

---

---

# Outcome Measures for Acute/Subacute Cervical Sensorimotor Complete (AIS-A) Spinal Cord Injury During a Phase 2 Clinical Trial

John D. Steeves,<sup>1</sup> Daniel P. Lammertse,<sup>2</sup> John L.K. Kramer,<sup>1</sup> Naomi Kleitman,<sup>3\*</sup> Sukhvinder Kalsi-Ryan,<sup>4</sup> Linda Jones,<sup>5</sup> Armin Curt,<sup>6</sup> Andrew R. Blight,<sup>7</sup> and Kim D. Anderson<sup>8</sup>

<sup>1</sup>ICORD, University of British Columbia and Vancouver Coastal Health, Vancouver, BC, Canada; <sup>2</sup>Craig Hospital, Englewood, Colorado, USA; <sup>3</sup>National Institutes of Health/National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA; <sup>4</sup>Toronto Western Hospital, Lyndhurst Centre, University of Toronto, Toronto, Ontario, Canada; <sup>5</sup>Geron Corporation, Menlo Park, California, USA; <sup>6</sup>Balgrist Hospital and EMSCI Network, University of Zurich, Zurich, Switzerland; <sup>7</sup>Acorda Therapeutics, Hawthorne, New York, USA; <sup>8</sup>The Miami Project to Cure Paralysis, University of Miami, Miami, Florida, USA

Effective treatment after cervical spinal cord injury (SCI) is imperative as so many activities of daily living (ADLs) are dependent on functional recovery of arm and hand actions. We focus on defining and comparing neurological and functional endpoints that might be used during acute or subacute Phase 2 clinical trials involving subjects with cervical sensorimotor complete SCI (ASIA Impairment Scale [AIS-A]). For the purposes of this review, the trial would examine the effects of a pharmaceutical small molecule, drug, biologic, or cell transplant on spinal tissue. Thus, neurological improvement is the intended consequence and is most directly measured by assessing neurological impairment (eg, motor aspects of the International Standards Neurological Classification of Spinal Cord Injury [ISNCSCI]). However, changes in neurological function, even if statistically significant, may not be associated with a clear functional impact (ie, a meaningful improvement in individual activity, such as independent self-care ADLs). The challenge is to measure improvement as precisely as possible (change in impairment), but to define a clinically meaningful response in the context of functional improvement (impact on activity limitations). The principal comparisons focused on elements of the ISNCSCI assessment, including upper extremity motor score and motor level. Personal activity capabilities were also examined at various time points. The data suggest that an improvement of 2 or more motor levels after cervical sensorimotor complete SCI may be a clinically meaningful endpoint threshold that could be used for acute and subacute Phase 2 trials with subjects having sensorimotor complete cervical SCI. **Key words:** activity, function, impairment, ISNCSCI, neurology, responder analysis, SCIM

With an increasing number of therapeutic interventions under development for the treatment of acute or subacute spinal cord injury (SCI), it is important to examine what type of clinical study endpoint can be reliably used to assess and validate whether an intervention has beneficial effects for the targeted spinal cord function and/or underlying biological mechanism. For the purposes of this article, the trial would examine the effects of a pharmaceutical small molecule, drug, biologic, or cell transplant on SCI outcomes.

Currently, no gold standard exists for the treatment of SCI, which would enable investigators

to determine the relative benefits and merits for any novel experimental treatment. In brief, clinical trial outcome measures vary in terms of the precise targets they assess, but an outcome measurement tool should be

1. specific for the therapeutic target,
2. sensitive for detecting small changes in target activity,
3. accurate in terms of low variability from measurement to measurement, and
4. reliably used by different examiners.

---

\*The views expressed here are those of the authors and do not represent those of the National Institutes of Health or the US Government.

**Table 1.** *International Classification of Functioning, Disability and Health (ICF)*

Level or domain	Body Structure /Function	↔ Activity ↔	Participation
What change in outcome is measured?	Improvement (or impairment), ie, neurological or physiological	Functional capacity (or limitation)	Quality of life (or restriction)
Outcome tools (examples)	Motor levels, electrophysiology	SCIM, FIM	SF-36, CHART
What factors influence outcomes within each domain?	<ul style="list-style-type: none"> <li>• CNS integrity</li> <li>• Neural circuits (sensory/motor)</li> </ul>	<ul style="list-style-type: none"> <li>• CNS integrity</li> <li>• Neural circuits</li> <li>• Adaptive behaviors</li> <li>• Rehabilitation</li> <li>• Psychology</li> <li>• Mobility</li> </ul>	<ul style="list-style-type: none"> <li>• CNS integrity</li> <li>• Neural circuits</li> <li>• Adaptive behaviors</li> <li>• Rehabilitation</li> <li>• Psychology</li> <li>• Mobility</li> <li>• Community/family support</li> <li>• Finances</li> <li>• Work/school</li> </ul>
Continually increasing number of independent variables →			

Note: SCIM = Spinal Cord Independence Measure; FIM = Functional Independence Measure; SF-36 = Short Form Health Survey; CHART = Craig Handicap Assessment and Reporting Technique; CNS = central nervous system.

We focus on defining and comparing primary and secondary neurological and functional endpoints that might be used during acute or subacute Phase 2 clinical trials involving subjects with cervical sensorimotor complete SCI (ASIA Impairment Scale [AIS-A]). The need for an effective treatment for cervical SCI is critical because so many activities of daily living (ADLs) are dependent on arm and hand function.

Neurological changes focus on the assessment of motor impairment as defined by the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI).<sup>1</sup> Functional ADLs can be evaluated with global assessments, such as the Spinal Cord Independence Measure (SCIM),<sup>2</sup> or specific examinations of upper limb function, such as the Graded Redefined Assessment of Strength, Sensibility and Prehension (GRASSP).<sup>3,4</sup>

A useful framework for conceptualizing trial outcome measurements is the *International Classification of Functioning, Disability and Health* (ICF). The ICF is a taxonomic structure or biopsychosocial model (endorsed by the World Health Organization [WHO])<sup>5</sup> categorizing factors and issues that influence the various domains of function due to a disabling condition across a continuum from Body Structure/Function

through Activity to Participation (**Table 1**). Within the Activity domain of ICF, there are more independent variables influencing a measure of recovery than within the Body Structure/Function domain. The Participation domain has even more potentially uncontrollable (independent) variables that can influence the outcomes for an individual living with SCI.

