A NEW SCORE BASED ON THE INTERNATIONAL STANDARDS FOR NEUROLOGICAL CLASSIFICATION OF SPINAL CORD INJURY FOR INTEGRATIVE EVALUATION OF CHANGES IN SENSORIMOTOR FUNCTIONS.

Lukas Grassner, MD, PhD[a,b,c,d], Daniel Garcia-Ovejero, PhD[e], Orpheus Mach, MA[d], Elisa Lopez-Dolado, MD, PhD[f], Eduardo Vargas, MD[g], Monica Alcobendas-Maestro, MD, PhD[f], Ana Esclarin, MD, PhD[f], Ludwig Sanktjohaneser, MD[d], Christof Wutte, MD[d,h,i], Johannes Becker, MD[d], Sara Lener, MD, PhD[a], Sebastian Hartmann, PhD, MD[a], Pierre-Pascal Girod, MD[a], Nikolaus Kögl, MD[a], Christoph J Griessenauer, MD[j,k], Marios Papadopoulus, MD[l], Fred Geisler, MD, PhD[m], Claudius Thome, MD[a], Eduardo Molina-Holgado, PhD[e], Joan Vidal, MD[n,o,p], Armin Curt, MD[q], Giorgio Scivoletto, MD[r], James Guest, MD[s,t], Doris Maier, MD[d], Norbert Weidner, MD[u], Rüdiger Rupp, Dr-Ing[u], John LK Kramer, PhD[v,x,y], Angel Arevalo-Martin, PhD[e]*.

[a] Department of Neurosurgery, Medical University Innsbruck, Innsbruck, Austria

[b] Institute of Molecular Regenerative Medicine, Paracelsus Medical University, Salzburg, Austria.

[c] Spinal Cord Injury and Tissue Regeneration Center Salzburg (SCI-TReCS), Paracelsus Medical University, Salzburg, Austria.

[d] Center for Spinal Cord Injuries, Trauma Center Murnau, Germany.

[e] Neuroinflammation Laboratory, Hospital Nacional de Paraplejicos, SESCAM, Toledo, Spain.

[f] Department of Physical Medicine and Rehabilitation, Hospital Nacional de Paraplejicos, SESCAM, Toledo, Spain.

[g] Sexual and Fertility Unit, Hospital Nacional de Paraplejicos, SESCAM, Toledo, Spain.

[h] Department of Neurosurgery, Trauma Center Murnau, Murnau, Germany

[i] Institute of Biomechanics, Trauma Center Murnau, Murnau, Germany and Paracelsus Medical University, Salzburg, Austria

[j] Department of Neurosurgery, Geisinger Health System, Danville, PA, USA.

[k] Department of Neurosurgery, Paracelsus Medical University, Salzburg, Austria.

[l] Academic Neurosurgery Unit, St. George's University of London, London, UK.

[m] Department of Medical Imaging, University of Saskatchewan, Saskatoon, SK, Canada.

[n] Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, 089M6, Badalona, Barcelona, Spain.

[o] Universitat Autonoma de Barcelona, 08M93, Bellaterra, Cerdanyola del Vallès, Spain

[p] Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Barcelona, Spain

[q] Spinal Cord Injury Center, Balgrist University Hospital, Zürich, Switzerland

[r] Spinal Unit and Spinal Rehabilitation (SpiRe) Lab, IRCCS Fondazione S. Lucia, Rome, Italy

[s] The Miami Project to Cure Paralysis, Miller School of Medicine, The University of Miami, Miami, FL, United States

[t] The Department of Neurological Surgery, Miller School of Medicine, The University of Miami, Miami, FL, United States

[u] Spinal Cord Injury Center, Heidelberg University Hospital, Germany.

[v] International Collaboration on Repair Discoveries (ICORD), University of British Columbia, Vancouver, British Columbia, Canada.

[x] Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, British Columbia

[y] Department of Anesthesiology, Pharmacology, and Therapeutics, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada.

***** corresponding author

Running title: Integrated Neurological Change Score

Keywords: spinal cord injuries, neurologic examination, neurologic recovery, INCS

ABSTRACT

Sensorimotor function of patients with spinal cord injury (SCI) is commonly assessed according to the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI). From the ISNCSCI segmental motor and sensory assessments, upper and lower extremity motor scores (UEMS and LEMS), sum scores of pin prick (PP) and light touch (LT) sensation, the neurological level of injury (NLI) and the classification of lesion severity according to the American Spinal Injury Association Impairment Scale (AIS) grade are derived. Changes of these parameters over time are widely used to evaluate neurological recovery. However, evaluating recovery based on a single ISNCSCI scoring or classification variable may misestimate overall recovery. Here, we propose an Integrated Neurological Change Score (INCS) based on the combination of normalized changes between two-time points of UEMS, LEMS, and total PP and LT scores. To assess the agreement of INCS with clinical judgement of meaningfulness of neurological changes, changes of ISNCSCI variables between two time-points of 88 patients from an independent cohort were rated by 20 clinical experts according to a 5-categories Likert Scale. As for individual ISNCSCI variables, neurological change measured by INCS is associated to severity (AIS grade), age and time since injury, but INCS better reflects clinical judgment about meaningfulness of neurological changes than individual ISNCSCI variables. In addition, INCS is related with changes in functional independence measured by the Spinal Cord Independence Measure (SCIM) in patients with tetraplegia. INCS may be a useful measure of overall neurological change in clinical studies.

INTRODUCTION

Spinal cord injuries (SCI) impair diverse neurological functions, including (but not limited to) motor and sensory abilities. The International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI)[1] represents the gold standard to assess sensorimotor impairment. The main aim of ISNCSCI examination is to determine the level and severity of a SCI based on the examination of the strength of upper and lower extremity key muscles and light touch and pin prick sensation in all dermatomes of the body. Additionally, sparing of sacral sensory and motor function is examined to classify completeness of injury. The changes of ISNCSCI scoring and classification variables between assessments at two different time points are widely used to evaluate neurological recovery after SCI. However, in spite that each of the ISNCSCI scoring and classification parameter contains valuable information by itself, every ISNCSCI variable focuses on a different aspect of recovery.

In clinical studies ~~typically~~ only one ISNCSCI parameter is frequently used as endpoint and needs to be carefully selected according to the characteristics of the participants. For example, differences between two examinations in the Upper Extremity Motor Score (UEMS), which measures ~~the strength of~~ five key muscle~~s~~ functions bilaterally and ranges from 0 (no motor function detectable by ISNCSCI) to 50 (no motor deficits detectable by ISNCSCI), may be useful to assess meaningful recovery for individuals with tetraplegia. However, it should not be used as a surrogate marker of overall recovery since it does not imply that motor improvement occurs in other muscular segments and, obviously, it is meaningless for patients without upper extremity impairment (e.g., lesions at T2 and below). Similarly, improvements in the Lower Extremity Motor Score (LEMS, also ranging from 0 to 50) may be considered as a general sign of recovery. But improvements in LEMS do not necessarily reflect other motor and/or sensory improvements, as occurs, for example, in patients with central cord syndrome, in which lower extremities are less affected than upper extremities ~~LEMS deficits may be mild~~ and a greater UEMS recovery may be expected [2].

Sensory functions are assessed within ISNCSCI as light touch (LT) and pin prick (PP) sensation, respectively. Both sensory modalities are bilaterally tested at key points of the 28 dermatomes that correspond to the innervation from the second cervical segment of the spinal cord (C2) to the lowest sacral segment (S4-5). The total LT and PP scores range from 0 (total absence of LT or PP sensation) to 112 (normal bilateral LT or PP sensation). Although the recovery of sensory functions, mainly PP, are predictive of motor recovery [3,4], they are not the definitive solution to estimate overall neurological changes since motor and sensory pathways are not perfectly paired.

The neurological level of injury (NLI) is defined as the most caudal segment of the spinal cord with intact LT and PP sensation and muscle antigravity strength (key muscle motor score≥3), provided that there is normal sensory and motor function in all rostral segments. Although it may be useful to evaluate the expansion or, conversely, the regression of a lesion, considerable amount of sensorimotor function may be recovered in incomplete lesions without significant changes in NLI, as frequently observed in patients with central cord syndrome [2].

In many trials, the conversion of the American Spinal Injury Association Impairment Scale (AIS grade) has been used because of the integrative nature of this parameter and its applicability in all patients. Patients are classified as AIS A when the lesion is complete; i.e., when no sensory or motor function is preserved at the most caudal spinal cord segments (S4-5) [1]. When sensory function is preserved but there is no motor function at S4-5 or three levels below the ~~NLI~~motor level on a given side, patients are classified as AIS B, also referred to as sensory incomplete motor complete. When the previous conditions are not met, patients are classified as motor incomplete, being AIS C if less than half of the muscles below the NLI have a motor score ≥3 (full range of motion against gravity) or as AIS D otherwise. When patients recover normal sensory and motor functions in the tested segments they are classified as AIS E. Although AIS grade classification integrates sensory and motor scores, it has several drawbacks. First of all, AIS grade conversion is poorly sensitive (which leads to requiring large numbers of participants in clinical studies) and

meaningful neurological recovery can be still achieved without AIS grade conversion [5]. Also, AIS grade conversion is subjected to a floor effect, since AIS A grade patients may worsen but cannot convert into a less functional grade. Additionally, a ceiling effect may also exist, since conversion from AIS D to E is unlikely [6].

Here we propose a new score that integrates the differences between two time points in UEMS, LEMS, and total PP and LT scores into a single value, which is highly sensitive and follows the clinical meaningfulness of neurological changes rated by clinician experts in SCI.

MATERIALS AND METHODS.

Data.

