity'² achieved with costlier, heavier or more challenging assistive devices, or associated with medical shortcomings, is considered to be of a lower value, and is scored lower. This approach is in accord with the revised draft of the WHO International Classification of Impairments, Disabilities and Handicaps (ICIDH).² Some items of the original SCIM have been rephrased to improve reliability. The new version, the Catz-Itzkovich SCIM or SCIM II, appears in Appendix 1.

*Correspondence: A Catz, Medical Director, Department IV, Spinal Rehabilitation, Loewenstein Rehabilitation Hospital, 278 Achuzta St., P.O. Box 3 Raanana 43100, Israel

Although both versions have been found to be reliable and valid,^{3,4} the studies were performed in a single clinical department with a limited number of subjects and by a multidisciplinary team made up especially for the purposes of the research. All used tests of concurrent criterion-related validity (correlation with scores on the Functional Independence Measure) and discriminative construct validity (comparison of sensitivity to functional changes with the Functional Independence Measure).^{5,6}

The present study was conducted as a part of the European Standardization of Outcome Measurement in Rehabilitation (Pro-ESOR) project. To further validate the SCIM II, we examined its components in a larger population, at two rehabilitation centers, in the routine work setting.

Patients and methods

The study population consisted of 202 inpatients with SCLs (138 males, 64 females) treated at Loewenstein Rehabilitation Hospital and Sheba Medical Center in Israel. Seventy-eight patients had tetraplegia and 124 paraplegia; 65 had complete or almost complete lesions (ASIA grade A or B) and 137 had incomplete lesions (ASIA C or D). Ages ranged from 14 to 94 years (mean 46.78, SD 18.38). Ninety-seven lesions were caused by trauma; causes in the remainder were spinal stenosis, disc protrusion, benign tumors, myelitis, vascular impairments, multiple sclerosis, congenital spine anomalies, syringomyelia, decompression sickness, or motor neuron disease.

A staff nurse evaluated all the patients with the SCIM II. One hundred and seventy-one patients underwent a second evaluation 2–4 weeks after the first.

Statistical analysis

Statistical analysis was done to determine the extent to which the construct of the SCIM II allows for the detection of the true functional level (ability) of any SCL patient (construct validity), and to examine the similarity of findings among subgroups (reliability). First, we examined the assignment of the SCIM II tasks to unidimensional subscales. Thereafter, we compared the observed scores for the tasks within the identified subscales to the expected values as estimated by the Rasch model,⁷ and then further analyzed them. All the analyses were performed using the BIGSTEPS computer program.

In disability scales, unidimensionality implies that the definitions of successful performance of the various tasks express functional ability only, and not other properties. Subscale unidimensionality was examined by factor analysis, a mathematical method that uses row item scores to identify hidden factors that might generate variations in scoring. The absence of more than one hidden factor proves unidimensionality.

The Rasch model was used for comparison with the observed scores because it possesses characteristics of a valid and reliable scale. These include, besides unidimensionality, unique and independent informativeness of each item, applicability of item testing, compliance of patient total ability with ability to perform single tasks (internal consistency), clarity of task categories, discrimination between estimated levels of task (item) difficulty and between patients with different estimated functional levels (ability), hierarchical arrangement of tasks and categories, and similarity of scoring across subgroups (reliability). Scales that fit the model also possess all these characteristics and are therefore valid and reliable. A less than optimal fit may be due to a problem with any of these characteristics, and each of them then needs to be further examined.

The main variables of the Rasch model, namely, person ability and item difficulty, are estimated in units called 'logits'. Each logit is the natural logarithm of the odds of success (derived from the percentage of patients who complied with the requirements of each item category) in performing a step (transition from one item category to the next category of a higher level of difficulty) on the way to achieving independence. Tasks of average difficulty are assigned 0 logits. Tasks with above-average difficulty get a positive logit score, and tasks with below-average difficulty get a negative logit score. Person ability is defined as 0 logits when the individual patient has a 50% chance of succeeding in an item or a step of an average difficulty. A person with a logit score of 2.0 has a 0.5 probability of 'passing' an item with a difficulty level of 2.0 logits.

