

## ORIGINAL ARTICLE

# International standards for neurological classification of spinal cord injury: classification skills of clinicians versus computational algorithms

C Schuld<sup>1</sup>, S Franz<sup>1</sup>, HJA van Hedel<sup>2,3</sup>, J Moosburger<sup>4</sup>, D Maier<sup>5</sup>, R Abel<sup>6</sup>, H van de Meent<sup>7</sup>, A Curt<sup>2</sup>, N Weidner<sup>1</sup>, EMSCI study group<sup>8</sup> and R Rupp<sup>1</sup>

**Study design:** This is a retrospective analysis.

**Objectives:** The objective of this study was to describe and quantify the discrepancy in the classification of the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) by clinicians versus a validated computational algorithm.

**Settings:** European Multicenter Study on Human Spinal Cord Injury (EMSCI).

**Methods:** Fully documented ISNCSCI data sets from EMSCI's first years (2003–2005) classified by clinicians (mostly spinal cord medicine residents, who received in-house ISNCSCI training by senior SCI physicians) were computationally reclassified. Any differences in the scoring of sensory and motor levels, American Spinal Injury Association Impairment Scale (AIS) or the zone of partial preservation (ZPP) were quantified.

**Results:** Four hundred and twenty ISNCSCI data sets were evaluated. The lowest agreement was found in motor levels (right: 62.1%,  $P=0.002$ ; left: 61.8%,  $P=0.003$ ), followed by motor ZPP (right: 81.6%,  $P=0.74$ ; left 80.0%,  $P=0.27$ ) and then AIS (83.4%,  $P=0.001$ ). Sensory levels and sensory ZPP showed the best concordance (right sensory level: 90.8%,  $P=0.66$ ; left sensory level: 90.0%,  $P=0.30$ ; right sensory ZPP: 91.0%,  $P=0.18$ ; left sensory ZPP: 92.2%,  $P=0.03$ ). AIS B was most often misinterpreted as AIS C and vice versa (AIS B as C: 29.4% and AIS C as B: 38.6%).

**Conclusion:** Most difficult classification tasks were the correct determination of motor levels and the differentiation between AIS B and AIS C/D. These issues should be addressed in upcoming ISNCSCI revisions. Training is strongly recommended to improve classification skills for clinical practice, as well as for clinical investigators conducting spinal cord studies.

**Sponsorship:** This study is partially funded by the International Foundation for Research in Paraplegia, Zurich, Switzerland. *Spinal Cord* (2015) 53, 324–331; doi:10.1038/sc.2014.221; published online 9 December 2014

## INTRODUCTION

The International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI)<sup>1</sup> are widely accepted among clinicians and researchers for classification of the location, severity and extent of a human spinal cord injury (SCI). ISNCSCI has been well investigated<sup>2</sup> throughout its >40-year history. The severity classification was fundamentally revised during this time period. In detail, the American Spinal Injury Association (ASIA) Impairment Scale (AIS), which superseded<sup>3</sup> the original Frankel Scale,<sup>4</sup> introduced several changes aiming at a more consistent and objective description, on the basis of the scores of a standardized examination of myotomes and dermatomes.<sup>5</sup>

The six-point manual muscle test<sup>6</sup> adopted for ISNCSCI can only be performed in the myotomes of arms and legs. Thus, for all other segments (cervical C2–C4, thoracic T2–L1 and sacral S2–S5) the graded assessment of key muscles is difficult and has not been formally included in the protocol. Consequently, the precise motor

classification of individuals with thoracic or very high cervical lesions relies on inference from the sensory scores. We recently identified the determination of motor level and motor incompleteness in these areas as most difficult classification tasks.<sup>7</sup>

The anatomical constraints together with the complex AIS definitions lead to a large set of classification rules. A rater must either memorize all ISNCSCI classification rules correctly or look them up in the available ISNCSCI booklet<sup>8</sup> or refer to a condensed version on the backside of the assessment form. The necessity and value of training on the ISNCSCI assessment has been addressed in several previous publications.<sup>7,9–11</sup>

The ISNCSCI evaluation consists of two parts, which require different types of skills.<sup>3,12</sup> The physical assessment of an individual consists of segmental sensory (light touch, pinprick) and motor (manual muscle test) evaluations, including an anorectal examination. The segmental level and magnitude of these sensorimotor scores forms the basis for the subsequent classification part. Accurate classification

<sup>1</sup>Spinal Cord Injury Center, Heidelberg University Hospital, Heidelberg, Germany; <sup>2</sup>Spinal Cord Injury Center, Balgrist University Hospital, Zurich, Switzerland; <sup>3</sup>Rehabilitation Center for Children and Juveniles, University Children's Hospital Zurich, Affoltern am Albis, Switzerland; <sup>4</sup>SRH Klinikum, Sektion Querschnittlähmung, Karlsbad-Langensteinbach, Germany; <sup>5</sup>Department of Spinal Cord Injury, BG-Trauma Center, Murnau, Germany; <sup>6</sup>Klinik Hohe Warte, Bayreuth, Germany; <sup>7</sup>Department of Rehabilitation Medicine, Radboud University, Nijmegen Medical Centre, Nijmegen, The Netherlands and <sup>8</sup>European Multicenter Study on Human Spinal Cord Injury

Correspondence: C Schuld, Spinal Cord Injury Center, University Hospital Heidelberg, Schlierbacher Landstraße 200a, Heidelberg University Hospital, 69118 Heidelberg, Germany.

E-mail: Christian.Schuld@med.uni-heidelberg.de

Received 10 March 2014; revised 3 November 2014; accepted 11 November 2014; published online 9 December 2014

is then based on specific ISNCSCI rules that use the collected scores to determine all functional levels (motor/sensory and neurological level of injury (NLI)), as well as the AIS.

