

## REVIEW

# Clinical trials in spinal cord injury: lessons learned on the path to translation. The 2011 International Spinal Cord Society Sir Ludwig Guttman Lecture

DP Lammertse<sup>1,2</sup>

After three decades of clinical research on interventions to improve neurological outcomes in persons with spinal cord injury (SCI), the promise of preclinical discovery has yet to be translated into a consensus standard of care treatment. Nonetheless, SCI researchers remain hopeful that advances in preclinical discovery coupled with improved clinical trial performance will yield effective restorative treatment. This historical review of key studies in SCI over the past 30 years illustrates the progress that has been achieved in establishing a high standard in the conduct of clinical research while providing important lessons for improving trial design, conduct and reporting. Through application of these lessons, the performance of SCI trials can be improved, thereby shortening the pathway to successful translation and the development of effective therapies.

*Spinal Cord* (2013) 51, 2–9; doi:10.1038/sc.2012.137; published online 20 November 2012

**Keywords:** spinal cord injury; clinical trials; neurological recovery

## INTRODUCTION

## The era of clinical trials in SCI

The search for effective treatments to improve neurological outcomes for persons with SCI entered the era of clinical trials over three decades ago. By some measures, we have seen much progress during the ensuing thirty years. A seemingly exponential growth of knowledge regarding the underlying mechanisms of injury has led to an increasing number of clinical trials, giving encouragement to scientists, clinicians and patients. At the same time, it must be acknowledged that the ongoing effort to translate promising approaches from the laboratory to the bedside has not produced an FDA-approved treatment for improving neurological function, nor has a consensus standard of care been achieved.<sup>1,2</sup> The path to translation, the process of taking a good idea from the laboratory to a proven effective and safe treatment, has not reached its goal. This has led some to question whether we are at a critical juncture similar to that faced by stroke researchers in the 1990s, which prompted the Stroke Therapy Academic Industry Roundtable initiative to make recommendations for improving preclinical studies,<sup>3</sup> as well as clinical trial methodologies on the basis of analyses of the shortcomings of prior research efforts.<sup>4</sup> In our own field, the Spinal Cord Outcomes Partnership Endeavor (SCOPE; [www.scope-sci.org](http://www.scope-sci.org)) was born of a similar concern that an organized effort to promote best research practices both in the pre-clinical and clinical realm would be required to find the most efficient path to successful translation. The goal of this communication is to explore what has been learned about the conduct and reporting of interventional trials over these thirty years through the prism of a review of key SCI studies. The emphasis will not be on molecules, cells, mechanisms, or the details of trial outcomes but rather on the process of SCI research, how it is communicated to the scientific (and general) public, the ensuing debate in the court of scientific opinion and the lessons thus learned.

Insights gained through this discussion can inform future research efforts to shorten the timeline to achievement of the goal of finding effective treatments.

## Before we begin: comments on the nature of bias

Bias in clinical research can be defined as ‘...the systematic tendency of any factors associated with the design, conduct, analysis and interpretation of the results of clinical trials to make the estimate of a treatment effect (therapeutic benefit) deviate from its true value.’<sup>5</sup> All of us harbor biases derived from education, experience and value systems whether we are scientists, clinicians, trial participants, patients, journalists or lay persons. Bias is inherently the enemy of objective science whether it is found in the design, conduct, interpretation or reporting of clinical trials. Good science requires active efforts to eliminate sources of bias at all levels of the research endeavor. Bias may also be introduced in the process of informal collegial discourse when opinion *gestalts* of ‘thought leaders’ are substituted for the difficult task of a careful reading of the source literature. Most clinicians have opinions about the efficacy and safety of methylprednisolone treatment in acute SCI, yet in my experience, few have actually read the published reports in detail or the counterpoint literature that followed. No one has the time to digest all of the medical literature being produced, even in the relatively small and circumscribed field of SCI, but when controversy and conflicting interpretations arise, and key management decisions hang in the balance, we owe it to our patients and the field to go to the source before making judgments. As the history that follows unfolds, it is my hope to tempt the reader to put biases aside and with an open mind re-visit the literature of SCI clinical trials that has been published in the past decades.

<sup>1</sup>Department of Research, Craig Hospital, Englewood, CO, USA and <sup>2</sup>Department of Physical Medicine and Rehabilitation, University of Colorado Denver, Aurora, CO, USA  
Correspondence: Dr DP Lammertse, Department of Research, Craig Hospital, 3425 South Clarkson Street, Englewood, CO 80113, USA.  
E-mail: [dlammertse@craighospital.org](mailto:dlammertse@craighospital.org)

Received 28 August 2012; revised 1 October 2012; accepted 7 October 2012; published online 20 November 2012