The goodness of fit of an item (task) is expressed by a mean square fit statistic (Table 1), or the weighted average of the squared item residuals, which are the differences between the score for the observed task and the corresponding expected score according to the Rasch model. Two types of fit statistics can be computed for each item. The 'outfit' statistic is sensitive to unexpected scorings of patients whose ability level is far from the item difficulty level. The 'infit' statistic, which is more informative, gives more weight to the residuals of patients whose ability level is close to the item difficulty level.⁷

Compliance with the model's characteristics can be demonstrated by a high, but not very high, probability of predicting success in a task by the success in other tasks of the same subscale. As the Rasch model is probabilistic and not deterministic, some failure of the model to predict the observed values is expected. When the discrepancy is optimal, the expected value of the mean square fit statistic is 1; a range of 0.8–1.2 for infit or of 0.6–1.4 for outfit is considered acceptable. A mean square value below 0.6 or 0.8, respectively, indicates significant 'overfit', that is, the item contributes no new information beyond that contributed by other items, or the patient's response to the task depends on his/her responses to other tasks of the scale. In both cases, the response is too predictable,

Table 1a Goodness of fit of SCIM II Self-Care tasks to Rasch Model

Task	Task difficulty	Difficulty SE	Emergent categories N	Disordered steps N	Item fit–Infit	(MSq) Outfit	Task-total r
1 Feeding	–4.00	0.17	4	1	1.20	1.70	0.64
2a Bathing – upper body	0.14	0.17	4	0	0.99	1.02	0.77
2b Bathing – low body	2.63	0.16	4	0	0.73	0.63	0.85
3a Dressing – upper body	0.32	0.14	4	0	0.92	0.85	0.84
3b Dressing – lower body	3.09	0.16	4	0	0.70	0.61	0.81
4 Grooming	–2.17	0.12	3	2	1.28	9.90	0.74

Table 1b Goodness of fit of SCIM II Respiration and Sphincter Management tasks to Rasch Model

Task	Task difficulty	Difficulty SE	Emergent categories N	Disordered steps N	Item fit–Infit	(MSq) Outfit	Task-total r
5 Respiration	–6.37	0.29	4	0	0.35	9.40	0.59
6 Bladder Management	1.89	0.09	2	2	0.94	0.71	0.72
7 Bowel Management	1.62	0.16	3	0	0.93	9.90	0.74
8 Use of Toilet	2.86	0.09	4	2	1.11	0.60	0.68

Table 1c Goodness of fit of SCIM II Room and Toilet Mobility tasks to Rasch Model

Task	Task difficulty	Difficulty SE	Emergent categories N	Disordered steps N	Item fit–Infit	(MSq) Outfit	Task-total r
9 Bed Mobility	–2.10	0.14	6	1	0.85	0.81	0.92
10 Bed–WChair Transfers	0.36	0.25	3	0	0.90	1.14	0.79
11 WChair–Toilet Transfers	1.74	0.30	3	0	1.07	1.28	0.72

Table 1d Goodness of fit of SCIM II Indoors and Outdoors Mobility tasks to Rasch Model

Task	Task difficulty	Difficulty SE	Emergent categories N	Disordered steps N	Item fit–Infit	(MSq) Outfit	Task-total r
12 Indoor Mobility	–1.59	0.14	6	2	0.63	0.53	0.81
13 Moderate Distance Mobility	–0.27	0.14	7	2	0.52	0.42	0.83
14 Outdoor Mobility	0.62	0.16	7	0	1.03	1.10	0.82
15 Stair Management	2.08	0.21	4	1	1.02	5.45	0.59
16 WChair–Car Transfers	–0.84	0.16	4	0	1.41	3.05	0.72

Table 1e Goodness of fit of SCIM II Sphincter Management (only) tasks to Rasch Model

Task	Task difficulty	Difficulty SE	Emergent categories N	Disordered steps N	Item fit–Infit	(MSq) Outfit	Task-total r
6 Bladder Management	–0.22	0.09	2	3	0.95	1.05	0.74
7 Bowel Management	–0.54	0.16	3	0	0.95	0.88	0.62
8 Use of Toilet	0.76	0.09	4	2	1.11	0.90	0.77

Task difficulty is given in logit units. SE=standard error, N=number, MSq=mean square, R=correlation coefficient

and hence, the item is probably redundant. A mean square value above 1.2 for infit or 1.4 for outfit indicates significant 'misfit', that is, noncompliance with the model's characteristics because of certain factors:

(1) To examine the possibility that misfit is due to an impaired relationship between a task and other components of the scale, correlations between task scores and the total subscale score are computed, omitting extreme scores (Table 1). A low or

negative correlation means that the task may not be related to the scale. A very high correlation indicates a presumably redundant task.

