

---

---

# Interrater Reliability of the Pediatric Neuromuscular Recovery Scale for Spinal Cord Injury

Andrea L. Behrman, PT, PhD, FAPTA,<sup>1,2</sup> Shelley A. Trimble, PT,<sup>3†</sup> Laura C. Argetsinger, DPT, PT,<sup>3</sup> MacKenzie T. Roberts, DPT, PT,<sup>3</sup> MJ Mulcahey, PhD, OTR/L,<sup>4</sup> Lisa Clayton, BS,<sup>2,5</sup> Mary E. Gregg, MS,<sup>6</sup> Doug Lorenz, PhD,<sup>6</sup> and Elizabeth M. Ardolino, PhD, PT<sup>7</sup>

<sup>1</sup>Department of Neurological Surgery, University of Louisville, Louisville, Kentucky; <sup>2</sup>Kentucky Spinal Cord Injury Research Center, University of Louisville, Louisville, Kentucky; <sup>3</sup>Spinal Cord Medicine Program, Frazier Rehab Institute, Louisville, Kentucky; <sup>4</sup>Jefferson College of Rehabilitation Science, Thomas Jefferson University, Philadelphia, Pennsylvania; <sup>5</sup>Kosair Charities Center for Pediatric NeuroRecovery, University of Louisville, Louisville, Kentucky; <sup>6</sup>Department of Biostatistics, University of Louisville, Louisville, Kentucky; <sup>7</sup>Doctor of Physical Therapy Program - Austin, University of St. Augustine for Health Sciences, Austin, Texas

**Background:** In synergy with the mounting scientific evidence for the capacity of recovery after spinal cord injury (SCI) and training, new evidence-based therapies advancing neuromuscular recovery are emerging. There is a parallel need for outcome instruments that specifically address recovery. The Pediatric Neuromuscular Recovery Scale (Pediatric NRS) is one example with established content validity to assess neuromuscular capacity within task performance. **Objective:** The objective of this study was to determine interrater reliability of the Pediatric NRS to classify motor capacity in children after SCI. **Methods:** Pediatric physicians (3), occupational therapists (5), and physical therapists (6) received standardized training in scoring the scale, then rated video assessments of 32 children post SCI, 2-12 years of age, 78% non-ambulatory. Interrater reliability was analyzed using Kendall coefficient of concordance for individual Pediatric NRS items and overall score. **Results:** The interrater reliability coefficient was determined to be near 1 for the overall Pediatric NRS score (ICC = 0.966; 95% CI, 0.89-0.98). Twelve of 16 individual items exhibited high concordance coefficients (Kendall's  $W \geq 0.8$ ) and four items demonstrated concordance coefficients,  $< 0.8$  and  $> 0.69$ . Interrater reliability was equivalent among groups defined by age and neurological level, but lower among non-ambulatory individuals. **Conclusion:** Strong interrater reliability was demonstrated by pediatric clinicians who scored children with SCI using the Pediatric NRS. **Key words:** outcomes, pediatrics, recovery, reliability, spinal cord injury

In concert with the mounting scientific evidence for recovery after SCI via neurotherapeutic interventions,<sup>1-5</sup> there is a growing need for capacity-based and recovery-based measures that can accurately assess neurological and functional recovery. While several adult measures have been developed and validated,<sup>6</sup> including the Neuromuscular Recovery Scale (NRS),<sup>7-11</sup> only a few have been designed and validated specifically for the pediatric SCI population. The Pediatric Measure of Participation (PMoP)<sup>12,13</sup> and the Pediatric Spinal Cord Injury Activity Measure (PEDI-SCI AM)<sup>14</sup> are self-reported outcome instruments that evaluate a child's participation

(PMoP) and what a child can or cannot do (PEDI-SCI AM) in performing mobility, daily routines, and ambulation. While providing important information about child- and parent-reported outcomes, the PMoP and PEDI-SCI AM were not developed to assess functional improvement only as a result of neurological recovery. Instead, like many instruments, they were developed and validated to evaluate functional change as a result of recovery, rehabilitation, compensation, and restorative interventions. Likewise, the Spinal Cord Independence Measure (SCIM-III),<sup>15-18</sup> a well-known disability rating scale of physical function validated in children with SCI,<sup>19</sup> is a

†Deceased

*Supplementary material:* The online version of this article contains eAppendix A.

*Corresponding author:* Andrea L. Behrman, PT, PhD, FAPTA, Department of Neurological Surgery, 220 Abraham Flexner Way, Suite 1521, University of Louisville, Louisville, KY 40202; phone: 502-582-7451; email: andreaBehrman@catholichealth.net

*Top Spinal Cord Inj Rehabil* 2019;25(2):121-131  
© 2019 Thomas Land Publishers, Inc.  
www.thomasland.com  
doi: 10.1310/sci2502-121

capacity measure used to assess functional change, regardless of whether that change is a result of spontaneous recovery or rehabilitation strategies.

