ISCoS
The International
Spinal Cord Society

ARTICLE

Reliability of the International Spinal Cord Injury Upper Extremity Basic Data Set

S Tongprasert[1] · A Namchandee[1] · K Sotthipoka[1] · P Kammuang-lue[1]

Abstract

Study design It is a psychometrics study.

Objective To assess the inter-rater reliability of the International Spinal Cord Injury Upper Extremity Basic Data Set (ISCI-UE).

Setting Maharaj Nakorn Chiang Mai Hospital, Chiang Mai, Thailand.

Methods Individuals with subacute and chronic cervical spinal cord injury (SCI) were recruited. One examiner rated five parts of the ISCI-UE, including the ability to reach and grasp, the shoulder function classification, utilization of adaptive devices used to enhance upper-extremity function, complications affecting upper-extremity function, and upper extremity/hand reconstructive surgery. A second blinded examiner repeated the procedures within 1 day. Quadratic weighted kappa was calculated to determine the inter-rater reliability.

Results Sixty participants were included in the study. Fifty-two patients were men, and the mean (SD) age of participants was 42.9 (14.3) years. The median (interquartile range) time since injury was 9.5 (1–53) months. A total of 117 upper limbs were assessed. The inter-rater reliability was substantial, with almost perfect agreement in all items (ability to reach and grasp = 0.98; shoulder function classification = 0.97; use of assistive devices = 0.89; complications = 0.74; and surgery = 1).

Conclusion The International Spinal Cord Injury Upper Extremity Basic Data Set (ISCI-UE) has very good inter-rater reliability for evaluating individuals with cervical SCI.

Introduction

The International Spinal Cord Injury (SCI) Upper Extremity Basic Data Set (ISCI-UE) [1] is one of several international SCI data sets developed to promote standardized data collection for SCI patients, which is aimed at improving SCI examination, treatment, and rehabilitation and comparison among SCI centers for research. Upper-extremity function is critical for individuals with tetraplegia. Based on previous studies [2–4], evidence suggests that regaining even partial function can lead to greater independence and enhancement in the quality of life (QoL). A survey of the needs of individuals with SCI indicates that improvement of hand

function was the highest priority for people with tetraplegia compared to other needs in improving their QoL [3]. In this respect, adequate therapy of the upper extremities is vital for individuals with tetraplegia, and relevant for the improvement of upper-extremity function.

Upper-extremity function can be assessed through two main methods: basic manual muscle testing in accordance with the International Standards for Neurological Classification of SCI (ISNCSCI) [5] and using novel assessment tools developed to evaluate the efficacy of therapies and surgery for restoring hand function. These novel assessments focus on primary hand functions (i.e., lateral pinch, pulp pinch, and palmar grasp) in individuals with tetraplegia and other patient populations. These instruments can be classified as assessment area according to International Classification of Functioning, Disability, and Health (ICF) domain. The tests that assess body function and structure domain are six-minute arm test (6-MAT), graded refined assessment of strength, sensibility, and prehension (GRASSP), and handheld myometer. The activity domain

✉ P Kammuang-lue
   pratchayapon.k@cmu.ac.th

1   Department of Rehabilitation Medicine, Faculty of Medicine,
     Chiang Mai University, Chiang Mai, Thailand

SPRINGER NATURE

should be assessed by box and block test (BBT), capabilities of upper-extremity instrument (CUE), Jebsen hand function test (JHFT), modified functional reach test (mFRT), Sollerman hand function test (SHFT), tetraplegia hand activity questionnaire (THAQ), and Van Lieshout test-short version (VLT-SV). The information regarding how to select these measurements can be assessed on the following websites: Academy of Neurologic Physical Therapy [6], National Institute of Neurological Disorders and Stroke (NINDS) common data elements [7], and Spinal Cord Injury Research Evidence (SCIRE) [8].

The ISCI-UE was developed for standardized data collection and reporting of a minimal amount of information about upper-extremity status in individuals with SCI. However, it is important to emphasize that the basic data sets are not intended to be used as measures [9]. The data sets represent the result of important methods used to systematically record highly relevant clinical information at the point of care, and the collection and recording must be consistent across practitioners in order to pool data or compare outcomes of care across facilities or programs.