The greater the number of independent variables within a clinical trial, the more difficult it is to be certain that any measured change is due to the effects of the therapeutic being examined. Thus, if the designated or known target of an experimental treatment is the central nervous system (CNS), it is more relevant and precise to use a neurological outcome measurement. Ultimately, for regulatory approval of a treatment, statistically significant differences between the experimental and control group must also be shown to be clinically significant. Can we link a change in a neurological outcome (ICF: Body Structure/Function) to a functional or clinically meaningful change (ICF: Activity)?

To achieve such goals, it is essential to

1. Understand the degree of spontaneous recovery (change in neurological impairment and related change in personal

activity) within the target population to be enrolled in the clinical study (eg, sensorimotor complete cervical spinal cord injury).

2. Investigate the strength of correlation between various neurological measures (eg, ISNCSCI) and the relevant activity items of SCIM or GRASSP.
3. Develop a responder definition (threshold) that could be used to define a clinically meaningful improvement that is justified by an association between an impact for improved independent activities with specific enhanced neurological criteria.
4. Determine the best method and statistical limitations for detecting significant differences in response between the experimental and control arms of a study (comparison of average values or proportion of responders within each study group).
5. Consider the pragmatic limitations (eg, necessary sample size, duration of study, demands of undertaking chosen outcome measures) for acute or subacute studies of SCI.

Because SCI therapies conceived as acute stage treatments are likely to be administered within hours or days after SCI, it is important that the selected outcome tools have the ability to accurately and sensitively track meaningful changes across a broad timeframe. There are also important differences in the goals of Phase 2 and Phase 3 clinical trials. A Phase 2 study (proof of concept) is an exploratory evaluation of efficacy, with the objective of demonstrating appropriate target activity of the therapeutic, as well as indicating the potential effect size and variability of an experimental therapy in comparison to a useful control group. Information is gained regarding choice of optimal clinical endpoints for a larger pivotal Phase 3 confirmatory trial of efficacy. Combined Phase 1/2 trials, where safety and bioactivity of the therapeutic are evaluated together, can often occur when the Phase 1 trial does not involve healthy subjects and is restricted to people having the clinical disorder. Nevertheless, the data from such a combined Phase 1/2 trial must be able to satisfy the essential outcomes for each respective trial phase. The ongoing evolutions of imaging and electrophysiological techniques

hold great potential as outcome measurement tools for early phase trials. Nevertheless, further refinements and validation studies are necessary to achieve the needed sensitivity and reliable accuracy for widespread adoption.

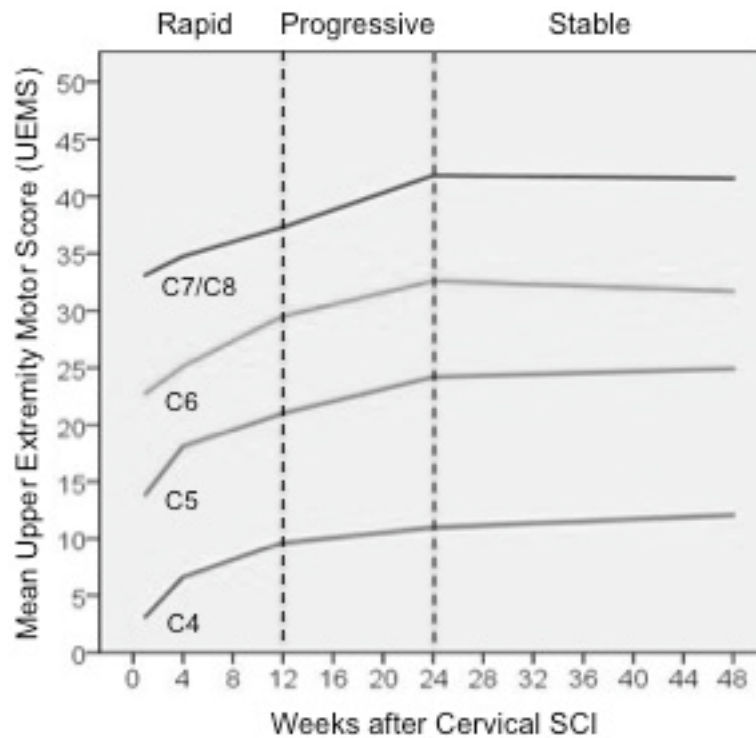
Admittedly, an investigator does not have to declare a primary clinical endpoint for a Phase 2 study (unless they are considering the Phase 2 data as part of the registration application for product approval to a regulatory agency). However, Phase 2 studies are important for establishing and validating the primary outcome to be used in the planned pivotal Phase 3 study. In summary, acute or subacute Phase 2 SCI (neurological) trials should employ outcome measures (eg, neurological endpoints) that are aligned with the specific therapeutic target, especially where achieving the neurological endpoint has been associated with a meaningful improvement in the independent capacity for ADLs (eg, self-care).

### Analysis Approach

Relevant primary literature, reviews, and databases were reviewed. Analyses and comparisons of outcomes after cervical sensorimotor complete (AIS-A) SCI were undertaken and guided by using

1. SYGEN clinical trial database,
2. National Spinal Cord Injury Database from the NIDRR Spinal Cord Injury Model Systems program,
3. EMSCI (European Multicenter study of Spinal Cord Injury) database,
4. “Outcomes Following Spinal Cord Injury” (Consortium for Spinal Cord Medicine: Clinical Practice Guidelines), and
5. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel.

