

ORIGINAL ARTICLE

Agreement of repeated motor and sensory scores at individual myotomes and dermatomes in young persons with spinal cord injury

L Krisa¹, J Gaughan², L Vogel^{3,4}, RR Betz¹ and MJ Mulcahey^{1,5}

Study design: A prospective repeated measures multicenter study to determine reliability at individual spinal levels when applied to young persons with spinal cord injury (SCI).

Objectives: To evaluate intra- and inter-rater agreement of repeated motor and sensory scores at individual spinal levels.

Setting: Shriners Hospitals for Children—Philadelphia and Chicago, USA

Methods: A total 189 youth with complete and incomplete SCI underwent four neurological exams by two different raters. Agreement between and within raters for each myotome and dermatome was evaluated for complete and incomplete SCI separately. Intraclass correlation coefficients and 95% confidence intervals were calculated.

Results: Overall, both intra- and inter-rater agreement resulted in moderate-to-high agreement among myotomes. Subjects with complete SCI had moderate agreement for light touch (LT) and pin prick (PP) testing, whereas subjects with incomplete SCI had >60.0% of dermatomes resulting in poor agreement for PP testing.

Conclusion: Overall, moderate-to-high agreement was found for muscle strength comparisons and moderate-to-poor agreement was found for PP and LT.

Spinal Cord (2013) 51, 75–81; doi:10.1038/sc.2012.127; published online 13 November 2012

Keywords: spinal cord injury; pediatric; ASIA standards; reliability study

INTRODUCTION

The International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) is the standard method for the evaluation of neurological impairment following spinal cord injury (SCI).¹ The sensory and motor exams are both used to classify the neurological level, motor level and sensory level. The neurological level is the most caudal segment of the cord with normal motor and sensory function bilaterally, whereas the motor level and sensory level are the most caudal segments of the cord where there is normal bilateral motor and sensory function, respectively. The ISNCSCI has undergone several revisions since the first edition in 1982, predominantly to improve reliability in regards to both examination and classification.^{1–5} With the most recent revision⁶ clarity on terminology and testing techniques has been provided.

Although a number of studies have examined the reliability of both the examination and the classification of the ISNCSCI, most evaluated the reliability of summed scores,^{5,7–9} providing little insight into the repeatability of scores at individual myotomes and dermatomes or which dermatomes and myotomes may be more prone to reliability issues. In addition to summing motor scores, Savic *et al.*⁵ analyzed individual myotomes in 22 patients with SCI and found substantial to almost perfect agreement for all the muscles. To date, only two studies focused their findings of the ISNCSCI on individual myotomes and dermatomes. Jonsson

*et al.*² evaluated the inter-rater reliability of the ISNCSCI (1992 version) by assessing the degree of agreement between sensory and motor scores at individual myotomes and dermatomes in adults with incomplete SCI. Mulcahey *et al.*¹⁰ looked at intra-rater agreement of repeated motor and sensory scores at individual myotomes and dermatomes in the pediatric population with complete SCI.

Purpose

The aim of this study was to evaluate intra- and inter-rater reliability of the ISNCSCI motor and sensory scores at individual myotomes and dermatomes, respectively, in the young persons with SCI. Based on our previous work,¹⁰ we hypothesized that, regardless of the American Spinal Injury Association Impairment Scale designation, there would be moderate-to-high agreement for motor scores at individual levels for both inter- and intra-rater reliability. Conversely, for incomplete injuries, we expected agreement of sensory scores at individual levels to be moderate-to-high for intra-rater reliability and only moderate for inter-rater reliability.

MATERIALS AND METHODS

This was a prospective repeated measures multicenter study completed in adolescents and youth.

¹Shriners Hospitals for Children, Philadelphia, PA, USA; ²Temple University School of Medicine, Philadelphia, PA, USA; ³Shriners Hospitals for Children, Chicago, IL, USA; ⁴Rush Medical College Chicago, Chicago, IL, USA and ⁵Jefferson School of Health Professions, Thomas Jefferson University, Philadelphia, PA, USA
Correspondence: Dr MJ Mulcahey, Department of Occupational Therapy, School of Health Professions, Thomas Jefferson University, 901 Walnut Street, Philadelphia, PA 19107, USA.

E-mail: maryjane.mulcahey-hershey@jefferson.edu

Received 23 January 2012; revised 29 June 2012; accepted 25 July 2012; published online 13 November 2012

Table 1 Characteristics of sample with complete (n=97) and incomplete (n=92) SCI

Characteristic	Complete SCI		Incomplete SCI	
	Number	Percent	Number	Percent
Gender				
Male	56	57.7	55	59.7
Female	41	42.3	37	40.3
Etiology of injury				
Non-traumatic	9	9.2	27	29.3
Motor vehicle accident	56	57.7	29	31.5
Sports	21	21.6	29	31.5
Violence	11	11.5	7	7.7
Neurological level group				
C1-C4	27	27.8	47	51.1
C5-T1	15	15.4	16	17.4
T2-T7	34	35.0	16	17.4
T8-L2	21	21.8	13	14.1
Age at exam				
6-11	28	28.8	19	20.6
12-15	25	25.7	23	25.0
16-21	44	45.5	50	54.4

Abbreviation: SCI, spinal cord injury.

Subjects or samples

The sample consisted of 189 subjects, 97 with complete SCI and 92 with incomplete SCI (Table 1) between 6 and 21 years of age (average = 14.5 years). The average time between injury and exam was 4.3 years (range = 0.5–20 years). Each of the participants had a minimum of two neurological exams performed by one examiner, but a maximum of four neurological examinations (two exams each performed by two different examiners) 1–4 days apart. To minimize variation in repeated test scores resulting from actual neurological changes, there were no subjects with newly acquired (<3 months) injuries and no evidence of concomitant brain injury that would interfere with cognitive ability to participate in the examination.