(2) To examine the possibility that misfit is due to difficulty in testing a task or to unclear phrasing of a category, the number of emergent categories and disordered steps, and the threshold locations between categories, are computed (Table 1, Figure 1). 'Emergent categories' are those actually and properly used by the scorers; 'disordered steps' are transitions to a non-adjacent or lower category

Figure 1 Item response probability curves. The probability of item response (scoring) is plotted against the degree of functional ability needed to perform a task, which is expressed as the difference between patient overall ability and the difficulty level of the task in logit scores. It is frequently evenly distributed, as in the case of bathing upper body, with threshold locations between categories ordered in the order of the category scores (a). However, for some tasks, examiners preferred scores of lower or higher categories and not the middle scores, thereby disrupting the order to threshold locations (b–d)

during scoring of a higher-level ability; the ‘threshold location’ is the logit value of the chance of success in a certain task in which the probability of choosing one of two adjacent task-categories is the same. Disordered steps correspond to disordered threshold locations. When the steps are ordered, the threshold locations between higher categories are higher than between lower categories. For example, the threshold location between categories scored 3 and 4 should be higher than the threshold between categories scored 2 and 3. This may not be true when categories are omitted or steps are disordered as a consequence of the scorers’ refraining from using certain task-categories or using them for patients who can comply with categories of a different difficulty level, either because they do not understand them or they find them difficult to test. The order of threshold locations between categories and the number of emergent categories (the categories with the highest

probability of response for some ability level), are assessed by plotting the probability of choosing each item category against the degree of functional ability needed to perform a certain task. The latter is expressed as the difference between overall patient ability and the difficulty level of the task.

- (3) To examine the possibility that misfit is due to an impairment in the discriminative ability of a scale or subscale between persons with different abilities and between items of different difficulties, the distribution of the variable estimates is checked along a logit scale (Figure 2).
- (4) To examine the hierarchical nature of the subscales, namely to detect their sensitivity in identifying different levels of task difficulty and patient ability (‘separation capacity’), statistically distinct levels of task difficulty or patient ability (‘strata’) are defined (Table 2). When a subscale has 3 or more task-difficulty or patient-ability strata, it is considered to have a good separation capacity.

Differences between strata may be due to true variations in person ability or item difficulty, or to incidental measurement errors. The extent of the true variations is expressed as the 'item separation index' (G) and the 'person separation index' (S), respectively, which is the ratio of the true spread of difficulty or of ability and its measurement error. The true item spread is expressed by the standard deviations of the measurements, corrected for the measurement errors; the measurement errors are derived from their squared standard errors.⁷ The higher the separation index, the more sensitive the subscale in identifying different levels of task difficulty or patient ability. G or S values higher than 2 correspond to more than 3 difficulty strata in a single subscale.

- (5) To demonstrate the capacity of the task-categories to discriminate between patient ability levels in each subscale, a histogram of the category thresholds is plotted (Figure 3). Each bar on the histogram shows the number of category-thresholds for each ability level ('threshold imprint') and reflects the number of transitions that can be made between the subscale categories by ability level. Each bar may also be viewed as a measure of the corresponding increase in the total subscale score for each change in ability level at a certain degree of ability. The distribution of the threshold imprint and the category density at certain ability levels represents the attention addressed by the category definitions to those levels.
- (6) To examine the possibility that misfit is due to differential task behavior by age, gender or examination subgroups, the item residuals, which reflect the distance of the observed scorings from those predicted by the Rasch model, are compared between paired subgroups (younger or older than 48 years, male or female, first or second examination). The means of the item residuals are compared by independent *t*-tests, controlling for the expected proportion of falsely rejected hypotheses (false discovery rate).⁸

Results

Unidimensionality of subscales based on factor analysis

Factor analysis showed that two of the SCIM II subscales, Self-Care and Respiration and Sphincter Management, have a single underlying construct. That is, one hidden factor was found to generate variations in their scoring, and this factor explained 72.56% and 59.97% of the response variance, respectively. However, for the Mobility subscale, two hidden factors generated variations in scoring, explaining 43.57% and 42.34% of the response variance, respectively. Most of the mobility tasks were associated with either one or the other of the extracted factors, and they could be assigned accordingly to two subscales. This created a subdivision similar to the original one in the SCIM II

questionnaire (room/toilet, indoors/outdoors). The only exception was wheel chair-car transfer, which was associated with both factors. After this subdivision, a single factor explained 89.58% of the response variance in the first Mobility subscale and 80.16% in the second.