## THE METHYLPREDNISOLONE STORY

### NASCIS 1: prevailing dogma is overturned

At the outset of this history of methylprednisolone (MP), one must confront the question: how could an established *de facto* 'standard of care' fall from favor? Was this 'standard of care' even if in part, medicolegally based, worthy—should it still be a standard? While I will endeavor to address the first question, I will defer on the second: I will not presume to add my opinions to the long list of authorities that have weighed in on one side or the other on the subject.<sup>6–9</sup> It could be justifiably claimed that the clinical trial era of SCI research began in February of 1979 with the enrollment of the first patient in a multicenter double-blind randomized control trial (RCT) of MP that was the first of three National Acute SCI Study (NASCIS) trials to be performed under the leadership of the NASCIS Group.<sup>10</sup> The NASCIS 1 trial, a test of 'high dose' (1000 mg bolus and daily thereafter for 10 days) vs 'standard dose' (100 mg bolus and daily thereafter for 10 days) MP began in an era when corticosteroid treatment of acute SCI was widespread and without firm clinical evidence basis (although the majority of pre-clinical laboratory studies had been supportive). It was also an era in which clinical research methodology and outcome measurement was in relatively uncharted territory. A placebo control was not included in NASCIS 1 because of ethical and medicolegal concerns of the clinical investigators—sufficient equipoise had not been established in an era of widespread clinical use. The initial report of the NASCIS 1 trial, which enrolled 330 patients, was published in 1984.<sup>10</sup> Analysis of neurological recovery measures did not show significant difference of outcome between the treatment groups, the lack of a 'dose effect' posing a challenge to the presumed efficacy of steroids in acute SCI. Safety analyses showed an elevated early-case fatality rate, which while not statistically significant, was of sufficient concern in the context of lack of demonstrated efficacy, that it prompted patient accrual to be discontinued before the planned study termination. On the basis of these results, the routine use of corticosteroid management in acute SCI was abandoned by many clinicians—NASCIS 1 had overturned prevailing dogma. Even with the disappointment of the first trial, research in the laboratory continued, resulting in a more refined understanding of MP mechanism of effect. Inhibition of lipid peroxidation was determined to be an important factor, one which would require even higher doses of drug than that tested in NASCIS 1, setting the stage for the second trial.

### NASCIS 2: the promotion of a new 'standard of care'

NASCIS 2 began in May 1985, less than a year and a half after publication of NASCIS 1, eventually enrolling 487 patients in a placebo-controlled RCT of MP (30 mg kg<sup>-1</sup> bolus, followed by 5.4 mg kg<sup>-1</sup> h<sup>-1</sup> in a 24-h regimen) or naloxone initiated after randomization within 12 h of injury.<sup>11</sup> The results of the first trial provided equipoise enabling the use of a placebo control group. In the initial publication of NASCIS 2 results in 1990, the investigators concluded that neurological recovery was improved in patients who received MP (but not naloxone) within 8 h of injury and that the treatment was relatively safe, with similar mortality and morbidity between the treatment groups. The study was indeed a landmark achievement: the first placebo-controlled multicenter RCT—the 'gold standard' in clinical trial methodology—of an intervention targeting the devastating neurological effects of SCI, and more importantly, the first study to report improved outcomes associated with treatment. The dissemination of these results was quickly followed by widespread adoption of the 'NASCIS 2 Protocol' in many parts of the world.

For many clinicians, new *de facto* standard of care had been established, and seeds of controversy had been sown.

### NASCIS 3: refinement of the 'standard of care'

The last of the three NASCIS trials, starting from the premise that 24 h of MP initiated within 8 h had been shown to be effective and safe, was designed to answer an unresolved question about the duration of treatment. Would 48 h of treatment be associated with additional improvement beyond that which could be achieved with the 24-h regimen, and would the longer treatment regimen be reasonably safe? On the basis of positive results reported in NASCIS 2, the investigators did not use a placebo control group—all three randomized treatment groups received a 30 mg kg<sup>-1</sup> bolus of MP before randomization, with the 'standard of care' control group receiving the full 24-h NASCIS 2 protocol, the 48-h group receiving an additional 24 h of MP, and the third group receiving 48 h of tirilazad (a 24-amino steroid that had been shown in preclinical studies to have promise as a neuroprotective agent). NASCIS 3 began in December 1991, 7 months after the publication of NASCIS 2 and enrolled 499 patients. The trial results were published in 1997 with the investigators, concluding that patients who receive MP within 3 h of injury can be treated with the 24-h regimen, but those whose treatment was not initiated until 3–8 h after injury should receive 48 h of MP, albeit with added risk of severe sepsis and pneumonia (the 48 h tirilazad treatment was not associated with better outcomes than the 24-h MP group).<sup>12</sup> NASCIS 3 thus proposed a refinement of the *de facto* standard of care that had been established by NASCIS 2.

### NASCIS Innovations: a high performance standard, indeed

Taken together, the NASCIS trials set a high standard, enrolling many participants (1316 total in the three studies) while utilizing randomized parallel control group design—with a placebo control group in the second study—and sophisticated statistical analyses. These studies established the use of summed motor and sensory index scores as measures of neurological function in SCI interventional trials, and documented the use of 'approved' primary outcome examiners. The third NASCIS study was the first to move beyond the measurement of impairment and incorporate a measure of activities—the Functional Independence Measure and its subscales—and in so doing, began to address the questions raised by SCI clinicians regarding the distinction between statistical significance of an abstract impairment outcome and clinical significance of an outcome: the functional impact of treatment. The manifest challenges of a large-scale multicenter investigation appeared to be well-met with high data collection rates and low patient drop-out rates.