This was a multicenter study, involving Philadelphia (n = 121) and Chicago (n = 68) Shriners Hospitals for Children, and was approved and reviewed by the Institutional Review Boards at participating sites. Written informed consent was obtained from parents or legal guardians of all subjects under the age of 18 years old, and participants between ages of 7 and 18 completed informed assents. We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during the course of this research. Although this study and the articles by Mulcahey et al.,¹¹ Vogel et al.,¹² Samdani et al.¹³ and Chafetz et al.¹⁴ reported on distinct types of reliability outcomes and used different stratifications and statistical methods for analyses, the subjects were drawn from the same sample.

Procedure

Repeated measures of the standardized ISNCSCI motor and sensory exams were administered by multiple testers to study intra- and inter-rater reliability. In total, seven raters who were trained in the examination techniques of the

Table 2 ICC values indicating intra-rater reliability for myotomes (M) and dermatomes (PP = test of discrimination and LT = test for light touch) in subjects with complete SCI

	Right			Left			
	M	PP		LT	LT		PP
		0.6 (0.52-0.68)	0.71 (0.64-0.77)	C2	0.58 (0.5-0.66)	0.68 (0.6-0.75)	
		0.66 (0.59-0.74)	0.75 (0.69-0.81)	C3	0.79 (0.73-0.84)	0.65 (0.58-0.73)	
		0.56 (0.48-0.65)	0.71 (0.64-0.77)	C4	0.76 (0.7-0.82)	0.69 (0.62-0.76)	
	0.89 (0.86-0.92)	0.75 (0.69-0.81)	0.7 (0.63-0.76)	C5	0.8 (0.75-0.85)	0.85 (0.81-0.89)	0.84 (0.79-0.88)
	0.92 (0.89-0.94)	0.72 (0.65-0.78)	0.71 (0.65-0.78)	C6	0.78 (0.72-0.83)	0.86 (0.82-0.89)	0.92 (0.89-0.94)
	0.98 (0.98-0.99)	0.86 (0.82-0.9)	0.79 (0.73-0.84)	C7	0.86 (0.82-0.9)	0.93 (0.9-0.95)	0.96 (0.95-0.97)
	1 (1-1)	0.93 (0.9-0.95)	0.9 (0.88-0.93)	C8	0.87 (0.84-0.91)	0.87 (0.83-0.9)	1 (1-1)
	0.98 (0.97-0.99)	0.9 (0.86-0.92)	0.87 (0.83-0.9)	T1	0.86 (0.82-0.89)	0.9 (0.87-0.92)	0.99 (0.99-0.99)
		0.9 (0.86-0.92)	0.89 (0.86-0.92)	T2	0.82 (0.77-0.87)	0.91 (0.89-0.94)	
		0.92 (0.9-0.94)	0.87 (0.84-0.9)	T3	0.81 (0.76-0.86)	0.92 (0.89-0.94)	
		0.9 (0.87-0.92)	0.88 (0.85-0.91)	T4	0.91 (0.88-0.93)	0.87 (0.83-0.9)	
		0.89 (0.86-0.92)	0.9 (0.87-0.92)	T5	0.88 (0.84-0.91)	0.84 (0.8-0.88)	
		0.9 (0.87-0.93)	0.87 (0.84-0.91)	T6	0.85 (0.8-0.89)	0.95 (0.93-0.96)	
		0.95 (0.93-0.96)	0.9 (0.87-0.92)	T7	0.91 (0.88-0.93)	0.84 (0.79-0.88)	
		0.89 (0.86-0.92)	0.88 (0.84-0.91)	T8	0.83 (0.78-0.87)	0.88 (0.84-0.91)	
		0.93 (0.9-0.95)	0.95 (0.93-0.96)	T9	0.93 (0.91-0.95)	0.95 (0.93-0.96)	
		0.89 (0.85-0.92)	0.9 (0.87-0.93)	T10	0.92 (0.9-0.94)	0.94 (0.92-0.95)	
		0.9 (0.87-0.92)	0.89 (0.86-0.92)	T11	0.87 (0.84-0.91)	0.88 (0.84-0.91)	
		0.61 (0.53-0.69)	0.76 (0.69-0.81)	T12	0.75 (0.69-0.81)	0.79 (0.74-0.84)	
		0.77 (0.71-0.82)	0.66 (0.59-0.74)	L1	0.85 (0.8-0.88)	0.79 (0.74-0.84)	
	0.82 (0.77-0.86)	1 (-.-)	0.75 (0.68-0.8)	L2	0.75 (0.68-0.8)	1 (-.-)	0.78 (0.72-0.83)
	0.82 (0.77-0.86)	0.75 (0.68-0.8)	0.68 (0.61-0.75)	L3	0.7 (0.63-0.77)		1 (-.-)
	0.99 (0.99-0.99)			L4			
	0.99 (0.99-0.99)			L5			
	0.99 (0.99-0.99)		1 (-.-)	S1	0.53 (0.44-0.61)		
		0.57 (0.49-0.66)	0.79 (0.73-0.84)	S2	0.45 (0.36-0.54)		
				S3			
				S4/5			
		Not calculated *	High	Moderate	Poor		
Muscle (20)		3 (15%)	12 (60%)	5 (25%)			
PP (56)		12 (21.5%)	12 (21.5%)	23 (41%)	9 (16%)		
LT (56)		2 (4%)	6 (11%)	32 (57%)	16 (28%)		

Abbreviations: ICC, intraclass correlation coefficient; SCI, spinal cord injury. Cells without values (*) indicate that the ICC could not be calculated either because the variation between subjects was not significant or there was inadequate number of subjects. Shaded cells indicate poor agreement (<0.75). Values in bold indicate high agreement (>0.90). Cells with horizontal lines reflect levels without key muscles.