Two sources of data were employed: 1) the Sygen clinical trial database and 2) the database generated by the Autoantibodies in Spinal Cord Injury study (ASCI, Clinicaltrials.gov register number NCT02493543).

The Sygen clinical trial was a phase III multicenter study that enrolled 797 patients from 1992 to 1998 and was designed to test the efficacy of GM-1 ganglioside therapy in acute traumatic SCI [7]. Efficacy was evaluated by determining the sensory and motor function according to the International Standards for the Neurological Classification of Spinal Cord Injury (ISNCSCI) at 1, 4, 8, 16, 26 and 52 weeks after SCI. Since the study concluded that there were no significant differences between treated and placebo groups at 26 and 52 weeks after injury, the database has been used to characterize the temporal course and extent of spontaneous recovery in the first year after SCI [4,6,8,9]. From this database, the 403 patients without missing data at 4 and 52 weeks on AIS grade, NLI, UEMS, LEMS, and sum PP and LT scores were selected. Demographical and clinical characteristics at 4 and 52 weeks of these patients are shown in Table 1. As later explained, the score of neurological change was developed with these data. Detailed information on the changes in ISNCSCI variables between 4 and 52 weeks of the 403 participants is shown in Supplementary Figure 1. Noteworthy, Sygen trial participants were evaluated according to a previous ISNCSCI

version [10]. AIS grades as well as total PP and LT parameters from the Sygen trial were recalculated according to the current ISNCSCI version[1].

The ASCI study began on 2014 and was designed to determine the profile of autoantibodies related with recovery in the acute and subacute phases of SCI. Participants were recruited at the Center for Spinal Cord Injuries (Trauma Center Murnau, Germany) or at Hospital Nacional de Paraplejicos (Toledo, Spain). Among other clinical data, sensorimotor function of the patients was evaluated according to ISNCSCI at 30 and 120 days after injury. ASCI database includes clinical, demographical and other data from 88 people with traumatic SCI. Results from a subset of 62 participants about the origin of the autoantibodies that increase after SCI have been already reported [11]. Demographical and clinical characteristics of these patients are shown in Table 2. As later explained, these data were used to evaluate the adjustment of the score with the clinical meaningfulness of neurological changes rated by clinician experts in SCI and with functional recovery evaluated by the Spinal Cord Independence Measure (SCIM) III. ISNCSCI assessments at 30 and 120 days after injury of all 88 participants are shown in Supplementary Table 1.

Development of the Integrated Neurological Change Score (INCS).

~~To develop t~~The Integrated Neurological Change Score (INCS) has been developed attending to the changes on ISNCSCI assessments between 4 and 52 weeks after injury ~~we employed~~from the subset of 403 patients from the Sygen trial database without missing data ~~on ISNCSCI assessments~~ at both time points ~~4 and 52 weeks after injury~~. The first ISNCSCI assessments in the Sygen trial were made in the first week after injury, being 94.3% of them performed in the first 72 hours [10,12]. Although more neurological changes would be expected by using first week assessments instead of those at 4 weeks after injury,~~This~~ early ISNCSCI assessments may be rather unreliable [4,13]. Thus, to optimize the tradeoff between reliability and amount of neurological changes, the second~~time interva~~ assessment~~l~~ –at 4 weeks– was chosen~~, 4 to 52 weeks, was chosen~~ because of its higher reliability and because a significant amount of neurological change can still occur until 52 weeks~~to allow as much recovery as~~

~~possible while maximizing the confidence on the ISNCSCI evaluations.(very early assessments are less reliable [4,12]).~~ Among the 403 selected participants, 22 (5.4%) converted at least 2 AIS grades and 111 (27.5%) one AIS grade. Worsening also occurred, 8 patients (2%) converted to a lower AIS grade. Detailed information about improvement and worsening of ISNCSCI parameters between 4 and 52 weeks of these 403 patients is provided in Supplementary Figure 1.

The following steps were followed to develop INCS:

1) Being X_t the set of UEMS, LEMS, total PP and total LT for a patient at a given time after injury (t),

$$X_t = \{UEMS_t, LEMS_t, totalPP_t, totalLT_t\},$$

the difference between the values at 4 and 52 weeks after injury for each of these variables was simply calculated as,

$$(1)\ \Delta x_i = x_{it2} - x_{it1},$$

where x_{it1} is the value of the variable in the i[th] position of X_t at the first assessment time (t1; 4 weeks after SCI) and x_{it2} is the value of the same determination at the second assessment time (t2; 52 weeks after SCI).

2) Raw differences in UEMS, LEMS, and total PP and LT scores between 4 and 52 weeks after injury were normalized to the maximum achievable gain or loss for each variable and each patient following the equation,

$$(2)\ \Delta x_{ir} = \begin{cases} \dfrac{\Delta x_i}{\max(x_i) - x_{it1}}\ if\ \Delta x_i \geq 0 \\ \dfrac{\Delta x_i}{x_{it1}}\ if\ \Delta x_i < 0 \end{cases}$$

Thus, for example, a patient with UEMS=25 at 4 weeks and UEMS=35 at 52 weeks would have a raw gain of 10 UEMS points and a normalized gain of 0.4 (the result of dividing the raw gain, 10, by the maximum possible improvement, 50-25=25). Of note, all the resulting normalized differences (Δx_{ir}) are in the same scale, in the range from -1 to +1, being -1 the maximum potential worsening for a given variable and +1 its maximum

potential improvement. Thus, Δx_{ir} quantifies the fraction of lost UEMS, LEMS, total PP or total LT function that is recovered when the change is positive or the fraction of preserved function that is lost if the change is negative. Δx_{ir} have characteristics of interval scale variables since their values are ordered, meaningful (for example $\Delta UEMS_r = 0.5$ means that a patient has recovered half of his/her lost upper extremity function evaluated by ISNCSCI) and regularly spaced (a patient with $\Delta UEMS_r = 0.5$ recovered the double of lost upper extremity function than a patients with $\Delta UEMS_r = 0.25$).

3) The above calculated normalized differences (Δx_{ir}, equation (2)) were subjected to Singular Value Decomposition (SVD)-based principal component analysis (PCA) without centering nor scaling. Mean centering would specify that the origin of the multivariate model is the centroid determined by the mean values of the variables. In this way, the mean values of the normalized differences of UEMS, LEMS, PP and LT would be the point of zero information, and the loading vectors would capture the deviances from the mean. For capturing the neurological change of a patient with SCI, the point of zero information should rather be the all-zero record, the absence of any change. Thus, by not mean-centering the data, any deviation from no change is considered as information and is captured in the loading vectors. Indeed, after long debates about their usefulness and comparisons with the standard centered PCA, non-centered SVD-based PCA has been used in several fields, from microarray analysis to ecology or satellite remote sensing, to name a few [14–17]. Scaling (dividing by the standard deviation to avoid that variables with higher absolute values may have a greater weight in the final loading vectors) is unnecessary since the normalized differences of UEMS, LEMS, total PP and total LT, are already in the same range of values, from -1 to +1.

Two different SVD-based PCA were performed: one for the 312 individuals with tetraplegia and another one for the 91 individuals with paraplegia. Participants were classified with paraplegia if there were no upper extremities deficits evaluable by ISNCSCI (UEMS=50) at both 4 and 52 weeks, and with tetraplegia otherwise. For the

patients with tetraplegia we included the normalized differences of all the four ISNCSCI variables specified above; for the patients with paraplegia we included all variables but UEMS, since the difference in UEMS between 4 and 52 weeks is always zero (UEMS=50 at both times).

Since we aimed to obtain SVD-derived components loadings that could be universally applied to any patient with SCI, we estimated their distribution in the population by bootstrapping 100,000 times with stratification by AIS grade at 4 weeks. For this, we employed the "boot" package for R language for statistical computing [18]. As a result, for each relative difference in UEMS, LEMS, PP and LT we obtained 100,000 loading estimates for patients with tetraplegia and another 100,000 loading estimates for each relative difference in LEMS, PP and LT for patients with paraplegia. The loading estimates for the first principal component (PC1) followed a normal distribution both for patients with paraplegia and tetraplegia. Therefore, their mean values were selected as the final loading estimates and 95% confidence intervals were calculated to confirm that these values were significantly different from zero (Supplementary Figure 2).

The resulting first principal component (PC_1) for tetraplegic or paraplegic patients is, thus, a linear combination of each normalized difference (Δx_{ir}) included in the SVD-based PCA without centering nor scaling and its corresponding mean estimated loadings:

$$(3)\ PC_1 = \sum_{i=1}^{p} \emptyset_{i1} \Delta x_{ir},$$

where $\emptyset_{i1}$ is the i[th] element of the PC_1 loadings vector ($\emptyset_1$). All the loadings in the equation above have a positive sign both for patients with tetraplegia and paraplegia. Thus, for each patient, PC_1 is a linear combination of all Δx_{ir} indicating overall improvement or worsening.

4) Since all the normalized differences (Δx_{ir}; equation (2)) are in the range -1 to +1 and all the loadings calculated in the step before ($\emptyset_{i1}$ in equation (3)) have a positive sign, the maximum achievable PC_1 score for any patient is just the sum of its PC_1 loadings.

Consequently, the PC_1 score of each patient can be normalized to the maximum potential PC_1 as follows,

$$(4)\ PC1_{normalized} = \frac{PC_1}{\sum_{i=1}^{p} \emptyset_{i1}} = INCS$$

By definition, the values of the normalized PC1 will also range between -1 and +1. We propose to use this value as the Integrated Neurological Change Score (INCS) of UEMS, LEMS, PP and LT for patients with tetraplegia and all these variables but UEMS for patients with paraplegia. The relative contribution of each variable to INCS is shown in Figure 1A-B.

Calculation of INCS.

