ORIGINAL ARTICLE

Interrater Reliability of the International Standards for Neurological Classification of Spinal Cord Injury in Youths With Chronic Spinal Cord Injury

Mary Jane Mulcahey, PhD, OTR\L, John P. Gaughan, PhD, Ross S. Chafetz, DPT, Larry C. Vogel, MD, Amer F. Samdani, MD, Randal R. Betz, MD

ABSTRACT. 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;92:1264-9.

Objectives: To evaluate the interrater reliability of the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) in children with chronic spinal cord injury (SCI), and to define the lower age limit at which the examinations have clinical utility.

Design: Repeated measures, multicenter reliability study.

Setting: Two U.S. pediatric specialty hospitals with recognized SCI programs.

Participants: Children (N=236) with chronic SCI.

Interventions: Subjects underwent 4 examinations by 2 raters: sensory tests (pin prick [PP] and light touch [LT]), a motor test, and a test of anal sensation (AS) and anal contraction (AC).

Main Outcome Measures: A 2-way general linear model analysis of variance was used for analysis. Intraclass correlation coefficients (ICCs) and 95% confidence intervals were calculated for PP, LT, motor, AS, and AC.

Results: No child younger than 6 years completed the examination. When examined as a function of age, interrater reliability for motor, PP, LT, AS, and AC was moderate (ICC=.89) to high (ICC=.99). There was poor reliability for AS (ICC=.49) in subjects with complete injuries but moderate reliability for all other variables. There was moderate to high reliability for classification of type (tetraplegia/paraplegia) and severity (complete/incomplete) of injury across age groups.

Conclusions: The ISNCSCI does not have utility for children younger than 6 years. For children older than 6 years, interrater reliability of PP, LT, and motor examinations is high.

Key Words: Paraplegia; Quadriplegia; Rehabilitation; Neurological impairment; Tetraplegia.

THE INTERNATIONAL STANDARDS for Neurological Classification of Spinal Cord Injury (ISNCSCI)[1,2] provide a method for the neurologic evaluation of persons after spinal cord injury (SCI) and for the classification of the neurologic consequence of the injury. The neurologic assessments, which include the motor, sensory, and anorectal examinations, provide the basis for classifying the neurologic level, motor scores and motor level, sensory scores and sensory level, the zone of partial preservation, and the degree of impairment or severity of the SCI according to the American Spinal Injury Association (ASIA) Impairment Scale (AIS).[1]

The sensory, motor, and anorectal examination techniques and the classification methodology of the ISNCSCI have been well described[1,2] and are summarized here. The sensory examination involves testing of 28 dermatomes on the right and left side of the body for sensitivity to pin prick (PP) and light touch (LT). The motor examination is completed through the testing of 10 muscles bilaterally. The strength of each muscle is graded on an ordinal scale from 0 (complete paralysis) to 5 (normal active movement, full range of motion against full resistance). These scores are summed across myotomes and sides of the body to generate a single total motor (TM) score. There has been a recommendation to separate the upper and lower extremity motor (UEM and LEM, respectively) scores because they can better predict functional activities on the motor FIM.[3]

The anorectal examination involves the evaluation of sensation and contraction of the external anal sphincter. For this, the examiner applies pressure with the index finger to the rectal wall to test for anal sensation (AS) and, to test for anal contraction (AC), makes the request of the person being examined to squeeze as if holding a bowel movement. Classification of the motor, sensory, and neurologic levels are determined based on the motor and sensory examination findings. The sensory level is defined as the most caudal dermatome with bilateral normal sensation to both PP and LT. Likewise, the

List of Abbreviations

AC	anal contraction
AIS	ASIA Impairment Scale
AS	anal sensation
ASIA	American Spinal Injury Association
CI	confidence interval
ICC	intraclass correlation coefficient
ISNCSCI	International Standards for Neurological Classification of Spinal Cord Injury
LEM	lower extremity motor
LT	light touch
PP	pin prick
SCI	spinal cord injury
TM	total motor
UEM	upper extremity motor

From the Shriners Hospitals for Children, Philadelphia, PA (Mulcahey, Betz, Samdani, Chafetz); Shriners Hospitals for Children, Chicago, IL (Vogel); and Biostatistical Consulting Center, Temple University School of Medicine, Philadelphia, PA (Gaughan).

Presented in part to the American Academy of Physical Medicine and Rehabilitation, November 6, 2010, Seattle, WA.

Supported by Shriners Hospitals for Children (grant no. 8956).

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Correspondence to Mary Jane Mulcahey, PhD, OTR/L, Shriners Hospitals for Children, 3551 N Broad St, Philadelphia, PA 19140, e-mail: *mmulcahey@shrinenet.org*. Reprints are not available from the authors.

0003-9993/11/9208-01057$36.00/0
doi:10.1016/j.apmr.2011.03.003

single motor level is the most caudal level where there is bilateral strength of at least 3/5, assuming the remaining rostral key muscles are all 5/5. The neurologic level is the lowest spinal segment where sensory and motor functions are normal on both sides.