Table 3 ICC values indicating intra-rater reliability for myotomes (M) and dermatomes (PP = test of discrimination and LT = test for light touch) in subjects with incomplete SCI

Right			Left			
M	PP	LT		LT	PP	M
	0.46 (0.38-0.56)	0.41 (0.33-0.51)	C2	0.5 (0.42-0.6)	0.61 (0.52-0.69)	
	0.65 (0.58-0.73)	0.59 (0.51-0.67)	C3	0.47 (0.39-0.57)	0.48 (0.39-0.57)	
	0.56 (0.48-0.65)	0.54 (0.45-0.63)	C4	0.38 (0.29-0.48)	0.55 (0.46-0.64)	
0.97 (0.96-0.98)	0.59 (0.51-0.68)	0.7 (0.63-0.77)	C5	0.63 (0.55-0.71)	0.69 (0.62-0.76)	0.9 (0.87-0.93)
0.91 (0.88-0.93)	0.69 (0.62-0.76)	0.75 (0.69-0.81)	C6	0.75 (0.69-0.81)	0.73 (0.66-0.79)	0.92 (0.89-0.94)
0.97 (0.96-0.98)	0.83 (0.78-0.87)	0.67 (0.6-0.75)	C7	0.74 (0.68-0.8)	0.77 (0.71-0.82)	0.94 (0.93-0.96)
0.96 (0.94-0.97)	0.82 (0.77-0.86)	0.81 (0.75-0.85)	C8	0.84 (0.79-0.88)	0.85 (0.8-0.89)	0.97 (0.96-0.98)
0.96 (0.94-0.97)	0.84 (0.8-0.88)	0.68 (0.61-0.75)	T1	0.73 (0.66-0.79)	0.78 (0.72-0.84)	0.97 (0.96-0.98)
	0.76 (0.69-0.81)	0.8 (0.75-0.85)	T2	0.62 (0.54-0.7)	0.72 (0.65-0.78)	
	0.82 (0.77-0.87)	0.77 (0.71-0.82)	T3	0.75 (0.68-0.81)	0.76 (0.7-0.82)	
	0.8 (0.74-0.85)	0.77 (0.71-0.83)	T4	0.72 (0.65-0.78)	0.83 (0.78-0.87)	
	0.73 (0.67-0.8)	0.77 (0.7-0.82)	T5	0.74 (0.68-0.8)	0.81 (0.76-0.86)	
	0.71 (0.64-0.78)	0.77 (0.71-0.82)	T6	0.8 (0.74-0.85)	0.78 (0.73-0.84)	
	0.74 (0.68-0.8)	0.74 (0.67-0.8)	T7	0.79 (0.73-0.84)	0.79 (0.74-0.84)	
	0.76 (0.7-0.82)	0.81 (0.76-0.86)	T8	0.72 (0.65-0.78)	0.75 (0.69-0.81)	
	0.76 (0.7-0.82)	0.78 (0.72-0.83)	T9	0.81 (0.76-0.86)	0.78 (0.72-0.83)	
	0.73 (0.67-0.8)	0.77 (0.72-0.83)	T10	0.83 (0.78-0.87)	0.75 (0.69-0.81)	
	0.7 (0.63-0.77)	0.77 (0.7-0.82)	T11	0.81 (0.75-0.85)	0.75 (0.69-0.81)	
	0.69 (0.62-0.76)	0.76 (0.7-0.82)	T12	0.79 (0.73-0.84)	0.63 (0.55-0.71)	
	0.62 (0.54-0.7)	0.79 (0.73-0.84)	L1	0.75 (0.69-0.81)	0.53 (0.44-0.62)	
0.93 (0.9-0.95)	0.68 (0.61-0.75)	0.72 (0.65-0.78)	L2	0.73 (0.67-0.8)	0.67 (0.59-0.74)	0.94 (0.92-0.95)
0.95 (0.93-0.96)	0.73 (0.67-0.79)	0.69 (0.62-0.76)	L3	0.75 (0.69-0.81)	0.64 (0.57-0.72)	0.92 (0.9-0.94)
0.98 (0.97-0.98)	0.48 (0.39-0.57)	0.75 (0.69-0.81)	L4	0.71 (0.64-0.78)	0.68 (0.61-0.75)	0.89 (0.85-0.92)
0.89 (0.86-0.92)	0.69 (0.61-0.76)	0.77 (0.71-0.83)	L5	0.75 (0.68-0.81)	0.63 (0.55-0.71)	0.9 (0.87-0.93)
0.84 (0.8-0.88)	0.59 (0.51-0.68)	0.71 (0.64-0.78)	S1	0.72 (0.66-0.79)	0.6 (0.51-0.68)	0.94 (0.91-0.95)
	0.65 (0.57-0.72)	0.76 (0.7-0.82)	S2	0.73 (0.66-0.79)	0.57 (0.48-0.65)	
	0.54 (0.46-0.63)	0.69 (0.62-0.76)	S3	0.69 (0.61-0.76)	0.61 (0.53-0.69)	
	0.61 (0.53-0.69)	0.75 (0.68-0.81)	S4/5	0.7 (0.62-0.76)	0.58 (0.5-0.67)	
	Not calculated*	High	Moderate	Poor		
Muscle (20)		15 (75%)	5 (25%)			
PP (56)			20 (36%)	36 (64%)		
LT (56)			29 (52%)	27 (48%)		

Abbreviation: ICC, intraclass correlations coefficient; SCI, spinal cord injury. Cells without values (*) indicate that the ICC could not be calculated either because the variation between subjects was not significant or there was inadequate number of subjects. Shaded cells indicate poor agreement (<0.75). Values in bold indicate high agreement (>0.90). Cells with horizontal lines reflect levels without key muscles.