The databases and published literature were mined for relevant and comparative data on the natural history of cervical sensorimotor complete SCI, which includes the temporal aspects of neurological and functional change over the first year after injury. The principal comparisons focused on elements of the ISNCSCI assessment, including upper extremity motor score and motor level.<sup>6-9</sup> In some instances, activity measures (eg, SCIM and GRASSP) have also been assessed at various time points after cervical sensorimotor



**Figure 1.** The average upper extremity motor score (UEMS) and rate of change in UEMS for different initial motor levels of cervical sensorimotor complete (AIS-A) SCI. Note that the average recovery in UEMS is independent of the initial motor level; each subgroup recovers between 8 and 12 motor points in the UEMS, with the most rapid improvement being within the first 3 months after SCI.

complete SCI. The relationship between changes in motor elements of the ISNCSCI assessment and corresponding changes in SCIM were examined over the first year after cervical AIS-A SCI.<sup>9</sup>

The data were evaluated for what could realistically be called a functionally meaningful change for people living with SCI and might be defined as a responder criterion. The data were also considered with different statistical models for comparing clinical trial outcomes including the more classical method of “comparison of means” (computing group averages) for a defined clinical endpoint versus a comparison in the proportion of subjects within each study arm who achieved a predefined threshold of change for a clinical endpoint (ie, responders). Preliminary findings were presented and discussed at a workshop within a combined meeting of the American Spinal Injury Association (ASIA) and the International Spinal Cord Society (ISCoS) in Washington, DC, in June 2011.

### Body Structure and Function

The ISNCSCI is a widely adopted clinical assessment of neurological impairment after SCI and has been used as a tool in acute SCI clinical trials to assess the benefits of therapeutic drugs.<sup>10,11</sup> Changes in AIS grade might be relatively insensitive to detecting a subtle but significant therapeutic benefit.<sup>7,8</sup> However, other aspects of the ISNCSCI exam may be more sensitive, responsive, and reliable to accurately describe a treatment effect.<sup>8,9</sup>

The principal neurological outcomes (components of ISNCSCI) responsive to change over time after cervical sensorimotor SCI were identified as the change in upper extremity motor score (UEMS; muscle strength from 0 to 5 within 5 key muscles of the arm and hand bilaterally), and/or motor level, which is defined as the most caudal spinal level as indexed by the key muscle group for that level having a muscle strength of at least 3 (full range of contraction against

gravity alone) while all of the more rostral key muscles are normal (5/5). Thus, these neurological endpoints might be potentially useful as clinical trial endpoints.<sup>8</sup> Although motor recovery is most dramatic within the first 3 months after SCI, it can continue to improve for many months after the initial injury.<sup>6,8,12</sup>

**Figure 1** describes the pattern of upper extremity motor recovery over the first year after cervical AIS-A SCI using the EMSCI dataset. The data represents the change in the UEMS for several initial motor levels of cervical SCI (C4-C8), and the change in slope represents the relative rate of recovery for different periods after SCI (1-3 months, 3-6 months, 6-12 months). A similar number of UEMS points (~10 over the first year) are recovered, regardless of initial motor level. Similar results were obtained from the SYGEN dataset.<sup>8</sup>

Measuring changes in the total number of motor points does not provide an appreciation for the distribution of the motor points across the various spinal cord segments and thus the potential functional significance cannot be immediately understood.<sup>13</sup> On the other hand, a change in spinal motor level does imply a functional impact, as the ability of the key muscle for that spinal segment must be able to contract against gravity over the full range of motion ( $\geq 3/5$ ). **Table 2** provides a summary for the change in cervical motor level at approximately 6 (n=99) and 12 months (n=81) after injury. Once again the recovery of the number of motor levels is not dependent on the initial motor level. Fully 65% to 70% of individuals with C4-C7 cervical AIS-A SCI will spontaneously recover one or more motor levels over the first year after spinal injury. By

**Table 2.** Proportion (%) of people living with cervical (C4-C7) sensorimotor complete (AIS-A) SCI spontaneously recovering 1 or 2 upper extremity motor levels at different times after injury

Time from baseline	1 motor level <sup>a</sup>	$\geq 2$ motor levels <sup>a</sup>
24 weeks (n=99)	64%	21%
48 weeks (n=81)	68%	26%

<sup>a</sup>Motor level recovery on either right or left side, whichever is greater.

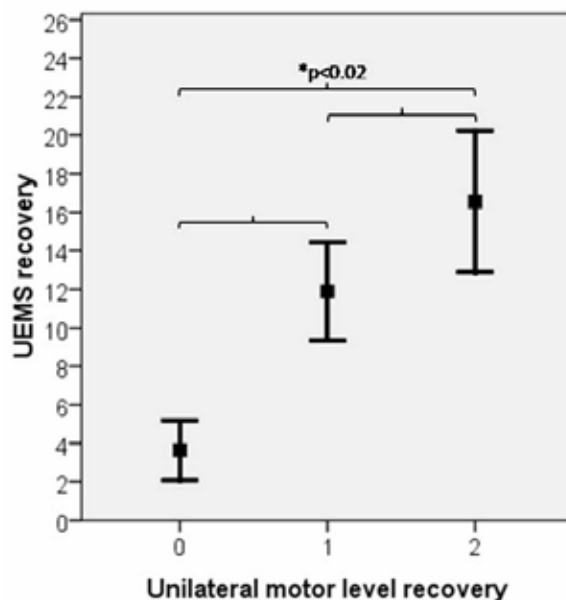
“spontaneous,” we mean with the current standard of care and without the benefit of any specialized or augmented rehabilitation. Similar findings were noted for comparable subjects within the Sygen dataset.<sup>6,8</sup> Because a large proportion of patients with cervical SCI will improve by at least one motor level (64% to 68% over the first year), the use of a single motor level improvement as a trial endpoint will have potential ceiling effects, requiring a large subject population to demonstrate a significant difference between study groups. Conversely, the number of patients spontaneously recovering 2 or more motor levels over the first 6 or 12 months after injury is a more modest proportion (~20% and 25%, respectively).