Fin Biering-Sørensen and an international working group developed the ISCI-UE and adjusted for appropriateness according to the comment from the members of the Executive Committee of the International SCI Standards and Data Sets and the members of the International Spinal Cord Society (ISCoS) Executive and Scientific Committees and American Spinal Injury Association (ASIA) board members. Following this, the international organizations and societies (around 40) and individuals were invited to review the data set. The data set was posted on the ISCoS and ASIA websites to allow for further comments and suggestions. The data set was then reviewed again and approved by the members of the ISCoS Executive and Scientific Committees and ASIA board members. The data set was then used to collect information on the training cases to confirm its appropriateness for clinical use [1, 10, 11]. The objective of the data set is to establish terminology that is commonly understood among health-care providers for evaluating the hand and upper-extremity function of individuals with SCI. The following minimal data elements were required for the evaluation of upper-extremity function of individuals with SCI: date performed, the ability to reach and grasp, shoulder function classification, and use of assistive devices. In addition, SCI-related complications affecting upper-extremity function like pain, spasms, contractures and edema, and upper extremity/hand reconstructive surgery were evaluated.

Using the ISCI-UE in a clinical setting requires additional work for data set evaluation, according to the recommendations of the Executive Committee of the International SCI Standards and Data Sets [9]. For all international data sets, reliability is a crucial element that allows for a meaningful comparison of data from different resources. The aim of this study was to determine the inter-rater reliability of the ISCI-UE data set.

Methods

Participants

Individuals with tetraplegia who received treatment at Maharaj Nakorn Chiang Mai Hospital, Chiang Mai, Thailand, from July 2015 to July 2016 were enrolled in the study. The inclusion criteria were as follows: (1) age—18–65 years and (2) neurological level—T1 or above. The exclusion criteria were as follows: (1) concomitant upper-extremity injury at the time of injury (e.g., brachial plexus injury, fracture, and amputation), (2) diseases or injuries of upper extremities prior to the onset of SCI, and (3) history of severe traumatic brain injury with inconsistent orientation to person, time, and place (Rancho Los Amigos scale <7). Individuals who agreed to participate in the study provided informed consent.

Measures

The International SCI Core Data Set was used to record descriptive characteristics, which included the following: age, gender, time post injury, cause of injury, and neurologic status [12]. Neurological examinations were performed by a physician trained in the use of International Standards for Neurological Classification (ISNCSCI). The core data set and ISNCSCI were collected by the same physician. The ISCI-UE for each patient was completed by two additional raters.

The ISCI-UE was used to assess the hand and upper-extremity function in the SCI population. The data set was composed of five variables: ability to reach and grasp, shoulder function classification, use of assistive devices, SCI-related complications affecting upper-extremity function, and upper extremity/hand reconstructive surgery. The ability to reach and grasp was classified into five levels (1–5). Shoulder function was classified into four levels (A–D). Scores from the ability to reach and grasp (numerical values from 1 to 5) were coupled with scores from shoulder function classification (letter values from A to D), and this defined the entire upper-extremity function [1, 10]. There were 20 patterns of the score (1–5 plus A–D). For example, a patient with central cord syndrome had right/left ability to reach and grasp scores 3/3 and right/left shoulder function scores A/B, then this patient had basic right hand upper-extremity function 3A and basic left hand upper-extremity function 3B. The third section (use of assistive devices) includes devices used to enhance upper-extremity

function. Complications affecting upper-extremity function are classified as minimal, moderate, or extensive. Finally, upper extremity/hand reconstructive surgery is included if it was specifically performed to restore arm and/or hand function. These last three domains have 4, 3, and 3 response items [1, 10]. The data sheet and the syllabus are available on the ISCoS website (www.iscos.org.uk).

Raters

The raters included a physiatrist experienced in SCI and a physician training in rehabilitation medicine. After reviewing the ISCI-UE data sheet and syllabus, the raters were standardized by scoring the training cases of the ISCI-UE data set [13]. Then, a discussion was conducted between the four investigators (two experienced SCI physiatrists and two physicians). Any difference of scoring and comments in each variable were discussed. Both raters' scores were nearly the same; therefore, this result ensures the consistency of raters in data collection. Raters performed separate assessments of the same SCI participant. The time interval between repeat ratings is less than 1 day in order to prevent changes to the attributes under examination. Also, raters were blinded against each other's ratings.

Statistical analysis

The descriptive statistics, including frequency, percentage, mean, standard deviation, median, and interquartile range, were used to present the following demographic data: age, gender, time post injury, severity of SCI, causes of injury, level of ability to reach and grasp, and level of shoulder function.