Although the reliability of the ISNCSCI assessment has been investigated,<sup>13–15</sup> an analysis for the reliability of the ISNCSCI classification, by typical clinicians on a typical cohort of SCI patients, has not been done. Therefore, the aim of this investigation was to describe and quantify any discrepancies in the ISNCSCI classification by typical SCI clinicians versus validated computational ISNCSCI algorithms.<sup>16</sup> This study was conducted as a part of the quality management system (ISO 9001:2008) of the European Multicenter Study on Human Spinal Cord Injury (EMSCI - <http://emsci.org>).<sup>17</sup>

## MATERIALS AND METHODS

Data sets from the first years of the EMSCI network (May 2003–April 2005; 2 years after project launch in 2001) were used to compare manual versus computerized ISNCSCI classification.<sup>16</sup> ISNCSCI data sets in EMSCI contain classification variables, as well as all segmental scores of light touch, pinprick, muscle function and anorectal examination. During the indicated time period, all ISNCSCI examinations and subsequent classifications were manually performed by SCI clinicians. The classification was carried out either directly after the examination or at the latest before entering the data set into the database. In the worst case there was only a timespan of a few days between examination and classification.

EMSCI data evaluation was computerized in terms of ISNCSCI classification in April 2005.<sup>18</sup> As part of this computerization, all ISNCSCI data sets, including the previously manually classified datasets were recalculated by the computer program. A backup of the manually classified data sets was used for this study. This reclassification process is repeated on ISNCSCI revision changes or bug fixes in the classification software.

The training levels of the clinicians performing the ISNCSCI assessments and classification within the time frame of this study was retrospectively surveyed by phone or by electronic mail and summarized in Table 1.

A flowchart of data processing together with the procedure for data set selection is provided in Figure 1. All data sets from the Spinal Cord Injury Center at the Heidelberg University Hospital were excluded, as the computerized ISNCSCI evaluation was developed in-house and used at the site before the official start of computerized classification throughout the EMSCI network. In addition, all ISNCSCI data sets containing not testable (NT) segmental scores in either light touch, pinprick, motor or anorectal examination were excluded (Box 'Exclusion of fragmentary data set (examination)' in Figure 1) owing to missing instructions on how to deal with this issue in the 2003 reference manual.

EMSCI's strict inclusion and exclusion criteria allow only single-event traumatic or ischemic SCIs. Concomitant peripheral nerve lesions above the level of injury (that is, plexus brachialis impairment), preinjury polyneuropathy, multiple SCIs and severe traumatic brain injuries are the exclusion criteria. Overall, these criteria reduce the probability for the presence of concomitant non-SCI-related neurological impairments above the level of injury. This is particularly relevant, because information in the comment box of the ISNCSCI worksheet, where non-SCI related issues are usually documented, is not used in the computational classification process.

Examinations and manual scoring were based on the ISNCSCI 2003 reference manual,<sup>12</sup> which has now been superseded.<sup>13</sup> Even though the basic classification rules have remained unchanged, some clarifications were published by Waring *et al.*<sup>19</sup> and the revised ISNCSCI booklet<sup>8</sup> (2011) and subsequently integrated into the computerized classification.

These clarifications led to the implementation of two additional rules into the classification algorithms: (1) Motor levels, sensory levels and the single neurological level were set to C1, if segment C2 was already impaired,<sup>8</sup> and (2) if there is no spared function below the sensory or motor level in a person with a complete injury the sensory or motor level should be listed in the zones of partial preservation (ZPP) block.<sup>19</sup>

All data sets affected by these rules were excluded from the analysis to avoid any bias of the results in favor of the computational algorithms. In detail, all cases in which clarification (1) applied were excluded for all analyses (gray shaded box 'Exclusion of NLI C1' in Figure 1) and cases in which rule (2) applied were excluded for the ZPP analyses only (gray shaded box 'Exclusion of data sets where respective levels match ZPP' in Figure 1). Besides these clarifications affecting the technical ISNCSCI implementation, the fundamental classification rules were not revised. However, some rules were rewritten to make the wording more precise.<sup>1</sup> The motor level determination was clarified (for areas where there are no myotomes to test), as well as the correct use of the according reference levels throughout the AIS classification process. In the recent standard, it is explicitly stated that to distinguish between sensory incomplete (AIS B) and motor incomplete lesions (AIS C/D) the motor level on each side is used as reference level, whereas the NLI is used to discriminate between AIS C and AIS D.