To assess recovery of function after SCI, the reference for comparison is specifically how a function or task was performed or accomplished prior to injury. As examples, the Graded Redefined Assessment of Strength, Sensibility and Prehension (GRASSP)<sup>20,21</sup> and Capabilities of the Upper Extremity Test (CUE-T)<sup>22,23</sup> are disease-specific (SCI) capacity instruments that evaluate upper extremity function as compared to typical preinjury function, and compensation is disallowed. The adult NRS applies a similar approach.<sup>7-11,24</sup>

Assessment of recovery in children, especially those injured prior to achieving functional milestones and independence, is difficult. Age-appropriate comparison to a typically developed function is one approach to measuring recovery. This approach may have potential challenges due to variation in function among typically developing children and the lack of empirical data that suggest neurological and functional recovery of children with paralysis mimics typical growth and development. Nonetheless, to capture the broad continuum of function that typifies childhood, and to optimize the sensitivity of a scale to detect recovery or decline in function, it must contain items that assess small incremental changes in capacity to execute movements during functional activities. Such changes also are unaided by compensation or assistive devices and reflect “typical” biomechanical and functional patterns that are expected within a given age group. Moreover, because of the heterogeneity of SCI impairment and the wide range of functional performance that individuals are capable of, an instrument for children with SCI must contain items that provide information about children who have significant, moderate, mild, and no impairment. The Pediatric Neuromuscular Recovery Scale (Pediatric NRS) was developed for the SCI population and accounts for these specific dimensions.

The Pediatric NRS was developed by a team of clinicians and researchers<sup>25</sup> based upon the adult NRS.<sup>10</sup> The highest score for each task of the NRS indicates performance of the task with preinjury movements, comparable to typical function, and

without use of compensation. Subscores for a task represent a hierarchy of recovery stages based on a sequential progression from no recovery to defined recovery. Recovery may be indicated, for example, by greater control of body/limb segments in a cephalo-caudal direction, an increase in performance duration, or greater percent body weight load bearing. Initial content validity of the Pediatric NRS was established.<sup>25</sup> Further psychometric testing is needed on this scale to support potential clinical use.

The purpose of this study was to determine the interrater reliability in the scoring of the Pediatric NRS for children with SCI aged 1 to 12 years representing all levels of injury severity. We hypothesized that the Pediatric NRS would demonstrate good agreement amongst physician (MD), occupational therapist (OT), and physical therapist (PT) raters who were provided an introductory course in the use and scoring of the Pediatric NRS. We also explored the relationship between patient age, injury level, and ambulatory status with interrater reliability.

## Methods

### Participants

All children with SCI in this study were recruited using a database “Human Locomotor Research Center Database for Potential Research Volunteers” (IRB#06.0647) at the University of Louisville. Inclusion criteria included children with single, non-progressive, acquired SCI, currently age 1-12 years, medically stable and asymptomatic for medical complications that might interfere with testing, and medically approved for weight-bearing and study participation. Children with hip flexion contractures  $>20^\circ$ , no history of participation in standing program, lack of medical approval for participation with subluxed or dislocated hips, and ventilator dependence were excluded. Use of these criteria established first, medical clearance for lower extremity weight-bearing; second, overall safety for participation; and third, lower extremity hip range allowing for upright, assisted stance, and stepping. All participants and legal authorized representatives signed informed consent to participate with children older than 7 years also signing assent.

Clinicians were recruited using an Institutional Review Board–approved study flyer (IRB#13.0261) posted on rehabilitation websites and emailed to clinicians. The principal investigator sent emails and held conference calls to delineate study requirements and consent participants. All clinicians who were recruited as raters signed an informed consent to participate. Each clinician (physician, occupational or physical therapists) was required to have 2 years of experience treating children with SCI. At least five were recruited from each category.

### **International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI)**

Physical therapists performed the ISNCSCI<sup>26</sup> examinations on pediatric participants with SCI for classification of injury severity by American Spinal Injury Association Impairment Scale (AIS) level. Therapists completed standardization training with online Wee-STEP training and Instep training (<http://asia-spinalinjury.org/learning/>). ISNCSCI testing and Pediatric NRS assessment occurred on the same day. Children under age 6 were not tested for injury classification due to known unreliability of testing.<sup>27</sup>

### **Videos of children undergoing Pediatric NRS assessment**

Videotaping accommodated multiple raters in different geographic locations. Thirty-two children with SCI were tested using the Pediatric NRS at one pediatric SCI rehabilitation center by pediatric PTs and assisted by another PT and technicians. The administration of the instrument was professionally taped, edited, and produced with lateral and frontal views. Recruitment of children with SCI continued until there were 32 accurate videos available for the raters to score. During clinical testing with the Pediatric NRS, the evaluation ceases when a child fails to meet the recovery-based criteria for performance of a task. To prevent informing the raters of the testing outcome, testers continued to examine children on at least one and up to several items beyond the usual point of cessation (ie, failure). Item videos were then edited with varying outcome points

to facilitate opportunities for clinical decision-making and scoring by the raters. The videos were reviewed and approved by the PI and PTs who participated in developing the instrument.

