Traversing the Translational Trail for Trials

John D. Steeves, John L.K. Kramer, and Jose Zariffa

ICORD (International Collaboration On Repair Discoveries), University of British Columbia and Vancouver Coastal Health, Blusson Spinal Cord Centre, Vancouver General Hospital, Vancouver, BC, Canada

The principles of spinal cord injury clinical trial programs are briefly reviewed as one example of the challenge faced by most human studies of neurologically directed therapeutic interventions, including rehabilitation strategies. Different trial protocols are reviewed, as are inclusion/exclusion criteria for study subjects, the choice of clinical endpoints, and the statistical approaches that might be used in a trial program. Potential factors that might confound the accurate interpretation of trial data are also identified. Regardless of the specific therapeutic target or the rehabilitation strategy to be evaluated, there are many unresolved issues that will have to be answered before specific and effective prescriptions can be delivered. **Key words:** *clinical trials, confounding variables, human study protocols, responder analysis, spinal cord injury*

The title may be difficult to say accurately and quickly, but it is easier than trying to complete a pivotal trial to validate an experimental treatment as having a clinically meaningful benefit for a neurological disorder. Clinical trials for neurological disorders do not readily lend themselves to using a simple and direct outcome measure (clinical endpoint) to detect whether a drug, cell transplant, or rehabilitation strategy provides a meaningful (functional) benefit to patients. The central nervous system (CNS) is the most heterogeneous tissue of the body with hundreds of different cell phenotypes, each of which is capable of responding to a particular treatment in some specific and perhaps unpredictable or incomplete manner. Brain and spinal cord disorders are exceedingly complex as they disrupt multiple control systems both internal and external to the CNS. The fact that each functional pathway has multiple feedback and feed-forward loops does not make a clinical outcome easy to interpret in terms of its locus of action. Thus, we currently have an imperfect understanding of how to best conduct human CNS studies or evaluate novel therapies. Nevertheless, some forethought and precision in the conduct of CNS clinical trials can eliminate some of the uncertainty and provide increased confidence in the results.

In this review, the focus is on rehabilitation strategies, and spinal cord injury (SCI) will be used as a model CNS disorder. However, many of the principles outlined here also apply to other CNS disorders, as well as to any prospective study of a drug treatment or cell transplant. Although it may seem obvious, clearly defining the biological target(s) of your intervention is fundamental. Is the rehabilitation strategy directly affecting the CNS alone or does it also influence an end organ (eg, muscle)? A detailed understanding of the therapeutic target(s) is important to designing an effective trial protocol and selecting an appropriate outcome measurement tool. **Table 1** provides an incomplete and unprioritized list of the overall rehabilitation goals for SCI; similar lists can be generated for other CNS disorders. **Table 1** highlights that the reduction of an aggravating medical challenge can be as important as an improvement in a functional activity. It also underscores that there are many different therapeutic targets after SCI, each of which will likely require specific interventions and unique outcomes measures to accurately gauge efficacy.

Regardless of the specific therapeutic target or the rehabilitation strategy to be evaluated, there are many unresolved issues that will have to be answered for each CNS intervention. **Table 2** outlines some of the undetermined concerns.

Top Spinal Cord Inj Rehabil 2012;18(1):79–84

doi: 10.1310/sci1801-79

Table 1. Selected goals for activity-dependent rehabilitation training after spinal cord injury (SCI)

Improve	Reduce
1. Limb function	1. Spasticity
a) range of motion	2. Pain
b) walking gait	3. Pressure sores
c) grasp and skilled hand movements	4. Infections
2. Balance and transfers	5. Recurring health complications
3. Muscle strength and bone density	6. Anxiety
4. Bladder / bowel function and sexual health	7. Depression
5. Cardiovascular / respiratory function and stamina	
6. Immune responses	
7. Physical and psychosocial health	
8. Activities of daily living (ADLs)	
9. Community integration and participation	
10. Quality of life (QoL)	

Randomized Control Trials (RCTs)

Along with surgical decompression and stabilization of the spinal column, rehabilitation training, involving both physiotherapy and occupational therapy, are the current treatment strategies for facilitating recovery after spinal damage. The increased neural activity within the injured spinal cord resulting from rehabilitation efforts may facilitate neural plasticity and the functional reorganization of appropriate CNS pathways. Much of current rehabilitation practice has relied on clinical experience, small studies, and case reports (for extensive review of SCI rehabilitation evidence, see SCIRE at www.scireproject.com). Therefore, are randomized control trials (RCTs) necessary? In brief, no; but RCTs usually provide the strongest evidence for a treatment effect that most everyone will accept with a high degree of confidence.[1] If the treatment effect of an experimental intervention was overwhelmingly strong, it might be judged as self-evident and an RCT would not be required. The use of parachutes to prevent death related to a gravitational challenge is one tongue-in-cheek example.[2] Nevertheless, there are no such dramatic *prima facie* CNS rehabilitation examples. Even if a rehabilitation program were boldly self-evident, like the use of parachutes, subsequent RCTs trials, with an experimental and control group, would still be needed to validate each successive iteration of the product (eg, steerable airfoil parachutes as opposed to first-generation parachutes with little or no steering capability).