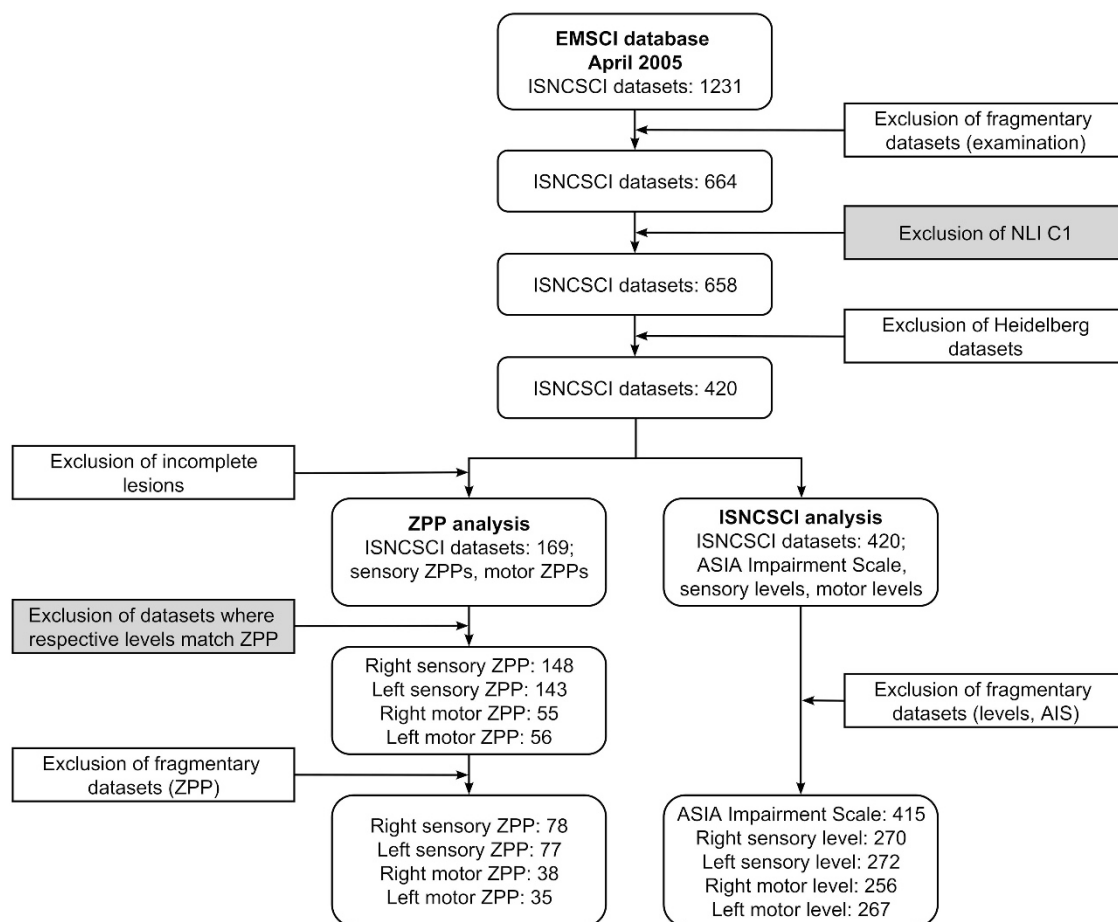
Besides these clarifications directly or indirectly relevant to this work, the recent ISNCSCI booklet (2011) introduced several additional changes, among them are altered positions for motor testing, the use of non-key muscles to distinguish in borderline cases between AIS B and AIS C, the renaming of the anorectal sensory test to deep anal pressure and a clarification that the ZPPs are not referenced from the NLI but instead from the corresponding sensory or motor level.<sup>1</sup>

All included data sets were computationally reclassified using the latest validated EMSCI ISNCSCI implementation of ISNCSCI's current revision

**Table 1** Participating SCI centers and the number of patients included in this work

SCI center	Number of included ISNCSCI data sets	ISNCSCI personnel and training levels
Uniklinik Balgrist, Spinal Cord Injury Center, Zurich, Switzerland	36	ISNCSCI was conducted by the residents under supervision of a consultant neurologists
Hôpital Raymond-Poincaré, Paris/Garches, France	21	One resident with specialization in rehabilitation and physical medicine performed most of the ISNCSCI classifications
SRH Klinikum Sektion Querschnittlähmung, Karlsbad-Langensteinbach, Germany	78	The chief physician of the SCI center, a specialized neurologist, conducted most of the examinations and classifications himself. He participated in a formal ISNCSCI instructional course
UMC St. Radboud, Nijmegen, The Netherlands	65	The senior specialist in rehabilitation and physical medicine performed all ISNCSCI examinations and classifications and was trained during his specialization in the Netherlands
Berufsgenossenschaftliche Unfallklinik Zentrum für Rückenmarkverletzte, Murnau, Germany	136	One resident participated in a formal ISNCSCI instructional course and shared his knowledge with the senior physicians and the remaining physicians. Residents performed and scored ISNCSCI
Klinik für Querschnittgelähmte, Bayreuth, Germany	84	A senior neurologist conducted all ISNCSCI examinations and classifications

Abbreviations: ISNCSCI, International Standards for Neurological Classification of Spinal Cord Injury; SCI, spinal cord injury.



**Figure 1** Data processing flowchart accounting for excluded data sets at every processing stage. Gray shaded boxes describe processing steps that are necessary owing to the ISNCSCI clarifications published by Waring *et al.*<sup>19</sup> 2009. The fork is necessary because the excluded data sets for the ZPP analysis are exclusively relevant for this analysis. Excluding them globally would decrease the overall sample size for general ISNCSCI analysis (right fork). AIS, ASIA Impairment Scale; ASIA, American Spinal Injury Association; EMSCI, European Multicenter Study on Human Spinal Cord Injury; ISNCSCI, The International Standards for Neurological Classification of Spinal Cord Injury; NLI, neurological level of injury; ZPP, zone of partial preservation.

published in 2011.<sup>1</sup> The latest version of the calculator was used in favor of the version implemented during the study period (2003–2005), because the comprehensive validation of the calculator's algorithms was performed later (2009–2010). The validation was performed by the first and the last author of this study. In an iterative approach, the algorithms were tuned on the basis of over 5000 not testable-free data sets of the EMSCI database until human ISNCSCI experts and the computer program agreed on the classification results. Public sources of correct data sets such as the reference manual (Appendix B),<sup>12</sup> the booklet<sup>8</sup> and a collection of difficult cases<sup>20</sup> were additionally and successfully used for validation. To our best knowledge, the current EMSCI algorithms implement ISNCSCI correctly so that differences between the computer and the clinician can be interpreted as the clinician's error. The details of the algorithms and the validation process are published elsewhere.<sup>16</sup>

Both manual (clinician) and computational (computer) classification methods were statistically tested for differences in the following ISNCSCI variables: right/left sensory level, right/left motor level, AIS, sensory and motor zones of partial preservation for right and left side. At the time of the assessments (2003–2005), the single NLI was not listed in the ISNCSCI assessment sheet (REV2001)<sup>12</sup> and was therefore not included in our analysis.

The degree of concordance is presented as raw percentages and histograms and confirmed statistically using Wilcoxon's matched pair test (all levels and ZPPs) and Bowker's test (AIS). All analyses were performed with Statistica 9.1 (StatSoft, Inc., Tulsa, OK, USA) and IBM SPSS Statistics 21 (IBM Corporation, Chicago, IL, USA). Significance level was set to  $\alpha=0.01$ . This more rigorous significance level was chosen in favor of the often used level of

0.05 owing to the retrospective character of this study and the comparatively large sample sizes. In general, a lower significance level reduces the risk of false positive tests.