INCS has been calculated for all the patients by using a self-developed code in R. INCS is calculated by:

1) Computing the normalized differences (Δx_{ir}) of UEMS, LEMS, total PP and total LT, by applying equations (1) and (2),
2) Applying to the normalized differences resulting from the previous step the corresponding normalized loadings for patients with tetraplegia or paraplegia, as stated in equation (4). Specifically, INCS is calculated for patients with tetraplegia as,

   $INCS_{tetra} = 0.3342\Delta UEMS_r + 0.2644\Delta LEMS_r + 0.1867\Delta PP_r + 0.2145\Delta LT_r$,

   and for patients with paraplegia as,

   $INCS_{para} = 0.3666\Delta LEMS_r + 0.2922\Delta PP_r + 0.3411\Delta LT_r$

A calculator of INCS implemented in Microsoft Excel is provided (Supplementary Material 1).

Rating of clinical meaningfulness of neurological changes by SCI experts.

A group of 20 clinical SCI experts (including ~~physical rehabilitation specialists~~physiatrists, spine surgeons and neurologists) evaluated the overall recovery of the 88 patients included in the ASCI study database based on the ISNCSCI assessments at 30 and 120 days after injury. All raters have been actively assessing for years the neurological function of patients according to ISNCSCI. The majority of them are also certified by the EMSCI-ISNCSCI training course at Heidelberg University Hospital [19]. The parameters presented to them include the AIS grade, NLI, UEMS, LEMS, and total PP and LT scores. ~~All raters were experienced in SCI evaluation, the majority of them certified by the EMSCI-ISNCSCI training course at Heidelberg University Hospital [18].~~ Each expert rated independently, based on his/her own criteria and experience, the overall recovery of patients on a 5-point Likert scale: 1) strong worsening, 2) moderate worsening, 3) non-significant change, 4) moderate improvement or 5) strong improvement. None of the raters had previous information about how the score was developed (see below) to avoid biases in their ratings. In addition, the raters were blind to the calculated INCS score for any of the 88 patients at all time.

Statistical analysis.

All statistical analysis were performed using R language for statistical computing [18] in R Studio. Before comparing more than two groups with one-way analysis of variance (ANOVA), normality of residuals was evaluated by Shapiro's test and homocedasticity by Levene's test using the "car" package[20]. When Shapiro's test indicated that it could not be assumed that the residuals were normally distributed, non-parametric Kruskal-Wallis rank sum test was employed instead of one-way ANOVA (data on figures 1C, 2A-F). The level of significance was established at a p-value<0.05. Dunn's post-hoc test was performed after significant results.

To compare the frequencies shown in Table 4, Fisher's exact test was employed.

Regression linear models were used to assess the relationship of INCS with demographical and clinical variables as well as with SCIM III (Tables 3 and 5).

The ability of INCS or the individual ISNCSCI variables to score overall patient recovery in agreement to expert clinicians was evaluated by linear discriminant analysis (LDA). Receiver Operating Characteristic (ROC) curves and Area Under the ROC Curve (AUC) from LDA models were created using the pROC library [21] following one versus all categories methodology.

RESULTS.

The influence of AIS grade, age, lesion level and timing on the spontaneous neurological recovery evaluated is still detected after integrating ISNCSCI variables into INCS.

For patients with tetraplegia INCS is a linear combination of normalized changes in UEMS, LEMS, sum PP and sum LT scores, and for patients with paraplegia is a linear combination of all those variables except normalized changes in UEMS. As shown in Figure 1A, motor function represents closely a 60% of INCS final score for patients with tetraplegia and roughly a 40% for patients with paraplegia (Figure 1B). Figure 1C shows the calculated INCS between 4 and 52 weeks after injury for the 403 Sygen trial patients employed to developed the INCS algorithm. Of note, in spite that INCS formula differs between patients with paraplegia and tetraplegia, INCS values are not different between paraplegia and tetraplegia among AIS B, AIS C and AIS D patients (Figure 1C).

Spontaneous neurological recovery is largely affected by severity (AIS grade), being patients with motor incomplete lesions (AIS C or AIS D) expected to better recover than patients with motor complete lesions (AIS A or AIS B) [4,22–24]. Accordingly, both AIS C and AIS D patients exhibit significantly higher INCS values (better neurological recovery) than AIS A or AIS B patients (Figure 1D-E). Neurological recovery has been also shown to be more frequent among AIS A patients with tetraplegia than among AIS A patients with paraplegia [4,22,24]. In agreement, the median INCS value among AIS A patients is significantly higher in

patients with tetraplegia compared to those with paraplegia (Figure 1C; median values, 0.078 vs 0.005 respectively; Mann-Whitney U test p-value=$6x10^{-9}$).

Even though spontaneous recovery might occur more than a year after SCI, most of it occurs in the first 2-3 months after injury [4,22,25–27]. Accordingly, INCS is higher when calculated between the first weeks after injury than between later time points (Figure 1F).

Age has been reported to negatively affects neurological and functional recovery after SCI [22,23,28–30]. Several cut-off values above which patients are expected to have a poorer neurological outcome has been established. Among these, 50 years has been proposed as the limit to separate patients with a better recovery (younger than 50) from those with a poorer recovery (older than 50) [31–35]. We have explored by multiple linear regression whether INCS is significantly associated to age including as covariates initial AIS grade, tetraplegia vs paraplegia, ~~gender~~sex and treatment (GM-1 vs placebo). We have tested age cut-off values ranging from 45 to 65 years old and a cut-off of 48 years yields the model with the higher adjusted coefficient of determination (r^2; 0.5206), higher age effect and most significant age p-value (Table 3). This model also confirms AIS grade as the strongest predictor of neurological change evaluated by INCS while, in agreement with previous reports about the lack of differences on neurological recovery by sex [36], it does not identify ~~gender~~sex as significantly associated to INCS (Table 3). As expected, treatment with GM-1 is not significantly related with INCS between 4 and 52 weeks (Table 3).

Overall, our data show that the firmly established effects of AIS grade, age, time since injury and lesion level (tetraplegia vs paraplegia) on spontaneous neurological recovery are maintained after integrating UEMS, LEMS, sum PP and sum LT into INCS.

INCS outperforms individual ISNCSCI variables in predicting the clinical judgment about the neurological change of patients with SCI.

The neurological recovery of 88 SCI patients from the ASCI study database was classified by expert raters as i) strong worsening, ii) moderate worsening, iii) non-significant change, iv) moderate improvement or v) strong improvement based on the changes in their

sensorimotor function evaluated by ISNCSCI between 30 and 120 days after injury (Supplementary Table 1). The overall interrater agreement was assessed by the statistic Fleiss kappa [37], resulting in a value of 0.56 ($p<0.001$). Accordingly to Landis and Koch (1977), the agreement was substantial for the categories "non-significant change" (kappa 0.7, $p<0.001$) and "strong improvement" (kappa 0.65, $p<0.001$), moderate for "moderate worsening" (kappa 0.45, $p<0.001$) and "strong worsening" (kappa 0.41, $p<0.001$), and fair for "moderate improvement" (0.30, $p<0.001$). A common classification was developed by the majority vote (Supplementary Table 1), resulting in 1 participant rated as "strong worsening", 3 as "moderate worsening", 47 as "non-significant change", 13 as "moderate improvement" and 24 as "strong improvement".

The differences between two given time points (Δ) of any of the six variables evaluated by raters (AIS grade, NLI, UEMS, LEMS, PP or LT sum scores) are commonly used as indicators of recovery. In spite that they are useful to evaluate specific aspects of sensorimotor recovery, none of them captures the classification by expert clinicians of overall change when used individually. For instance, the distribution of AIS grade conversions across the rating categories is significantly different by Fisher's exact test ($p=0.006$), but 19% of patients classified as showing a "non-significant change" converted at least one AIS grade (a common criterion of change), while 85% and 39% of patients classified as "moderate" or "strong improvement" respectively do not experience any AIS grade conversion (Table 4). Thus, defining overall recovery based on AIS grade conversion may lead to the classification of patients without significant recovery as having recovery and the inclusion of many patients with a significant overall recovery as having no recovery. To note, the magnitude of the neurological recovery associated to a single AIS grade conversion is considered different when a patient converts into motor incomplete (from AIS B to AIS C) than when conversion is within motor incomplete patients (for example AIS C to AIS D) [22]. However, when considering all the other five variables, raters classified the 1 patient converting form AIS B to AIS C as "moderate improvement" and the 2 others as "strong improvement", while the 10 patients converting from AIS C to AIS D were classified as "strong improvement" (Table 4).

The change in NLI between the two times of assessment is not different between the five rating categories (p-value>0.05; Figure 2A).

Using UEMS to measure recovery is only applicable to tetraplegic patients. Still, in our dataset it detects recovery, but it fails to detect worsening, since ΔUEMS values in "strong/moderate worsening" categories are similar to those participants rated as "non-significant change" (Figure 2B).

LEMS do not change (ΔLEMS=0) in "strong" and "moderate worsening" categories as well as in 44 out of the 47 participants rated as "non-significant change". Values of ΔLEMS are significantly higher only in the rating category "strong improvement" (Figure 2C). Therefore, in our dataset ΔLEMS may be adequate to evaluate strong improvements, but it does not reflect worsening (Figure 2C).

Values of ΔPP are significantly higher in the group of "strong improvement" and "moderate improvement" compared to "non-significant change" and "moderate worsening groups (Figure 2D), while ΔLT is significantly higher among the "strong improvement" group compared to "non-significant change" (Figure 2E).

