

Agreement in Neurological Classification of Spinal Cord Injury using the 2002 Standards

Ralph J. Marino, MD (1); Steven Kirshblum, MD (2); Linda Jones, PT, MS (3); Joseph Tal, PhD (4)

(1) Thomas Jefferson University, Philadelphia, PA; (2) Kessler Institute for Rehabilitation, West Orange, NJ; (3) Proneuron Biotechnologies, Englewood, CO; (4) TechnoStat Ltd., Kfar Saba, Israel

ABSTRACT

Background: Despite its widespread use, the accuracy of classification of spinal cord injury (SCI) using the 2002 International Standards for Neurological Classification (the *Standards*) has not been evaluated. That is the purpose of this study.

Design: Pre and Post test classification of sample cases by eight physicians attending an investigators meeting for the Proneuron *Phase II autologous incubated macrophage study for treatment of acute SCI patients*.

Methods: Subjects determined sensory and motor levels, Complete vs. Incomplete, and ASIA Impairment Scale (AIS) grade on 15 sample cases. Five cases were identical to those used by Donovan (1990). The other ten cases were developed by one author (RJM) to illustrate known areas of difficulty. A general discussion of classification followed. The test was repeated on the following day. Statistical analysis included percent correct responses (according to 2002 *Standards*) for all cases and the 5 Donovan cases. Cohen's kappa statistic (chance-corrected agreement) was calculated for AIS grades.

Results: Pretest average agreement per case was 95.4% for sensory and 92.5% for motor level, with a range of 80-100% except for one case (44% for motor level). Agreement for completeness was 92.5%. Agreement for AIS grade was 77.5%, with a kappa of 0.68. For the five Donovan cases, pretest scores were 93.8% for sensory level, 92.5% for motor level, and 95% for completeness. Agreement for AIS grade was 87.5%, with a kappa of 0.81. Posttest results showed improved agreement on Donovan AIS grades, 95% with a kappa of 0.92. Borderline AIS grades, motor level in regions without a key muscle, and cases with associated injuries were most problematic.

Conclusion: Agreement in classification of SCI by experienced clinicians is improved using the 2002 *Standards* compared to earlier versions, and is generally within desired range for outcome measures or clinical trials. A structured approach to classification should enhance accuracy.

INTRODUCTION

The *International Standards for Classification of Spinal Cord Injury (the Standards)*, developed by the American Spinal Injury Association (ASIA) are widely used for assessing and classifying patients with spinal cord injury (SCI) in both research and clinical settings. Previous studies on reliability of classification using the *Standards* are limited and demonstrated areas of confusion. Past versions of the *Standards* were revised to improve reliability of classification. The latest version of the *Standards* to be tested for reliability was the 1992 edition. Subsequently the *Standards* were revised in 1996 and in 2000; the booklet was reprinted in 2002 with corrections to the 2000 edition. The second edition of the Reference Manual (2003) was published last year. No studies have been published that examine the reliability of classification of SCI using the current edition of the *Standards* (2002).

METHODS

Proneuron Biotechnologies has initiated a clinical trial titled "A Phase II, Randomized, Controlled Study to Evaluate the Safety and Efficacy of Autologous Incubated Macrophages for the Treatment of Patients with Complete Spinal Cord Injury". The *Standards* are used to determine eligibility for the study (level and completeness) and as outcomes (motor/sensory scores and ASIA Impairment Scale (AIS) grade). To ensure accuracy of classification of SCI by the investigators, a training session was conducted. Subjects consisted of eight physician investigators. On day 1, subjects determined sensory and motor levels, Complete vs. Incomplete, and AIS grade on 15 sample cases. Five cases were identical to those used by Donovan (1990), and later by Priebe (1991). The other ten cases were developed by one author (RJM) to illustrate known areas of difficulty.

A general discussion of classification followed the pretest, but answers to specific cases were not given. On the following day, the subjects were asked to classify the same 15 cases.