ISNCSCI¹⁵ participated in the study; four raters had more than 3 years of experience conducting the ISNCSCI, and the remaining three had <1 year of experience. Before the study, reliability among all raters was determined for examination. Each subject had up to four repeated examinations performed by two raters (the two raters, each performed two examinations on the same subject, provide the platform for analysis of within and between rater reliability). Raters were not randomized. The sensory examinations were performed according to the standard ISNCSCI methods¹⁵ unless the subject was inconsistent at a spinal level. In this case, 8 out of 10 trials had to be identified correctly for sharp/dull discrimination (pin prick (PP)) and light touch (LT) appreciation to receive a score of 2. This stringent sensory testing protocol was employed to ensure accuracy of the data. Scoring was performed by the trained raters, and data were entered into a secure database by two independent assistants blinded to the study.

Data analysis

All data were ranked transformed before analysis. The intraclass correlation coefficient (ICC) as a measure of reliability was calculated separately for complete and incomplete SCI groups in addition to intra- and inter-rater reliability based on the repeated measures design using the components of variance model (2, 1 for inter-rater ICC and 3, 1 for intra-rater ICC) to estimate parameters of interest for ICC calculations (that is, inter- and intra-

rater).¹⁶ Also, 95% confidence intervals (CIs) were calculated for all ICC values.

For reliability studies, sample size consideration is not based on statistical methodology but rather on the desired precision of the reliability estimates. The sample in this study represents a convenience sample of individuals willing to undergo two to four neurological exams. Although this study contains 189 subjects, there was not an equal distribution of subjects with an SCI at each spinal level; therefore, sample size at each level was too small to analyze intra- and inter-rater reliability at each dermatome and myotome when subjects were separated based on their level of injury. Therefore, all subjects with complete SCI were analyzed together as were all subjects with incomplete SCI, regardless of their neurological level, motor level or sensory level. All spinal segments, including those scored absent or normal, were tested and used for data analysis. Both inter- and inter-rater reliability were analyzed for both groups.

RESULTS

The results are shown in Tables 2–5, presenting ICC values with 95% CIs for each study parameter. ICC values >0.90 reflect high agreement, values from 0.75 to 0.90 reflect moderate agreement and values <0.75 are reflective of poor agreement.^{7,9} The 95% CI provides an indication of precision of the coefficients; a narrow CI reflects good precision (10–20% of the estimated true value) and a

Table 4 ICC values indicating inter-rater reliability for myotomes (M) and dermatomes (PP = test of discrimination and LT = test for light touch) in subjects with complete SCI

Right				Left		
M	PP	LT		LT	PP	M
	0.59 (0.51-0.67)	0.7 (0.63-0.77)	C2	0.57 (0.49-0.65)	0.66 (0.59-0.73)	
	0.66 (0.58-0.73)	0.74 (0.68-0.8)	C3	0.78 (0.72-0.83)	0.64 (0.57-0.72)	
	0.59 (0.52-0.68)	0.69 (0.62-0.76)	C4	0.74 (0.68-0.8)	0.68 (0.6-0.75)	
0.89 (0.86-0.92)	0.76 (0.7-0.81)	0.7 (0.63-0.77)	C5	0.81 (0.76-0.86)	0.86 (0.82-0.9)	0.83 (0.78-0.87)
0.91 (0.89-0.94)	0.72 (0.66-0.79)	0.7 (0.63-0.77)	C6	0.79 (0.73-0.84)	0.85 (0.81-0.89)	0.91 (0.89-0.94)
0.98 (0.98-0.99)	0.86 (0.82-0.89)	0.79 (0.74-0.84)	C7	0.87 (0.83-0.9)	0.92 (0.9-0.94)	0.96 (0.95-0.97)
1 (1-1)	0.93 (0.91-0.95)	0.91 (0.88-0.93)	C8	0.88 (0.84-0.91)	0.86 (0.83-0.9)	1 (1-1)
0.99 (0.98-0.99)	0.9 (0.87-0.93)	0.87 (0.83-0.9)	T1	0.86 (0.82-0.9)	0.89 (0.86-0.92)	0.99 (0.98-0.99)
	0.89 (0.86-0.92)	0.89 (0.86-0.92)	T2	0.82 (0.77-0.86)	0.91 (0.88-0.93)	
	0.92 (0.89-0.94)	0.87 (0.84-0.91)	T3	0.8 (0.75-0.85)	0.91 (0.89-0.94)	
	0.9 (0.87-0.92)	0.9 (0.87-0.92)	T4	0.91 (0.88-0.93)	0.88 (0.85-0.91)	
	0.89 (0.86-0.92)	0.9 (0.87-0.92)	T5	0.87 (0.83-0.9)	0.84 (0.79-0.88)	
	0.9 (0.87-0.92)	0.87 (0.83-0.9)	T6	0.84 (0.8-0.88)	0.95 (0.93-0.96)	
	0.94 (0.93-0.96)	0.89 (0.86-0.92)	T7	0.91 (0.88-0.93)	0.84 (0.79-0.88)	
	0.89 (0.85-0.92)	0.87 (0.84-0.91)	T8	0.83 (0.78-0.87)	0.88 (0.84-0.91)	
	0.94 (0.92-0.95)	0.94 (0.92-0.96)	T9	0.93 (0.91-0.95)	0.95 (0.93-0.96)	
	0.89 (0.85-0.92)	0.91 (0.89-0.94)	T10	0.92 (0.9-0.94)	0.95 (0.93-0.96)	
	0.9 (0.87-0.92)	0.89 (0.86-0.92)	T11	0.9 (0.87-0.93)	0.88 (0.85-0.91)	
	0.73 (0.66-0.79)	0.81 (0.76-0.86)	T12	0.8 (0.75-0.85)	0.78 (0.73-0.83)	
	0.76 (0.7-0.82)	0.69 (0.62-0.76)	L1	0.84 (0.8-0.88)	0.78 (0.72-0.83)	
0.82 (0.77-0.86)	1 (-.)	0.74 (0.67-0.8)	L2	0.74 (0.67-0.8)	1 (-.)	0.78 (0.73-0.84)
1 (-.)	0.74 (0.67-0.8)	0.67 (0.6-0.74)	L3	0.69 (0.62-0.76)		1 (-.)
1 (-.)			L4			
1 (-.)		0.6 (0.52-0.68)	L5			
1 (-.)		0.99 (-.)	S1	0.52 (0.43-0.6)		
	0.56 (0.48-0.65)	0.77 (0.71-0.83)	S2	0.43 (0.35-0.53)		
			S3			
			S4/5			
	Not calculated*	High	Moderate	Poor		
Muscle (20)	3 (15%)	13 (65%)	4 (20%)			
PP (56)	12 (21.5%)	12 (21.5%)	22 (39%)	10 (18%)		
LT (56)	7 (13%)	8 (14%)	26 (46%)	15 (27%)		