In addition, if subjects have a very stable baseline in terms of the outcome to be measured, as is

Table 2. Specific activity-dependent rehabilitation prescriptions require answers to these questions

1. When to start a physical and occupational rehabilitation intervention after onset of the central nervous system disorder?
2. Which rehabilitation regimen is best for each subtype or stage of the disorder?
3. What duration is required for each rehabilitation session?
4. What frequency of sessions is required per day or per week?
5. How many weeks or months are required to recover a specific functional capacity or activity of daily living?
6. How to best assess rehabilitation effort (eg, cardiovascular activity)?
7. How long will each benefit be sustained after acquisition?
8. What type of ongoing maintenance programs are required to sustain a benefit?
9. What are the reasons for lack of compliance or maintenance of a program by a patient?
10. Are particular training regimens generalizable to other similar activities or are they task specific? Can multiple functional activities / modalities (simultaneously or concurrently) be trained without one learned behavior suppressing or extinguishing another?

sometimes the case with subjects in chronic stages of a disorder, then each subject might serve as his or her own control. An example is a crossover design protocol.[3,4] Each subject is randomly assigned to receive either the experimental treatment or the placebo control treatment first, followed by the alternate treatment (after an appropriate washout period). The possibility of residual effects (carryover influence) of the preceding treatment is always a concern, which is why the washout period must be sufficiently long to guard against this possibility. A well-conducted crossover trial can reduce the number of subjects required to show statistical significant differences. However, the functional capacity of the subject must be stable (ie, chronic SCI) or spontaneous improvement could mask any benefit of therapy, leading to a type II (β) error (false negative).

Subject Selection

Most everyone believes in evidence-based clinical practices with the strongest level of evidence being provided by RCT studies. However, it can be difficult to achieve these demanding goals when the number of available and appropriate participants to enroll in a CNS rehabilitation study is low. Many CNS disorders either do not have a high incidence or they have a variety of different identifiable subtypes (heterogeneity). If all severities of a disorder are enrolled and analyzed as a single group, it is likely that a treatment will be judged as not significant, as a benefit in mild forms of the disorder are often cancelled out (offset) by no benefit to subjects with severe forms of the disorder. Thus, within a reasonable timeframe, CNS rehabilitation trials often have difficulty enrolling the necessary number of appropriate homogeneous subjects (with similar neurological impairments). The result has been numerous studies involving heterogeneous subjects and/or underpowered studies where it is difficult to draw a definitive conclusion. Even when a sufficient number of homogeneous subjects can be enrolled (eg, stroke), the primary clinical endpoint for the study (outcome measure) can be challenging and often dependent on a relatively indirect and/or subjective outcome (eg. cognitive test) using an imprecise ordinal scale.[5]

Spontaneous Recovery

Neurological disorders are rarely static, showing either spontaneous improvement or deterioration, which can vary from slow to rapid. Ideally, it is preferred to have subjects who are stable in terms of their neurological and functional status before any treatment is administered to determine whether the therapy has benefit for the subjects. Thus, after sudden trauma (eg, stroke, SCI, etc), chronic time points (several months after damage) are known to be less variable and preferred for gathering stable, accurate baseline data. However, therapeutically altering any biological system weeks, months, or years after damage is usually more difficult, which is why early interventions are encouraged. The desire to treat patients while they may be also spontaneously improving is not being criticized here, but it does create difficulties when trying to determine how much of any recovery is due to a specific treatment and how much is due to other influences. Therefore, it is important to understand the degree of spontaneous functional "plasticity" and/or acquisition of compensatory behaviors after each type of CNS trauma.[6] This spontaneous recovery must be tracked over a sufficiently long time period until stability is achieved. The collected natural history data are important to establishing a reasonable threshold (clinical endpoint) for demonstrating whether a therapeutic intervention has a benefit.