### **Instrument**

The Pediatric NRS was developed and validated<sup>25</sup> for use in children with SCI who are 1-12 years old. The 13-item scale uses a 12-point rating to evaluate tasks performed in supine, sitting, standing, walking, or during transitional movements with three items that test standing and stepping in the body weight support treadmill environment (BWST). The scoring system rates the task based on performance (1A-4C), with 1A being unable to complete and 4C indicating the task is fully recovered per scale criteria. An overall phase classification is calculated based on summative, individual scores for each task using an algorithm. There are three age categories for the instrument due to age-appropriate differences in motor development and therefore functional performance: 1-2, 3-5, and 6-12 years<sup>25</sup> (see “Phase Cards for Pediatric NRS,” eAppendix A).

### **Online introductory course on the Pediatric NRS for raters**

The raters completed standardized training for administration and scoring of the Pediatric NRS via an online introductory course made available through the NeuroRecovery Training Institute (<http://neurorti.evidencinmotion.com/>). Each rater was provided with their own course, an instructional manual for conducting and rating the Pediatric NRS, and a complete set of color-identified rating cards for each age group. The raters were given 2 weeks to complete the online instructional course. Course advancement was accomplished by each rater providing correct electronic responses to questions following each of the eight learning modules with practice via electronic score sheets and scoring. An introduction module provided context for development and purpose of the Pediatric NRS as a measurement instrument. Each additional module covered two to three items on the scale with video demonstration of how to conduct the

item with a typically developing child of varying ages and accompanied by detailed explanation for set-up, verbal cues, and scoring. The course required approximately 6 hours to complete and culminated with a final assignment and scoring of two pediatric participants with SCI undergoing testing with the Pediatric NRS. An explanation of the scoring outcomes was provided with questions/answers. Fifteen raters completed the online course over an average period of  $16 \pm 5$  days and a range of 11-30 days. Once all raters completed the online course, the process of reviewing and rating online videos began.

The study coordinator uploaded videos periodically and notified the raters via either a message board on the website or an email message when videos were available to review. Videos were released two or three at a time for raters to review within a 2-week period. If needed, raters were given more time to complete reviews. The raters completed an electronic score sheet on each video. Each rater sent the original completed score sheets to the study coordinator. The score sheets were de-identified by the study coordinator. If the raters had questions related to the video, the study coordinator would send only the question to the principal and co-investigator so that the rater remained anonymous. Only questions related to logistics were answered and shared with all raters.

All raters completed the ratings within a 14-month period. Seventeen raters agreed to participation, and 14 completed the study. Availability and time constraints prevented participation by one MD and completion by another MD and an OT.

### Data analysis

Given the ordinal scaling of Pediatric NRS items, the nonparametric, rank-based Kendall coefficient of concordance ( $W$ ) was used to evaluate interrater reliability of the 16 items (13 tasks including three upper extremity items with right and left scores). Confidence intervals for the Kendall statistics were calculated nonparametrically using a bootstrap resampling scheme over 10,000 iterations. To estimate interrater reliability of the summary score, a two-way intraclass correlation coefficient and 95%

bootstrapped confidence interval were calculated. We additionally explored how reliability for the summary score was affected by age, injury level, and ambulatory status by calculating reliability coefficients for groups of participants defined by these factors. All analyses were conducted in the open-source R software package (R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, v. 3.3.0).

## Results

### Child participant and rater characteristics

Demographic and clinical characteristics of the 32 pediatric participants are presented in **Table 1**. Ages ranged from 2 to 12 years with an average of 6 years 3 months, and there were 17 boys and 15 girls. There was representation from each AIS class, although 14 participants under age 6 did not have an AIS rating.<sup>27</sup> The majority of child participants ( $n = 25$ ) were non-ambulatory. Injury levels ranged from C1 to L1.

The characteristics of the 14 raters are in **Table 2**. All raters were female. The raters had an average of 11 years practice experience, 10 years pediatric experience, and 7 years pediatric SCI experience.

### Interrater reliability

We found strong interrater reliability for the summary score (ICC = 0.96; 95% CI, 0.89-0.98). For the individual Pediatric NRS items, 12 of 16 items exhibited concordance coefficients 0.80 or higher, indicating reasonable interrater reliability among the 14 raters (**Table 3**). The interrater reliability of the summary score was consistent across age groups and groups defined by neurological level (**Table 4**). Interrater reliability was lower for non-ambulatory individuals than ambulatory individuals.

Seven items (Supine to Sit, Sit Outside Base of Support [BOS], Reach Overhead [right], Sit to Stand, Dynamic Stand, Walk, Step Adaptability) exhibited lower concordance ( $< 0.80$ ) and/or overly wide confidence intervals. Diagnostic Bland-Altman plots for the Supine to Sit, Sit Outside BOS, Dynamic Stand, and Walk items