Thus, individual Δ values of AIS grade, NLI, UEMS, LEMS, PP or LT are not appropriate surrogate indicators of overall changes in sensorimotor function. Contrary to any individual ISNCSCI parameter, INCS values are distributed in regular intervals from the worst to the best categories of recovery, with mean values of -0.33, -0.23, 0.03, 0.28 and 0.51 for "strong worsening", "moderate worsening", "non-significant change", "moderate improvement" and "strong improvement", respectively (Figure 2F).

The overall accuracy of the classification among the five categories mentioned above after a linear discriminant analysis (LDA) run with INCS as the single independent variable is 77.3%, while using AIS grade, NLI, UEMS, LEMS, sum PP or sum LT as independent variables are 59%, 53.4%, 68%, 75%, 68% and 62.5%, respectively. And, more importantly, INCS is able to better predict all the categories by linear discriminant analysis, as shown by ROC-AUCs for every rating category (Figure 2G-K).

Association of INCS with functional recovery.

In a subset of 75 out of the 88 patients from the ASCI database, the Spinal Cord Independence Measure (SCIM) III was determined at two different times. The first evaluation was performed at 30 days after injury and the second at discharge (170±18 days after injury). Among these, 41 patients suffered from tetraplegia (16 AIS A, 4 AIS B, 8 AIS C and 13 AIS D) and 34 from paraplegia (28 AIS A, 4 AIS B and 2 AIS C; no AIS D patients with paraplegia). Functional recovery was determined as the difference between SCIM III values at 30 days after injury and discharge. Figure 3A shows that increase of the total SCIM score between 30 days after injury and discharge strongly correlate with INCS between 30 and 120 days after injury. By SCIM sub-scales, self-care (items 1-4) and mobility (items 9-17) correlate with INCS (Figure 3B, D), while respiration and sphincter management (items 5-8) do not. As shown in Figure 3A, B and D, correlation of SCIM differences and INCS is more obvious in patients with tetraplegia (red dots in Figure 3) than in those with paraplegia (orange dots). Actually, by multiple linear regression INCS is significantly associated with the increase of the total SCIM score as well as with the increase in self-care and mobility subscales in patients with tetraplegia, but not in patients with paraplegia (Table 5). Interestingly, age was also a factor associated to increase in total SCIM, self-care and mobility among patients with tetraplegia, while time between first and second SCIM assessments (closely time of rehabilitation) was a factor significantly associated to gaining in functional independence in patients with paraplegia (Table 5).

Overall, our results show that recovery of neurological function measured as INCS is related to functional motor recovery in patients with tetraplegia.

DISCUSSION.

Our results support INCS as an unbiased and integrative quantification of the change in sensorimotor function of patients with SCI that parallels the clinical judgments of experienced clinicians. As expected for a measure of sensorimotor function change, INCS is

dependent on the time of assessment since injury and is strongly associated to severity (intial AIS grade). In addition, INCS is associated to improvements in SCIM in patients with tetraplegia. Overall, INCS could be a useful readout measure for clinical studies.

Interpreting INCS.

As stated above, INCS values range between -1 and +1. Among all the possible values, only -1, and +1 have an absolute and unique interpretation: -1 is the maximum theoretical worsening for a patient according to ISNCSCI and +1 means full recovery of a patient. All the other values have not a unique interpretation. For example, INCS=0 most probably would reflect that the patient did not experience any change in UEMS, LEMS, total PP nor total LT, but it could be achieved also by balancing the gaining and loss in these variables. Of note, INCS=+1 will always indicate full recovery, but the opposite may not be true: some patients with full recovery might not have an INCS=+1. This will happen in cases with minor neurological deficits after SCI, like patient #48 in Supplementary Table 1. This patient was at the first assessment AIS D C6, with UEMS=49 and LEMS, PP and LT fully preserved, and at the second assessment converted into AIS E. Since LEMS, PP and LT already present their maximum value at the initial assessment, differences in LEMS, PP and LT were zero, only the differences in UEMS accounted for the score and, thus, INCS=0.33. In spite this patient fully recovered, 12 out of the 20 clinicians rated the neurological change of this patient as “non-significant change”, and the other 8 as “moderate improvement” (Supplementary Table 1). Thus, for these rare situations in which full recovery is not considered a significant improvement, INCS prevents overestimating the recovery.

Even though only INCS values of +1 and -1 are univocally interpretable, INCS provides a quantification of the change in overall sensorimotor function that reflects expert opinion about the clinical meaningfulness of the change and, as discussed below, could be useful as an endpoint for clinical studies.

INCS as an endpoint for clinical trials.

INCS has been developed in an unsupervised manner with the 403 patients included in the Sygen trial with no missing ISNCSCI data at 4 and 52 weeks after injury. As unsupervised, the algorithm to develop INCS was blind to any classification of recovery of these patients. However, when INCS is calculated for the 88 patients from the ASCI database, whose changes were graded by experts, it accurately agrees with their recovery/worsening rating (Figure 2 G-K). Besides this, INCS outperforms any single ISCNSCI variable in capturing overall changes. For instance, the commonly used criterion of improvement, AIS grade conversion, is a weak indicator of expert opinion about overall recovery (Figure 2 G-K). In fact, 19% of patients rated as showing no overall neurological improvement convert in their AIS grade, while 84% and 42% of patients respectively classified as experiencing moderate or significant improvement did not convert their AIS grade (Table 4). This is in line with a recent survey among physicians and patients suggesting that neurological recovery may be underestimated when using only AIS grade conversion [5]. Anyway, in spite of the known drawbacks of AIS grade conversion as endpoint in SCI studies, a gaining of at least two AIS grades is commonly established as the cut-off value for classifying patients with good recovery. However, due to low responsiveness this criterion may require a large number of patients to reach enough statistical power since the percentage of those who convert decreases as time goes by after injury [4]. Indeed, among the 403 patients from the Sygen trial employed to develop INCS, only 5.4% recovered by at least two AIS grades between 4 and 52 weeks after injury. Something similar is observed among patients from the ASCI study, where only 4.5% (4 out of 88) converts at least two AIS grades between 30 and 120 days after injury. A less restrictive criterion of recovery, of at least one AIS grade conversion, will increase the fraction of patients among the good recovery group, but at the expense of widening the limits of what is considered by patients and physicians as meaningful recovery to moderate changes [5]. In addition, when studying therapeutic interventions in SCI, detecting worsening may be as important as detecting improvement. However, attending to AIS grade conversion, worsening cannot be detected among AIS A

patients, while this is still possible with INCS, ~~.~~ rendering it appropriate to additionally evaluate safety in early stage clinical trials.

Other ISNCSCI variables, as UEMS, are only useful to determine recovery or worsening in patients with tetraplegia. LEMS detects strong improvements (ROC-AUC 0.94; Figure 2 K), but it was not as sensitive to detect moderate improvement and worsening among the 88 participants included in the ASCI database.

Important to note, the expected INCS values for spontaneous recovery are highly dependent on lesion severity and time since injury (initial AIS grade; Figure 1C-F, Table 3). Thus, caution should be taken when comparing INCS across different initial AIS grades in non-chronic patients, when spontaneous recovery is still expected. For example, while the median INCS value between 4 and 52 weeks for AIS C patients is around 0.5, this value is extremely rare among AIS A patients: ~~—~~ only 5 out of the 249 AIS A patients reached this value (Figure 1C). Interestingly, among these 5 outliers (1 with paraplegia and 4 with tetraplegia) 2 converted into AIS C and 3 into AIS D, and all of them presented significant increments in LEMS, PP and LT sum scores (median normalized differences 0.42, 0.56 and 0.84 respectively). In addition, the 4 AIS A patients with tetraplegia also presented an exceptional recovery of upper extremities motor function (median normalized difference 0.67), one of them exhibiting full recovery according to ISNCSCI. Further work is warranted to determine the precise distribution of INCS values expected by spontaneous recovery in order to ~~define~~ establish cut-off values that could ~~define~~ be used in clinical trials to define recovery or worsening according to time interval since injury and initial AIS grade. Limitations may be found for clinical trials starting within the first days after injury, since some ISNCSCI assessments may be unreliable [4] and, thus, cut-offs may be established over misestimated INCS values. Also, in this phase a difference of hours in the baseline assessment results in different recovery [39], so many intervals differing in hours should be established.

The assessment of recovery after SCI may be centered in the evaluation of the neurological outcome (regardless of its utility for daily life activities), measured by ISNCSCI,

or in the evaluation of the ability to perform daily life activities (regardless of whether they are achieved by neurological restoration or by compensation/training), measured, for example, by the Spinal Cord injury Independence Measure scale (SCIM) [40,41]. Thus, it could be interesting to use INCS together with differences in SCIM over time, to have an estimation of the corresponding functional independence. In this regard, INCS is a predictive variable of SCIM changes in patients with tetraplegia, suggesting that neurological recovery and restoration of motor functions underlies the improvement in the ability to perform daily life activities in this subset of patients (Table 5). In patients with paraplegia INCS and SCIM are not associated, which could indicate that restoration of daily life activities in these patients are not linked to neurological recovery but to training of the items evaluated by SCIM during the rehabilitation process. Supporting this, the time elapsed between the first and second SCIM evaluations, which closely reflects the time of rehabilitation, is a predictive variable of SCIM improvement (Table 5).

Also, INCS could be used in combination with the Spinal Cord Ability Ruler (SCAR) [42]. Without entering into details, SCAR reflects the motor performance of patients from C5 to C8 spinal segments evaluated according to ISNCSCI and the performance along a series of ordered and selected daily life activities evaluated according to SCIM [42]. Thus, INCS and SCAR could complement each other: SCAR is an interval scale that provides a clinical meaningful quantification of the current volitional performance of patients (with the exception of central cord syndrome) while INCS reflects the overall change in sensorimotor function between two ISNCSCI assessments.