## RESULTS

The database sample for this study contained 420 eligible ISNCSCI data sets of 185 patients treated in six SCI centers (Table 1). Table 1 also lists the involved clinicians performing ISNCSCI examinations and classifications in these centers at that time, together with an overview of their training levels. The differences between manual and computational ISNCSCI scaling, scoring and classification are summarized in Table 2. The lowest agreement was found for motor levels (right: 62.1%, left: 61.8%) followed by motor ZPP (right: 81.6%, left 80.0%) and then AIS (83.4%). The differences in the motor levels (right:  $P=0.002$ ; left:  $P=0.003$ ) and the AIS ( $P=0.001$ ) were significant. Sensory levels showed the best concordance (right: 90.8%; left: 90.0%), as did sensory ZPP (right: 91.0%; left: 92.2%). Histograms for all level variables are presented in Figures 2a–d. Positive differences on each  $x$  axis indicate that the clinician determined a more rostral level than the computer implementation. While sensory levels (Figure 2a), sensory ZPPs (Figure 2c) and motor ZPPs (Figure 2d) show a symmetrical distribution of errors around the correct level, motor levels (Figure 2b) are skewed toward positive

**Table 2 Agreement between computational and manual scoring, scaling and classification: manual results got subtracted from computational results, and thus, for example, a difference of one segment means that the manually assigned level was one segment rostral or caudal to the level determined by the computational algorithm**

	Identical (%)	Within $\pm 1$ segment (%)	Within $\pm 2$ segments (%)	P-value	N (data sets)	Mean difference	s.d.	Range
Right sensory level	90.8	5.9	1.1	0.657	270	0.07	1.89	-6. 25
Left sensory level	90.0	5.9	1.5	0.300	272	-0.00	1.81	-13 to 25
Right motor level	62.1	25.8	5.5	0.002	256	0.11	2.00	-14 to 12
Left motor level	61.8	27.0	6.0	0.003	267	-0.03	2.20	-14 to 12
ASIA Impairment Scale (AIS)	83.4	14.0	2.4	0.001 <sup>a</sup>	415	0.00	0.51	-3 to 2
Right sensory ZPP	91.0	6.4	1.3	0.176	78	0.09	0.56	-2 to 4
Left sensory ZPP	92.2	5.2	0.0	0.028	77	0.21	1.03	0 to 8
Right motor ZPP	81.6	13.2	0.0	0.735	38	-0.16	1.44	-8 to 3
Left motor ZPP	80.0	11.5	5.7	0.271	35	0.34	1.78	-1 to 10

Abbreviation: ZPP, zone of partial preservation.

<sup>a</sup>Bowker's Test

differences, which means a deviation by clinicians to classify a more rostral level.

The motor level agreement was analyzed in more detail (Figure 3) by taking into account that only myotomes within the arms (Figure 3b) and legs (Figure 3d) are included in ISNCSCI's manual muscle examination. Different error patterns are revealed for all other myotomes (Figures 3a and c). The skewness in the overall motor level agreement (Figure 2b) is unambiguously caused by the misclassifications of motor levels within the spinal segments of arms C5-T1 and legs T2-L1 (Figures 3b and d, respectively).

The correct classification of AIS grades C (concordance 54.5%) or B (concordance 66.7%) appeared to be more difficult in comparison with AIS grades A (concordance 93.4%) or D (concordance 86.0%; Figure 4). AIS B is most commonly misinterpreted as AIS C and vice versa (AIS B as AIS C: 29.4%; Figure 4b and AIS C as AIS B: 38.6%; Figure 4c).

## DISCUSSION

In this retrospective analysis of neurological data sets, obtained in the early years of EMSCI (2003–2005), the differences in classification of ISNCSCI variables by clinicians and computational algorithms were compared in a large European cohort of traumatic and ischemic SCI subjects. This kind of analysis was previously performed exclusively in the artificial setting of ISNCSCI instructional courses.<sup>7,10–21</sup> In the framework of instructional courses, the effects of a formal ISNCSCI training are typically assessed by the same test before (pre-test) and after the training (post-test). In both tests, the attendees rate and classify a predefined set of ISNCSCI cases. The main outcome measure is the change of the percentage of correct answers in pre-testing compared with post-testing. Table 3 summarizes the results of our own instructional courses together with the results of the present study. Besides the previously published result<sup>7</sup> that with training classification skills improve substantially, two conclusions can be drawn from the comparison of the instructional course results with this study's results. First, the classification skills of the clinicians were much better than those of the course participants at the stage of the pre-test, and, second, the error pattern is slightly lower but in the same range as the post-tests results (Table 3).

With regard to the primary aim of this study focusing on the comparison of clinicians versus computational data evaluation, the manual determination of both the motor levels and the AIS were identified as the most challenging steps. This corresponds to previously published results on the effects of standardized formal training in the ISNCSCI classification part<sup>7,10</sup> and a conference proceeding

describing the differences between human and computational AIS classification in a North American cohort.<sup>22</sup>