**Table 1.** Demographics of pediatric individuals with SCI

Subject	Sex	Age (y)	AIS <sup>a</sup> /Paraplegia AIS <sup>a</sup> /Tetraplegia	Injury/ neurological level	Ambulatory
1	M	2	Paraplegia	L1	Y
2	M	2	Tetraplegia	C4	N
3	M	2	Paraplegia	L5	Y
4	M	3	Paraplegia	T2	N
5	M	3	Paraplegia	T5	N
6	F	3	Tetraplegia	C4	N
7	M	3	Paraplegia	T10	N
8	M	3	Paraplegia	T12	N
9	F	4	Tetraplegia	C5	N
10	M	4	Paraplegia	T11	N
11	F	5	Paraplegia	T4	N
12	F	5	Tetraplegia	C5	N
13	F	5	Tetraplegia	C5	N
14	M	5	Paraplegia	T2	N
15	M	6	A-Paraplegia	T8	N
16	M	6	B-Tetraplegia	C6	N
17	F	6	A-Paraplegia	T2	N
18	F	6	A-Paraplegia	C1	N
19	M	7	D-Paraplegia	C1	Y
20	F	7	B-Paraplegia	T10	N
21	M	8	C-Tetraplegia	C2	N
22	M	8	B-Paraplegia	C1	N
23	M	8	D-Tetraplegia	C2	Y
24	F	8	A-Paraplegia	C2	N
25	F	9	A-Tetraplegia	C7	N
26	F	9	B-Paraplegia	C1	N
27	F	9	C-Paraplegia	T3	N
28	M	10	B-Tetraplegia	C5	N
29	M	11	D-Tetraplegia	C1	Y
30	F	12	D-Paraplegia	T4	Y
31	F	12	A-Paraplegia	C5	N
32	F	12	D-Tetraplegia	C3	Y
Summary	17M, 15F	6 (3)	6A; 5B; 2C; 5D; 14 NT 20 Para 12 Tetra	18C; 12T; 2L	25N; 7Y

Note: Summary values for each column indicate either mean (*SD*) or the number in each category, as appropriate for the data. M = male, F = female; Injury levels: C = cervical, L = lumbar, T = thoracic. American Spinal Injury Association Impairment Scale (AIS): A, B, C, and D. Paraplegia and Tetraplegia noted when participant is too young for AIS testing. Y = Yes, N = No.

<sup>a</sup> Neurological level based on physical examination with International Standards for Neurological Classification of SCI or injury level (most cephalad involved segment) via medical record if <6 years old.

are shown in **Figure 1**. Such plots are typically a method to describe agreement (or disagreement) between two quantitative measurements of the same variable, such as Tukey mean-difference plot or limits of agreement. The elliptical shape

observed in these plots indicates that low and high functioning patients (at the left and right extremes of the graph) were easier for raters to score, while moderate functioning patients (the middle of the graph) were more challenging for raters to assess.

**Table 2.** Demographics of raters

Rater	Sex	Age (y)	PT/OT/MD	Practice experience (y)	Pediatric experience (y)	Pediatric SCI experience (y)
1	F	43	PT	16	5	5
2	F	41	OT	16	16	16
3	F	35	PT	11	11	8
4	F	29	PT	5	4.5	4
5	F	36	MD	5	5	5
6	F	29	OT	5	5	4
7	F	51	PT	20	15	15
8	F	34	PT	7	7	4
9	F	55	MD	27	27	0
10	F	26	OT	2	2	2
11	F	35	OT	12	11	10
12	F	24	OT	1.5	1.5	1
13	F	44	PT	18	18	17
14	F	35	MD	6	9	0
Summary	14 F	37 (9)	6/5/3	11 (7)	10 (7)	7 (6)

*Note:* The summary values are either mean (*SD*) or the total number, as appropriate for the data. MD = medical doctor; OT = occupational therapist; PT = physical therapist; y = years.

The Supine to Sit and Sit Outside BOS items exhibited a full distribution of scores across the horizontal axis, indicating that the participants exhibited a full range of abilities in these items. The broad dispersion along the vertical axis denotes significant variability in ratings among examiners in individual pediatric participants, the source of the lower concordance coefficients for these items. Reach Overhead and Sit to Stand also exhibited this pattern (Bland-Altman plots not shown).

Dynamic Stand and Walk exhibited large gaps in performance along the horizontal axis, indicating a lack of breadth in participant ability. Participants either performed very poorly on the task and scored low on these items or performed very well and scored high, resulting in little variation in scores over which to measure rater disagreement, the source of the lower concordance coefficients for these items. For example, Step Adaptability exhibited scarce between-participant variability (Bland-Altman plot not shown). Of the 448 total ratings for this item (32 participants x 14 raters), phase level 4A was unrepresented, phase level 3C received only a single rating, and levels 3A, 3B,

and 4B were scored in only 3, 5, and 6 instances, respectively.

## Discussion

This is the first report of interrater reliability for scoring motor capacity in children with SCI using the Pediatric NRS among newly trained raters including physicians and physical and occupational therapists with pediatric experience. Interrater reliability ranged from strong (12/16 items) to good (4/16 items) when raters scored professional videos of children age 2-12 years with SCI undergoing Pediatric NRS assessment. Strong interrater reliability also was found for the overall phase summary score. An online training course with demonstration videos and accompanying instructional manual for conducting and scoring the Pediatric NRS was sufficient for raters to exhibit good to strong interrater reliability for scoring.