ANALYSIS

Statistical analysis included percent correct responses (according to 2002 *Standards*) for all cases and the 5 Donovan cases. Cohen's kappa statistic (chance-corrected agreement) was calculated for AIS grades. Correct responses were determined by the lead author.

RESULTS

Agreement for sensory and motor levels, completeness, and AIS grades are found in figures 1 and 2.

Figure 1: PI agreement for all cases

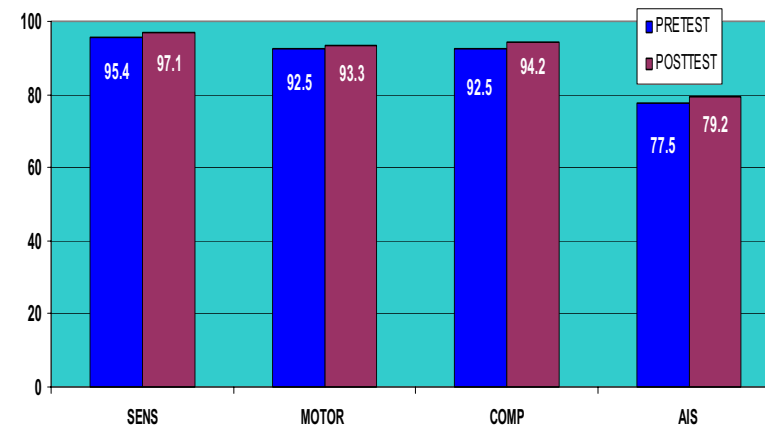
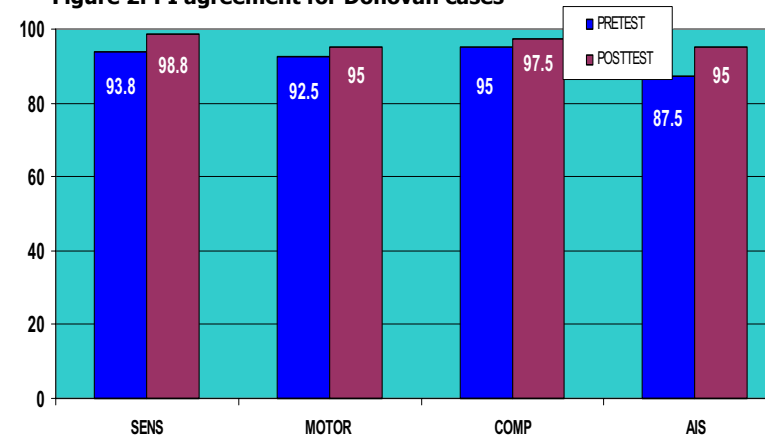


Figure 2: PI agreement for Donovan cases



For all cases, agreement was excellent for sensory and motor levels and for completeness, and good for AIS grades. For the Donovan cases, agreement was excellent except for the pretest AIS grading. There was a small improvement from pretest to posttest.

Results for determining AIS grade are found in tables 1 and 2. Rows indicate "correct" responses and columns indicate rater responses. Of the 15 cases, 5 were AIS grade A, 3 were grade B, 5 were grade C, and 2 were grade D. Raters generally were highly accurate at determining if injuries were complete (39/40 ratings correct on pretest), but had some difficulty with motor complete vs. incomplete and distinguishing AIS C from D. For the Donovan cases, agreement was 87.5% on the pretest and 95% on the posttest.

In evaluating the cases and errors, mistakes in classification occurred most often where rules are somewhat complicated, or where there may be disagreement with the current definitions. This occurs for motor levels at transition zones where no muscle is available to test, namely C3-C5 and T12-L2. For AIS grade, the borderline cases where more or less than 50% of muscles are at least grade 3, depending on where one starts counting, were problematic. Cases with other neurological deficits, such as paraplegia with a brachial plexus injury, also created disagreement in classification.