### Public dissemination: the beginning of controversy

Yet, in spite of widely praised methodology and trial conduct (acknowledged as well by critics), the NASCIS trials have engendered unprecedented controversy, which had its beginnings in the public dissemination of NASCIS 2 results on 30 March 1990 with a press release to mass media, including newspaper and television outlets.<sup>13</sup> Full publication of the results in the *New England Journal of Medicine* would not occur for another 6 weeks, requiring a partial lifting of the journal's information embargo (this being before the era of advance online publication). In the public media, the early release of trial results before publication was portrayed as an ethical imperative prompted by the concern that any delay in widespread adoption of early MP treatment would be harmful to public health—and potentially cost individual patients recovery that they would not achieve without treatment. A week later, in response to the volume of

inquiry from clinicians regarding the details of the treatment protocol (including the unprecedented magnitude of the dosing regimen) and potential safety concerns to be discussed with patients, the National Institutes of Health sent a mass fax to hospitals in the US briefly summarizing the protocol and the trial results—the details still being left for print publication on 17 May 1990.

### The stakes are high

Having, for the first time, a *de facto* standard of care had far reaching ramifications, not only for newly injured patients and clinicians, but also for translational scientists. The mass media dissemination effort had created a public perception of efficacy and safety, the result of a single published study that had not been vetted by the usual give-and-take debate within the scientific community. Although the number of such legal filings cannot be readily established, reports of malpractice claims pursued on behalf of patients who had not received MP treatment began to surface. Clinical trial equipoise was shifted: for some time after the publication of NASCIS 2, RCTs in acute SCI dispensed with placebo control groups in deference to the *de facto* standard of care. Preclinical research followed suit, adding MP comparators and study of MP-new drug interactions to laboratory investigations. Ironically, the equipoise shift prevented a US-based attempt to mount a confirmatory placebo-controlled RCT with the result that the US Food and Drug Administration (FDA) has never added acute SCI to the indications list for MP. With the stakes so high, the intense scientific scrutiny that was to follow came as no surprise. With the publication of NASCIS 2, the debate began in earnest, beginning with letters to the editor but gaining momentum over the years with numerous articles in the medical literature deconstructing the NASCIS 2 and 3 data and analysis, calling into question the interpretation and conclusions of the original publications.<sup>6,8,14,15</sup> Several smaller studies attempted to replicate aspects of the NASCIS trials, producing equivocal or conflicting results.<sup>16,17</sup> At the same time, the safety of MP treatment was being challenged by the publication of a number of non-randomized case series reports, suggesting an association between treatment and immune system compromise, pneumonia, mechanical ventilation days, gastrointestinal complications and myopathy.<sup>18–21</sup>

### The Great Debate: the quality of NASCIS evidence

While the debate over the validity of NASCIS 2 and 3 conclusions may have been motivated by many factors including timing of the public dissemination, and feelings that a far-reaching standard had been imposed without proper vetting, the substance of the argument revolved around the quality of the evidence. While a thorough review of the various positions is well beyond the scope of this communication, the interested reader is encouraged to consult the abundant literature on the topic that has been published over the past 20 years.<sup>6–9,14,15</sup> A few key points should be considered in the context of lessons, which may be taken from this experience. The primary end point of the trials was a change in neurological function on the basis of a very complex measurement scheme with expanded motor (70 point)/sensory (58 point) scores (with data from only the right extremities used in the analysis), 5 motor/5 sensory categorizations, 3 broad categorizations of completeness and 5 completeness-by-level categorizations—yet there was no categorization of discreet segmental level. Beyond the unilateral expanded motor/sensory scores, this scheme is difficult and somewhat counter-intuitive for clinicians familiar with the International Standards of Neurological Classification of SCI, which evolved from the American Spinal Injury Associations standards that were in existence at the time of NASCIS

2 and 3. While the findings should stand on their own, difficulties in understanding the measurement scheme has been a potential confounder in acceptance of the trial conclusions. In addition, the precise criteria by which to judge change in neurological function was not specified, leaving any change in the myriad combinations within the measurement scheme as the possible basis for conclusions—there was not a precise *a priori* definition of the primary end point. Numerous analyses of the various measurement possibilities were presented—but not all; a number of the analyses were positive, but not all—thus leaving critics to claim biased, selective reporting. A primary hypothesis at the outset of NASCIS 2 and 3 was that treatment effect would be influenced by time of initiation, but left the determination of the critical time points (for example, MP initiated within 8 h in NASCIS 2) to *post hoc* analysis of the data, leaving some critics to discount the value of the conclusions. Questions were also raised regarding the representativeness of the sample, noting that patients with normal motor exams were enrolled and that their distribution across the treatment groups was not balanced—and was skewed in the direction favoring the conclusions. Finally, the landmark NASCIS 2 trial was criticized for making efficacy claims based purely on a measure of impairment: the functional impact of a 5–6-point improvement on an expanded motor score was not determined. In response to this concern, NASCIS 3 incorporated the use of the Functional Independence Measure, showing greater improvement in Self Care and Sphincter subscales for the 48-h MP group at 6 months, but the difference was not demonstrated at the 12-month follow-up.<sup>22</sup> The debate thus focused attention on the question of *how much* change in *what type* of measure should be demonstrated to support the establishment of a standard of care.