Rasch analysis was then applied to each of the four subscales proved to be unidimensional by factor analysis.

Goodness of fit to the Rasch model

An acceptable goodness of fit between tasks within the subscales (infit 0.8–1.2, outfit 0.6–1.4) was found for most of the tasks in all subscales (Table 1). Analysis yielded an infit misfit for grooming and wheel chair-car transfer, and an infit overfit for bathing and dressing lower body, for respiration, and for indoor and moderate distance mobility. An outfit misfit was found for feeding, grooming, respiration, bowel management, stair management, and wheel chair-car transfer, and an outfit overfit for indoor mobility and moderate distance mobility. Removal of the respiration task from the Respiration and Sphincter Management subscale improved the goodness of fit of the rest of the subscale tasks (Table 1e).

Relationship of total patient and single-task performance ability

Most of the task scores showed moderate to high (but not very high) correlations with patient total scores ($r=0.64$ – 0.85). Respiration and stair management showed a lower correlation with the total score ($r=0.59$), and bed mobility, a rather high correlation ($r=0.92$) (Table 1).

Usability of categories and order of threshold locations between them

Although most of the task categories were properly used and classified as emergent categories, some caused disordered steps (Table 1). The probability of compliance with the various categories was mainly evenly distributed throughout the degrees of functional ability needed for success in task performance, with frequently ordered threshold locations (Figure 1a). However, for mobility for various distances (Appendix 1; 12, 13, 14), most of the categories with the odd scores were omitted. For several other tasks, examiners preferred scores of lower or higher categories and not the middle scores. This tendency was prominent in scoring feeding, grooming, bladder management, and use of the toilet (Figure 1b,c). As a consequence, the thresholds between categories in these items were disordered. These included the threshold between 1 and 2 for feeding; between 2 and 3 and 3 and 4 for grooming, between 8 and 12 and 12 and 15 for bladder management; and between 5 and 6 for bed mobility (Figure 1b–d).

Subscale discrimination of ability and difficulty and hierarchical nature

The degrees of the patient ability and task difficulty were mostly evenly distributed in the subscales Self-Care and Mobility (Table 1, Figure 2a–c). Feeding was found to be the easiest Self-Care task, and bed and indoor mobility the easiest Mobility tasks, with logit scores of -4.00 ($SE=0.17$), -2.10 ($SE=0.14$) and -1.59 ($SE=0.14$), respectively. The hardest were dressing, with a logit score of 3.09 ($SE=0.16$), and wheel chair–toilet transfer and climbing stairs, with logit scores 1.74 ($SE=0.30$) and 2.08 ($SE=0.21$), respectively. The similarity in the difficulty of dressing and bathing indicated suboptimal discrimination. The items of these subscales were distributed into 9–21.31

statistically distinct level of item difficulty (strata), and conformed to item separation indices of 6.50 – 15.73 , with reliability values of 0.98 – 1 . Patient ability had 4.09 – 4.72 strata, with 2.82 – 3.29 separation indices and reliability of 0.89 – 0.92 . In the Respiration and Sphincter Management subscale, respiration was found to be much easier than the other tasks (logit score -6.37 vs 1.62 – 2.86) (Figure 2d). When respiration was excluded from the analysis, the degrees of difficulty of the other tasks were almost evenly distributed, with a range of 0.54 – 0.76 . Bowel management was a little easier than bladder management, and both of these were easier than use of the toilet (Figure 2e). The sphincter management tasks were distributed into 6.39 difficulty strata, with an item separation index of 4.54 ,

Figure 2

(continued on page 402)

and a reliability of 0.95. However, the patient ability separation index and reliability were rather low, 1.23 and 0.60 respectively, with only 1.97 levels of ability (Table 2).

Discrimination ability of task categories: distribution of category thresholds along ability levels

In the Self-Care and Mobility subscales, the number of category thresholds was almost evenly distributed throughout the various levels of ability (Figure 3a–c). In the Respiration and Sphincter Management subscale, task categories were crowded mainly in the higher degrees of difficulty (Figure 3d). However, when respiration was excluded from the analysis, the distribution of the category thresholds became almost even, albeit through a narrow range of ability levels (Figure 3e).

Differential task behavior by age, gender and examination subgroups

A significant difference in fit to the Rasch model between the first to the second examination was found

for feeding, bathing lower body, and bowel management ($P < 0.05$). No significant differences were found between the examinations for the other tasks or between age and gender subgroups during the first examination.