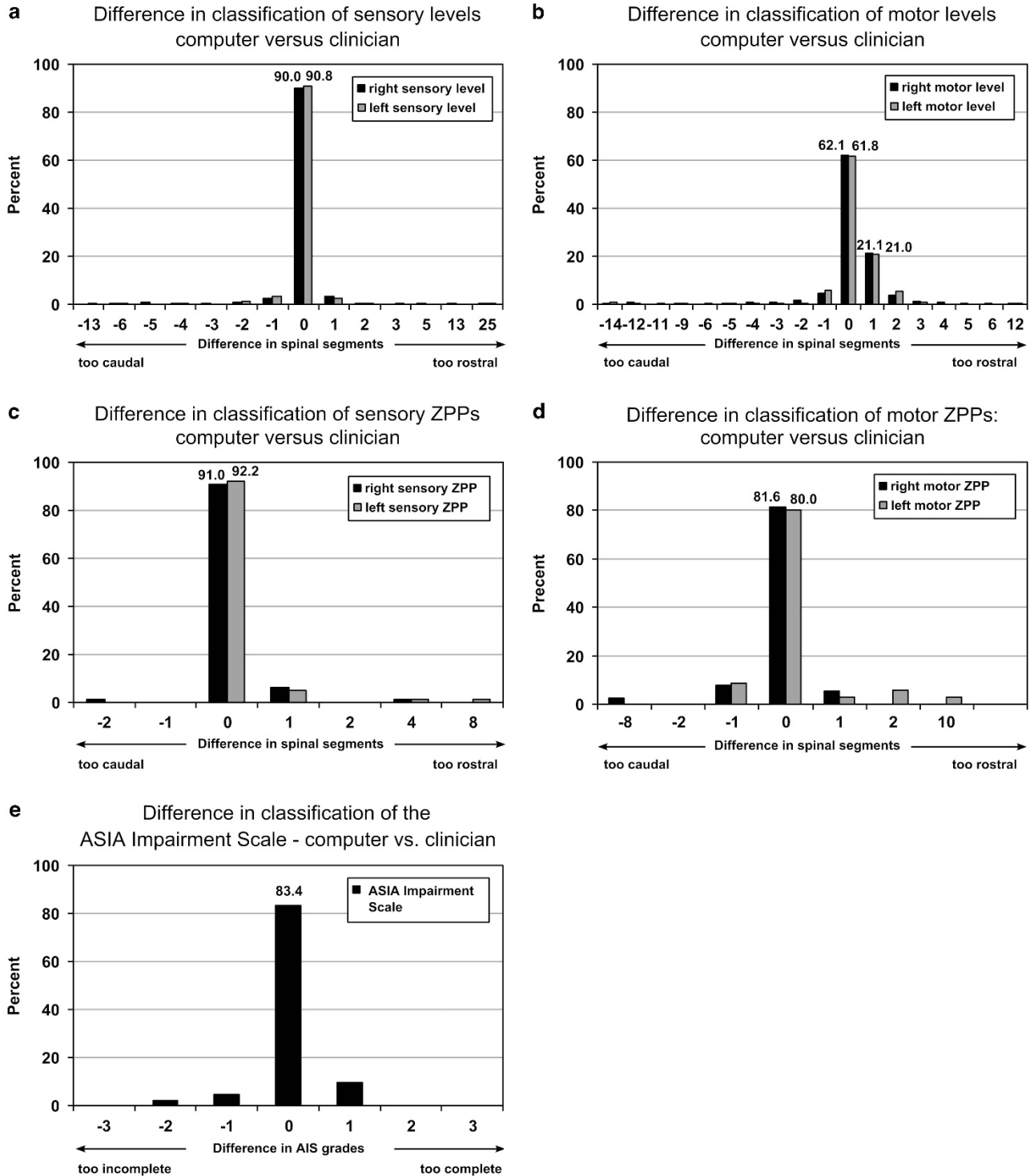
Our results show that motor levels are frequently (26.4%; Figure 2b) classified one segment rostral to the correct level, indicating that the ISNCSCI definition of motor level was not applied properly. ISNCSCI defines motor level as the most caudal key muscle with a manual muscle test score of at least 3 out of 5 or better, provided that all rostral key muscles are judged to be intact and unimpaired (5/5). The motor level can be different for the right and left side of the body. Some of the clinicians' motor level classification errors might be attributed to a misleading motor level definition of the front side of the old superseded ISNCSCI examination sheet (revision 2003). On this sheet, a brief statement right next to the boxes of the motor level defines all neurological levels as 'the most caudal segment with normal function', which is correct for the sensory levels but not for the motor levels (see the correct definition above). The current worksheet (REV 02/13) corrects this issue and provides an updated motor level definition on page 2. Determination of motor levels seems to pose a general problem for the raters. We found different error patterns (Figure 3) between myotomes on arms and legs compared to not in ISNCSCI examined myotomes in high cervical (C2–C4), thoracic (T2–L1) and low sacral (S2–S5) region. Please refer to Kramer *et al.*<sup>23</sup> and Steeves *et al.*<sup>24</sup> for further considerations and discussions regarding motor levels.

Another finding from this study is that AIS grades B and C are susceptible to being misleadingly exchanged (Figure 4), indicating that the determination of motor incompleteness is the most challenging step of the AIS classification. Therefore, determination of motor levels and motor incompleteness should be emphasized in ISNCSCI training and documentation. Instructive examples should be provided, which will help understand respective rules better and to reduce the likelihood of misconceptions. With the recent activities,<sup>5,25</sup> ASIA's International Standards Committee already provided instructions for better interpretation together with some example cases regarding these issues. In fact, this issue was a highlighted topic in the 2011 revision<sup>1</sup> and the accompanying references publication.<sup>5</sup>

On the basis of the results of this study, training on ISNCSCI classification is strongly recommended for clinical practice, as well as for research. Training programs are available online (International Standards Training e-Learning Program (InSTeP)<sup>3</sup> developed by ASIA) and are usually parts of research networks<sup>7,10</sup> and clinical trials.<sup>26,27</sup> In addition, for research, one should consider accrediting trial investigators as ISNCSCI raters, only after they have successfully