### Sorting through bias...overturning a standard of care?

Mounting criticism of the NASCIS trial conclusions and growing safety concerns culminated in the publication of the 2002 Neurosurgery recommendations on pharmacological therapy after acute cervical SCI, which classified methylprednisolone as ‘...an option in the treatment of patients with acute spinal cord injuries that should be undertaken only with the knowledge that the evidence suggesting harmful side effects is more consistent than any suggestion of clinical benefit’...damned by faint praise, indeed.<sup>8</sup> The *coup de grâce* for MP may have been the publication in 2008 of the 8th edition of the Advanced Trauma Life Support Student Course Manual by the American College of Surgeons, which stated that there was insufficient evidence to support the routine use of steroids in SCI.<sup>23</sup> Many trauma centers in the United States responded by removing MP therapy from the routine standing orders for treatment of acute SCI. Within a few months, a dramatic shift occurred in MP utilization in the trauma facilities in Colorado, where I practice—a change from a routine, to a rare event. What was once a *de facto* ‘standard of care’ had become distinctly the exception. Indeed, the remarkable intensity and duration of this debate has left us with a ‘hung jury’ in the court of scientific debate, but one in which a ‘retrial’ appears unlikely.<sup>24</sup>

### GM-1 GANGLIOSIDE: THE SYGEN EXPERIENCE

#### 1991—a new option worthy of study

Just 13 months after the landmark NASCIS 2 study was reported, another promising therapy entered the arena with the 1991 publication of the initial GM-1 Ganglioside trial results.<sup>25</sup> This single-center randomized double-blinded placebo-controlled pilot study, which enrolled 37 patients between January 1986 and May 1987, examined the efficacy and safety of intravenous monosialotetrahexosylganglioside (GM-1) in promoting recovery of neurological

function in patients with acute SCI. Study medication was initiated within 72 h of treatment and given in a single daily injection for 18–32 doses (presumably limited by transfer from the trauma study center to the SCI rehabilitation facility). Neurological recovery was measured by the change from enrollment to 1-year follow-up in the Frankel scale and the American Spinal Injury Association motor index score. The investigators reported a significant difference between the treatment and control groups favoring GM-1 in both of the defined efficacy outcomes and no differences in adverse events, thus setting the stage for a larger multicenter trial.

### 2001 Multicenter Sygen Study—unfulfilled promise

On the basis of promising results of the single-center pilot study, a multicenter trial of GM-1 (now given the proprietary name Sygen) was undertaken, which enrolled 760 participants with American Spinal Injury Association Impairment Scale (AIS) A, B, C, D injuries at 28 neurotrauma centers between 1992 and 1997—the largest randomized clinical trial ever conducted for treatment of SCI.<sup>26</sup> The Sygen study compared two doses of the intravenous study drug (100 and 200 mg) to placebo initiated within 3 days of SCI given daily for 56 days. The *a priori* primary efficacy outcome was the achievement of ‘marked recovery’ (two grades of improvement) at 26 weeks on the Modified Benzel Scale, which added gradations of ambulatory ability to the AIS. This was the first SCI trial of a restorative intervention to employ a responder definition, which dichotomizes the outcome (responder: yes or no) on the basis of the amount of change felt to have clear functional impact. The trial also measured American Spinal Injury Association Motor and Sensory index scores and the proportion of participants who recovered normal bowel and bladder function. Analysis of the primary outcome revealed an overall response rate that was higher than expected, but failed to show a significant difference between treated and control group participants although there was a trend favoring treatment in the subgroup analysis of those with AIS B classification. There was a significant difference between the treatment and control groups in the time at which marked recovery was achieved: the patients receiving GM-1 recovered function earlier but the control group ‘caught up’ over the course of follow-up. Had the *a priori* end point been measured at the conclusion of treatment (8 weeks), the investigators could have claimed significant benefit. The secondary analyses of sensory/motor scores and bowel/bladder outcomes generally showed trends favoring treatment that did not achieve statistical significance. Overall the trends suggesting benefit were restricted to the less severely injured. Analysis of safety data failed to show significant differences between the groups.

The Sygen Trial, even though it failed to demonstrate efficacy was nonetheless a landmark achievement that has provided guidance to the field. Compared with the trials that preceded it, the Sygen trial was an advance in scientific rigor: a placebo-controlled double-blinded RCT with a clearly defined *a priori* end point. The trial provided training in the performance of the International Standards of Neurological Classification of SCI examinations, which formed the basis of primary and key secondary outcome determination. This study was the first of its type to perform reliability assessments of the trained examiners.<sup>27</sup> The Sygen trial, which began enrollment shortly following the publication of the second NASCIS trial, required full NASCIS 2 protocol MP as an inclusion criterion, in an effort to control a plausible independent variable, which at the time, was a routine *de facto* ‘standard of care’. Thus the trial was not simple investigation of Sygen’s effect on neurological recovery but rather a test of whether Sygen could add an increment of improvement

beyond that due to MP. NASCIS 2 had indeed cast a long shadow. It should also be noted that the multicenter Sygen Trial early enrollment was complicated by the availability of compassionate use drug. After publication of the initial GM-1 trial with its promising results, the study sponsor received many requests from clinicians asking for open-label drug for the treatment of SCI. In response, a compassionate use protocol was developed allowing for Sygen (GM-1) treatment with local Institutional Review Board approval. As a result, for the first 6 months of the multicenter trial, potential participants could choose open-label compassionate use GM-1, or enroll in the RCT with a 1 in 3 risk of placebo assignment. The well-publicized access to open-label Sygen by a professional athlete who sustained SCI in competition was followed by a drop in trial accrual, which forced the sponsor to withdraw availability of the compassionate use alternative. This experience highlights the challenges to the conduct of rigorous scientific investigation in settings, where open-label alternatives coexist with the inevitable public (and clinical) biases favoring unproven ‘treatment’ over the risk of placebo.