CONCLUSION

INCS is not intended to substitute pre-existing recovery measures. INCS has been developed to fill the existing gap in assessing with a single value the overall change in sensorimotor function evaluated by ISNCSCI with the aim of being useful as a readout for clinical studies.

ACKNOWLEDGEMENTS.

This work was funded by the Ministry of Science and Innovation of Spain; Grant identification number PID2020-120652RB-I00 to DGO and AAM.

CONFLICTS OF INTEREST.

The authors declare that there is no conflict of interest.

ETHICS STATEMENT.

The study protocol and the informed consent sheet of the ASCI study were evaluated and approved by the Ethics Committee of the Toledo Health Care Area and by the Ethics Committee of the Bavarian Medical Board. The study follows and adheres to the World Medical Association Declaration of Helsinki and is registered at the public database Clinicaltrial.gov (registration number NCT02493543).

REFERENCES.

1. Rupp, R., Biering-Sørensen, F., Burns, S.P., Graves, D., Guest, J., Jones, L., Read, M.S., Rodriguez, G.M., Schuld, C., Tansey, K.E., Walden, K., and Kirshblum, S. (2021). International Standards for Neurological Classification of Spinal Cord Injury: Revised 2019. Top. Spinal Cord Inj. Rehabil. 27, 1–22.
2. Wirz, M., Zörner, B., Rupp, R., and Dietz, V. (2010). Outcome after incomplete spinal cord injury: Central cord versus Brown-Sequard syndrome. Spinal Cord 48, 407–414.
3. Oleson, C. V., Burns, A.S., Ditunno, J.F., Geisler, F.H., and Coleman, W.P. (2005). Prognostic value of pinprick preservation in motor complete, sensory incomplete spinal cord injury. Arch. Phys. Med. Rehabil. 86, 988–992.
4. Fawcett, J.W., Curt, A., Steeves, J.D., Coleman, W.P., Tuszynski, M.H., Lammertse,

D., Bartlett, P.F., Blight, A.R., Dietz, V., Ditunno, J., Dobkin, B.H., Havton, L.A., Ellaway, P.H., Fehlings, M.G., Privat, A., Grossman, R., Guest, J.D., Kleitman, N., Nakamura, M., Gaviria, M., and Short, D. (2007). Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: Spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. Spinal Cord 45, 190–205.

5. ter Wengel, P.V., Post, M.W.M., Martin, E., Stolwijk-Swuste, J., Hosman, A.J.F., Sadiqi, S., Vandertop, W.P., and Öner, F.C. (2020). Neurological recovery after traumatic spinal cord injury: what is meaningful? A patients' and physicians' perspective. Spinal Cord 58, 865–872.
6. Steeves, J.D., Lammertse, D., Curt, A., Fawcett, J.W., Tuszynski, M.H., Ditunno, J.F., Ellaway, P.H., Fehlings, M.G., Guest, J.D., Kleitman, N., Bartlett, P.F., Blight, A.R., Dietz, V., Dobkin, B.H., Grossman, R., Short, D., Nakamura, M., Coleman, W.P., Gaviria, M., and Privat, A. (2007). Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP panel: Clinical trial outcome measures. Spinal Cord 45, 206–221.
7. Geisler, F.H., Coleman, W.P., Grieco, G., and Poonian, D. (2001). The Sygen® multicenter acute spinal cord injury study. Spine (Phila. Pa. 1976). 26, S87–S98.
8. Tong, B., Jutzeler, C.R., Cragg, J.J., Grassner, L., Schwab, J.M., Casha, S., Geisler, F., and Kramer, J.L.K. (2018). Serum Albumin Predicts Long-Term Neurological Outcomes After Acute Spinal Cord Injury. Neurorehabil. Neural Repair 32, 7–17.
9. Warner, F.M., Jutzeler, C.R., Cragg, J.J., Tong, B., Grassner, L., Bradke, F., Geisler, F., and Kramer, J.K. (2019). The Effect of Non-Gabapentinoid Anticonvulsants on Sensorimotor Recovery After Human Spinal Cord Injury. CNS Drugs 33, 503–511.
10. Geisler, F.H., Coleman, W.P., Grieco, G., and Poonian, D. (2001). Measurements and recovery patterns in a multicenter study of acute spinal cord injury. Spine (Phila. Pa. 1976). 26, S68–S86.
11. Arevalo-Martin, A., Grassner, L., Garcia-Ovejero, D., Paniagua-Torija, B., Barroso-

Garcia, G., Arandilla, A.G., Mach, O., Turrero, A., Vargas, E., Alcobendas, M., Rosell, C., Alcaraz, M.A., Ceruelo, S., Casado, R., Talavera, F., Palazón, R., Sanchez-Blanco, N., Maier, D., Esclarin, A., and Molina-Holgado, E. (2018). Elevated Autoantibodies in Subacute Human Spinal Cord Injury Are Naturally Occurring Antibodies. Front. Immunol. 9, 2365.

12. Geisler, F.H., Coleman, W.P., Grieco, G., and Poonian, D. (2001). Recruitment and early treatment in a multicenter study of acute spinal cord injury. Spine (Phila. Pa. 1976). .
13. Burns, A.S., Lee, B.S., Ditunno, J.F., and Tessler, A. (2003). Patient selection for clinical trials: The reliability of the early spinal cord injury examination. J. Neurotrauma 20, 477–482.
14. Noy-Meir, I. (1973). Data Transformations in Ecological Ordination: I. Some Advantages of Non-Centering. J. Ecol. 61, 329–341.
15. Cadima, J., and Jolliffe, I. (2009). On relationships between uncentred and column-centred principal component analysis. Pakistan J. Stat. 25, 473–503.
16. Alexandris, N., Gupta, S., and Koutsias, N. (2017). Remote sensing of burned areas via PCA, Part 1; centering, scaling and EVD vs SVD. Open Geospatial Data, Softw. Stand. 2, DOI 10.1186/s40965-017-0028-1.
17. Alter, O., and Golub, G.H. (2006). Singular value decomposition of genome-scale mRNA lengths distribution reveals asymmetry in RNA gel electrophoresis band broadening. Proc. Natl. Acad. Sci. U. S. A. 103, 11828–11833.
18. R Development Core Team 3.0.1. (2019). R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. .
19. Schuld, C., Wiese, J., Franz, S., Putz, C., Stierle, I., Smoor, I., Weidner, N., and Rupp, R. (2013). Effect of formal training in scaling, scoring and classification of the International Standards for Neurological Classification of Spinal Cord Injury. Spinal Cord 51, 282–288.
20. Fox, J., and Weisberg, S. (2019). An R Companion to Applied Regression Third

Edition. Sage Publications.

21. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., and Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12, 77.

22. Kirshblum, S., Snider, B., Eren, F., and Guest, J. (2021). Characterizing Natural Recovery after Traumatic Spinal Cord Injury. J. Neurotrauma 39, 1267–1284.

23. Bravo, P., Labarta, C., Alcaraz, M.A., Mendoza, J., and Verdú, A. (1996). An assessment of factors affecting neurological recovery after spinal cord injury with vertebral fracture. Paraplegia 34, 164–166.

24. Coleman, W.P., and Geisler, F.H. (2004). Injury severity as primary predictor of outcome in acute spinal cord injury: Retrospective results from a large multicenter clinical trial. Spine J. 4, 373–378.

25. Steeves, J.D., Kramer, J.K., Fawcett, J.W., Cragg, J., Lammertse, D.P., Blight, A.R., Marino, R.J., Ditunno, J.F., Coleman, W.P., Geisler, F.H., Guest, J., Jones, L., Burns, S., Schubert, M., Van Hedel, H.J.A., and Curt, A. (2011). Extent of spontaneous motor recovery after traumatic cervical sensorimotor complete spinal cord injury. Spinal Cord 49, 257–265.

26. Kirshblum, S., Botticello, A., Lammertse, D.P., Marino, R.J., Chiodo, A.E., and Jha, A. (2011). The impact of sacral sensory sparing in motor complete spinal cord injury. Arch. Phys. Med. Rehabil. 92, 376–383.

27. Kirshblum, S., Millis, S., McKinley, W., and Tulsky, D. (2004). Late neurologic recovery after traumatic spinal cord injury. Arch. Phys. Med. Rehabil. 85, 1811–1817.

28. Pollard, M.E., and Apple, D.F. (2003). Factors associated with improved neurologic outcomes in patients with incomplete tetraplegia. Spine (Phila. Pa. 1976). 28, 33–39.

29. Scivoletto, G., Morganti, B., and Molinari, M. (2004). Neurologic recovery of spinal cord injury patients in Italy. Arch. Phys. Med. Rehabil. 85, 485–489.

30. Cifu, D.X., Seel, R.T., Kreutzer, J.S., and McKinley, W.O. (1999). A multicenter investigation of age-related differences in lengths of stay, hospitalization charges, and

outcomes for a matched tetraplegia sample. Arch. Phys. Med. Rehabil. 80, 733–740.

31. Oleson, C. V., Marino, R.J., Leiby, B.E., and Ditunno, J.F. (2016). Influence of Age Alone, and Age Combined With Pinprick, on Recovery of Walking Function in Motor Complete, Sensory Incomplete Spinal Cord Injury. Arch. Phys. Med. Rehabil. 97, 1635–1641.

32. Engel-Haber, E., Zeilig, G., Haber, S., Worobey, L., and Kirshblum, S. (2020). The effect of age and injury severity on clinical prediction rules for ambulation among individuals with spinal cord injury. Spine J. 20, 1666–1675.

33. Penrod, L.E., Hegde, S.K., and Ditunno, J.F. (1990). Age effect on prognosis for functional recovery in acute, traumatic central cord syndrome. Arch. Phys. Med. Rehabil. 71, 963–968.