Discussion

For a disability scale to be valid, its scoring should be affected only by the patient's ability and by the difficulty of the tested tasks, and not by any other property. The scale's items should be arranged in a hierarchical order, so easier tasks are scored higher and more disabled patients are scored lower. For the scale to be reliable, similar scores should be elicited in various relevant patient groups and settings. To achieve validity and reliability, proper conditions for observing the performance should be available, and the definitions of the scoring criteria should be clearly understood by the scorers.

In the present study of the SCIM II, the unidimensionality of the subscales was proven by

Figure 2 Task difficulty calibration in subscales. Distribution of the person ability and item difficulty along a logit scale. (a) Self care. F=Feeding, G=grooming, BU=Bathing upper body, DU=Dressing upper body, BL=Bathing lower body, DL=Dressing lower body. (b) Room and toilet mobility. B=Bed mobility, BW=Bed–Wheelchair transfers, WT=Wheelchair–toilet transfers. (c) Indoors and outdoors mobility. I=Indoor mobility, W=Wheelchair–car transfers, M=mobility to moderate distances, O=Outdoor mobility, S=Stair management. (d) Respiration and sphincter management. R=Respiration, BO=Bowel management, BL=Bladder management, UT=Use of toilet. (e) Sphincter management. BO=Bowel management, BL=Bladder management, UT=Use of toilet

Table 2 Separation between degrees of item difficulty and patient ability

Subscale	Item			Patient		
	G	Strata (4G + 1)/3	Reliability $G^2/(1 + G^2)$	S	Strata (4S + 1)/3	Reliability $S^2/(1 + S^2)$
Self-Care	15.73	21.31	1.00	3.21	4.61	0.91
Respiration and Sphincter Management	20.92	28.23	1.00	1.74	2.65	0.75
Mobility (room and toilet)	6.50	9.00	0.98	2.82	4.09	0.89
Mobility (indoors and outdoors)	7.23	9.97	0.98	3.29	4.72	0.92
Sphincter Management only	4.54	6.39	0.95	1.23	1.97	0.60

G = item separation index, S = person separation index

factor analysis and validity and reliability were confirmed by tests for goodness of fit to the Rasch model. The goodness of fit indicated that most of the scale's items are uniquely and independently informative, and that for most of the items, the score for patient total ability and ability to perform single tasks are satisfactorily correlated. It also indicated that item testing is usually applicable; the task categories are clearly phrased; the subscales can reasonably discriminate between estimated levels of task difficulty and between patients with different estimated functional levels; the tasks and categories are in hierarchical order; and the scoring across subgroups is similar.

Further analyses generally supported these findings. The many emergent categories, the few disordered steps, and the frequently ordered threshold locations all confirmed the clarity of phrasing and the applicability of task testing.

The even distribution of the patients' abilities and most of the task difficulties, and the almost even distribution of the task categories for the abilities and difficulties, indicated that the scale and its components address similar attention to the various conditions that affect functional performance of patients with a SCL. Analysis of the discrimination ability of the subscales showed that the categories, which define the requirements for performing the tasks, are generally arranged according to degree of difficulty. The values of the separation indices confirmed the hierarchical nature of the scale. Analysis of the distribution of category thresholds along ability levels indicated that the category definitions for the most part equally addressed the various degrees of difficulty.

From all these factors we can infer the relatedness of individual ability and item difficulty for any SCL patient or functional condition. In the present series, they indicated, to a large extent, that the relationship in our particular patient group and items is likely to be true for the same items with other SCL patients, and for the same patients with other items of a similar type.

The similarity of the SCIM II task scores in most of the comparisons of age, gender and examination subgroups verified the overall reliability of the scale.

According to these findings, the SCIM II components are capable of reliably identifying the functional ability of patients with a SCL with any disability and

challenged by any difficulty of daily tasks, without risk of bias by properties unrelated to the functional level.

In spite of the generally favorable results, the analysis yielded some drawbacks of the scale. The tasks dressing and bathing, and the tasks indoor and moderate distance mobility, showed a slight overfit, or overlap of difficulty degree, which may indicate redundancy of the items. The task bed mobility was highly correlated with its total subscale's score, which may also indicate task redundancy. Bathing also showed a significant difference in scoring between the first and second examinations, pointing to possible item unreliability. Removing bathing from the scale would probably save scorers time and effort without affecting the patient's functional evaluation. This should be carefully considered, while keeping in mind that some patients do well with dressing but not with bathing, and that problems in performing an important daily task may be overlooked in those specific patients. Similar consideration should be given to moderate distance mobility.