It is clear that a change in UEMS will influence the number of motor levels recovered, but is the increase in the motor score uniform for each improvement in motor level? As can be seen in **Figure 2**, the answer to this question is no. On average over the first year, a 1 motor level improvement is associated with a 6-point increase in UEMS ( $6.18 \pm 3.58$ ), whereas a 2 motor level improvement relates to an average UEMS increase of only 2 additional motor points ( $8.36 \pm 4.48$ ). Thus a small change in UEMS could mean a significant functional improvement from 1 to 2 motor levels, but also underscores that a change in the overall UEMS might not reflect an individual's functional capacities after cervical SCI.<sup>8,9</sup>

## Activity Measures

Previous panels have described a variety of activities associated with preserved or recovered function at different spinal segmental levels for the cervical cord. Whiteneck et al performed one of the more comprehensive clinical evaluations in 1999 (**Table 3**). It was entitled, Outcomes Following Traumatic Spinal Cord Injury: Clinical Practice Guidelines for Health-Care Professionals, and was funded by the Paralyzed Veterans Association (available at: <http://www.pva.org>).

It is clear from **Table 3** that a change in motor level function can mean an enhanced capacity for independent activity. Recently, a number of new assessment tools have been developed to track activities after SCI. For example, GRASSP



**Figure 2.** Average motor score (upper extremity motor score [UEMS]) recovered over the first year (~1 to 48 weeks) after cervical (C4-C7) sensorimotor complete (AIS-A) SCI as a function of the number of motor levels recovered over the same time period (error bars = 95% confidence interval). Although all comparisons of UEMS as a function of the number of recovered motor levels were statistically significant, only the recovery of 2 motor levels had a significant impact on regaining independent activities for self-care as noted by the SCIM (see Figure 4).

is a recent outcome measure and still in its first version.<sup>3</sup> GRASSP continues to be refined as a focused assessment of upper extremity impairment and function. In brief, within a single protocol, GRASSP is attempting to combine and extend elements of ISNCSCI along with a number of previous activity measures of hand and arm function (Figure 3). GRASSP assesses and tracks subtle differences in sensory, motor, and integrated function of the upper extremity (eg, prehension), but it seems most sensitive and responsive for tracking changes in motor strength and various prehensile abilities of the hand according to innervation. At this time, GRASSP has completed psychometric testing and was found to have good construct validity and high interrater and test-retest reliability.<sup>3,4</sup> When an appropriate number of subjects have been assessed, it will be interesting to correlate changes in GRASSP with neurological outcomes.




The SCIM is one of the better documented activity measures specifically developed for SCI,<sup>2,14-20</sup> and it continues to gain widespread acceptance as a general (global) functional assessment of

activities after SCI. SCIM is currently in its third version and includes 3 subscales of self-care (20 points), respiration and sphincter management (40 points), and mobility (20 points). The most valid SCIM items relating to upper extremity motor recovery after cervical sensorimotor complete SCI are found within the domain of self-care. The principal activities tracked within the self-care subscale relate to the capacity and independence for feeding (3 points), bathing (upper and lower body, 3 points each), dressing upper and lower body, 4 points each), and personal grooming (3 points). SCIM has undergone a number of psychometric evaluations and the self-care subscale is particularly notable for high interrater reliability and internal consistency.<sup>19</sup>

The relationship between a change in SCIM and a neurological outcome has only been reported once where it was suggested that the SCIM could detect segmental changes within the cervical region.<sup>21</sup> The possible correlations have been recently examined closely<sup>9</sup> with a focus on changes in ISNCSCI motor level as they relate to changes in the SCIM self-care subscore over the first year

**Table 3.** Capacity for activities of daily living (ADLs) by individuals with different cervical segmental levels of motor function after cervical SCI

	Function		Capability		
	C1-C3	C4	C5	C6	C7-C8
Bowel					
Bladder					
Bed Mobility					
Transfers					
Pressure Relief					
Eating					
Dressing					
Grooming					
Bathing					
Wheelchair					
Car transport					
Daily Home Care	24 hr attendant	18-24 hr attendant	6-12 hr assistance	4 hr housework	1 hr housework

	Total Assist		Partial Assist		Independent
--	--------------	--	----------------	---	-------------

Note: Modified from Whiteneck et al, 1999. Outcomes Following Traumatic Spinal Cord Injury: Clinical Practice Guidelines for Health-Care Professionals as provided by Consortium for Spinal Cord Medicine through the Paralyzed Veterans Administration (available at: [www.pva.org](http://www.pva.org))

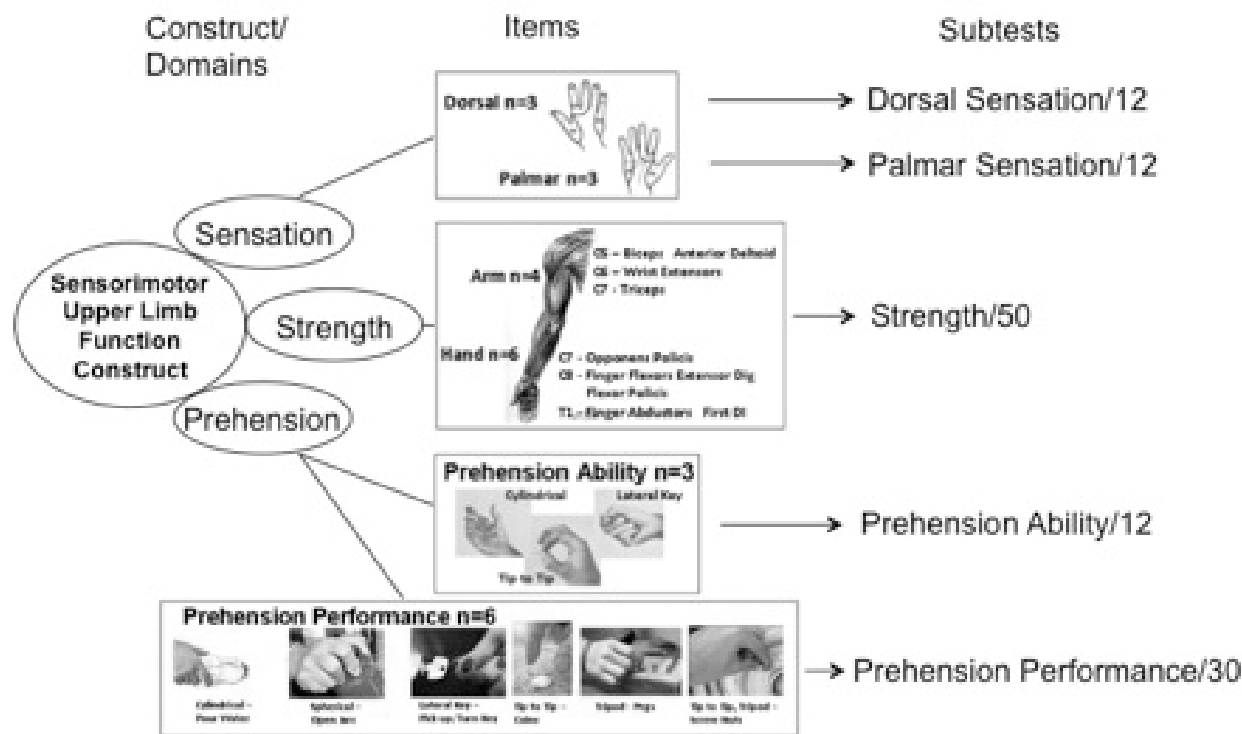
after cervical sensorimotor complete (AIS-A) SCI (**Figure 4**). There is a significant difference in SCIM self-care subscore between individuals recovering 1 versus 2 or more motor levels. In brief, the recovery of 2 motor levels is correlated with a significant increase in functional activities (as assessed by the SCIM self-care subscore). Thus, a 2 motor level improvement in cervical cord function could be a useful clinical endpoint for early phase clinical trials of an experimental therapeutic targeted to the CNS, especially if the desire is to establish an achievable neurological outcome that is correlated with a meaningful improvement in functional activity.