34. Burns, S.P., Golding, D.G., Rolle, W.A., Graziani, V., and Ditunno, J.F. (1997). Recovery of ambulation in motor-incomplete tetraplegia. Arch. Phys. Med. Rehabil. 78, 1169–1172.

35. Scivoletto, G., Morganti, B., Ditunno, P., Ditunno, J.F., and Molinari, M. (2003). Effects on age on spinal cord lesion patients' rehabilitation. Spinal Cord 41, 457–464.

36. Scivoletto, G., Morganti, B., and Molinari, M. (2004). Sex-related differences of rehabilitation outcomes of spinal cord lesion patients. Clin. Rehabil. 18, 709–713.

37. Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. Psychol. Bull. 76, 378–382.

38. Landis, J.R., and Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159–174.

39. Evaniew, N., Sharifi, B., Waheed, Z., Fallah, N., Ailon, T., Dea, N., Paquette, S., Charest-Morin, R., Street, J., Fisher, C.G., Dvorak, M.F., Noonan, V.K., Rivers, C.S., and Kwon, B.K. (2020). The influence of neurological examination timing within hours after acute traumatic spinal cord injuries: an observational study. Spinal Cord 58, 247–254.

40. Catz, A., Itzkovich, M., Agranov, E., Ring, H., and Tamir, A. (1997). SCIM - Spinal cord

independence measure: A new disability scale for patients with spinal cord lesions. Spinal Cord 35, 850–856.

41. Itzkovich, M., Gelernter, I., Biering-Sorensen, F., Weeks, C., Laramee, M.T., Craven, B.C., Tonack, M., Hitzig, S.L., Glaser, E., Zeilig, G., Aito, S., Scivoletto, G., Mecci, M., Chadwick, R.J., El Masry, W.S., Osman, A., Glass, C.A., Silva, P., Soni, B.M., Gardner, B.P., Savic, G., Bergström, E.M., Bluvshtein, V., Ronen, J., and Catz, A. (2007). The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study. Disabil. Rehabil. 29, 1926–1933.

42. Reed, R., Mehra, M., Kirshblum, S., Maier, D., Lammertse, D., Blight, A., Rupp, R., Jones, L., Abel, R., Weidner, N., Curt, A., and Steeves, J. (2017). Spinal cord ability ruler: An interval scale to measure volitional performance after spinal cord injury. Spinal Cord 55, 730–738.

Table 1: Demographical and clinical characteristics of the 403 patients from the Sygen trial database employed to developed INCS.

	4 weeks	52 weeks
Age		
Mean (range): years	31.8 ± 1.3 (11-69)	
≤ 35: n (%)	261 (64.8 %)	
36-55: n (%)	113 (28 %)	
≥ 56: n (%)	29 (7.2 %)	
~~Gender~~Sex: n (%)		
Females	89 (22 %)	
Males	314 (78 %)	
AIS grade: n (%)		
AIS A	249 (61.8 %)	214 (53.1 %)
AIS B	44 (10.9 %)	48 (11.9 %)
AIS C	65 (16.1 %)	34 (8.4 %)
AIS D	45 (11.2 %)	106 (26.3 %)
AIS E	0 (0%)	1 (0.2 %)
Level of lesion: n (%)		
C1-C4	156 (38.7 %)	117 (29 %)
C5-T1	157 (39 %)	170 (42.2 %)
T2-T10	87 (21.6 %)	105 (26.1 %)
T11-L2	3 (0.7 %)	6 (1.5 %)
Below L2	0 (0 %)	5 (1.2 %)

Abbreviations: AIS, American Spinal Injury Association Impairment Scale.

Table 2: Demographical and clinical characteristics of the 88 patients from the ASCI study database.

	30 days	120 days
Age: years		
Mean (range)	42.4 ± 3.3 (19-79)	
≤ 35: n (%)	36 (41 %)	
36-55: n (%)	30 (34 %)	
≥ 56: n (%)	22 (25 %)	
~~Gender~~Sex: n (%)		
Females	12 (13.6 %)	
Males	76 (86.4 %)	
AIS grade: n (%)		
AIS A	50 (56.8 %)	41 (46.6 %)
AIS B	10 (11.4 %)	12 (13.6 %)
AIS C	14 (15.9 %)	9 (10.2 %)
AIS D	14 (15.9 %)	25 (28.4 %)
AIS E		1 (1.2 %)
Level of lesion		
C1-C4	26 (29.5 %)	22 (25 %)
C5-T1	25 (28.4 %)	26 (29.5 %)
T2-T10	27 (30.7 %)	25 (28.4 %)
T11-L2	10 (11.4%)	14 (15.9 %)
Below L2		1 (1.2 %)

Abbreviations: AIS, American Spinal Injury Association Impairment Scale.

Table 3: Multiple linear regression between INCS at 4-52 weeks and clinical and demographical variables.

	Coefficient (95% CI)	*p* value
Intercept	0.039 (-0.017 to 0.096)	0.17
Initial AIS grade		
A	Reference	
B	0.100 (0.040 to 0.159)	**0.0009**
C	0.385 (0.334 to 0.436)	**$<2 \times 10^{-16}$**
D	0.484 (0.425 to 0.543)	**$<2 \times 10^{-16}$**
Initial Neurological level of injury		
Paraplegia	Reference	
Tetraplegia	0.056 (0.012 to 0.100)	**0.011**
~~Gender~~Sex		
Female	Reference	
Male	0.015 (-0.027 to 0.059)	0.47
Age		
< 48 years	Reference	
≥ 48 years	-0.057 (-0.107 to -0.006)	**0.025**
Treatment		
Placebo	Reference	
GM-1	0.014 (-0.022 to 0.050)	0.45

Abbreviations: AIS, American Spinal Injury Association Impairment Scale

Table 4: Distribution of AIS grade improvements across rating categories of neurological change.

Δ AIS grade	A AIS grade	AIS grade conversion	Strong worsening	Strong worsening	Moderate worsening	Moderate worsening	Non-significant change	Non-significant change	Moderate improvement	Moderate improvement	Strong improvement	Strong improvement
-1	-1	B-A	0	0	1 (33 %)	1	0	0	0	0	0	0
0	0	A-A	1 (100 %)	1	2 (67 %)	2	38 (81 %)	32	11(84%)	4	10 (41.6 %)	1
0	1	B-B	0	0	0	0	7 (15 %)	5	2 (16 %)	0	12 (50 %)	0
0	2	C-C	0	0	0	0	2 (4 %)	0	0	3	1 (4.2 %)	1
0	3	D-D	0	0	0	0	0	1	0	4	1 (4.2 %)	8
1		A-B		0		0		6		1		0
1		B-C		0		0		0		1		2
1		C-D		0		0		0		0		10
1		D-E		0		0		1		0		0
2		A-C		0		0		2		0		0
2		B-D		0		0		0		0		1
3		A-D		0		0		0		0		1

Abbreviations: AIS, American Spinal Injury Association Impairment Scale.

Table 5: Multiple linear regression between difference in SCIM from 30 days after injury to discharge and INCS, clinical and demographical variables.

		Patients with tetraplegia		Patients with paraplegia	
		Coefficient (95% CI)	*p* value	Coefficient (95% CI)	*p* value
ΔTotal SCIM	Intercept	51.8 (18.9 to 84.8)	**0.003**	39.5 (9.1 to 69.9)	**0.01**
	INCS	42.9 (13.5 to 72.3)	**0.005**	-19.5 (-95.6 to 56.6)	0.60
	Basal AIS grade				
	A	*Reference*		*Reference*	
	B	2 (-25.5 to 29.5)	0.88	1.9 (-21.9 to 25.7)	0.87
	C	19.6 (-3.6 to 42.9)	0.09	4.4 (-44.3 to 53.2)	0.85
	D	11.4 (-14.5 to 37.3)	0.37		
	~~Gender~~Sex				
	Female	*Reference*		*Reference*	
	Male	-7.8 (-28.7 to 13)	0.44	11.9 (-9.8 to 33.8)	0.26
	Age	-0.6	**0.02**	-0.15 (-0.58 to 0.26)	0.44
	ΔTime	-0.01 (-0.16 to 0.12)	0.81	0.08 (0.003 to 0.17)	**0.04**
ΔItems 1-4	Intercept	11.2 (2.9 to 19.5)	**0.009**	9.5 (-0.83 to 19.9)	0.06
	INCS	14.6 (7.2 to 22)	**0.0003**	5.4 (-20.7 to 31.2)	0.68
	Basal AIS grade				
	A	*Reference*		*Reference*	
	B	2.5 (-4.4 to 9.4)	0.46	-0.7 (-8.9 to 7.3)	0.84
	C	5.3 (-0.5 to 11.2)	0.07	-8.4 (-25 to 8.1)	0.30
	D	0.6 (-5.8 to 7.1)	0.84		
	~~Gender~~Sex				
	Female	*Reference*		*Reference*	
	Male	-1.7 (-7 to 3.5)	0.50	4.3 (-3.1 to 11.7)	0.24
	Age	-0.16 (-0.28 to -0.03)	**0.01**	-0.09 (-0.23 to 0.05)	0.20
	ΔTime	-0.01 (-0.04 to 0.02)	0.57	0.02 (-0.0008 to 0.05)	0.056
ΔItems 5-8	Intercept	23 (8.1 to 37.9)	**0.003**	20.3 (5 to 35.5)	**0.01**
	INCS	12 (-1.1 to 25.3)	0.07	-21.3 (-59.5 to 16.8)	0.26
	Basal AIS grade				
	A	*Reference*		*Reference*	
	B	-2.5 (-14.9 to 9.8)	0.67	2.1 (-9.8 to 14)	0.71
	C	5.6 (-4.8 to 16.1)	0.28	3.8 (-20.6 to 28.2)	0.75
	D	-0.35 (-12 to 11.3)	0.95		
	~~Gender~~Sex				
	Female	*Reference*		*Reference*	
	Male	-1.3 (-10.7 to 8)	0.77	2.2 (-8.6 to 13.2)	0.67
	Age	-0.21 (-0.43 to 0.006)	0.056	0.006 (-0.2 to 0.2)	0.95
	ΔTime	0.007 (-0.05 to 0.07)	0.82	0.02 (-0.01 to 0.06)	0.25
ΔItems 9-17	Intercept	17.5 (4.7 to 30.3)	**0.008**	9.7 (-1.1 to 20.5)	0.07
	INCS	16.2 (4.8 to 27.6)	**0.006**	-3.4 (-30.6 to 23.8)	0.79
	Basal AIS grade				
	A	*Reference*		*Reference*	
	B	2 (-8.6 to 12.7)	0.69	0.53 (-7.9 to 9)	0.89
	C	8.6 (-0.4 to 17.6)	0.06	9.1 (-8.3 to 26.5)	0.29
	D	11.1 (1 to 21.2)	**0.03**		
	~~Gender~~Sex				
	Female	*Reference*		*Reference*	
	Male	-4.7 (-12.8 to 3.3)	0.23	5.4 (-2.4 to 13.2)	0.16
	Age	-0.2 (-0.4 to -0.01)	**0.03**	-0.07 (-0.2 to 0.07)	0.32
	ΔTime	-0.01 (-0.07 to 0.04)	0.61	0.03 (0.005 to 0.06)	**0.02**