Abbreviation: ICC, intraclass correlations coefficient; SCI, spinal cord injury.

Cells without values (*) indicate that the ICC could not be calculated either because the variation between subjects was not significant or there was inadequate number of subjects. Shaded cells indicate poor agreement (<0.75). Values in bold indicate high agreement (>0.90). Cells with horizontal lines reflect levels without key muscles.

wide CI reflects weak precision. Myotomes and dermatomes with no values indicate that the ICC could not be calculated either because the variation between subjects was not significant or there was inadequate number of subjects.

Intra-rater reliability

Myotomes. In general, there was moderate-to-high agreement in myotomes. In the subjects with complete SCI, there was high agreement in 60% (12/20) of myotomes with 25% (5/20) having moderate agreement and the remaining 15% (3/20) not able to be calculated; ICC values ranged from 0.78 to 1.0 (Table 2). Out of 20 myotomes in the incomplete SCI group, 75% (15/20) had high agreement with the remaining 25% of myotomes (5/20) having moderate agreement. ICC values ranged 0.84–0.97 (Table 3).

Dermatomes. The PP with the complete SCI group demonstrated ICC ranges from 0.56 to 1 with 21.5% (12/56) of dermatomes having high agreement and 41% (23/56) resulting in moderate agreement (Table 2). For LT, the majority of dermatomes, 57% (32/56), showed moderate agreement with ICC values that ranged 0.45–1.0. As shown in Table 3, subjects with incomplete SCI have ICC values that range from 0.46 to 0.85 for PP, with 64% (36/56) of dermatomes having poor agreement. LT values ranged from 0.38 to 0.83, with 52% (29/

56) showing moderate agreement and 48% (27/56) having poor agreement.

Inter-rater reliability

Myotomes. The agreement of myotomes in subjects with complete SCI was high in 65% (13/20) of myotomes with moderate agreement in 20% (4/20) of myotomes. The ICC values ranged from 0.78 to 1.0, and 15% (3/20) of myotomes were unable to be calculated (Table 4). Table 5 provides ICC values that range from 0.87 to 0.98 in subjects with incomplete SCI, with 85% (17/20) of myotomes having high and 15% (3/20) of myotomes having moderate agreement.

Dermatomes. The agreement for PP scores in subjects with complete SCI was high in 21.5% (12/56) of dermatomes, with 39% (22/56) and 18% (10/56) having moderate to poor agreement, respectively. ICC values were unable to be calculated for 21.5% (12/56) of dermatomes, and ICC values ranged 0.56–1.0 (Table 4). This population also had 14% (8/56) high agreement for LT scores in addition to a 46% (26/56) and 27% moderate to poor agreement, respectively. ICC values ranged from 0.43 to 0.94, and 13% (7/56) of dermatomes were unable to be calculated. Table 5 illustrates the agreement of PP scores to be moderate in 39% (22/56) of subjects with incomplete SCI, with the remaining 61% (34/56) having poor agreement. On the other hand, the majority of dermatomes, 61% (34/56), had moderate

Table 5 ICC values indicating inter-rater reliability for myotomes (M) and dermatomes (PP = test of discrimination and LT = test for light touch) in subjects with incomplete SCI