### Relationship Between Measures of Body Structure/Function and Activity

Most clinicians and regulatory agencies require that treatments demonstrate functional benefits to the target patient population. Quantitative

or semi-quantitative biochemical, anatomical, or neurological changes in a beneficial direction would be viewed as evidence of biological activity and only suggestive of therapeutic efficacy. Such outcomes might gain acceptance as clinical trial endpoints, but only if there is a demonstrated positive correlation with measures of improved functional activities of daily living (in principle a neurological impairment measure should really only have a positive and moderate correlation with a functional activity measure, otherwise they would be measuring the same constructs).

As reviewed above, examination of the natural history (patterns of recovery) for neurological motor function (ISNCSCI) and functional activity (SCIM) suggest that recovery of 2 motor levels after a cervical sensorimotor complete AIS-A SCI relates to a significant and potentially clinically meaningful change in the SCIM self-care subscore. Thus regaining 2 motor levels may be a reasonable and pragmatic clinical endpoint for this patient



**Figure 3.** The GRASSP is a multi-modal test based on a construct of "sensorimotor upper limb function." The GRASSP consists of sensory testing with Semmes Weinstein monofilaments, muscle strength testing with traditional motor grading, and prehension testing with 2 types of assessments. The test renders 5 subtest scores for each extremity. The 5 scores provide a profile of the upper limb. Although the prehension tests look at grasping, the measure considers appropriate movement secondary to adequate innervation. Therefore, the benefits of GRASSP are to assess body structure and function and have the opportunity to establish the influence of impairment on function.

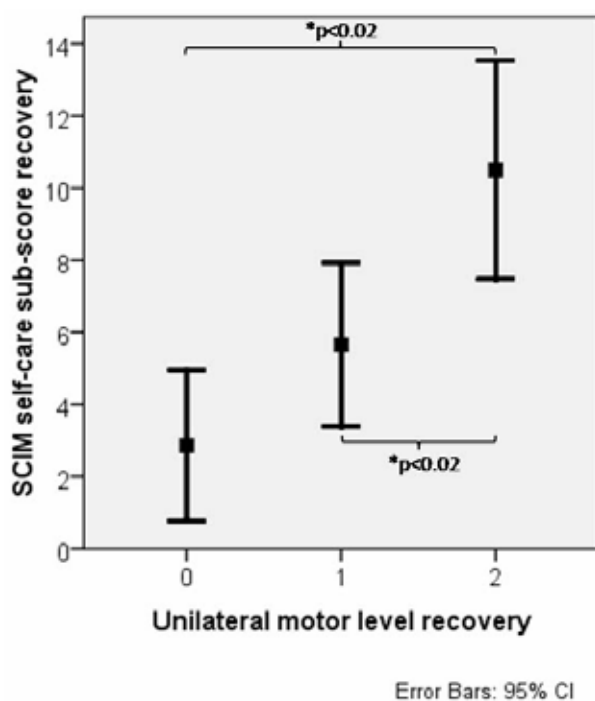
population in an acute or subacute Phase 2 trial where the therapeutic target is the CNS. **Table 4** provides a few summary examples to illustrate changes in cervical motor level.<sup>8</sup>

Whether it is spontaneous or therapeutically induced, recovery over the course of an acute or subacute SCI clinical trial can be influenced by a variety of contributing factors, only one of which is the treatment intervention. For example, it is recognized that neuroplasticity (changes in the strength of synaptic connections or axonal sprouting) within any residual and functionally subliminal neural circuit that is preserved after the initial mechanical trauma can contribute to recovery of function below the level of spinal damage. Likewise, any observed improvement in measures of ADLs could be the result of skills acquired due to rehabilitation efforts or

development of compensatory behaviors that enable completion of the task.

Rehabilitation training is now a standard of care provided to all patients, including those enrolled in a clinical trial, though these standards can vary significantly from place to place and over time. Thus rehabilitation training and compensatory strategies, although beneficial and perhaps essential to consolidating any therapeutic-induced effect, can also be viewed as confounding factors in a clinical trial when a therapeutic benefit is only viewed on the basis of an activity outcome. Conversely, given the complex nature of neurological disorders such as SCI, when a patient does not functionally recover, we do not know why. It could mean the therapeutic is without effect or it could be for a myriad of reasons or mechanisms that we do not know or appreciate. Negative





**Figure 4.** Average recovery in the SCIM self-care subscore as a function of the number of motor levels recovered over the first year (~1 to 48 weeks) after cervical (C4-C7) sensorimotor complete (AIS-A) SCI (n=66). There is no statistical significance in the SCIM self-care subscore when no (0) or 1 motor level is recovered, but there is a statistically significant difference in the SCIM self-care subscore between recovery of 0 and 2, as well as 1 or 2 motor levels ( $P < .02$ ). This significant increase in SCIM self-care subscore when 2 motor levels are recovered is likely to reflect a clinically meaningful benefit to people living with cervical sensorimotor complete SCI (modified from Kramer et al<sup>9</sup>).