Abbreviations: CI, confidence interval; AIS, American Spinal Injury Association Impairment Scale. ΔTime, time interval (days) between first and second SCIM evaluations. Models' adjusted R^2 and p-values: total SCIM tetraplegia: 0.43, p=0.001; total SCIM paraplegia: 0.07, p=0.25; items 1-4 tetraplegia: 0.53, p=0.0001; items 1-4 paraplegia: 0.14, p=0.12; items 5-8 tetraplegia: 0.11, p=0.16; items 5-8 paraplegia: 0.06, p=0.26; items 9-17 tetraplegia: 0.58, p=0.00002; items 9-17 paraplegia: 0.58, p=0.07.

FIGURE LEGENDS

Figure 1: Overall recovery measured as INCS is dependent on severity and time after injury. (A) Relative contribution of normalized UEMS, LEMS, sum PP and sum LT to INCS for patients with tetraplegia or (B) these variables but UEMS for patients with paraplegia. (C) INCS between 4 and 52 weeks after injury for the 91 patients with paraplegia and the 312 patients with tetraplegia employed to develope the score. Kruskal-Wallis test shows that INCS vary significantly across AIS grades both in patients with paraplegia and tetraplegia. (D) Post-hoc Dunn's tests confirm that AIS C and D patients with paraplegia exhibit a greater recovery (higher INCS values) than AIS A patients. (E) The same difference is observed in patients with tetraplegia, where in addition AIS C and D patients exhibit higher INCS values than AIS B patients. (F) Cumulative INCS starting at 1 week after injury up to 52 weeks were calculated for every time interval for the Sygen trial participants. Individuals were grouped by initial AIS grade and tetraplegia or paraplegia. Only complete time series –participants with UEMS, LEMS, total PP and total LT scores recorded at all time points– were included. Median values of INCS at each time point are displayed. Time series of AIS B and D patients with paraplegia were composed by less than 3 participants and are not shown. Numbers above the boxplots in (C) or at the side of temporal curves in (F) indicate the sample size. * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Figure 2: Alignment of individual ISNCSCI variables or INCS with the clinical judgment of neurological change. Specialists in SCI rated the overall neurological change between 30 and 120 days after injury of 88 SCI patients recruited for the ASCI study. Each patient was assigned to a definitive category based on majority vote. The differences (Δ) between 30 and 120 days in the neurological level of injury

(NLI; panel A), upper extremity motor score (UEMS; panel B), lower extremity motor score (LEMS; panel C), total pin prick test (PP; panel D) and total light touch test (LT; panel E) are significantly different between rating categories in all cases but NLI (Kruskal-Wallis p-value<0.05). ΔUEMS is significantly higher among patients classified as showing "moderate" or "strong improvement" vs. those rated as "non-significant change" (B), while ΔLEMS is higher only among patients classified as "strong improvement" (C). ΔPP is significantly higher among patients rated as experiencing strong or moderate improvement vs. those classified as "non-significant change" (D), while ΔLT failed to reach statistical significance between "moderate improvement" and "non-significant change" (E). INCS values between 30 and 120 days after injury are evenly distributed and significantly different between rating categories (F; Kruskal-Wallis p-value<0.05). INCS outperforms individual differences in AIS grade, NLI, UEMS, LEMS, PP or LT in correctly classifying the patients among rating categories by linear discriminant analysis (G-K). Text inside the panels represent AUC-ROC values and 95% confidence intervals (among parenthesis). Confidence intervals could not be estimated for "strong worsening" category (constituted by a single patient). K-W stands for Kruskal-Wallis. * Dunn's test p<0.05; ** Dunn's test p<0.01; *** Dunn's test p<0.001.

Figure 3: Neurological recovery measured as INCS is related to functional recovery measured by differences in SCIM over time. (A) Total SCIM differences between 30 days after injury and discharge significantly correlate with INCS between 30 and 120 days after injury. This correlation is also observed in subscales of self-care (items 1-4) (B) and mobility (items 9-17) (D), but not in respiration and sphincter

management (items 5-8) (C). Shadowed grey areas depicts the 95% confidence interval of regression lines. Abbreviations: par, paraplegia; tet, tetraplegia.

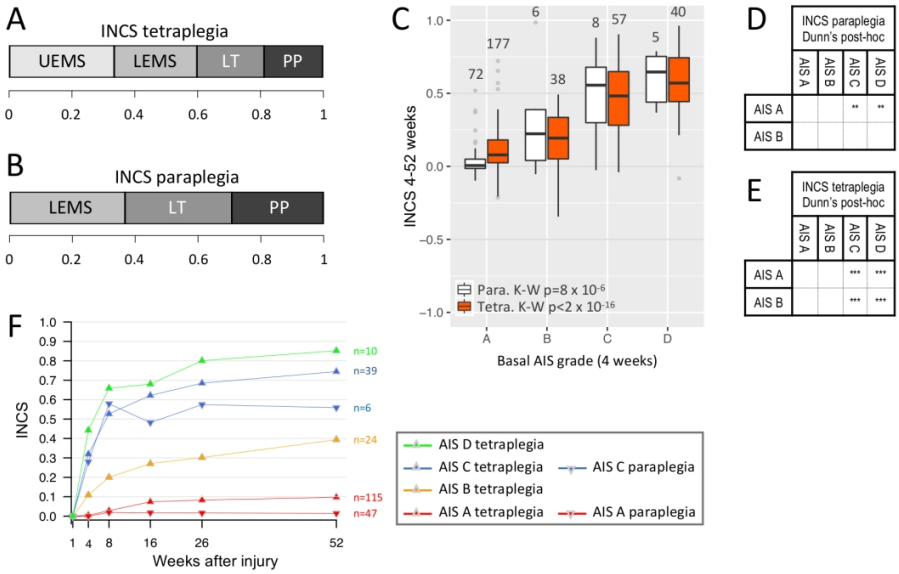


Figure 1: Overall recovery measured as INCS is dependent on severity and time after injury. (A) Relative contribution of normalized UEMS, LEMS, sum PP and sum LT to INCS for patients with tetraplegia or (B) these variables but UEMS for patients with paraplegia. (C) INCS between 4 and 52 weeks after injury for the 91 patients with paraplegia and the 312 patients with tetraplegia employed to develope the score. Kruskal-Wallis test shows that INCS vary significantly across AIS grades both in patients with paraplegia and tetraplegia. (D) Post-hoc Dunn's tests confirm that AIS C and D patients with paraplegia exhibit a greater recovery (higher INCS values) than AIS A patients. (E) The same difference is observed in patients with tetraplegia, where in addition AIS C and D patients exhibit higher INCS values than AIS B patients. (F) Cumulative INCS starting at 1 week after injury up to 52 weeks were calculated for the Sygen trial participants. Individuals were grouped by initial AIS grade and tetraplegia or paraplegia. Only complete time series –participants with UEMS, LEMS, total PP and total LT scores recorded at all time points– were included. Median values of INCS at each time point are displayed. Time series of AIS B and D patients with paraplegia were composed by less than 3 participants and are not shown. Numbers above the boxplots in (C) or at the side of temporal curves in (F) indicate the sample size. * p<0.05; ** p<0.01; *** p<0.001.