Wheelchair-car transfer showed misfit and multidimensionality, indicating that its score was influenced by properties other than patient ability. This was probably due to the difficulty in assessing this task by observation because of the lack of a car inside the hospital. Respiration also showed misfit, with a low correlation with its subscale's total score. In addition, respiration was found to be of much lesser difficulty than the other tasks in its subscale. We suggest that respiration does not belong to the same area of function as these other tasks. They were assigned to the same subscale for convenience, as they were scored by the same members of the multidisciplinary team that did the SCIM assessment. Stair management also showed some misfit and a low correlation with its subscale's total score.

The tasks feeding and grooming showed minor misfit. Scorers either refrained from using their middle categories or used them incompatibly with the probability of success in the tasks (ie, threshold locations between the scores were disordered). Feeding also showed a significant difference in scoring between the first and second examinations, indicating a reliability problem. Bowel management showed a minor misfit and a problem of reliability, also

Figure 3 Threshold imprint for self-care subscale. The thresholds for transit between categories of compliance with the subscale tasks, and their density, are almost evenly distributed throughout the difficulty degrees of the Self-Care (a) and Mobility (b,c) tasks, and the tasks of Sphincter Management and Use of the Toilet (e). When the tasks of Respiration and Sphincter Management are analyzed together, categories are crowded mainly in the higher degrees of difficulty (d)

demonstrated in a previous study.³ The tasks mobility for various distance, bladder management, use of the toilet, and bed mobility showed a good fit to the Rasch model, but the examiners omitted most of the categories with the odd scores, or preferred lower or higher scored categories, disordering the threshold

locations between them. Finally, the sphincter management and use of toilet tasks showed low capacity of separation between levels of patient ability.

In all the items with problems of reliability or separation ability, or in which examiners avoided using certain task categories or scored some im-

Table 3 Main suggested changes in SCIM II for the formulation of the new version

Task	Suggested change
1. Feeding	Rephrasing, to refer to pouring and drinking as well; eliminating a category, thereby reducing the maximum score for the item, which was found to be the easiest task
2. Bathing	No change (the similarity of difficulty with dressing was addressed by rephrasing the last item)
3. Dressing	Rephrasing, to refer also to clothes with buttons and zippers; adding two categories, thereby increasing the maximum score of the item, which was found to be the most difficult task
4. Grooming	Rephrasing, to combine two categories, thereby reducing the maximum score of the item, which was found to be easy
5. Respiration	Rephrasing, to refer also to the possible need of oxygen, assisted coughing, or intermittent assisted ventilation
6. Bladder management	Rephrasing, to refer to combinations of intermittent catheterizations with or without need for assistance for applying a drainage instrument
7. Bowel management	Minor rephrasing, to drop 'proper timing', which may be interpreted variably
8. Use of toilet	Rephrasing, to combine the three categories dealing with partial assistance into two, without reducing the maximum score of the item
9. Bed mobility and sore prevention	Rephrasing, to score by summing equally weighted activities of this task, and combine categories without reducing the maximum score of the item
10. Bed-wheelchair transfer	Minor rephrasing, to refer to the use of adaptive devices and to the possibility that the patient does not require a wheelchair
11. Toilet-tub-wheelchair transfer	Minor rephrasing, to refer to the use of adaptive devices and to the possibility that the patient does not require a wheelchair
12. Indoors mobility	Adding an indication that the item refers to mobility on even surfaces
13. Moderate distance mobility	Adding an indication that the item refers to mobility on even surfaces
14. Outdoors mobility	Adding an indication that the item refers to mobility on even surfaces
15. Stair management	No change
16. Wheelchair-car transfer	Minor rephrasing, to refer to the possibility that the patient does not require a wheelchair

properly, rephrasing or removal of the categories had to be considered. Therefore, a third version of SCIM has been formulated, with changes based on the findings of this study and on comments of our staff and of colleagues in several countries. The main changes suggested for the new version are detailed in Table 3. This version will be subjected to an international validation, which will contribute to further refinement and improvement of the scale and resolve potential cross-cultural differences that may affect the utility of the scale outside Israel.