The analytical methods utilized in this study were modeled from reliability studies on the musculoskeletal and bowel function data sets, and included percentage agreement and Cohen's kappa for assessing reliability [9, 14–17]. Quadratic weighted kappa (κ) was used to determine reliability because it provided an estimation of agreement corrected for chance [17]. The κ values were interpreted based on Landis and Koch, where a score $\leq 0 =$ poor, $0.01–0.20 =$ slight, $0.21–0.40 =$ fair, $0.41–0.60 =$ moderate, $0.61–0.80 =$ substantial, and $0.81–1.0 =$ almost perfect agreement [16, 17]. The authors gave a positive rating for reliability when $\kappa \geq 0.70$ in a sample size of at least 50 patients [9, 18]. Percentage agreements with an a priori minimum of 90% for variables with two categories and 80% for variables with three or more categories were considered satisfactory [9]. Percentage agreement values were used if a high prevalence existed in some variables and none in the extreme variables. For instance, kappa may inaccurately reflect agreement. It may be more appropriate to consider the percentage of agreement to supplement the kappa coefficients such as sub-

analysis of the ability to reach and grasp, shoulder function classification, and whole upper-extremity function. Data collection and data management were conducted by another researcher. Statistical analysis was conducted using SPSS 22.0 for Windows.

Results

Demographics

The two raters separately classified 60 individuals with SCI from traumatic and non-traumatic causes. The characteristics of participants are listed in Table 1. A total of 52 (87%) individuals were male, with a mean (SD) age of 42.9 (14.3) years, and median (interquartile range) time since injury of 9.5 (1–53) months. The two most common causes of injury were transport-related activities (55%) and falls (33%). The distribution on the ASIA impairment scale (AIS) was as follows: A, 22%; B, 12%; C, 18%; D, 45%; and E, 3%. One-third of the 60 injuries (34%) had a neurological level of C4.

Of the 120 upper extremities assessed, three samples were excluded due to concomitant upper-extremity injury at the time of injury as follows: elbow dislocation, anterior interosseous nerve injury, and amputation. Thus, a total of 117 upper extremities were analyzed. No sampled patients underwent upper-extremity surgery for the improvement of arm and/or hand function.

Inter-rater agreement

Ability to reach and grasp

The frequency of the level of hand–arm function (ability to reach and grasp) rate, as assessed by the two raters, is presented in Table 2. The percentage of agreement in the ability to reach and grasp was 91.5, and the weighted kappa coefficient was almost in perfect agreement at 0.98 (95% CI, 0.96–0.99).

The ability to reach and grasp was also analyzed by category by the AIS grade and neurological level. The percentage agreement was acceptable in all AIS subgroups as follows: A, 85%; B, 93%; C, 82%; D, 90%; and E, 100%. Likewise, a positive percentage agreement existed when categories with neurological level were high (C1–C4), 89%; mid (C5–C6), 92%; and low (C7–T1), 94% cervical levels.

Shoulder function classification

The frequency of the level of shoulder function rate, as assessed by the two raters, is presented in Table 3. The

Table 1 Characteristics of the study group

Age, mean (SD) (years)	42.93 ± 14.3
Gender, N (%)	
Male	52 (87)
Female	8 (13)
Time since injury, median (range) (months)	9.5 (1–53)
Neurological level, N (%)	
Cervical (C2–C3)	5 (8)
Cervical (C4)	20 (34)
Cervical (C5)	14 (23)
Cervical (C6)	12 (20)
Cervical (C7)	5 (8)
Cervical (C8)	3 (5)
Thoracic (T1)	1 (2)
American Spinal Injury Association Impairment Scale (AIS), N (%)	
AIS-A	11 (18)
AIS-B	7 (12)
AIS-C	11 (18)
AIS-D	27 (45)
AIS-E	2 (3)
Cause of injury, N (%)	
Sport	1 (2)
Assault	1 (2)
Transport-related activities	33 (55)
Falls	20 (33)
Others	5 (8)

Table 2 The frequency of the ability to reach and grasp scores from rater one and two

Ability to reach and grasp	Rater 2					
	1	2	3	4	5	Total
Rater 1						
1	**10**	0	0	0	0	10
2	1	**13**	0	0	0	14
3	0	2	**31**	2	0	35
4	0	0	1	**10**	3	14
5	0	0	0	1	**43**	44
Total	11	15	32	13	46	117

Bold numbers represent the frequency in which two raters assign exactly the same rating

prevalence of no complications or complications having only minimal impact on function.

Upper extremity/hand reconstructive surgery

Since no patients underwent upper-limb reconstructive surgery, the percentage agreement was 100%.