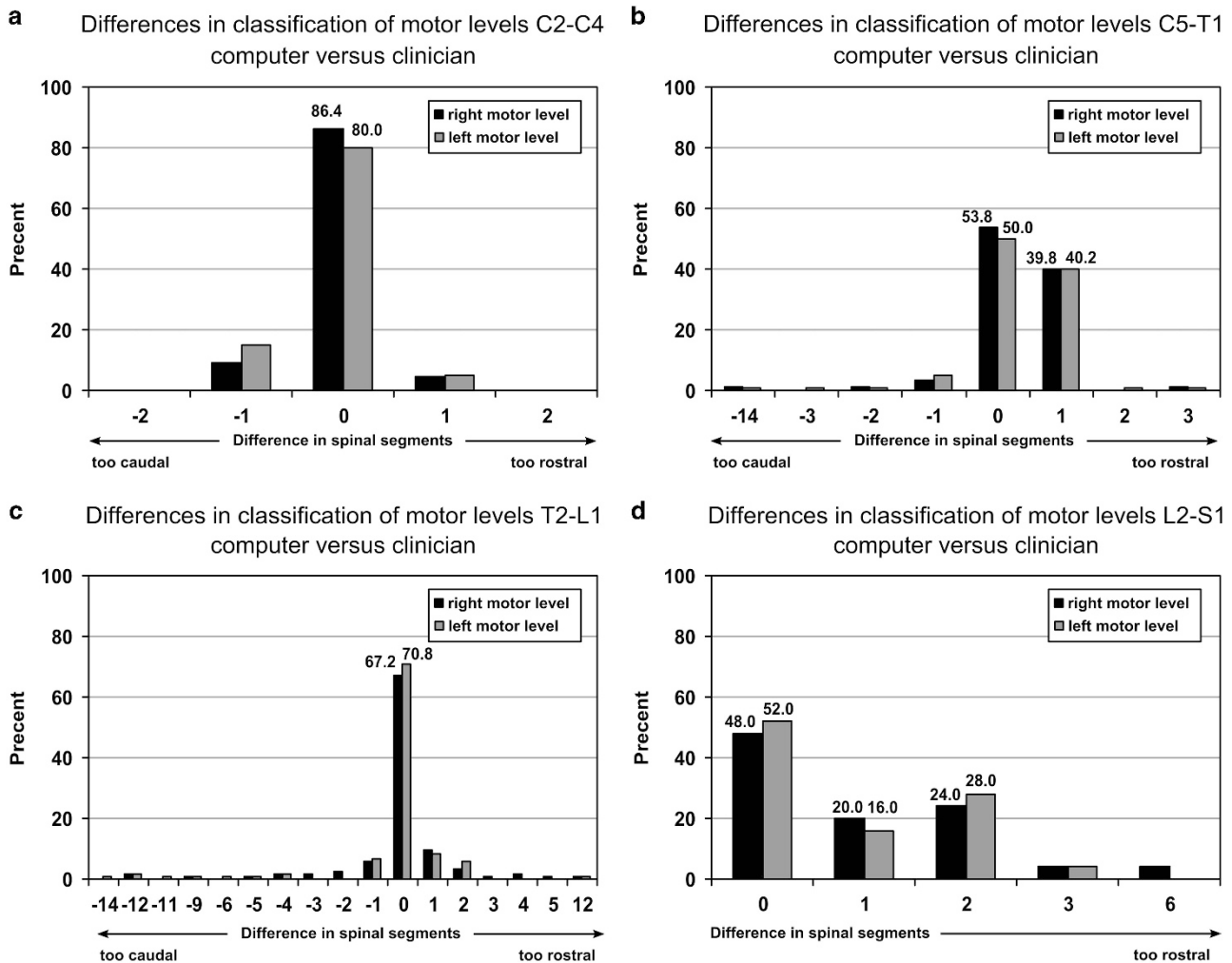
participated in workshops on the theory and practice of all aspects of the ISNCSCI assessment and classification. This is in line with the current guidelines of the 'International Campaign for Cures of Spinal Cord Injury Paralysis (ICCP)',<sup>28</sup>

Insights from this study may help to not only carefully interpret ISNCSCI data sets from previously conducted or currently running clinical trials, but also to plan future interventional trials in SCI. As outlined here, validated computer-aided



**Figure 2** Differences between clinicians and computational ISNCSCI scoring, scaling and classification for all derived ISNCSCI variables (Sensory levels (a), Motor levels (b), Sensory ZPPs (c), Motor ZPPs (d) and AIS (e)). Raw agreement in percent is displayed on the y axes of all subfigures. Positive differences on each x axis indicate that the SCI professional determined a more rostral level than the computer implementation. Subfigures a–d have the differences in levels on their x axes, subfigure e has the difference in AIS grades. ASIA, American Spinal Injury Association; ISNCSCI, The International Standards for Neurological Classification of Spinal Cord Injury; ZPPs, zones of partial preservation.





**Figure 3** Differences in motor level determination described as raw agreement. Positive differences on each x axis indicate that the SCI professional determined a more rostral level than the computer implementation. In subfigure (a), motor levels are within the high cervical region (C2–C4), and in subfigure (c) motor levels are within the thoracic region (T2–L1). In those segments by definition the motor level follows the sensory level, because the myotomes cannot be assessed. Subfigures (b) and (d) depict all testable myotomes on arms (b) and legs (d).

algorithms will likely eliminate errors in ISNCSCI classification by humans.<sup>16</sup>