SCI is similar to many other traumatic neurological disorders in terms of the time course for spontaneous recovery. The largest degree of recovery occurs within the first few weeks and months after the initial damage.[7-11] The magnitude of this change can be significant, with as much as half of any spontaneous improvement happening within the first 4 months. Stability in terms of a functional plateau may not be reached for a year or more. Thus, if rehabilitation training is begun within a few weeks after injury, it is critical to know how much improvement might be spontaneous. Only large databases can provide the necessary detailed information on the natural history of spontaneous recovery for a disorder (including the various subtypes). It may seem obvious, but different severities of CNS damage show differing magnitudes and rates of spontaneous recovery;

thus a patient with sensorimotor complete SCI will usually improve at a slower rate than a patient with a motor incomplete spinal injury.[8]

Clinical Endpoint Threshold

With knowledge of the degree of spontaneous recovery for the target (eg, CNS) of the experimental intervention, it is possible to consider what might be a reasonable threshold for demonstrating a therapeutic effect. Obviously, previous clinical trial successes can shape this decision. Currently, however, there are very few CNS clinical trials that have statistically demonstrated a true clinical benefit for a specific rehabilitation intervention. When this has occurred, the benefit often has a number of limitations on the timing and/or baseline capacities of the patient for realizing that benefit (eg, constraint-induced movement therapy after stroke[12]). There are many analyses of past clinical studies where it has been retrospectively concluded that the threshold for the clinical endpoint was set too high (ie, insensitive to detect a subtle but real effect[9]). For many years, it was advocated that any drug, cell transplant, or rehabilitation therapy after SCI should be capable of stimulating robust axonal regeneration the length of the spinal cord. It is now appreciated that this is an unrealistic and likely unnecessary goal to improve the functional outcomes and quality of life for people living with SCI.

Statistical Treatment of Data

Increasing the likelihood that patients with sensorimotor complete SCI can regain function over as little as 2 spinal cord segments can make a dramatic improvement in their ability to independently complete many self-care activities (eg, grooming, feeding).[13] It is reasonable to suggest that functional motor improvement over even one spinal cord segment could be clinically meaningful, but the natural history data shows that between 70% to 80% of patients with cervical complete SCI will spontaneously recover motor function over one segment after SCI.[10] Thus, if the threshold for a clinical trial endpoint is set as recovery of motor function by one cord segment, it may be impossible to statistically differentiate the experimental and control groups in terms of their response to the experimental therapy (ie, a ceiling effect). The only way to overcome such a situation is to enroll many more subjects, which can be prohibitively expensive and impractical. If functional motor improvement over 2 spinal cord segments is used as the clinical endpoint, the natural history data would show that only about 25% of subjects with cervical sensorimotor complete SCI would spontaneously achieve this outcome and statistical ceiling effects are unlikely to impact a study.

A clinical endpoint used to determine the efficacy of an experimental treatment can be measured on a variety of scales. The common efficacy goal is determining whether the effect of the novel rehabilitation treatment is significantly better than that of the control treatment (current standard of care). This goal can be achieved by a comparison of the average change (mean) in the measured variable between the experimental and control groups. Unless the sample size is considerable, however, a few individuals having a large or small change in the primary outcome can dramatically influence the average values.

Because many CNS disorders are highly variable in terms of spontaneous recovery or deterioration, it can be difficult to observe a treatment effect. Currently it is almost impossible to identify why a subject with a CNS disorder does not respond to a therapeutic or fails to spontaneously improve to the average level of any historical control subjects. There can be a myriad of underlying reasons for a lack of response, including a lack of motivation by the subject. Thus "responder analysis" involves dichotomizing a relatively continuous or ordinal variable into a binary determination; does the trial participant (experimental or control) respond (achieve) to the predefined primary outcome (Yes or No). Responder analysis is based on defining a threshold value above which a subject is considered to be a "responder," and below which a subject is considered to be a "nonresponder." Responder analysis may be most suited to those clinical disorders and trials where there is a limited amount of information to explain the various possible outcomes and/or there is no prior gold standard therapeutic with which to compare the current trial results.

Table 3. Possible confounding factors for a spinal cord injury rehabilitation trial that cannot always be controlled by a study investigator

Prior emergency or primary care, intensive care treatment and management
Surgical decompression and spine stabilization (and timing of surgery after SCI)
Damage to other organs or subsequent medical complications and treatments
Type and extent of standard rehabilitation training in which a subject engages during a trial, as well as the amount of compensatory behaviors acquired during that time

Thus, it may be more reasonable to determine the proportion (%) of patients that actually benefit from a therapy using "responder analysis."[14] The challenge then becomes by what percentage should the responders in the experimental group exceed the placebo control responders. Would a 10% difference between the groups cause clinicians to change their rehabilitation practices? Given the large number of novel interventions coming forward, these discussions need to be undertaken by rehabilitation researchers. The percentage difference in the number of responders between the experimental and control groups will also dictate the number of subjects necessary to power a study adequately (pragmatic concerns cannot be ignored).