### The Sygen Epilogue

With the publication of the multicenter Sygen trial results failing to show efficacy, the previous enthusiasm for this new option plummeted. The drug has never achieved regulatory approval for any use in the United States, and although it was once available for treatment of neurological conditions in several countries in Europe, the European Medicines Agency no longer lists Sygen as an approved treatment. Interestingly, the 2002 Neurosurgery recommendations stated that ‘treatment of patients with acute spinal cord injuries with GM-1 ganglioside is recommended as an option without demonstrated clinical benefit.’<sup>8</sup> In spite of its failures, the Sygen trial was a major milestone, not only in setting a new standard in the design and conduct of clinical trials in SCI, but also in the wealth of data that has been made available to the field for published analyses of recovery after SCI.<sup>27–29</sup>

### Fampridine-SR: a physiological approach... with potential application in chronic SCI

A few small trials published in the 1990s reported positive effects on a sensorimotor function, spasticity and pain in patients with SCI who had been given 4-aminopyridine (4-AP), a potassium channel blocker that improves conduction block over axon segments that are demyelinated or poorly myelinated after injury.<sup>30,31</sup> On the basis of these pilot trial results, several phase 2 trials were undertaken, culminating in the publication of a multicenter double-blind randomized placebo-controlled study of a sustained release version of 4-AP called fampridine-SR in 91 patients with chronic ( $\geq 18$  months post injury) motor incomplete SCI.<sup>32</sup> Several doses (25 and 40 mg b.i.d.) were compared with placebo over an 8-week treatment period. The prospectively defined primary outcome utilized a daily patient questionnaire that assessed the participant’s spasticity, bladder, bowel and sexual function, as well as overall physical well-being. The study also measured quality of life with a Subject Global Impression questionnaire (terrible-delighted seven-point scale), motor and sensory impairment with the International Standards of Neurological Classification of SCI and spasticity with the Ashworth scale. Analysis of the data did not show a difference in the overall positive response rates for the patient diary questionnaire, but did find significant improvement in the Subject Global Impression for patients in the 25-mg b.i.d. group. While there was a trend in the overall improvement in spasticity favoring the 25-mg b.i.d. group, *post hoc* analysis showed significant benefit for those participants with baseline Ashworth scores greater than the median score for the study



population (that is, moderate to severe spasticity). The higher dose group had disappointing efficacy outcomes and experienced a significantly greater incidence of adverse events (for example, dizziness, insomnia and nausea) and dropout rate. The phase 2 trial results were deemed to be supportive of further study and informed the key decisions on dose (25 mg b.i.d.), the primary outcome (improved spasticity) and an important eligibility criterion (Ashworth score at least 3 on the 1–5 scale) for the phase 3 trials that followed. Two phase 3 studies, each enrolling over 200 participants with chronic incomplete SCI (AIS B, C or D) were undertaken in 2002 with the intent to meet the FDA standard, which requires confirmatory evidence from separate trials for registration of a new drug. The trials were identical in design—double-blind, placebo-controlled, parallel group studies—and run concomitantly at two distinct sets of North American study centers. Enrollment and follow-up were completed in 2003 and while the complete results of the trials have yet to be published, information contained in press releases from the sponsor ('Acorda Therapeutics Reports Results of Fampridine-SR Clinical Trials' at [www.evaluatepharma.com/Universal/View.aspx?type=Story&id=51906](http://www.evaluatepharma.com/Universal/View.aspx?type=Story&id=51906)) indicates that neither study showed a significant difference in the primary end points of change baseline in Ashworth scale and Subject Global Impression. One of the phase 3 studies demonstrated a trend favoring treatment in the Ashworth scale measure ( $P=0.069$ ) when assessed across all observations during treatment according to the *a priori* plan. *Post hoc* analysis limited to the last Ashworth score obtained during treatment (a methodology commonly used in drug trials) showed significant improvement in the treated vs control group ( $P=0.006$ ). The other phase 3 trial also showed improvement in mean Ashworth scores during the treatment phase, but neither in the *a priori* nor *post hoc* analyses was there a difference between treatment and control. The phase 3 experience with fampridine-SR highlights the importance of robust RCT data (both trials would have appeared to be positive in the absence of the control comparator) and the consequences of choices made with regard to outcome measurement—both the measure itself and how it is applied. At least one of the trials would have been positive if the choice had been made for a single end point (change in mean Ashworth) assessed at the conclusion of the full-dose period. Nonetheless, even if those choices had been made, the results of the two trials taken together would still have lacked the confirmation required for registration of a new drug by the FDA—a requirement to ensure that such decisions are made on the basis of robust evidence. After the failure of the SCI trials for treatment of spasticity, the sponsor turned its attention to the study of the drug for treatment of patients with multiple sclerosis—specifically for improved walking ability, a function in that disorder that can perhaps be measured with more precision than spasticity in SCI. That endeavor met with more success and with the publication of two confirmatory phase 3 trials, FDA approval of the drug (now re-named dalfampridine) was achieved in 2010 (FDA 22 January 2010 'Dalfampridine' at <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm221277.htm>) for use in patients with multiple sclerosis, and thus available off-label for treatment of patients with SCI.