In conclusion, the findings of this study, which is based on analysis with the Rasch model using log transformation to convert ordinal scores into interval measures, confirm the validity and reliability of the Catz-Itzkovich SCIM. Until the future international version is validated, SCIM II is suitable for routine clinical use despite its flaws. The current version is sensitive to functional changes of SCL patients,⁴ rates functional achievements according to their importance for these patients, includes ADL functions relevant to SCL patients, is user friendly, and is superior to the original SCIM.³ Serial SCIM II examinations can adequately assess the SCL patient's primary ADL capabilities and changes over time. By focusing the rehabilitation efforts on the proper tasks, it can contribute to better care, improved outcomes, and greater client satisfaction.

References

- 1 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.
- 2 Halbertsma J *et al*. Towards a new ICIDH, *Disabil Rehabil* 2000; **22**: 144–156.
- 3 Catz A *et al*. The Catz-Itzkovich SCIM: a revised version of the Spinal Cord Independence Measure. *Disabil Rehabil* 2001; **23**: 263–268.
- 4 Catz A *et al*. The Second Version of the SCIM – A Disability Scale for Patients with Spinal Cord Lesions. Sensitivity to functional changes. Loewenstein Rehabilitation Hospital, Raanana, Ministry of Defence, Tel Aviv, 2000.
- 5 Hinderer SR, Hinderer KA. Quantitative methods of evaluation. In: DeLisa JA (ed). *Rehabilitation Medicine: Principles and Practice*. Philadelphia: JB Lippincott Company 1993.
- 6 Uniform Data Set for Medical Rehabilitation, Version 3.0. Data Management Service of the Uniform Data System for the Medical Rehabilitation and the Center for Functional Assessment Research. SUNY, Buffalo, New York, 1990.
- 7 Wright BD, Masters GN. *Rating Scale Analysis*. Chicago: Mesa Press, 1982.
- 8 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J R Stat Soc* 1995; **B57**: 289–300.

Appendix 1

שרותי בריאות
כללית

Patient Name: _____ ID: _____

Examiner Name: _____

Department IV

Medical Director: Dr. Amiram Catz

**LOEWENSTEIN HOSPITAL
REHABILITATION CENTER**

Affiliated with the Sackler Faculty
of Medicine, Tel-Aviv University

Tel: 972-9-7709090 Fax: 972-9-7709986

e-mail: amiramc@clalit.org.il

(The score for each function should be placed in the adjacent square, below the date)

SCIM-SPINAL CORD INDEPENDENCE MEASURE
Version II
Self-Care

DATE

--	--	--	--	--	--	--	--

1. Feeding (cutting, opening containers, bringing food to mouth, holding cup with fluid)

0. Needs parenteral, gastrostomy or fully assisted oral feeding

1. Eats cut food using several adaptive devices for hand and dishes; unable to hold cup

2. Eats cut food using only one adaptive device for hand; holds an adapted cup

3. Eats cut food without devices; holds a regular cup; needs assistance to open containers

4. Independent in all tasks without any adaptive device

2. Bathing (soaping, manipulating water tap, washing). **A-upper body; B-lower body**

A. 0. Requires total assistance

1. Requires partial assistance

2. Washes independently with adaptive devices or in a specific setting

3. Washes independently; does not require adaptive devices or a specific setting

B. 0. Requires total assistance

1. Requires partial assistance

2. Washes independently with adaptive devices or in a specific setting

3. Washes independently; does not require adaptive devices or a specific setting

3. Dressing (preparing clothes, dressing, undressing). **A-upper body; B-lower body**

A. 0. Requires total assistance

1. Requires partial assistance

2. Dresses independently with adaptive devices or in a specific setting

3. Dresses independently; does not require adaptive devices or a specific setting

B. 0. Requires total assistance

1. Requires partial assistance

2. Dresses independently with adaptive devices or in a specific setting

3. Dresses independently; does not require adaptive devices or a specific setting

4. Grooming (washing hands and face, brushing teeth, combing hair, shaving, applying makeup)

0. Requires total assistance

1. Performs only one task (e.g., washing hands and face)

2. Performs some tasks using adaptive devices; needs help to put on/take off devices

3. Independent with adaptive devices

4. Independent without adaptive devices

SUBTOTAL (0-20)

--	--	--	--	--	--	--	--

Respiration and Sphincter Management
5. Respiration

0. Requires assisted ventilation

2. Requires a tracheal tube and partially assisted ventilation

4. Breathes independently but requires much assistance in tracheal tube management

6. Breathes independently and requires little assistance in tracheal tube management

8. Breathes without a tracheal tube, but sometimes requires mechanical assistance for breathing

10. Breathes independently without any device

6. Sphincter Management - Bladder

0. Indwelling catheter

4. Residual urine volume > 100cc; no catheterization or assisted intermittent catheterization

8. Residual urine volume < 100cc; needs assistance for applying drainage instrument

12. Intermittent self-catheterization

15. Residual urine volume < 100cc; no catheterization or assistance in urine drainage required

7. Sphincter Management - Bowel

0. Improper or irregular timing or very low frequency (less than once in 3 days) of bowel movements

5. Proper and regular timing, but requires assistance (e.g., for applying suppository); rare accidents (less than once a month)