**Table 4.** Illustrative examples of motor level changes that could occur after cervical sensorimotor complete (AIS-A) SCI (based on ISNCSCI definition of motor level)

Change in motor level	Initial segmental motor scores and derived motor level (highlighted)				
	C5	C6	C7	C8	T1
	5	3	2	1	0
	Examples of subsequent segmental motor scores that change motor level				
	C5	C6	C7	C8	T1
	-1	4	2	2	1
0	5	5	2	1	0
+1	5	5	3	3	0
+2	5	5	5	3	0

*Note:* Motor level (above shaded squares) is defined as the most caudal spinal segment, as indexed by the key muscle group for that segment, having a muscle strength score of at least 3 (contraction against gravity alone) while all of the more rostral key muscles are normal (5/5). In these examples, initial sensory level is constant at C5 (table is modified from Steeves et al. *Spinal Cord*. 2011;49:257-265). ISNCSCI = International Standards Neurological Classification of Spinal Cord Injury.

outcomes are difficult or impossible to interpret (see statistical considerations below).

Although SCIM, as well as other activity assessment tools, might describe useful functional capabilities, they were not designed to measure the result of a neurological change, but rather the outcome of rehabilitation. Activity measures are not able to reliably dissociate improvement due to actual repair of damaged spinal cord tissue versus rehabilitation training or compensation movements. For the purposes of this discussion, neural tissue is the target of drugs and cell-based therapies; thus, neurological recovery is the intended consequence and is most directly measured by assessing neurological impairment (eg, motor elements of ISNCSCI). However, changes in neurological function, even if statistically significant, may not be associated with a clear “functional impact” (ie, a neurological change that is associated with a meaningful improvement in personal activities, such as improved independence in key self-care activities). The challenge for clinical science is to measure improvement as precisely as possible (impairment), while defining a clinically meaningful response in the context of functional improvement (impact on activity limitations).

Obviously, it is important to limit false-positive (type I) and false-negative (type II) errors, so an initial primary endpoint threshold should be set at a reasonably achievable, but clinically valid limit, that is consistent with the known baseline statistical parameters within the study population. This consideration is especially important for less common clinical disorders, such as SCI, where the number of eligible subjects for enrollment is limited. Is it necessary to define co-primary outcomes or can a single clinical endpoint (outcome) accurately indicate a functional benefit while more directly measuring the therapeutic effect on the target tissue? For acute or subacute cervical sensorimotor complete SCI, a useful and clinically meaningful primary endpoint might be improvement of 2 cervical motor levels.

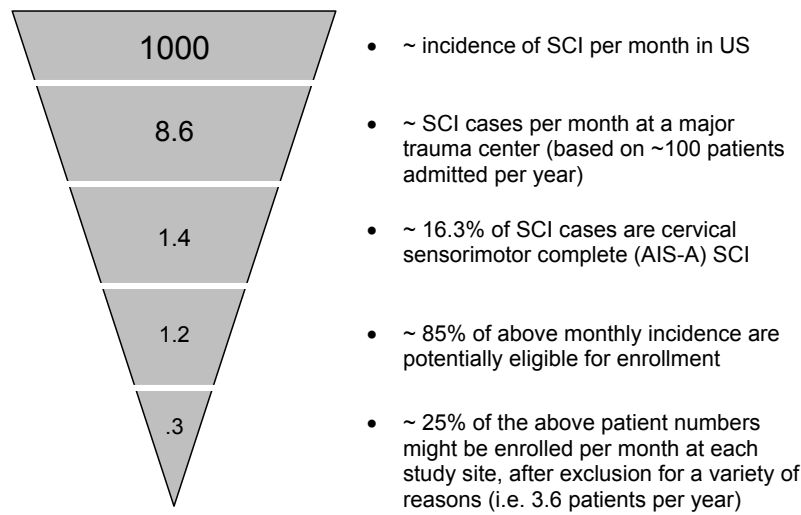
### Statistical Considerations

If the initial baseline classification of study subjects is inaccurate and/or the participants are heterogeneous in terms of the severity and level of SCI, subsequent statistical analysis is

severely compromised. A clinical endpoint used to determine the efficacy of an experimental treatment can be measured on a variety of scales. There are usually 2 main issues to address. The first is determining whether or not the effect of the experimental treatment is significantly better (statistically) than that of the control treatment. This goal can be achieved by a comparison of the average change (measurement of central tendency or comparison of means) in the measured variable between the experimental and control study arms. However, a few individuals having a large or small change in the primary outcome can dramatically influence these values (ie, over or under estimations). An experimental therapeutic may also only benefit a subset of the study population and thus measurements of central tendency may not capture the benefit in this subset of subjects.

The second issue to address is the clinical meaningfulness of the change, which typically hinges on its measured magnitude in the study population. Responder analyses provide one way to address these 2 issues in one step, by determining the proportion of people in each treatment group who attain a predefined threshold of clinically meaningful change (eg, minimum clinically important difference [MCID]) on the outcome measure. This approach has some statistical drawbacks,<sup>22</sup> but it deals with the problem of the potential overestimation of the population effect or the underestimation of the individual benefit. It does require a clear and well-justified definition of what we believe to be a clinically meaningful benefit, usually before the trial can be initiated.

Because there is a limited history of SCI clinical trials and no gold standard of treatment to compare any novel therapeutic, it is almost impossible to identify why a subject responds or does not respond to an experimental intervention.<sup>23</sup> Defining a responder threshold in this setting is a difficult task and requires prospective agreement by the trial investigators on what constitutes a clinical benefit for the planned study. It is important for the accuracy of the trial that the responder criterion (threshold) is as close as possible to a MCID for the study setting. Opinions will initially vary between and even within different constituencies (patient, physician, investigator, regulator, health insurer, or sponsor) on what is minimally important.