The AIS is based on sacral segments and the key muscle scores. The designation of complete or incomplete is based on sensory and motor findings at S4-5 and\or presence of voluntary contraction and\or deep pressure during anorectal examination.[2] If there is sacral sparing as evidenced by preservation of PP and LT sensation in the S4-5 dermatome, preservation of anorectal sensation (deep pressure), or volitional anal contraction, then the SCI is an incomplete injury. For incomplete injuries, the designation of AIS "B" refers to sensory incomplete, whereas designations of AIS "C" and "D" refer to motor incomplete with either gravity eliminated or against gravity motor function, respectively.

The ISNCSCI has undergone several revisions of both examination and classification techniques in adults with SCI.[4-9] In their current form,[10] they are the recommended method for measuring neurologic outcomes in persons with SCI.[11] Despite widespread use of the ISNCSCI in clinical practice and clinical trials, with few exceptions,[12,13] reliability studies on the motor and sensory examinations have been conducted only in adults, using relatively small samples. Until recently, there has been a serious void in studies that evaluate the use of the anorectal examination as an indicator of injury severity.[13,14] Despite the limited empirical work on the ISNCSCI examinations and classification techniques, as in adult practice, they are used routinely in children and adolescents as a way to diagnose their injury, prognosticate recovery, and define potential outcomes of rehabilitation.

This multicenter study had 3 main goals. The first was to evaluate the interrater reliability of the ISNCSCI examinations as functions of age and type of injury (tetraplegia/paraplegia) in children with chronic SCI. The second goal was to evaluate the reliability of the ISNCSCI classification in children with chronic SCI. The third goal was to use the data generated from the study to establish guidelines for use of the ISNCSCI with children. Results of intrarater reliability, which showed good to strong reliability of total motor and sensory scores on repeated examinations by the same rater, are published elsewhere.[12,13,15,16] To our knowledge, this study is the first to report the results of interrater (among raters) reliability for the ISNCSCI examinations and classification in children and adolescents.

METHODS

The design was a repeated measures, multicenter reliability study. The study protocol was approved by the institutional review boards of record at each participating center. Written informed consent was obtained from the legal guardian of each subject younger than 18 years. Children between 7 and 18 years of age also provided written informed assent. Subjects 18 years and older provided their own consent. The institutional review board-approved Health Insurance Privacy and Portability Act forms were also reviewed with subjects for their consent.

Sample

The sample was one of convenience consisting of youths between 3 months and 21 years of age. To minimize variation in repeated test scores resulting from actual neurologic changes, by study design, we enrolled youths with chronic SCI who were not changing neurologically. Children were recruited from 2 U.S. pediatric orthopedic SCI specialty hospitals. Children were excluded if they had any neurologic changes within the last 3 months, if they were receiving mechanical ventilation without an effective method to communicate responses during test sessions, or if neurologic comorbidities existed that would influence the neurologic examination (eg, brachial plexus injury, traumatic brain injury). The scores from the first of 4 examinations were used to classify subjects' level and severity of injury unless the first examination differed from the 3 subsequent examinations and all 3 subsequent examinations had the same scores; under this condition, subjects were classified according to the 3 subsequent examinations. Although this report and the articles by Vogel,[13] Samdani,[14] and Chafetz[15] and colleagues report on distinct types of reliability outcomes and used different stratifications and statistical methods for analysis, the subjects were drawn from the same sample.

Data Collection

Seven raters who were formally trained in the evaluation[16] and classification[17] methods of the ISNCSCI performed all evaluations. Formal training included a series of lectures on testing techniques, viewing of the testing technique video published by the ISNCSCI training packet, and hands-on practice with immediate feedback from the instructor. The training session was provided by an expert physical therapist external to the institutions participating in the study and who has conducted formal competency programs on the ISNCSCI examination and classification techniques for international clinical trials. Three raters carried out the examinations at 1 institution, and the 4 other raters conducted them at the second institution. All the subjects participated in 4 repeated ISNCSCI examinations conducted by 2 different raters who performed 2 examinations each, on 4 separate days. The time period between the 2 examinations ranged from 24 hours to 4 days (average time, 2.4 days). The examinations were conducted using the standardized techniques published by the ASIA.[1,2] As an effort to improve the standardization of the anorectal examinations, for this study, AS was tested by having the examiner gently apply pressure to the rectal wall a minimum of 3 times. If, without prompting, subjects accurately identified the examiner applying pressure, they were scored as having rectal sensation. If their response was inconsistent, a minimum of 8 of 10 accurate responses was required for designating the injury as incomplete. After sensation testing, subjects were asked to squeeze "as if to hold in a bowel movement." Standardization of examination[16] and scoring[17] among the raters was implemented before data collection. Scores were documented immediately throughout the examinations using the ISNCSCI form that was modified, with permission, for this study (fig 1). Data underwent double entry into a secure database by research assistants blinded to the study. Classification of neurologic level, motor level, sensory level, injury severity, and AIS were determined using the examination scores and standard techniques published by ASIA.[1] Classification was confirmed using a computer program.[18]

Data Analysis

Data were deidentified for analyses by the biostatistician who was blinded to the subjects' injury characteristics. As a result of failing the Wilk-Shapiro test of normality, data were transformed to normalized ranks to accommodate the nonnormality of ordinal scale measurements and allow use of parametric methods.[19,20] The use of rank transformation before analysis of variance is well established and has been described elsewhere as a way to bridge between parametric and nonpara-