#### **Autologous activated macrophages: an acute cell therapy is put to the test**

In 2000, the first patient was enrolled in a phase 1 open-label, non-randomized trial of autologous incubated macrophage therapy in which a patient's peripheral monocytes are incubated with skin tissue to produce activated macrophages, which are then injected into the

caudal boundary of the contused spinal cord within 2 weeks of injury.<sup>33</sup> The therapy is intended to improve neurological recovery by modulating the immune response to injury with the introduction of a growth-promoting macrophage phenotype into the cellular environment. Over the following 3 years, the study went on to enroll 8 participants with complete AIS A injuries, 3 of whom recovered to motor incomplete (AIS C) status over the 1-year follow-up period—a recovery rate felt to compare favorably with the expected rate derived from previously published analyses of existing databases. These results prompted the initiation of a multicenter phase 2 trial in 2003, which was the first RCT of a cell-based therapy for acute SCI. While both macrophage trials showed a reasonable safety profile (an important consideration, especially in a therapy that requires major surgery for cell implantation), the results of the multicenter trial failed to deliver on the promise of phase 1. Analysis of the primary outcome measure (improvement of the AIS from A to B or better) failed to show a significant difference between treatment and control recipients but notably there was a strong trend ( $P=0.053$ ) favoring the control group.<sup>34,35</sup> Notably, the multicenter RCT employed a 2:1 treatment-to-control randomization scheme on the basis of belief that the phase 1 trial had demonstrated some efficacy signal and that potential participants would be more inclined to accept randomization with the knowledge that they had a 67% chance of treatment group assignment. This decision, which resulted in a small ( $n=17$ ) control group, may have contributed to the much higher than expected AIS conversion rate (59%) in the control group. The conversion rate in the treatment group (27%) would have looked more favorable compared with historical controls in the absence of the control group findings, thus providing a strong argument for the use of concurrent parallel control group design. The reasons for the failure of this trial are potentially numerous but must begin with the acknowledgment that the therapy may be ineffective or at least ineffective when delivered in the manner defined in the study protocol. To this point, it should be noted that the injection protocol and total dose used in phase 1 had been significantly modified for phase 2 on the basis of interim pre-clinical animal studies of cell distribution and rat-to-human scale-up calculations, going from 4 injections totaling 4 million cells to 6 injections totaling 1.5 million cells. Cell therapy dose and delivery targeting are likely to be critical protocol parameters that warrant preclinical attention to the issues of cell distribution within the zone of targeted pathology and dose response assessment. The phase 2 macrophage trial also highlighted concerns with the use of AIS as an outcome measure: of the four individuals who experienced AIS A to C conversions, none regained any volitional motor function in the lower extremities; of the 17 participants who converted to incomplete status, only three regained any sensation in the lower extremity L1-S3 dermatomes. Interpretation of these International Standards of Neurological Classification of SCI data is dependent on rigorous examination and in this regard, it should be acknowledged that the multicenter macrophage trial achieved a high measurement standard, building on the example of the Sygen trial with training and reliability testing of primary outcome examiners. While the results of the multicenter macrophage trial were disappointing, the learning curve for future SCI cell therapy trials will hopefully benefit from the experience gained in the conduct of these clinical investigations.

#### **The SCI Locomotor Trial (SCILT): an RCT of a complex rehabilitation intervention**

The last decade of the 20th century saw a remarkable evolution in the theory and practice of SCI neurorehabilitation from compensatory

skill acquisition emphasizing the use of spared above-lesion function toward increasing emphasis on repetitive, task-specific motor learning that often called into play below-level neural circuits below the lesion in patients with incomplete injuries. The emergence of Body Weight Supported Treadmill Training (BWSTT) as an 'Activity-Based Therapy' was based on animal model studies that had demonstrated the positive effects of patterned hind limb activities in spinal cord transected quadrupeds during treadmill training.<sup>36</sup> Afferent hind limb load and proprioceptive input to infra-lesional locomotor central pattern generator circuits in the lumbar cord were posited to produce activity-dependent plasticity resulting in the training effect—the isolated spinal cord 'learned' to produce reciprocal stepping movements (albeit without regaining over-ground ambulation). Human pilot studies using body weight support over a treadmill and therapist (or robotic) assisted stepping soon followed, which demonstrated feasibility of BWSTT and suggested functional benefit.<sup>37,38</sup> The concept proved seductive: for the first time, rehabilitation clinicians contemplated the potential to provide a treatment fundamentally different than teaching compensatory activities, rather one that held promise to improve the neurological circuitry underlying the recovery of ambulation. Clinicians enthusiastically adopted the concept, willing to accept the incremental labor (multiple therapists per treatment session) and equipment (specialized support harness and treadmill) costs over traditional gait training because of the presumed superiority of BWSTT based largely on published small trials in chronic incomplete SCI that were either uncontrolled or relied on outcome comparisons with historical controls. This provided the philosophical context of the SCILT trial, a well-designed RCT, which was undertaken with the expectation that BWSTT would be shown superior to conventional over-ground gait training in the recovery of functional ambulation in patients with incomplete SCI (AIS B, C, D) enrolled within several months of injury.<sup>39</sup> The study included 146 participants from six SCI treatment centers between 2000 and 2003 who were unable at enrollment to ambulate over-ground without at least moderate assistance (scored  $\leq 3$  on 1–7 locomotor subscale of the FIM). In an attempt to narrow the focus of the trial to the specific variable of therapeutic technique, the study design equalized the intensity (up to 1 h per day, 5 days per week), duration (12 weeks) and post-injury initiation of treatment (within 56 days of SCI) between the BWSTT and conventional over-ground gait training groups. The primary outcome analysis for the AIS B and C participants compared the proportion who recovered to supervised or independent walking (FIM Locomotor score  $\geq 5$ ) at 6 months; for AIS D participants, the analysis compared the over-ground walking speed at 6 months with all primary outcome measurement performed by blinded examiners. Much to the surprise of SCI clinicians, the SCILT results published in 2006 failed to show any difference in walking outcomes between those treated with BWSTT and those who received conventional over-ground gait training.<sup>39</sup> Although *a priori* power analysis utilizing clinical data from the participating study centers was utilized to determine enrollment targets, an interim conditional power analysis done at the request of the Safety and Data Monitoring Committee using subject data showed that thousands of additional participants would be required to achieve statistical significance in the primary outcome measures—which resulted in stopping the trial. The vigorous debate that ensued centered on the validity of using comparator groups receiving equal intensity and duration of treatment.<sup>40–42</sup> Was the intensity and duration of the conventional over-ground gait training regimen representative of standard clinical practice? While the polarization of opinion