**Figure 5.** Illustrates the well-documented “funnel effect” for enrolling subjects to a clinical trial. The numbers are projections based on the incidence of SCI and cervical sensorimotor complete SCI within the United States. The numbers also reflect the incidence of SCI cases within a typical major US trauma center. The low levels of enrollment reflect the difficulties in acquiring adequate informed consent from an acute target population that has suffered a major traumatic injury (SCI). The difficulties are numerous and include the patient not being competent to understand the study protocol, as well as the risks and limitations of the trial procedures.

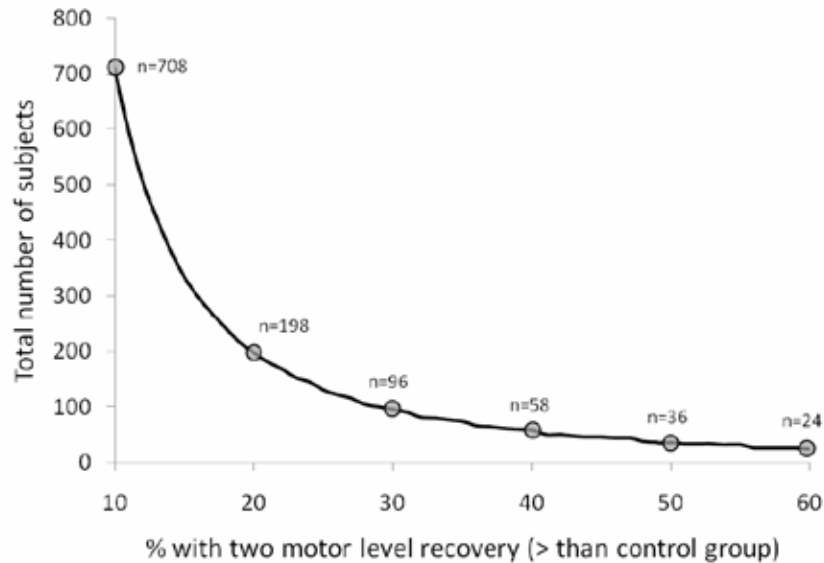
From the available data we have outlined above, it seems that a 2 motor level improvement would make a reasonable responder criterion that is clinically meaningful as determined by the positive relationship between a 2 motor level improvement and a significant increase in SCIM self-care subscore (Figure 4). As can be seen from Table 1, only about 20% to 25% of people with cervical sensorimotor complete SCI would be expected to recover 2 motor levels in a placebo group (over the first year after injury). The 2 motor level criterion will not capture those subjects exhibiting a more widely distributed pattern of motor recovery (ie, a small increase in motor points over many spinal cord segments caudal to the motor level), but the historical data suggest that this is a small minority of subjects.

The final question for this Phase 2 scenario is what percentage difference one expects between the responders in the experimentally treated cohort above that observed for responders within the placebo control group? As an example, using the values from Table 1, an investigator could decide on a 2 motor level improvement as the responder threshold and want 20% more of the experimentally treated subjects to respond above

the rate of placebo control responders with a trial end date at ~6 months. The natural history in Table 1 indicates that ~20% of the target population will spontaneously recover 2 motor levels within 6 months after cervical sensorimotor complete SCI. Thus, 40% of the experimentally treated group would have to respond (improve 2 motor levels), which would effectively double the proportion of responders beyond that suggested by the natural history of cervical sensorimotor complete SCI. Needless to say, the percentage difference in responder rates can be set at any suitable proportion, but there are pragmatic considerations for each percentage difference, not the least of which is the necessary number of subjects needed for the study (sample size).

### Pragmatic Considerations for Undertaking a Phase 2 Trial

All clinical trials must contend with a funnel effect for enrollment of subjects to a study (Figure 5). Typical enrollment rates across numerous medical disorders are almost always less than 10% of eligible participants.<sup>24</sup> There are many reasons for this limited enrollment rate, too numerous



**Figure 6.** Using a 2 motor level improvement after acute cervical sensorimotor complete (AIS-A) SCI, the graph illustrates the potential number of appropriate subjects to be enrolled for both the experimental and control groups (ie, total number of study subjects with a 1:1 allocation to each group). The number of required subjects varies depending on the anticipated percentage difference between the experimental and control group (for the number of subjects recovering 2 motor levels).

to list here, but it means that each year there are potentially only 1,200 people with acute or subacute SCI in the United States who are likely to participate in all SCI clinical studies.

If participants were randomized 1:1 between experimental and control groups, a typical power calculation, using the 2 motor level improvement response criterion with an expected 20% responder difference between groups, would suggest at least 198 subjects are required to complete the study (Figure 6). If only a 10% difference was expected between groups, it is likely that 708 subjects would have to enroll. This figure assumes no participant withdraws from the study before completion, but 15% of participants are likely to withdraw over the course of a study, so it would be prudent to recruit additional subjects. Thus, a single clinical phase 2 study could require participation from half of the annual cervical SCI target population in the United States!

Given these power calculations, the next concern would be how many study centers would be needed to recruit the desired number of participants over 1 year. It has been estimated that each major trauma

center caring might enroll no more than 3 to 4 subjects each year with acute or subacute cervical sensorimotor complete SCI (Figure 5). Thus, if a 10% difference was expected between groups, the number of hospital study centers might require approximately 200 hundred hospitals to participate. This is a large number of study sites to train, coordinate, and financially support by a sponsor. Therefore, it becomes easier to appreciate why a study might choose a 20% difference between the experimental and control groups.

In addition, the choice of trial outcome tools, endpoints, and number of repeated measures will make particular demands on the participants, examiners, study site, and sponsor. For example, a clinical trial measure that requires a lengthy assessment procedure can exhaust or annoy subjects, especially if it relies on repeated measurements over the trial period. If the outcome measure is technically demanding or requires a sophisticated interpretation of the outcome variables, it can lead to errors in the accuracy and consistency of the measurements (eg, electrophysiology, imaging). Thus, the necessary

degree of training in the chosen outcome tools and the ongoing recertification of trial personnel should be considered during the protocol development.