Right			Left			
M	PP	LT		LT	PP	M
	0.46 (0.37-0.56)	0.44 (0.35-0.54)	C2	0.5 (0.41-0.6)	0.6 (0.52-0.69)	
	0.64 (0.56-0.72)	0.57 (0.48-0.66)	C3	0.55 (0.46-0.64)	0.47 (0.37-0.56)	
	0.59 (0.5-0.68)	0.59 (0.5-0.68)	C4	0.43 (0.34-0.53)	0.54 (0.45-0.63)	
0.97 (0.96-0.98)	0.61 (0.53-0.7)	0.71 (0.64-0.78)	C5	0.64 (0.56-0.72)	0.69 (0.61-0.76)	0.9 (0.87-0.93)
0.9 (0.87-0.93)	0.67 (0.59-0.75)	0.77 (0.71-0.83)	C6	0.76 (0.7-0.82)	0.71 (0.64-0.78)	0.92 (0.89-0.94)
0.96 (0.95-0.97)	0.83 (0.78-0.88)	0.71 (0.64-0.78)	C7	0.77 (0.71-0.83)	0.76 (0.7-0.82)	0.94 (0.92-0.96)
0.96 (0.95-0.98)	0.83 (0.78-0.87)	0.85 (0.8-0.89)	C8	0.83 (0.78-0.88)	0.86 (0.81-0.9)	0.97 (0.96-0.98)
0.96 (0.94-0.97)	0.85 (0.81-0.89)	0.68 (0.61-0.76)	T1	0.75 (0.68-0.81)	0.81 (0.75-0.86)	0.97 (0.96-0.98)
	0.76 (0.7-0.82)	0.82 (0.77-0.87)	T2	0.65 (0.57-0.73)	0.71 (0.64-0.78)	
	0.82 (0.77-0.87)	0.79 (0.73-0.84)	T3	0.78 (0.72-0.84)	0.82 (0.77-0.87)	
	0.8 (0.74-0.85)	0.77 (0.71-0.83)	T4	0.73 (0.66-0.8)	0.86 (0.81-0.89)	
	0.75 (0.68-0.81)	0.78 (0.71-0.83)	T5	0.76 (0.69-0.82)	0.84 (0.79-0.88)	
	0.7 (0.62-0.77)	0.77 (0.71-0.83)	T6	0.8 (0.74-0.85)	0.82 (0.77-0.87)	
	0.74 (0.67-0.8)	0.73 (0.66-0.8)	T7	0.79 (0.73-0.85)	0.79 (0.74-0.85)	
	0.76 (0.69-0.82)	0.83 (0.78-0.87)	T8	0.71 (0.64-0.78)	0.76 (0.69-0.82)	
	0.76 (0.69-0.82)	0.79 (0.73-0.84)	T9	0.81 (0.76-0.86)	0.78 (0.72-0.83)	
	0.73 (0.65-0.79)	0.79 (0.73-0.84)	T10	0.82 (0.77-0.87)	0.75 (0.69-0.81)	
	0.68 (0.61-0.76)	0.77 (0.71-0.83)	T11	0.82 (0.76-0.86)	0.77 (0.7-0.82)	
	0.67 (0.6-0.75)	0.77 (0.71-0.83)	T12	0.78 (0.72-0.84)	0.64 (0.56-0.72)	
	0.62 (0.54-0.71)	0.82 (0.77-0.87)	L1	0.77 (0.7-0.83)	0.55 (0.46-0.64)	
0.93 (0.91-0.95)	0.68 (0.61-0.76)	0.73 (0.66-0.8)	L2	0.75 (0.69-0.81)	0.66 (0.59-0.74)	0.94 (0.91-0.95)
0.94 (0.92-0.96)	0.71 (0.64-0.78)	0.72 (0.65-0.79)	L3	0.78 (0.72-0.83)	0.7 (0.62-0.77)	0.92 (0.89-0.94)
0.98 (0.97-0.98)	0.47 (0.38-0.57)	0.78 (0.72-0.84)	L4	0.72 (0.65-0.79)	0.69 (0.61-0.76)	0.91 (0.88-0.93)
0.91 (0.88-0.93)	0.68 (0.6-0.75)	0.78 (0.72-0.84)	L5	0.79 (0.73-0.84)	0.65 (0.57-0.73)	0.9 (0.87-0.93)
0.87 (0.83-0.91)	0.6 (0.51-0.69)	0.72 (0.65-0.79)	S1	0.75 (0.68-0.81)	0.61 (0.53-0.7)	0.94 (0.91-0.96)
	0.67 (0.59-0.74)	0.77 (0.71-0.83)	S2	0.78 (0.72-0.84)	0.57 (0.48-0.66)	
	0.54 (0.45-0.63)	0.69 (0.62-0.76)	S3	0.7 (0.63-0.77)	0.6 (0.51-0.69)	
	0.61 (0.53-0.7)	0.74 (0.67-0.8)	S4/5	0.7 (0.62-0.77)	0.58 (0.5-0.67)	
	Not calculated*	High	Moderate	Poor		
Muscle (20)		17 (85%)	3 (15%)			
PP (56)			22 (39%)	34 (61%)		
LT (56)			34 (61%)	22 (39%)		

Abbreviation: ICC, intraclass correlations coefficient; SCI, spinal cord injury. Cells without values (*) indicate that the ICC could not be calculated either because the variation between subjects was not significant or there was inadequate number of subjects. Shaded cells indicate poor agreement (<0.75). Values in bold indicate high agreement (>0.90). Cells with horizontal lines reflect levels without key muscles.

agreement for LT scores, with 39% (22/56) having poor agreement for this test. ICC values range from 0.43 to 0.85.

DISCUSSION

The purpose of this study was to examine inter- and intra-rater reliability agreement of the ISNCSCI exam at every dermatome and myotome in youth and adolescents with complete and incomplete SCI. There were no dermatomes or myotomes that were statistically different between inter- and intra-rater reliability in subjects with incomplete SCI and only one myotome (L3-R), and one dermatome each for PP (T12-R) and LT (L5-R) showed differences for complete SCI subjects.