Items with greater variability of responses of raters across individuals (eg, Supine to Sit and Sit Outside BOS) require consideration. The task

**Table 3.** Interrater reliability of 14 raters for each of the 16 Pediatric Neuromuscular Recovery Scale items and summary score

Item	Coefficient	95% CI
Supine to Sit	0.75	(0.59, 0.85)
Sit Inside BOS	0.84	(0.73, 0.90)
Sit Outside BOS	0.69	(0.48, 0.83)
Object to Mouth: Right	0.87	(0.72, 0.96)
Object to Mouth: Left	0.84	(0.69, 0.93)
In-Hand Manipulation: Right	0.81	(0.68, 0.90)
In-Hand Manipulation: Left	0.86	(0.76, 0.92)
Reaching Overhead: Right	0.76	(0.63, 0.86)
Reaching Overhead: Left	0.80	(0.70, 0.87)
Sit to Stand	0.80	(0.64, 0.90)
Static Stand	0.84	(0.71, 0.92)
Dynamic Stand	0.81	(0.60, 0.94)
Walk	0.78	(0.54, 0.90)
Stand Adapt	0.92	(0.83, 0.97)
Step Retrain	0.98	(0.94, 1.00)
Step Adapt	0.80	(0.66, 0.90)
Summary score <sup>a</sup>	0.96	(0.89, 0.98)

Note: Values are Kendall's coefficient of concordance and nonparametric bootstrapped 95% confidence intervals unless otherwise noted. BOS = base of support.

<sup>a</sup> 2-way intraclass correlation coefficient.

**Table 4.** Interrater reliability of Pediatric Neuromuscular Recovery Scale summary scores by patient groups

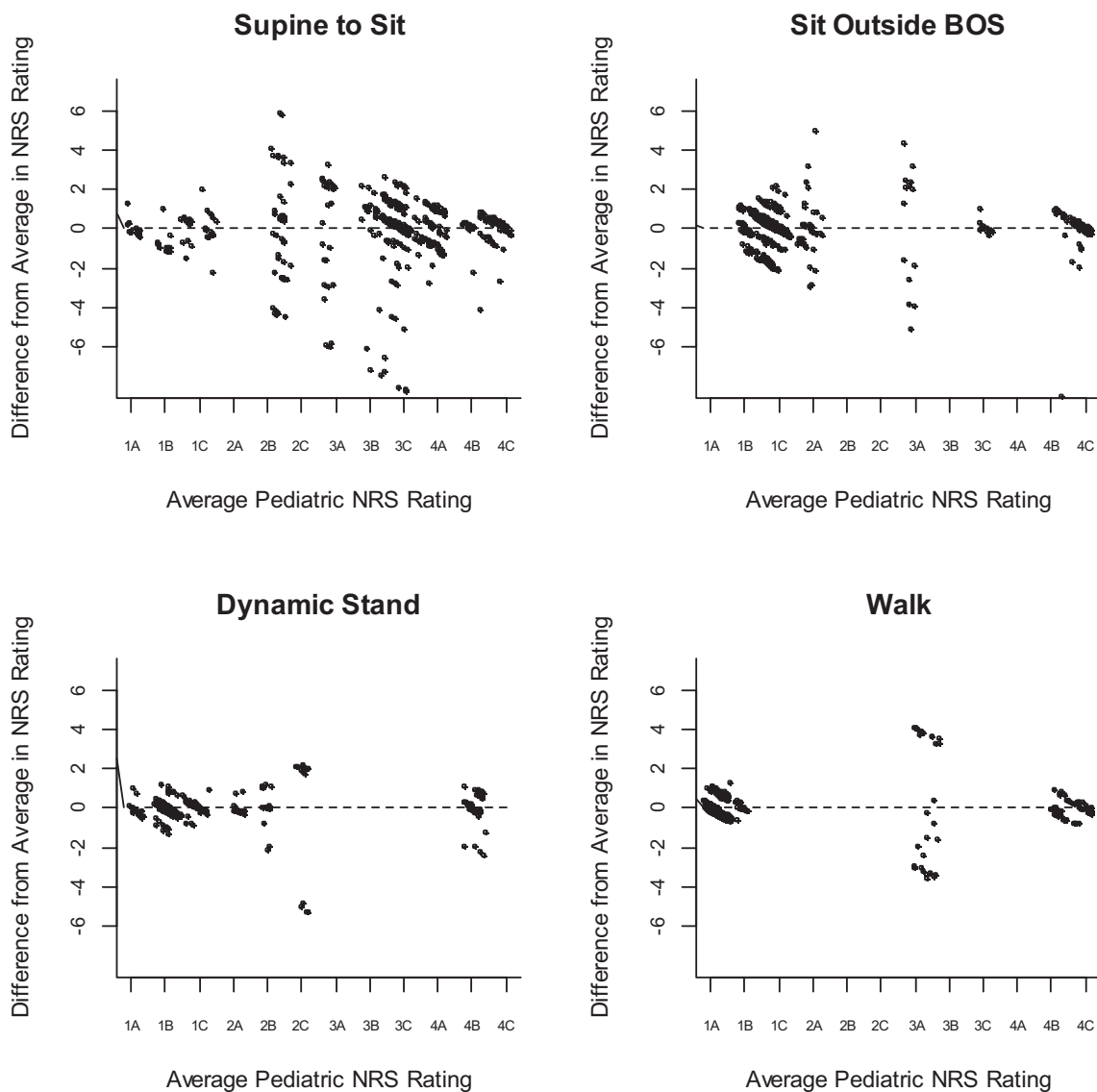
Patient group	ICC (95% CI)
All	0.96 (0.89, 0.98)
Age group	
2-4 ( <i>n</i> = 10)	0.95 (0.89, 0.99)
5-8 ( <i>n</i> = 14)	0.95 (0.91, 0.98)
9-12 ( <i>n</i> = 8)	0.96 (0.91, 0.99)
Neurological level	
Cervical ( <i>n</i> = 18)	0.98 (0.96, 0.99)
Thoracic/Lumbar ( <i>n</i> = 14)	0.97 (0.94, 0.99)
Ambulatory status	
No ( <i>n</i> = 25)	0.81 (0.71, 0.89)
Yes ( <i>n</i> = 7)	0.98 (0.95, 0.99)