10. Regular bowel movements, with proper timing, without assistance; rare accidents (less than once a month)

8. Use of Toilet (perineal hygiene, clothes adjustment before/after, use of napkins or diapers)

0. Requires total assistance

1. Partially undresses lower body, needs assistance in all remaining tasks

2. Partially undresses lower body and partially cleans self (after); needs assistance in adjusting clothes and/or diapers

3. Undresses and cleans self (after); needs assistance in adjusting clothes and/or diapers

4. Independent in all tasks but needs adaptive devices or special setting (e.g., bars)

5. Independent without adaptive devices or special setting

SUBTOTAL (0-40)

--	--	--	--	--	--	--	--

Mobility (room and toilet)

DATE

--	--	--	--	--	--	--	--

9. Mobility in Bed and Action to Prevent Pressure Sores

0. Requires total assistance
1. Turns in bed to one side only
2. Turns in bed to both sides but does not fully release pressure
3. Releases pressure when lying only
4. Turns in bed and sits up without assistance
5. Independent in bed mobility; performs push-ups in sitting position without full body elevation
6. Independent in bed mobility; performs push-ups in sitting position with full body elevation

--	--	--	--	--	--	--	--

10. Transfers: bed-wheelchair (locking wheelchair, lifting footrests, removing and adjusting arm rests, transferring, lifting feet)

0. Requires total assistance
1. Needs partial assistance and/or supervision
2. Independent

--	--	--	--	--	--	--	--

11. Transfers: wheelchair-toilet-tub (if uses toilet wheelchair - transfers to and from; if uses regular wheelchair - locking wheelchair, lifting footrests, removing and adjusting armrests, transferring, lifting feet)

0. Requires total assistance
1. Needs partial assistance and/or supervision, or adaptive device (e.g., grab-bars)
2. Independent

Mobility (indoors and outdoors)

12. Mobility Indoors

0. Requires total assistance
1. Needs electric wheelchair or partial assistance to operate manual wheelchair
2. Moves independently in manual wheelchair
3. Requires supervision while walking (with or without devices)
4. Walks with a walking frame or crutches (swing)
5. Walks with crutches or two canes (reciprocal walking)
6. Walks with one cane
7. Needs leg orthosis only
8. Walks without aids

--	--	--	--	--	--	--	--

13. Mobility for Moderate Distances (10 - 100 meters)

0. Requires total assistance
1. Needs electric wheelchair or partial assistance to operate manual wheelchair
2. Moves independently in manual wheelchair
3. Requires supervision while walking (with or without devices)
4. Walks with a walking frame or crutches (swing)
5. Walks with crutches or two canes (reciprocal walking)
6. Walks with one cane
7. Needs leg orthosis only
8. Walks without aids

--	--	--	--	--	--	--	--

14. Mobility Outdoors (more than 100 meters)

0. Requires total assistance
1. Needs electric wheelchair or partial assistance to operate manual wheelchair
2. Moves independently in manual wheelchair
3. Requires supervision while walking (with or without devices)
4. Walks with a walking frame or crutches (swing)
5. Walks with crutches or two canes (reciprocal walking)
6. Walks with one cane
7. Needs leg orthosis only
8. Walks without aids

--	--	--	--	--	--	--	--

15. Stair Management

0. Unable to climb or descend stairs
1. Climbs and descends at least 3 steps with support or supervision of another person
2. Climbs and descends at least 3 steps with support of handrail and/or crutch or cane
3. Climbs and descends at least 3 steps without any support or supervision

--	--	--	--	--	--	--	--

16. Transfers: wheelchair-car (approaching car, locking wheelchair, removing arm- and footrests, transferring to and from car, bringing wheelchair into and out of car)

0. Requires total assistance
1. Needs partial assistance and/or supervision
2. Independent with adaptive devices
3. Independent without adaptive devices

--	--	--	--	--	--	--	--

SUBTOTAL (0-40)

--	--	--	--	--	--	--	--

TOTAL SCIM SCORE (0-100)

--	--	--	--	--	--	--	--