As hypothesized, the agreement of myotomes in both complete and incomplete subjects (inter- and intra-rater reliability) was moderate-to-high, except for those myotomes where an ICC value could not be calculated because of a lack of variability between/within subjects. An ICC of 1 indicates complete agreement; this occurred only in subjects with complete SCI. These findings may have occurred because those muscles that were scored as a 5 (normal) and 0 (paralyzed) were also included in the results; previous reports^{10,17} have shown that agreement of repeated testing of unimpaired and completely paralyzed muscles is usually perfect, and for this reason, previous studies have excluded them from analysis.⁵ We choose not to exclude

myotomes and dermatomes that were unimpaired and completely paralyzed because this would lead to a variable number of myotomes and dermatomes at each level, potentially, leading to skewed data. In addition, when conducting a reliability study, it is important to test all scores, regardless of where they fall on the testing scale, to ensure adequate results.

As anticipated, the results of the sensory data are less straightforward. For both inter- and intra-rater reliability, there was no dermatomes with high agreement for sensory testing in subjects with incomplete SCI. There were more dermatomes with poor agreement (61% and 64%) for PP scoring when compared with LT scoring (39% and 48%) in both inter- and intra-rater reliability analyses, respectively. These findings are similar to those of Jonsson *et al.*² who found acceptable agreement in 49 out of 92 dermatomes tested for LT compared with only 26 out of 96 dermatomes tested for PP in adults with SCI. Interestingly, this is not the case in subjects with complete SCI. There was higher agreement for PP and lower agreement for LT.

Overall subjects with incomplete SCI had more dermatomes and myotomes with lower ICC values for both sensory (LT and PP) and muscle strength tests when compared with subjects with complete SCI. These differences were statistically significant in 80% of myotomes, and 78.5% and 77% of dermatomes for PP and LT,

respectively (data not shown). These findings are not surprising and support our original hypotheses. Cohen *et al.*³ found that classification of incomplete injuries is 'problematical in many areas' in contrast to classification of complete injuries.

Past studies^{18–20} have found natural recovery of motor and sensory function to occur within the first year following an SCI. However, the extent of disagreement on repeated sensory (LT and PP) scores reported in this study is not likely because of neurological improvements, given the short time interval of 2–4 days between all the ISNCSCI testing session and the chronicity of the injuries. Thus, the low ICC values must be attributed to other sources of variation, such as the subject, testing environment, the test itself or a combination of all three.

The current study examined the reliability of the ISNCSCI in children and youths with both complete and incomplete chronic SCI at individual myotomes and dermatomes. Past studies^{5,7–9} have used the overall motor and sensory scores to assess the reliability of the ISNCSCI; however, the summed score may be the same or different based on what is happening at the individual level. Therefore, summed motor or sensory scores are not an optimal method to detect motor and sensory changes that occur due to recovery or treatment nor are they optimal for detecting progression or remediation of symptoms. Although lower agreement values were expected in subjects with incomplete SCI, the degree of poor agreement on the sensory scores was unanticipated, particularly in light of the strict '8 out of 10' criteria set forth in the ISNCSCI manual.¹⁵ This raises some concern about the results of the ISNCSCI sensory examination in youths with incomplete SCI. Clinical trials that define improvements in sensory and/or motor function by changes in scores at individual spinal levels may not be indicative of the actual changes because of our finding indicating unacceptable ranges of agreement. Perhaps, a more robust examination is needed to better define the location and severity of the injury in subjects classified with an incomplete injury by the ISNCSCI. Currently, we are establishing neuroimaging criteria based on diffusion tensor imaging for evaluating the location and severity of SCI in children and youth.²¹ This technique shows promise in quantifying viable neural tissue within the injured spinal cord in SCI. Also, the poor-to-moderate agreement at the S4/5 dermatomes in subjects with incomplete injury is of concern. Samdani *et al.*¹³ reported that 40% of subjects who underwent the ISNCSCI examinations had no S4/5 sensation but did respond positively to deep anal pressure: concluding the importance of testing anal sensation to determine the American Spinal Injury Association Impairment Scale classification. However, studies^{22–24} have indicated the possibility that anal pressure is perceived by patients through an alternative pathway, which is not indicative of spinal cord integrity. With the validity of the test for anal pressure in question and the extent of poor reliability at the S4/5 dermatomes, misclassification of completeness is a modest possibility. A more accurate way of determining SCI severity is needed. Wietek *et al.*²⁴ used functional magnetic resonance imaging to study cortical activity during anorectal stimulation and showed cortical activation in areas similar to those found in healthy volunteers with less extensive activation. Therefore, brain imaging techniques may help to identify the functional pathways that might not be identified with conventional SCI testing.

Study limitations

There are limitations to this study. For example, even though we standardized the testing techniques, youths participating in the study had varying degrees of experience with the examination. Thus, youths

differed in their knowledge about the examination, and we do not know the effect of this difference on the examination results. Second, as a way to ensure sufficient number of subjects for analysis, we grouped subjects based on their severity of injury (complete vs incomplete) and not by age or level of injury. Based on our previous work, after 6 years of age, there is no indication that age introduces variability into the examination,¹¹ and thus, we did not feel groupings by age were appropriate. Although the strongest design would be to group subjects based on neurological level or motor level, even with 187 subjects, there would be insufficient numbers per level for analysis.¹¹

CONCLUSION

Inter- and Intra-rater agreement was moderate-to-high for myotome testing in both complete and incomplete SCI subjects. The agreement of sensory testing (PP and LT) was worse than motor scores with no high agreement for any dermatome in incomplete subjects. Complete subject agreement was slightly better, however, there were still <25% high agreements for either sensory modality. These results suggest that caution should be used when determining if a subject has made true improvements in the ISNCSCI overall scores (total motor and total sensory) or if this change was simply due to a degree of variation at the individual dermatome and myotome level.

DATA ARCHIVING

There were no data to deposit.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was funded by the Shriners Hospitals for Children Research Grant no. 8956.