Note: Values are two-way intraclass correlation coefficients, 95% CI.

performance itself may have been challenging to score based on descriptive criteria, particularly for those with recovery of partial capacity or mid-level functioning compared to those at the low or high ends of the recovery spectrum (see **Figure 1**). While this difficulty has been similarly reported for a pediatric observational assessment in infants and toddlers with SCI,<sup>28</sup> the preponderance of Pediatric NRS items and reliability outcomes did not reflect this challenge. The Supine to Sit graph (**Figure 1**) shows that participants with moderate levels of recovery exhibited the most disagreement (ie, cloud of points is vertically widest in the middle). In comparison, the Walk graph (**Figure 1**) shows a lack of distribution of recovery levels of capacity among the participants, with one or two exceptions. Thus, most of the participants could either walk or not walk.

Items with a distinctive spread in responses for recovery across the continuum of scores (eg, Dynamic Stand, Walk) may have occurred for several reasons. For example, 78% of the population undergoing the Pediatric NRS and rated for this study was non-ambulatory. A more diverse population representing the injury severity and resultant impairments may more aptly test the reliability of certain items of the Pediatric NRS. The scores for Dynamic Stand and Walk items were fairly distributed at either low or high ends of the Pediatric NRS. Thus, Walk scores indicating low recovery were likely from children with SCI who are non-ambulatory and those with high scores of recovery were ambulatory. These results may be due to the population tested, a limitation of the scale to detecting differences, or interpretation by raters.

The exclusion criteria for this study meant enrollment of only children with the current capacity to fully attempt each task without range of motion limitations, cleared for weight-bearing, and safety/feasibility to complete the assessment. Although other factors may limit patient completion of an assessment and require explanatory comments (eg, wrist in a cast), we focused interrater reliability on assessing those children who could safely attempt performance of the tasks with neuromuscular capacity being the only limiting factor. Clinical judgment may also determine relevance, feasibility, and safety in



**Figure 1.** Bland-Altman plots of four Pediatric Neuromuscular Recovery Scale (NRS) items: Supine to Sit, Sit Outside Base of Support (BOS), Dynamic Stand, Walk. The difference in each scorer's rating from the average for a given pediatric participant are plotted approximately vertically above the participant's average score (horizontal score) and is a measure of variability, that is, disagreement among raters. Points have been randomly staggered by a nominal amount to minimize overlap.

selection of this outcome measure or its specific items for use.

With the establishment of content validity<sup>25</sup> and determination of good to strong interrater reliability, the next steps to establish the psychometric properties of the Pediatric NRS are to examine its test-retest reliability, responsiveness to therapeutic intervention, and item validity (hierarchy) for children with SCI.

#### Study limitations

First, 32 children were selected to represent the population of children with SCI, the breadth of injury severity, and the range of recovery. Not all item levels of recovery, however, were illustrated by the study population with 78% non-ambulatory. This may have accounted for a lack of spread in Pediatric NRS scores for several items. Second,



Pediatric NRS evaluations were presented to raters via video. Video presentations may not have aptly captured a child's movements as if viewed in person. Video presentation though did allow the raters the opportunity to review a child's performance with lateral/front views several times and pause to consider rating decision. Video evaluation also removed any variability in the administration of the Pediatric NRS among clinicians, which may have upwardly biased our estimates of reliability. A study design utilizing in-person assessments of the Pediatric NRS was infeasible, requiring children to repeat the assessment for each of 14 raters, and it may have been subjected to fatigue or familiarity bias. Video review is also a common strategy used by clinicians to reassess scoring after conducting an in-person patient evaluation. Thus, video review is also a valid medium for scoring pediatric patient performance on a motor capacity scale.<sup>29</sup> Third, item rating variability may reflect differences in interpretation of scoring instructions. Though all raters completed and passed a standardized training course for scoring the Pediatric NRS, identifying scoring nuances may be helpful. Performance of children with partial paralysis, furthermore, may create a challenge for scoring. A larger scope of instructional video examples and explanations may further assist competency and standardization. Fourth, the Pediatric NRS was novel to raters before study participation. Examiners who gain experience in conducting the Pediatric NRS likely will demonstrate greater interrater reliability than those participating in this study as first-time raters.<sup>11</sup> Though raters in this study were clinician raters with experience in the field of pediatric SCI, we anticipate use of the Pediatric NRS by clinicians from entry-level to expert. As noted from our findings, familiarity in identifying nuances and even typical, compensatory patterns may enhance the learning process for use of the Pediatric NRS by clinicians and be particularly beneficial to the entry-level clinician. Lastly, the Pediatric NRS was developed to evaluate children age 1-12 with acquired SCI, accounting for developmental differences across ages.<sup>25</sup> Our eligibility criteria for participation as a child with SCI in this study was thus identical to this age range, age 1-12 years. The sample of children videotaped for the study were enrolled