Confounding Factors

There are many factors that could alter the valid conduct of a clinical study, such as the criteria for the inclusion or exclusion of potential trial participants, as well as ethics associated with the recruitment of subjects,[15] but they fall beyond the scope of this short review. A study investigator cannot always control some potential confounders, as they involve prior or necessary concomitant treatments (**Table 3**).

Nevertheless, there is a longer list of potential confounding factors that any study investigator can and should control (**Table 4**).

Trial Checklist

There are many elements to the design and execution of a valid rehabilitation study, and this summary cannot outline each detail, but investigators should consider the following points:

- Check to ensure appropriate inclusion/exclusion criteria and ethics for the clinical target.
- Avoid enrollment of heterogeneous subjects.
- To establish useful baseline data, accurately diagnose the severity (impairment) of the

Table 4. Possible confounding factors for a spinal cord injury rehabilitation trial that can be controlled by a study investigator

• Unsuitable study protocol design or statistical analysis of trial results
• Inappropriate inclusion and exclusion criteria for a subject to participate in trial, resulting in heterogeneity within the subject population
• Lack of "blinded" randomization of subjects and randomization of treatment allocation sequence
• Lack of appropriate control subjects, which match criteria of experimental treatment group
• Investigator or subject expectations and bias, because the investigators and/or subjects are not blinded to the treatment provided
• Inappropriate, insensitive, or unreliable clinical outcome tools and/or primary outcome measure (clinical trial endpoint)
• Lack of independent "blinded" assessments of trial outcome measures
• Poor intra- and inter-rater reliability of outcome assessments (lack of ongoing training of trial assessors)
• Lack of sufficient follow-up assessments to assure persistence of therapeutic benefit (sometimes at least 6 months after trial completion)

CNS disorder and functional capacity of each subject.
- Evaluate plans for ensuring appropriate control data.
- To establish enrollment strategies and determine the required number of participating study centers, understand the incidence or prevalence of the disorder being studied.
- To establish an appropriate threshold for clinical endpoints, learn the natural history (spontaneous recovery) of the disorder.
- To demonstrate significance for primary clinical endpoint, determine the statistical power (minimum sample size) that is necessary to complete the study.
- Standardize clinical protocols between participating centers and ensure ongoing coordination of efforts.
- Be prepared for subjects dropping out (for reasons too numerous to list).
- Account for contributions due to spontaneous recovery or compensatory behaviors.
- List all the potential confounding variables and how they will be tracked and reported.
- Randomized control trials (RCT) are optimal, but "blinded" assessments are critical.

REFERENCES

1. Harvey LA, Lin CW, Glinsky JV, De Wolf A. The effectiveness of physical interventions for people with spinal cord injuries: a systematic review. *Spinal Cord.* 2009;47:184-195.
2. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomized controlled trials. *Brit Med J.* 2003;327:1459-1461.
3. 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;45:232-242.
4. Herrero AJ, Menéndez H, Gil L, et al. Effects of whole-body vibration on blood flow and neuromuscular activity in spinal cord injury. *Spinal Cord.* 2011;49:554-559.
5. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcomes measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol.* 2007;6:1094-1105.
6. Dimyan MA, Cohen LG. Neuroplasticity in the context of motor rehabilitation after stroke. *Nature Rev Neurol.* 2011; 7:76-85.
7. Burns AS, Ditunno JF. Establishing prognosis and maximizing functional outcomes after spinal cord injury: a review of current and future directions in rehabilitation management. *Spine* 2001; 26:S137-145.
8. 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.
9. 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;45:206-221.
10. 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.
11. Zariffa J, Kramer JK, Fawcett JW, et al. Characterization of neurological recovery following traumatic sensorimotor complete thoracic spinal cord injury. *Spinal Cord.* 2011;49:463-471.
12. Wolf SL, Winstein CJ, Miller JP, et al. Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke. The EXCITE randomized clinical trial. *J Am Med Assoc.* 2006;296:2095-2104.
13. Steeves JD, Lammertse DP, Kramer JLK, et al. New 13 Outcome measures for acute/subacute cervical sensorimotor complete (AIS-A) spinal cord injury during a phase 2 clinical trial. *Top Spinal Cord Inj Rehabil.* 2012;18(1):1-14.
14. Burkhoff D, Parides M, Borggrefe M, Abraham WT, Kadish A. "Responder analysis" for assessing effectiveness of heart failure therapies based on measures of exercise tolerance. *J Card Fail.* 2009;15:108-115.
15. 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;45:222-231.