Discussion

To the authors' knowledge, the present study represents the first study evaluating inter-rater reliability of the ISCI-UE. The results provide strong evidence to support the ISCI-UE having satisfactory inter-rater reliability for the ability to reach and grasp as well as shoulder function. The percentage agreement for whole upper-extremity function was also acceptable (>80%). Inter-rater reliability for the use of assistive devices and complications affecting upper-extremity function exhibited almost perfect agreement and substantial agreement, respectively.

A particular strength of this study was the large number of cases included, which was due to the tertiary care center serving a large proportion of Northern Thailand. A comparison of the raters, each having different degrees of SCI experience (resident and SCI attending), was also carried out. Interpretation of the results of this study did not yield any significant difference in reliability between the raters, indicating that the data set was easy to understand and evaluate. When introducing a new measurement in clinical practice, the minimum comprehensive validation for inter-rater reliability should be investigated. In order to test the validity of this data set, content validity testing was conducted during data set development. However, the construct validity of the ability to reach and grasp, as well as shoulder function classification in ISCI-UE has not been studied. Although the ability to reach and grasp, which was adopted

percentage of agreement in the level of shoulder function was 93.2, and the weighted kappa was almost in perfect agreement at 0.97 (95% CI, 0.94–0.99). After performing sub-analysis, the level of agreement between raters was not different. The percentage agreements of AIS subgroups were A, 81%; B, 86%; C, 91%; D, 96%; and E, 100%. The percentage agreements, as classified by neurological level, were as follows: high (C1–C4), 94%; mid (C5–C6), 94%; and low (C7–T1), 89% cervical levels.

For whole upper-extremity function, a percentage agreement within one point was acceptable (84%).

Use of assistive devices

The percentage agreement in the use of assistive devices was acceptable (93%) and the weighted kappa was very good 0.89 (95% CI, 0.77–1.00).

Complications affecting upper-extremity function (e.g., pain, spasm, contractures, and edema)

SCI-related complications also exhibited an acceptable percentage agreement (89%). Weighted kappa was substantial at 0.74 (95% CI, 0.52–0.96). Also, there was a high

Table 3 The frequency of shoulder function scores from rater one and two

The shoulder function classification	Rater 2				
	A	B	C	D	Total
Rater 1					
A	**15**	0	0	0	15
B	1	**6**	0	0	7
C	0	2	**33**	5	40
D	0	0	0	**55**	55
Total	16	8	33	60	117

Bold numbers represent the frequency in which two raters assign exactly the same rating

from RULER SCI (RITZ upper-limb evidence-based rehabilitation in SCI), had been benchmarked with a spinal cord independence measure (SCIM), the correlation was not defined [19, 20].

The inter-rater reliability of this study was acceptable and correlated with the inter-rater reliability tests of other SCI basic data sets. For example, 9 out of 12 items on the International SCI Musculoskeletal Basic Data Set scored satisfactory for inter-rater reliability ($\kappa = 0.62–1.0$ and percentage agreement = 75–100%) [14]. Regarding the agreement of the International SCI Bowel Function Basic and Extended Data Sets, 47 of 52 items were confirmed between versions from both Italy and Denmark ($\kappa > 0.2$) [15], while the International SCI Spinal Column Injury Basic Data Set ($\kappa = 0.79–0.89$) [21] and the Korea version of the International SCI Basic Pain Data Set ($\kappa = 0.693–0.780$ and intraclass correlation = 0.731–0.858) were also in agreement with the version from these nations [22].

The benefits of the ISCI-UE include no requirement of equipment, no copyright, and no fee for its use; therefore, it is suitable for clinical use. However, this data set may not be sensitive enough for minimal clinical changes due to the limited data requirement for hand function evaluation.

Some issues were experienced during the analysis of this data set. First, the pathology was not in the order of the classification; for example, in the case of a tetraplegic patient who had finger flexor function without wrist extensor function. Such an individual should be scored for the ability to reach and grasp group 4, as the definition of this group is voluntary control of wrist <u>and</u> some extrinsic hand muscles. Another issue was encountered while evaluating chronic SCI patients with severe spasticity or joint contracture, one rater scored less than the true level. However, this data set also involved complications affecting upper-extremity function, therefore, assessors should address during this section.

One limitation of this study is that it did not involve evaluation of intra-rater reliability. However, if the inter-rater reliability is strong, intra-rater reliability is also likely to be strong [9]. The construct validity of the ability to reach and grasp, as well as shoulder function classification has not been tested; therefore, it should be studied in the future. Future studies could also involve an investigation of the inter-rater reliability between other SCI health-care providers, such as physical therapists and occupational therapists, to assess and determine their reliability.