based on convenience for recruitment and accessibility for participation. The sample reflected only the range of 2-12 years. Thus, we were unable to test the youngest of ages, age 1, for interrater reliability with the Pediatric NRS.

### **Knowledge translation**

The Pediatric NRS is intended to serve as a capacity-based measure for use in the context of assessing change relative to neurotherapeutic interventions, such as activity-based therapies (eg, neuromuscular electrical stimulation, locomotor training) in children, age 1-12 years, with acquired SCI. Assessment of neuromuscular recovery and capacity after injury is conducted in comparison to preinjury, task-specific movement patterns or developmentally-similar movements for pediatric populations. The online instructional course for the introduction to the scoring and conduct of the Pediatric NRS requires approximately 6 hours to complete. Descriptions of the items and score cards are readily available through publications. Content validity and interrater reliability (in experienced clinicians) have been established for the Pediatric NRS. Further psychometric testing of the Pediatric NRS is warranted to provide clear evidence of its utility in the clinic and/or in research for measuring responsiveness to neurotherapeutic and possibly biological interventions for upper extremity, trunk, and lower extremity neuromuscular capacity in children with acquired SCI, age 1-12 years.

### **Conclusion**

The Pediatric NRS demonstrated good to strong interrater reliability when scoring children with SCI aged 2-12 years via video review by novice, yet trained, raters (pediatric physicians and occupational and physical therapists having completed an introductory training course). Established content validity<sup>25</sup> and current findings for interrater reliability contribute to the relevant and necessary psychometric evaluation of the Pediatric NRS as a capacity- and recovery-based instrument for use with children with SCI. Further research should include methods to systematically assess intrarater reliability, validity of within-task hierarchy, and responsiveness to treatment.

### Conflicts of Interest

Dr. Ardolino reports grants from the Craig H. Neilsen Foundation during the conduct of the study and non-financial support from NeuroRecovery Training Institute outside the submitted work.

Dr. Behrman reports grants from the Craig H. Neilsen Foundation and from Kosair Charities during the conduct of the study and personal fees from NeuroRecovery Training Institute, other from NeuroRecovery Learning, Inc., and personal fees from Shriners Research Foundation outside the submitted work.

Dr. Mulcahey reports grants from Jefferson University (Philadelphia University + Thomas Jefferson University) during the conduct of the study and personal fees from *Topics in Spinal Cord Injury Rehabilitation*, grants from Craig H. Neilsen Foundation, and personal fees from Craig H. Neilsen Foundation outside the submitted work.

Dr. Roberts reports grants from NeuroRecovery Network during the conduct of the study.

Ms. Clayton, Ms. Gregg, and Dr. Lorenz have nothing to disclose.

### Funding Support

This work was supported by the Craig H. Neilsen Foundation (Award n. 260284).

### Acknowledgments

The authors thank Laura Westberg and Dori Goya for their administrative support in the execution of the award and preparation of this article. The authors especially acknowledge the contribution of Shelley A. Trimble, career pediatric physical therapist, who was instrumental in the development of the Pediatric NRS. Shelley passed away on November 15, 2016, due to metastatic breast cancer. We were privileged to work with Shelley and recognized the ongoing impact of her work in pediatric neurorecovery and specifically with the Pediatric NRS through this award. Last, and with much appreciation, we thank the children and their families who gave their time to participate in this study.