exposed by the SCILT debate is yet to be fully resolved, the message of this important trial may simply be that the intensity, duration and timing of rehabilitation interventions are critical independent variables that should not only be addressed in future clinical studies, but also draw the attention of providers grappling with the challenges of reconciling evidence-based best practices with cost effectiveness concerns.<sup>43</sup> The difficulties inherent in undertaking power calculations on the basis of effect size estimates and uncontrolled clinical data were also highlighted in the SCILT trial's gross underestimate of target enrollment.

#### Parting thoughts...lessons learned on the path to translation

Three decades into the era of clinical trials for improved neurological outcomes in SCI, there is as yet, no intervention that has achieved consensus standard of care status. There has been no regulatory approval for a therapeutic with the indication of improving neurological function after SCI. While it must be acknowledged that the failures of trials to date may be primarily because of inadequate potency of the tested interventions, clinical investigators have an absolute responsibility to improve the rigor and conduct of human trials to ensure that the true effects of treatment, positive and negative, are accurately detected and reported in an unbiased manner. If small effects go undetected because of poor trial design or conduct, important opportunities to leverage successful translation may be lost. The trials reviewed here have served to shape the evolving concept of the state-of-the-science, of best clinical research practices that should inform better (and hopefully more successful) trials in the future as follows:

- *The control of bias—trial design is important.* The experience of the trials reviewed here repeatedly confirms the critical importance of prospective randomized control parallel group design—with placebo controls and blinding utilized, whenever feasible. The treatment group outcomes of these trials would have been considered 'positive' without the strict test of a randomized control group comparison. Historical controls simply cannot provide an adequate comparator, considering the gamut of potential independent variables that may influence outcomes. This admonishment is especially noteworthy in the current era of unproven stem cell therapies promoted on the internet with quasi-experimental case series reports and patient testimonials.<sup>44</sup>
- *Confirmatory evidence is the key.* The demonstration of significant results in a single trial is generally not sufficient to establish a standard of care. Regulatory agencies such as the FDA impose a standard of 'robustness' that require confirmation of results by a second trial.
- *The primary outcome measurement choices are critical...and we need better measures.* The trials surveyed here demonstrate the challenges of choosing the best measure of success in SCI trials. Demonstration of statistical significance on an impairment measure will not be sufficient proof of clinical significance. Refinements of impairment and activities measures as well as the establishment of Minimal Clinically Important Difference criteria for these SCI outcomes remain a challenge for the field.<sup>45</sup> When to measure the primary end point is also critical: demonstration of short-term (for example, 8–12 week) efficacy and safety will not prove that the benefit will last, especially when testing a treatment during the first year after injury.
- *Key outcome examiners should be trained and be tested for inter-rater reliability.* The ability of a clinical trial to precisely define the true effect of an intervention depends on the reliability of the outcome

data. The trials reviewed here have all utilized training of study personnel responsible for the generation of key outcome data, some have also documented reliability testing of examiners.