175x120mm (300 x 300 DPI)

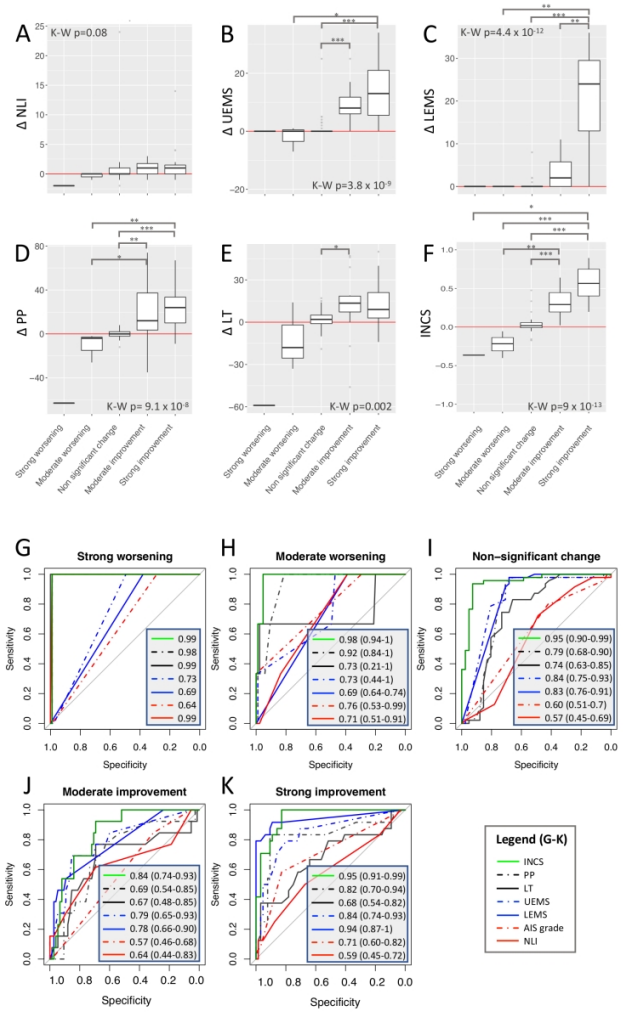


Figure 2: Alignment of individual ISNCSCI variables or INCS with the clinical judgment of neurological change. Specialists in SCI rated the overall neurological change between 30 and 120 days after injury of 88 SCI patients recruited for the ASCI study. Each patient was assigned to a definitive category based on majority vote. The differences (Δ) between 30 and 120 days in the neurological level of injury (NLI; panel A), upper extremity motor score (UEMS; panel B), lower extremity motor score (LEMS; panel C), total pin prick test (PP; panel D) and total light touch test (LT; panel E) are significantly different between rating categories in all cases but NLI (Kruskal-Wallis p-value<0.05). ΔUEMS is significantly higher among patients classified as showing "moderate" or "strong improvement" vs. those rated as "non-significant change" (B), while ΔLEMS is higher only among patients classified as "strong improvement" (C). ΔPP is significantly higher among patients rated as experiencing strong or moderate improvement vs. those classified as "non-significant change" (D), while ΔLT failed to reach statistical significance between "moderate improvement" and "non-significant change" (E). INCS values are evenly distributed and significantly different between rating categories (F; Kruskal-Wallis p-value<0.05). INCS outperforms individual differences in AIS grade, NLI, UEMS, LEMS, PP or LT in correctly classifying the patients among rating categories by linear

discriminant analysis (G-K). Text inside the panels represent AUC-ROC values and 95% confidence intervals (among parenthesis). Confidence intervals could not be estimated for "strong worsening" category (constituted by a single patient). K-W stands for Kruskal-Wallis. * Dunn's test $p<0.05$; ** Dunn's test $p<0.01$; *** Dunn's test $p<0.001$.

190x275mm (300 x 300 DPI)

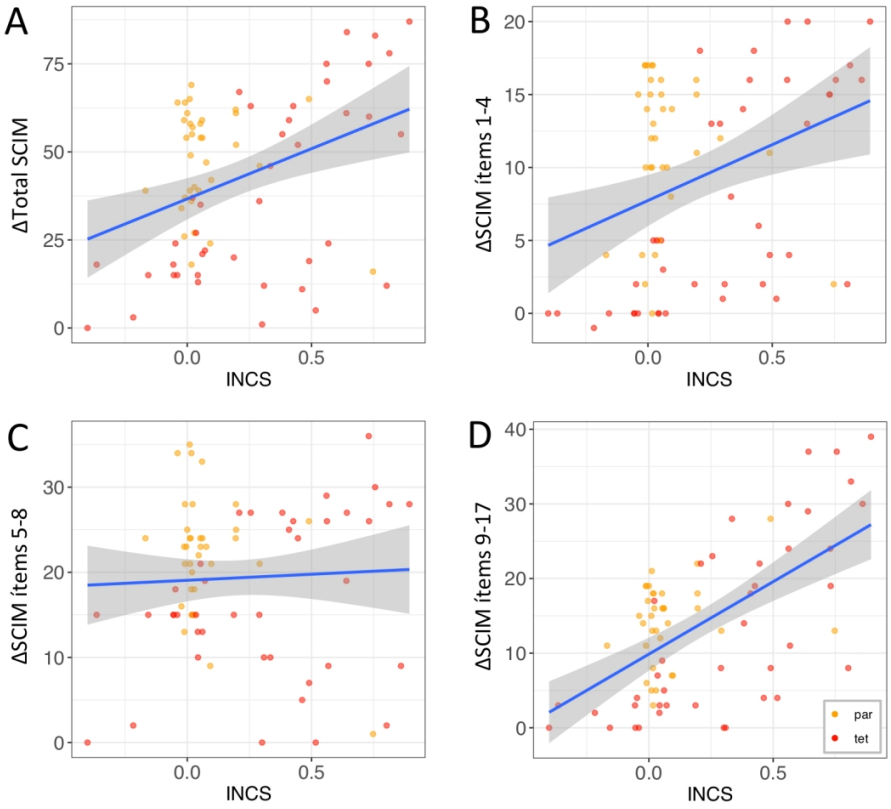
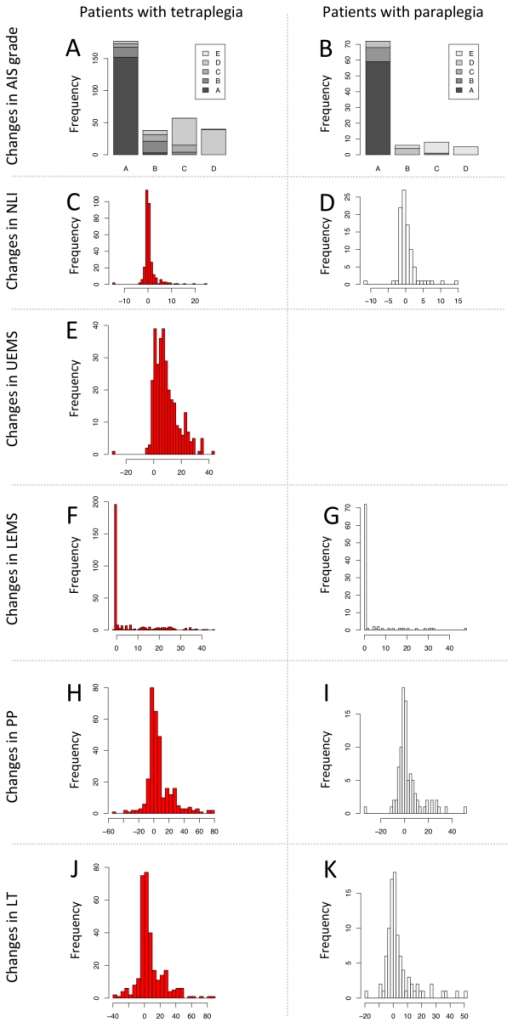


Figure 3: Neurological recovery measured as INCS is related to functional recovery measured by differences in SCIM over time. (A) Total SCIM differences between 30 days after injury and discharge significantly correlate with INCS between 30 and 120 days after injury. This correlation is also observed in subscales of self-care (items 1-4) (B) and mobility (items 9-17) (D), but not in respiration and sphincter management (items 5-8) (C). Shadowed grey areas depicts the 95% confidence interval of regression lines. Abbreviations: par, paraplegia; tet, tetraplegia.

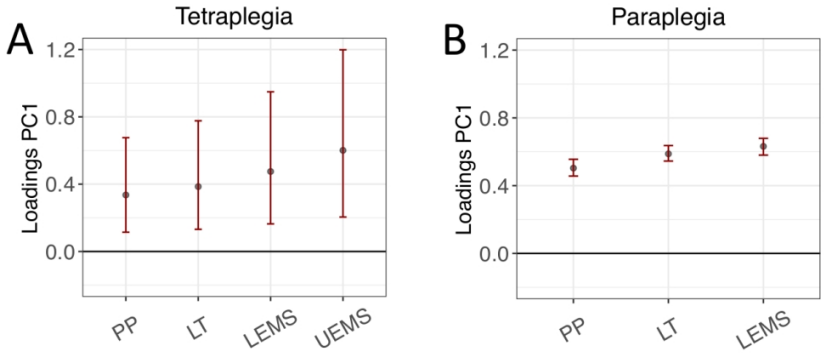
146x128mm (300 x 300 DPI)



Supplementary Figure 1: Changes in ISNCSCI parameters between 4 and 52 weeks of the 403 selected participants in the Sygen trial. (A, B) X-axis represents AIS grade at 4 weeks and bar graphs are colored according to AIS grade at 52 weeks. (C, D) X-axis represents the number of spinal segments recovered (positive values) or lost (negative values) between 4 and 52 weeks . (E-K) X-axis values represents the differences in the scores between 4 and 52 weeks.

Supplementary Figure 1: Changes in ISNCSCI parameters between 4 and 52 weeks of the 403 selected participants in the Sygen trial. (A, B) X-axis represents AIS grade at 4 weeks and bar graphs are colored according to AIS grade at 52 weeks. (C, D) X-axis represents the number of spinal segments recovered (positive values) or lost (negative values) between 4 and 52 weeks . (E-K) X-axis values represents the differences in the scores between 4 and 52 weeks.

190x275mm (300 x 300 DPI)



Supplementary Figure 2: Singular Value Decomposition-based Principal Component Loadings for the relativized differences of AIS grade, NLI, UEMS, LEMS, PP and LT were estimated by bootstrap resampling method for patients with tetraplegia (A) and paraplegia (B; mean values and 95% confidence intervals are shown).

127x62mm (300 x 300 DPI)