- 1 Kirshblum SC, Burns SP, Biering-Sorensen F, Donovan WH, Graves DE, Jha A *et al.* International standards for neurological classification of spinal cord injury. *J Spinal Cord Med* 2011; **34**: 535–546.
- 2 Jonsson M, Tollback A, Gonzales H, Borg J. Inter-rater reliability of the 1992 international standards for neurological and functional classification of incomplete spinal cord injury. *Spinal Cord* 2000; **38**: 675–679.
- 3 Cohen ME, Ditunno Jr JF, Donovan WH, Maynard Jr FM. A test of the 1992 International Standards for Neurological and Functional Classification of Spinal Cord Injury. *Spinal Cord* 1998; **36**: 554–560.
- 4 Cohen ME, Sheehan TP, Herbison GJ. Content validity and reliability of the international standard for neurological classification of spinal cord injury. *Top Spinal Cord Inj Rehabil* 1996; **1**: 15–31.
- 5 Savic G, Bergstrom EM, Frankel HL, Jamous MA, Jones PW. Inter-rater reliability of motor and sensory examinations performed according to American Spinal Injury Association standards. *Spinal Cord* 2007; **45**: 444–451.
- 6 Kirshblum SC, Waring W, Biering-Sorensen F, Burns SP, Johansen M, Schmidt-Read M *et al.* Reference for the 2011 revision of the International Standards for Neurological Classification of Spinal Cord Injury. *J Spinal Cord Med* 2011; **34**: 547–554.
- 7 Chafetz RS, Vogel LC, Betz RR, Gaughan JP, Mulcahey MJ. International standards for neurological classification of spinal cord injury: training effect on accurate classification. *J Spinal Cord Med* 2008; **31**: 538–542.
- 8 Marino RJ, Jones L, Kirshblum S, Tal J, Dasgupta A. Reliability and repeatability of the motor and sensory examination of the international standards for neurological classification of spinal cord injury. *J Spinal Cord Med* 2008; **31**: 166–170.
- 9 Mulcahey MJ, Gaughan J, Betz RR, Vogel LC. Rater agreement on the ISNCSCI motor and sensory scores obtained before and after formal training in testing technique. *J Spinal Cord Med* 2007; **30**(Suppl 1): S146–S149.
- 10 Mulcahey MJ, Gaughan J, Betz RR. Agreement of repeated motor and sensory scores at individual myotomes and dermatomes in young persons with complete spinal cord injury. *Spinal Cord* 2009; **47**: 56–61.
- 11 Mulcahey MJ, Gaughan JP, Chafetz RS, Vogel LC, Samdani AF, Betz RR. Interrater reliability of the international standards for neurological classification of spinal cord

- injury in youths with chronic spinal cord injury. *Arch Phys Med Rehabil* 2011; **92**: 1264–1269.
- 12 Vogel L, Samdani A, Chafetz R, Gaughan J, Betz R, Mulcahey MJ. Intra-rater agreement of the anorectal exam and classification of injury severity in children with spinal cord injury. *Spinal Cord* 2009; **47**: 687–691.
- 13 Samdani A, Chafetz RS, Vogel LC, Betz RR, Gaughan JP, Mulcahey MJ. The international standards for neurological classification of spinal cord injury: relationship between S4-5 dermatome testing and anorectal testing. *Spinal Cord* 2011; **49**: 352–356.
- 14 Chafetz RS, Gaughan JP, Vogel LC, Betz R, Mulcahey MJ. The international standards for neurological classification of spinal cord injury: intra-rater agreement of total motor and sensory scores in the pediatric population. *J Spinal Cord Med* 2009; **32**: 157–161.
- 15 American Spinal Injury Association. *Reference Manual for International Standards for Neurological Classification of Spinal Cord Injury*. American Spinal Injury Association: Chicago, 2003.
- 16 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–428.
- 17 Calhoun CL, Gaughan JP, Chafetz RS, Mulcahey MJ. A pilot study of observational motor assessment in infants and toddlers with spinal cord injury. *Pediatr Phys Ther* 2009; **21**: 62–67.
- 18 Waters RL, Adkins RH, Yakura JS, Sie I. Motor and sensory recovery following complete tetraplegia. *Arch Phys Med Rehabil* 1993; **74**: 242–247.
- 19 Waters RL, Yakura JS, Adkins RH, Sie I. Recovery following complete paraplegia. *Arch Phys Med Rehabil* 1992; **73**: 784–789.
- 20 Wu L, Marino RJ, Herbison GJ, Ditunno Jr JF. Recovery of zero-grade muscles in the zone of partial preservation in motor complete quadriplegia. *Arch Phys Med Rehabil* 1992; **73**: 40–43.
- 21 Mulcahey M, Samdani A, Gaughan J, Barakat N, Faro S, Betz RR et al. Diffusion tensor imaging in pediatric spinal cord injury: preliminary examination of reliability and clinical correlation. *Spine* 2012; **37**: E797–E803.
- 22 Finnerup NB, Gyldensted C, Fuglsang-Frederiksen A, Bach FW, Jensen TS. Sensory perception in complete spinal cord injury. *Acta Neurol Scand* 2004; **109**: 194–199.
- 23 Sherwood AM, Dimitrijevic MR, McKay WB. Evidence of subclinical brain influence in clinically complete spinal cord injury: discomplete SCI. *J Neurol Sci* 1992; **110**: 90–98.
- 24 Wietek BM, Baron CH, Erb M, Hinninghofen H, Badtke A, Kaps HP et al. Cortical processing of residual ano-rectal sensation in patients with spinal cord injury: an fMRI study. *Neurogastroenterol Motil* 2008; **20**: 488–497.