## Summary

Other outcome measures (clinical endpoints) and tools have not been included in this focused review (eg, autonomic nervous system function, pain, bladder/bowel function) and warrant a separate discussion. The above discussion is directed to some of the considerations in the choice of outcome measures for an acute or subacute SCI Phase 2 trial where the therapeutic target is the damaged cervical spinal cord and the anticipated outcome is an improvement in neurological motor function that is clinically meaningful. As discussed, the current evidence suggests that an improvement of 2 motor levels (ISNCSCI) within the injured cervical cord is related to a clinically meaningful recovery of independent activities of self-care as measured by the SCIM. In addition, an improvement of 2 motor levels after SCI is a

reasonable trial endpoint (based on the natural history of cervical SCI recovery) and can be adapted to different types of statistical analysis including responder analysis.

For any particular trial protocol, the specific outcome measures used and the threshold levels for demonstrating statistical significance and clinical meaningful benefit should be based on the therapeutic intervention being tested, the target tissue for the intervention, the expected effects of the treatment, and the practical limitations for a study involving a disorder with a low incidence.<sup>6-9,17,23,24</sup> The examples presented here provide a guide based on the known natural history of cervical sensorimotor complete SCI, as well as the pragmatic ramifications of the decisions made when designing an interventional SCI trial.

## Acknowledgments

Financial support and intellectual input have been received from members of SCOPE (Spinal Cord Outcomes Partnership Endeavor; www.scopesci.org).

## REFERENCES

1. Marino R, Barros T, Biering-Sorensen F, et al. International standards for neurological classification of spinal cord injury (6th ed). *J Spinal Cord Med*. 2003;26(Suppl 1):S50-S56.
2. Catz A, Iltzkovich M, Agranov E, Ring H, Tamir A. SCIM - spinal cord independence measure: a new disability scale for patients with spinal cord lesions. *Spinal Cord*. 1997;35:850-856.
3. Kalsi-Ryan S, Curt A, Fehlings MG, Verrier MC. Assessment of the hand in tetraplegia using the Graded Redefined Assessment of Strength Sensibility and Prehension (GRASSP): impairment versus function. *Top Spinal Cord Inj Rehabil*. 2009;14(4):34-46.
4. Kalsi-Ryan S, Beaton D, Curt A, et al. The Graded Redefined Assessment of Strength Sensibility and Prehension (GRASSP) - reliability and validity [published online ahead of print May 13, 2011]. *J Neurotrauma*.
5. World Health Organization. Towards a common language for functioning, disability and health. ICF. Geneva: World Health Organization; 2002. <http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf>
6. Fawcett JW, Curt A, Steeves JD, et al. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP Panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. *Spinal Cord*. 2007;45(3):190-205.
7. Steeves JD, Lammertse D, Curt A, 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(3):206-221.
8. Steeves JD, Kramer JK, Fawcett JW, et al. Extent of spontaneous motor recovery after traumatic cervical sensorimotor complete spinal cord injury. *Spinal Cord*. 2011;49:257-265.
9. Kramer JK, Lammertse DP, Shubert M, et al. Relationship between motor recovery and independence after sensorimotor complete cervical spinal cord injury. *Neural Rehabil Neural Repair*. Submitted for publication.
10. Bracken MB, Shepard MJ, Collins WF, et al. A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal cord injury. Results of the second National Acute Spinal Cord Injury Study. *N Engl J Med*. 1990;322:1405-1411.
11. Geisler FH, Coleman WP, Grieco G, et al. Measurements and recovery patterns in a multicenter study of acute spinal cord injury. *Spine*. 2001;26:S68-S86.
12. Ditunno JF Jr, Cohen ME, Hauck WW, Jackson AB,

- Sipski ML. Recovery of upper-extremity strength in complete and incomplete tetraplegia: a multicenter study. *Arch Phys Med Rehabil.* 2000;81:389-393.
13. Walker MD. Acute spinal-cord injury. *N Engl J Med.* 1991;324:1885-1887.
  14. Itzkovich M, Tamir A, Philo O, et al. Reliability of the Catz-Itzkovich Spinal Cord Independence Measure assessment by interview and comparison with observation. *Am J Phys Med Rehabil.* 2003;82:267-272.
  15. Catz A, Itzkovich M, Tesio L, et al. A multi-center International Study on the Spinal Cord Independence Measure, Version III: Rasch psychometric validation. *Spinal Cord.* 2007;45:275-291.
  16. Itzkovich M, Gelernter I, Biering-Sorensen F, et al. The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study. *Disabil Rehabil.* 2007;29:1926-1933.
  17. Anderson, K, Aito, S, Atkins, M, et al. Functional recovery measures for spinal cord injury: comparison by a multi-national work group. *J. Spinal Cord Med.* 2008;31:133-144.
  18. Glass CA, Tesio L, Itzkovich M, et al. Spinal Cord Independence Measure, version III: applicability to the UK spinal cord injured population. *J Rehabil Med.* 2009;41(9):723-728.
  19. Anderson KD, Acuff ME, Arp BG, et al. United States (US) multi-center study to assess the validity and reliability of the Spinal Cord Independence Measure (SCIM III). *Spinal Cord.* In press. doi: 10.1038/sc.2011.20.
  20. Bluvshstein V, Front L, Itzkovich M, et al. SCIM III is reliable and valid in a separate analysis for traumatic spinal cord lesions. *Spinal Cord.* 2011;49:292-296.
  21. van Hedel HJA, Curt A. Fighting for each segment: estimating the clinical value of cervical and thoracic segments in SCI. *J Neurotrauma.* 2006;23:1621-1631.
  22. Snapinn SM, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials.* 2007;8:31-36.
  23. Lammertse D, Tuszynski MH, Steeves JD, et al. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP Panel: clinical trial design. *Spinal Cord.* 2007;45(3):232-242.
  24. Tuszynski MH, Steeves JD, Fawcett JW, et al. Guidelines for the conduct of clinical trials for spinal cord injury (SCI) as developed by the ICCP Panel: clinical trial inclusion/exclusion criteria and ethics. *Spinal Cord.* 2007;45(3):222-231.