The present study concludes that the inter-rater reliability of the International Spinal Cord Injury Upper Extremity Basic Data Set is very high. However, further studies must be conducted to assess the construct validity of the data set before it is applied in research and clinical use.

Compliance with ethical standards

References

1. Biering-Sorensen F, Bryden A, Curt A, Friden J, Harvey LA, Mulcahey MJ, et al. International spinal cord injury upper extremity basic data set. Spinal Cord. 2014;52:652–7.
2. Anderson KD. Targeting recovery: priorities of the spinal cord-injured population. J Neurotrauma. 2004;21:1371–83.
3. Snoek GJ, IJzerman MJ, Hermens HJ, Maxwell D, Biering-Sorensen F. Survey of the needs of patients with spinal cord injury: impact and priority for improvement in hand function in tetraplegics. Spinal Cord. 2004;42:526–32.
4. Snoek GJ, IJzerman MJ, Post MW, Stiggelbout AM, Roach MJ, Zilvold G. Choice-based evaluation for the improvement of upper-extremity function compared with other impairments in tetraplegia. Arch Phys Med Rehabil. 2005;86:1623–30.
5. Kirshblum SC, Burns SP, Biering-Sorensen F, Donovan W, Graves DE, Jha A, et al. International standards for neurological classification of spinal cord injury (revised 2011). J Spinal Cord Med. 2011;34:535–46.
6. Spinal cord injury EDGE task force outcome measures recommendation [Internet]. 2015. http://www.neuropt.org/professional-resources/neurology-section-outcome-measures-recommendations/spinal-cord-injury. Accessed 25 April 2018.
7. Spinal cord injury [Internet]. 2018. https://www.commondataelements.ninds.nih.gov/SCI.aspx#tab=Data_Standards. Accessed 25 April 2018.
8. Outcome measures [Internet]. 2018. https://scireproject.com/outcome-measures/list-sci/. Accessed 25 April 2018.

9. Biering-Sorensen F, Alexander MS, Burns S, Charlifue S, DeVivo M, Dietz V, et al. Recommendations for translation and reliability testing of International Spinal Cord Injury Data Sets. Spinal Cord. 2011;49:357–60.

10. Biering-Sorensen F, Bryden A, Curt A, Friden J, Harvey LA, Mulcahey MJ, et al. International spinal cord injury upper extremity basic data set version 1.1. Spinal Cord. 2015;53:890.

11. Biering-Sorensen F, Charlifue S, DeVivo M, Noonan V, Post M, Stripling T, et al. International Spinal Cord Injury Data Sets. Spinal Cord. 2006;44:530–4.

12. DeVivo M, Biering-Sorensen F, Charlifue S, Noonan V, Post M, Stripling T, et al. International Spinal Cord Injury Core Data Set. Spinal Cord. 2006;44:535–40.

13. International SCI Upper Extremity Basic Data Set Version 1.1 Training Cases English. 2016. http://www.iscos.org.uk/international-sci-upper-extremity-data-sets. Accessed 16 Oct 2016.

14. Baunsgaard CB, Chhabra HS, Harvey LA, Savic G, Sisto SA, Qureshi F, et al. Reliability of the International Spinal Cord Injury Musculoskeletal Basic Data Set. Spinal Cord. 2016;54:1105–13.

15. Juul T, Bazzocchi G, Coggrave M, Johannesen IL, Hansen RB, Thiyagarajan C, et al. Reliability of the international spinal cord injury bowel function basic and extended data sets. Spinal Cord. 2011;49:886–91.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.

17. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85:257–68.

18. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.

19. Kalsi-Ryan S, Curt A, Verrier MC, Fehlings MG. Development of the graded redefined assessment of strength, sensibility and prehension (GRASSP): reviewing measurement specific to the upper limb in tetraplegia. J Neurosurg Spine. 2012;17(1 Suppl):65–76.

20. Starkey ML, Curt A. Clinical assessment and rehabilitation of the upper limb following cervical spinal cord injury. In: Reinkensmeyer DJ, Dietz J, editors. Neurorehabiliation technology. Switzerland: Springer; 2016. p. 107–38.

21. Lucantoni C, Krishnan RG, Gehrchen M, Hallager DW, Biering-Sorensen F, Dahl B. Reliability of the radiographic variables in the International Spinal Cord Injury Spinal Column Injury Basic Data Set compared with the AO classification. Spinal Cord. 2016;54:884–8.

22. Kim HR, Kim HB, Lee BS, Ko HY, Shin HI. Interrater reliability of the Korean version of the International Spinal Cord Injury Basic Pain Data Set. Spinal Cord. 2014;52:855–8.