Table 1: Pretest AIS agreement for all cases

Rows contain correct grade, columns rater grades. Green cells are correct responses, light blue are incorrect.

		ASIA: Pre				Group Total	
		A	B	C	D		
ASIA Key	A	Count	39		1		40
	B	Count	2	16	6		24
	C	Count	5	5	27	3	40
	D	Count			5	11	16
Group Total		Count	46	21	39	14	120

Table 2: Posttest AIS agreement for all cases

See table 1 legend for explanation

		ASIA: Post				Group Total	
		A	B	C	D		
ASIA Key	A	Count	38		2		40
	B	Count	2	13	8	1	24
	C	Count	3	4	32	1	40
	D	Count			4	12	16
Group Total		Count	43	17	46	14	120

Table 3: Agreement in Classification

	Donovan 1990	Priebe 1991	Marino 2004
Sensory Level	.81	.93	.99
Motor Level	.66	.85	.95
Frankel/AIS – agreement	.73	.81	.95
Frankel/AIS – kappa	NA	.67	.92

DISCUSSION

This evaluation of classification of SCI using the International Standards indicates that improvements have been made in interrater reliability and agreement. Table 3 compares testing across versions of the Standards. For the Donovan cases, agreement increased for sensory, motor and AIS grading.

Consistency in classification depends on both the quality of data collection and recording and the quality of classifying recorded data. Comments or notes are needed when there are associated injuries, such as a brachial plexus injury, which results in abnormal sensory and motor data not related to a SCI. It is imperative that muscle scores less than 5 are given only when neurological weakness is thought to be the cause, not when pain or disuse result in suboptimal patient effort. Muscles that are not testable should be labeled as such.

Accurate classification using a completed neurological form can best be obtained by proceeding in the following order, as described in the Reference Manual for the Standards (2003):

STEPS IN CLASSIFICATION

- Determine the sensory level.
- Determine the motor level.
- Determine whether lesion is complete or incomplete (sacral sparing definition)
- Determine AIS grade.
 - Is injury complete? YES= AIS A, else
 - Is injury motor complete? YES = AIS B, else
 - Is injury C or D - count key muscles below single neurological level; if at least half > grade 3, then D, else C.

It will not be possible to classify 100% of patients. Some patients will have associated injuries or conditions in critical places causing necessary scores to be missing, which has implications for clinical trials. While in clinical practice one wants to classify the SCI as accurately as possible, uncertainty is incompatible with clinical trials.

Inclusion/exclusion criteria should be established to exclude individuals who have critical information missing, or whose patterns of deficits are unusual and whose classification is controversial. For example, in study of an intervention for people with complete SCI, it would be prudent to exclude patients with cervical injuries who have no sacral sparing but have some movement in their legs. These patients are technically complete with an extended motor zone of partial preservation, but many clinicians think such patients have recovery similar to motor incomplete patients. If the primary outcome measure is change in motor score, one may not want to include in a trial individuals with muscles that are not-testable due to fractures, contractures, or other conditions.

CONCLUSION

The current version of the Standards has good to excellent agreement in classification except for unusual patterns of deficits. Trained examiners using a consistent approach to classification should achieve high levels of agreement. For clinical trials, inclusion/exclusion criteria can be used to eliminate unusual cases that may bias results.

REFERENCES

- Donovan WH, Wilkerson MA, Rossi D, Mechoulam F, Frankowski RF. A test of the ASIA guidelines for classification of spinal cord injuries. *J Neuro Rehab.* 1990;4:39-53.
- International Standards for Neurological and Functional Classification of Spinal Cord Injury.* Chicago, IL: American Spinal Injury Association; 2002.
- Priebe MM, Waring WP. The interobserver reliability of the revised American Spinal Injury Association standards for neurological classification of spinal injury patients. *Am J Phys Med Rehabil.* 1991;70(5):268-270.
- Reference Manual for the International Standards for Neurological Classification of Spinal Cord Injury.* Chicago, IL: American Spinal Injury Association; 2003.