---

### REFERENCES

- Behrman AL, Nair PM, Bowden MG, et al. Locomotor training restores walking in a nonambulatory child with chronic, severe, incomplete cervical spinal cord injury. *Phys Ther*. 2008;88(5):580-590.
- Kapadia N, Zivanovic V, Popovic MR. Restoring voluntary grasping function in individuals with incomplete chronic spinal cord injury: Pilot study. *Top Spinal Cord Inj Rehabil*. 2013;19(4):279-287.
- Mayr W, Krenn M, Dimitrijevic MR. Epidural and transcutaneous spinal electrical stimulation for restoration of movement after incomplete and complete spinal cord injury. *Curr Opin Neurol*. 2016;29(6):721-726.
- Rejc E, Angeli CA, Bryant N, Harkema SJ. Effects of stand and step training with epidural stimulation on motor function for standing in chronic complete paraplegics. *J Neurotrauma*. 2017;34(9):1787-1802.
- Angeli CA, Edgerton VR, Gerasimenko YP, Harkema SJ. Altering spinal cord excitability enables voluntary movements after chronic complete paralysis in humans. *Brain*. 2014;137(Pt 5):1394-1409.
- Jones LAT, Bryden A, Wheeler TL, et al. Considerations and recommendations for selection and utilization of upper extremity clinical outcome assessments in human spinal cord injury trials. *Spinal Cord*. 2018;56(5):414-425.
- Behrman AL, Velozo C, Suter S, Lorenz D, Basso DM. Test-retest reliability of the Neuromuscular Recovery Scale. *Arch Phys Med Rehabil*. 2015;96(8):1375-1384.
- Velozo C, Moorhouse M, Ardolino E, et al. Validity of the Neuromuscular Recovery Scale: A measurement model approach. *Arch Phys Med Rehabil*. 2015;96(8):1385-1396.
- Tester NJ, Lorenz DJ, Suter SP, et al. Responsiveness of the neuromuscular recovery scale during outpatient activity-dependent rehabilitation for spinal cord injury. *Neurorehabil Neural Repair*. 2016;30(6):528-538.
- Behrman AL, Ardolino E, Vanhiel LR, et al. Assessment of functional improvement without compensation reduces variability of outcome measures after human spinal cord injury. *Arch Phys Med Rehabil*. 2012;93(9):1518-1529.
- Basso DM, Velozo C, Lorenz D, Suter S, Behrman AL. Interrater reliability of the Neuromuscular Recovery Scale for spinal cord injury. *Arch Phys Med Rehabil*. 2015;96(8):1397-1403.
- Mulcahey M, Calhoun CL, Tian F, Ni P, Vogel LC, Haley SM. Evaluation of newly developed item banks for child-reported outcomes of participation following spinal cord injury. *Spinal Cord*. 2012;50(12):915-919.
- Mulcahey M, Slavin MD, Ni P, et al. The Pediatric Measure of Participation (PMoP) Short Forms. *Spinal Cord*. 2016;54(12):1183-1187.

14. Slavin MD, Mulcahey M, Calhoun Thielen C, et al. Measuring activity limitation outcomes in youth with spinal cord injury. *Spinal Cord*. 2016;54(7):546-552.
15. Catz A, Iitzkovich M, Agranoy E, Ring H, Tamir A. SCIM--spinal cord independence measure: A new disability scale for patients with spinal cord lesions. *Spinal Cord*. 1997;35(12).
16. Iitzkovich M, Gelernter I, Biering-Sorenson F, et al. The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study. *Disabil Rehabil*. 2007;29(24):1926-1933.
17. Anderson KD, Acuff ME, Arp BG, et al. United States (US) multi-center study to assess the validity and reliability of the Spinal Cord Independence Measure (SCIM III). *Spinal Cord*. 2011;49(8):880-885.
18. Bluvshstein V, Front L, Iitzkovich M, et al. SCIM III is reliable and valid in a separate analysis for traumatic spinal cord lesions. *Spinal Cord*. 2011;49(2):292-296.
19. Mulcahey M, Thielen CC, Sadowsky C, et al. Despite limitations in content range, the SCIM-III is reproducible and a valid indicator of physical function in youths with spinal cord injury and dysfunction. *Spinal Cord*. 2018;56(4):332-340.
20. Velstra IM, Fellinghauer C, Abel R, Kalsi-Ryan S, Rupp R, Curt A. The Graded and Redefined Assessment of Strength, Sensibility, and Prehension Version 2 provides interval measure properties. *J Neurotrauma*. 2018;35(6):854-863.
21. 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.
22. Marino RJ, Patrick M, Albright W, et al. Development of an objective test of upper-limb function in tetraplegia: The Capabilities of Upper Extremity Test. *Am J Phys Med Rehabil*. 2012;91(6):478-486.
23. Marino RJ, Kern SB, Leiby B, Schmidt-Read M, Mulcahey MJ. Reliability and validity of the Capabilities of Upper Extremity Test (CUE-T) in subjects with chronic spinal cord injury. *J Spinal Cord Med*. 2015;38(4):498-504.
24. Harkema SJ, Behrman AL, Barbeau H. *Locomotor Training: Principles and Practice*. New York: Oxford University Press, Inc.; 2011.
25. Ardolino EM, Mulcahey MJ, Trimble S, et al. Development and initial validation of the Pediatric Neuromuscular Recovery Scale. *Pediatr Phys Ther*. 2016;28(4):416-426.
26. Kirshblum SC, Waring W, Biering-Soerensen F, et al. Reference for the 2011 revision of the International Standards for Neurological Classification of Spinal Cord Injury. *J Spinal Cord Med*. 2011;34(6):547-554.
27. Mulcahey M, 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(8):1264-1269.
28. Calhoun CL, Gaughan JP, Chafetz RS, Mulcahey M. A pilot study of observational motor assessment in infants and toddlers with spinal cord injury. *Pediatr Phys Ther*. 2009;21(1):62-67.
29. Hansen L, Erhardtsen KT, Bencke J, Magnusson SP, Curtis DJ. The reliability of the Segmental Assessment of Trunk Control (SATCo) in children with cerebral palsy. *Phys Occup Ther Pediatr*. 2018;38(3):291-304.