### Limitations

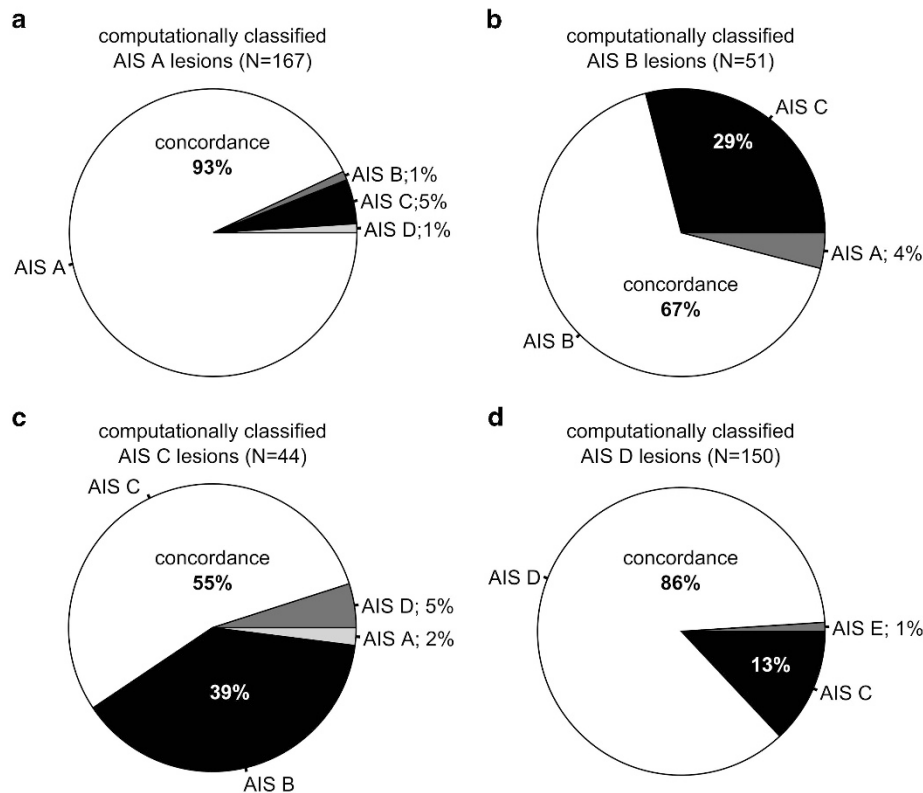
This study uses data sets that were obtained almost 10 years ago and were predominantly classified by residents with different classification skills and level of training. A correlation between training levels and classification errors cannot be performed, because this information was not documented during the study period. These issues must be carefully taken into account when interpreting the outcomes of this study. Over the past decade, ASIA's International Standards Committee has continuously worked on the clarification of the classification rules and on better education of the SCI community. It can be assumed that for large-scale clinical trials classification skills can be expected to be much better than in this cohort of typical, not explicitly (for SCI clinical trials) trained clinicians. However, for projects with smaller budgets such as SCI registries and pilot studies in SCI, the findings of this study are still important to (1) estimate the bias produced by manual classifications performed by nonoptimally trained clinicians and to (2) reinforce the important message that continuous training in ISNCSCI is needed. An effective way to

minimize human classification errors and to reclassify ISNCSCI examinations according to new revisions of the standards is to use computational ISNCSCI classification.

However, there are limitations in relying exclusively on a computational approach. Ideally, an ISNCSCI examiner is performing a preliminary manual classification during the examination. This helps avoid misclassification of non-SCI-related issues above the level of injury and to identify the most crucial dermatomes and myotomes for classification. The latter include those segments in which motor function is preserved more than three levels below the motor level, which is important for the accurate classification of AIS B versus an AIS C/D. As a consequence, the examiner can focus on these dermatomes and myotomes to ensure conclusive scoring results. Relying exclusively on computer classification skills might extinguish this very important neurological diagnostic skill over time. One should keep in mind that correct examination scores represent the critical prerequisite for a correct ISNCSCI classification. Therefore, future studies will systematically explore the effects related to the examination technique in the EMSCI sample.

In early versions of the EMSCI database it was not mandatory to fill in all ISNCSCI variables. This fact alone resulted in a low share

### Degree of concordance in ASIA Impairment Scale classification of computational algorithms versus clinicians



**Figure 4** Degree of concordance in AIS classification of computation algorithms versus clinicians. The correct classification is arranged by AIS (a–d). White sectors display concordance between the computer and the clinician. Accordingly, gray sectors display discordance. The gray shading encodes the percentage of discordance. AIS, ASIA Impairment Scale; ASIA, American Spinal Injury Association.

**Table 3** Comparison of clinicians' classification errors (fourth column) versus ISNCSCI raters before an instructional course (second column) and trained ISNCSCI rates after the instructional course (third column)

	EMSCI's ISNCSCI instructional course: pre-test error rate (%)	EMSCI's ISNCSCI instructional course: post-test error rate (%)	Clinicians classification errors (%)
Sensory levels	60.6	96.8	90.4
Motor levels	20.8	81.8	62.0
ASIA Impairment Scale	45.8	88.1	83.4
Sensory ZPPs	51.4	93.3	91.6
Motor ZPPs	50.5	93.3	80.8

Abbreviations: EMSCI, European Multicenter Study on Human Spinal Cord Injury; ISNCSCI, International Standards for Neurological Classification of Spinal Cord Injury; ZPP, zone of partial preservation.

Instructional course data in the second and the third column are taken from Schuld *et al*.<sup>7</sup>

of fully evaluated ISNCSCI examination sheets, especially with regard to the sensory and motor levels (63.4%, Figure 1). In contrast, ASIA Impairment Scales were determined conscientiously (98.8%, Figure 1). This work used data sets that were acquired before institution of formal ISNCSCI instructional courses in the EMSCI network. Since 2006, more than 250 participants have been trained in an ongoing ISNCSCI training program.<sup>7</sup> It is anticipated that the overall ISNCSCI skills increased in the EMSCI network, not only because of the ongoing training but also owing to the online training efforts of ASIA and the efforts of the International Standards Committee for better clarification of ISNCSCI variables, which

will hopefully lead to more consistent standards. The information of the worksheet's comment box is currently not incorporated into the computational algorithms, because data extraction from this free text information is a nontrivial task. A first step for including more clinical expertise into the algorithms might be to open the asterisk nomenclature currently used for the muscle grade 5 to all motor and sensory grade. This would identify situations in which the tested impairment may not be caused by the SCI. In computational classification, an asterisk denoted graded would be treated as intact without the need of information retrieval from the comment box.

## CONCLUSION

Clinicians commit classification errors when assessing the neurological impairment after SCI. The most difficult tasks are the correct determination of motor levels and motor incompleteness. We recommend more clarification and clearer examples with regard to these issues in upcoming ISNCSCI revisions. Training is strongly recommended for clinical practitioners as an effective means for reducing classification errors. It should be a prerequisite for accreditation of examiners in clinical trials. For consistent classification, data sets should be pre-analyzed by ISNCSCI experts and finally processed by computational ISNCSCI algorithms as a means of effective quality control.

## DATA ARCHIVING

There were no data to deposit.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank all patients and all centers (<http://emsci.org/members>) of the EMSCI network. This work was partially funded by the International Foundation for Research in Paraplegia (IFP), Zurich, Switzerland.

- Kirshblum SC, Burns SP, Biering-Sorensen F, Donovan W, Graves DE, Jha A *et al*. International standards for neurological classification of spinal cord injury (revised 2011). *J Spinal Cord Med* 2011; **34**: 535–546.
- Furlan JC, Fehlings MG, Tator CH, Davis AM. Motor and sensory assessment of patients in clinical trials for pharmacological therapy of acute spinal cord injury: psychometric properties of the ASIA Standards. *J Neurotrauma* 2008; **25**: 1273–1301.
- American Spinal Injury Association. *International Standards Training E Program (InSteP)*. Available at: <http://www.asia-spinalinjury.org/elearning/elearning.php>, (accessed on 24 June 2014).
- Frankel HL, Hancock DO, Hyslop G, Melzak J, Michaelis LS, Ungar GH *et al*. The value of postural reduction in the initial management of closed injuries of the spine with paraplegia and tetraplegia I. *Paraplegia* 1969; **7**: 179–192.
- 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.
- Daniels L, Worthingham C. *Muscle Testing: Techniques of Manual Examination*, 3rd edn. Saunders: Philadelphia, PA, USA, 1972.
- Schuld C, Wiese J, Franz S, Putz C, Stierle I, Smoor I *et al*. Effect of formal training in scaling, scoring and classification of the International Standards for Neurological Classification of Spinal Cord Injury. *Spinal Cord* 2013; **51**: 282–288.
- American Spinal Injury Association. *International Standards for Neurological Classification of Spinal Cord Injury Booklet*. American Spinal Injury Association: Atlanta, GA, USA, 2011.
- Burns AS, Lee BS, Ditunno JF, Tessler A. Patient selection for clinical trials: the reliability of the early spinal cord injury examination. *J Neurotrauma* 2003; **20**: 477–482.
- 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.
- Cohen ME, Ditunno JF, Donovan WH, Maynard FM. A test of the 1992 International Standards for Neurological and Functional Classification of Spinal Cord Injury. *Spinal Cord* 1998; **36**: 554–560.
- American Spinal Injury Association. *Reference Manual for the International Standards for Neurological Classification of Spinal Cord Injury*. American Spinal Injury Association: Chicago, IL, USA, 2003.
- Mulcahey MJ, Gaughan J, Betz RR, Johansen KJ. The International Standards for Neurological Classification of Spinal Cord Injury: reliability of data when applied to children and youths. *Spinal Cord* 2007; **45**: 452–459.
- 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**: S146–S149.
- Savic G, Bergström EMK, 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.
- Schuld C, Wiese J, Hug A, Putz C, van Hedel HJA, Spiess MR *et al*. Computer implementation of the international standards for neurological classification of spinal cord injury for consistent and efficient derivation of its subscores including handling of data from not testable segments. *J Neurotrauma* 2012; **29**: 453–461.
- Curt A, Schwab ME, Dietz V. Providing the clinical basis for new interventional therapies: refined diagnosis and assessment of recovery after spinal cord injury. *Spinal Cord* 2004; **42**: 1–6.
- Rupp R, Schweidler J, Curt A, Dietz V, Gerner HJ. An electronic tool for multicenter administration, assessment and analysis of clinical trials in spinal cord injury. *Biomedizinische Technik* 2005; **50**: 1460–1461.
- Waring WP, Biering-Sorensen F, Burns S, Donovan W, Graves D, Jha A *et al*. 2009 review and revisions of the international standards for the neurological classification of spinal cord injury. *J Spinal Cord Med* 2010; **33**: 346–352.
- Kirshblum SC. *Difficult cases in classification of SCI*. Presented at the American Paraplegia Society Meeting. Las Vegas, NV, USA, 2006.
- Liu N, Zhou MW, Krassioukov AV, Biering-Sorensen F. Training effectiveness when teaching the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) to medical students. *Spinal Cord* 2013; **51**: 768–771.
- Oleson C, Marino R. Spinal Cord Injury Classification: Comparison of Human and Computer Algorithm for ASIA Impairment Scale Grades. In: *37th Annual Scientific Meeting of the American Spinal Injury Association and the 50th Annual Scientific Meeting of the International Spinal Cord Society*. Grand Hyatt Washington Hotel, Washington, DC, USA, June 8, 2011. p 49. [http://www.iscos.org.uk/sitefiles/ISCOS\\_A\\_SIA\\_Bk\\_COMPLETE\\_lo-res.pdf](http://www.iscos.org.uk/sitefiles/ISCOS_A_SIA_Bk_COMPLETE_lo-res.pdf).
- Kramer JL, Lammertse DP, Schubert M, Curt A, Steeves JD. Relationship between motor recovery and independence after sensorimotor-complete cervical spinal cord injury. *Neurorehabil Neural Repair* 2012; **26**: 1064–1071.
- Steeves JD, Lammertse DP, Kramer JL, Kleitman N, Kalsi-Ryan S, Jones L *et al*. Outcome Measures for Acute/Subacute Cervical Sensorimotor Complete (AIS-A) Spinal Cord Injury During a Phase 2 Clinical Trial. *Topics in spinal cord injury rehabilitation* 2012; **18**: 1–14.
- Kirshblum SC, Biering-Sorensen F, Betz R, Burns S, Donovan W, Graves DE *et al*. International Standards for Neurological Classification of Spinal Cord Injury: cases with classification challenges. *J Spinal Cord Med* 2014; **37**: 120–127.
- Consortium for Spinal Cord Medicine. Outcomes following traumatic spinal cord injury: clinical practice guidelines for health-care professionals. *J Spinal Cord Med* 2000; **23**: 289–316.
- Jones LA, Lammertse DP, Charlifue SB, Kirshblum SC, Apple DF, Ragnarsson KT *et al*. A phase 2 autologous cellular therapy trial in patients with acute, complete spinal cord injury: pragmatics, recruitment, and demographics. *Spinal cord* 2010; **48**: 798–807.
- Steeves JD, Lammertse D, Curt A, Fawcett JW, Tuszynski MH, Ditunno JF *et al*. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP panel: clinical trial outcome measures. *Spinal Cord* 2007; **45**: 206–221.