- *Clearly define the design, primary outcome and analysis plan a priori.* The most compelling findings will be derived from trials that have clear prospective declaration of design, a primary outcome and analysis plan. The NASCIS trials, while landmark accomplishments, ultimately suffered in the court of scientific opinion because a precise definition of primary outcome was not felt to have been established at the outset. Registration of trials (for example, on [www.clinicaltrials.gov](http://www.clinicaltrials.gov)), which requires some disclosure of key trial elements before the initiation of enrollment is thus to be recommended and is now required for publication of results in many peer-reviewed journals.
- *Reporting of randomized trials should be guided by the Consolidated Standards for Reporting Trials statement recommendations.* The dissemination of clinical trial results should be free of biased reporting and contain sufficient information on the methodology and results to allow accurate interpretation by the reader. The trials reviewed here have reflected the evolution of reporting standards over the past several decades embodied in the recently published Consolidated Standards for Reporting Trials statement, which provides guidelines for the thorough and transparent reporting of parallel group randomized trials.<sup>46</sup>
- *The Rehabilitation variable.* Preclinical studies have demonstrated the benefits of rehabilitation intervention targeting the mechanism of activity-dependent plasticity. Animal studies have also demonstrated the potential interactions of ‘biological’ and ‘rehabilitation’ interventions leading to the assumption that trials of cells or drugs should at minimum control for the rehabilitation variable.<sup>47</sup> The important questions about *how much* of *what* rehabilitation treatment delivered *when* will make a difference in SCI outcomes remain largely unanswered and should be high priorities for the field.
- *‘...not if, but when...’.* In closing, I should state that while disappointed by the fact that a clear standard of care treatment for improving neurological outcome in SCI has not been achieved, the continued advances in the preclinical sciences combined with the efforts to improve the conduct of clinical trials embodied in this review give optimism for our eventual success. The question is not if, but when we will have effective treatments to reverse the effects of SCI paralysis.

- 1 Tator CH. Review of treatment trials in human spinal cord injury: issues, difficulties, and recommendations. *Neurosurgery* 2006; **59**: 957–987.
- 2 Kwon BK, Sekhorn LH, Fehlings MG. Emerging repair, regeneration, and translational research advances for spinal cord injury. *Spine* 2010; **35**: S263–S270.
- 3 Stroke Therapy Academic Industry Roundtable. Recommendations for standards regarding preclinical neuroprotective and restorative drug development. *Stroke* 1999; **30**: 2752–2758.
- 4 Fisher M. Recommendations for advancing development of acute stroke therapies: stroke therapy academic industry roundtable 3. *Stroke* 2003; **34**: 1539–1546.
- 5 Lammertse D, Tuszynski MH, Steeves JD, Curt A, Fawcett JW, Rask C *et al.* Guidelines for the conduct of clinical trials in spinal cord injury as developed by the ICCP panel: clinical trial design. *Spinal Cord* 2007; **45**: 232–242.
- 6 Nesathurai S. Steroids and spinal cord injury: revisiting the NASCIS 2 and NASCIS 3 trials. *J Trauma* 1998; **45**: 1088–1093.
- 7 Bracken MB. Steroids for acute spinal cord injury (Cochrane Review). In: *The Cochrane Library*. Issue 1 (John Wiley & Sons, Ltd.: Chichester, UK, 2004).
- 8 American Association of Neurological Surgeons/Congress of Neurological Surgeons. Pharmacological therapy after acute cervical spinal cord injury. *Neurosurgery* 2002; **50**: S63–S72.
- 9 Hall ED, Springer JE. Neuroprotection and acute spinal cord injury: a reappraisal. *NeuroRx* 2004; **1**: 80–100.

- 10 Bracken MB, Collins WF, Freeman DF, Shepard MJ, Wagner FW, Siltan RM *et al.* Efficacy of methylprednisolone in acute spinal cord injury. *JAMA* 1984; **251**: 45–52.
- 11 Bracken MB, Shepard MJ, Collins WF, Holford TR, Young W, Baskin DS *et al.* A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury. *N Engl J Med* 1990; **322**: 1405–1411.
- 12 Bracken MB, Shepard MJ, Holford TR, Leo-Summers L, Aldrich EF, Fazl M *et al.* Administration of methylprednisolone for 24 or 48 h or tirilazad mesylate for 48 h in the treatment of acute spinal cord injury. *JAMA* 1997; **277**: 1597–1604.
- 13 Leary W. Treatment is said to reduce disability from spinal injury. *The New York Times*, 31 March 1990. available at <http://www.nytimes.com/1990/03/31/us/treatment-is-said-to-reduce-disability-from-spinal-injury.html>.
- 14 Coleman WP, Benzel E, Cahill DW, Ducker T, Geisler F, Green B *et al.* A critical appraisal of the reporting of the national acute spinal cord injury studies (II and III) of methylprednisolone in acute spinal cord injury. *J Spinal Disord* 2000; **13**: 185–199.
- 15 Hurlbert RJ. Methylprednisolone for acute spinal cord injury: an inappropriate standard of care. *J Neurosurg* 2000; **93**: 1–7.
- 16 Otani K, Abe H, Kadoya S, Nakagawa H, Ikata T, Tominaga S *et al.* Beneficial effect of methylprednisolone sodium succinate in the treatment of acute spinal cord injury (in Japanese). [*Sekitsui Sekizui*] 1994; **7**: 633–647.
- 17 Pointillart V, Petitjean ME, Wiart L, Vital JM, Ticoipe M, Dabadie P. Pharmacological therapy of spinal cord injury during the acute phase. *Spinal Cord* 2000; **38**: 71–76.
- 18 Galandiuk S, Raque G, Appel S, Polk HC. The two-edged sword of large-dose steroids for spinal cord trauma. *Ann Surg* 1993; **218**: 419–427.
- 19 Gerndt SJ, Rodriguez JL, Pawlik JW, Taheri PA, Wahl WL, Michaels AJ *et al.* Consequences of high-dose steroid therapy for acute spinal cord injury. *J Trauma* 1997; **42**: 279–284.
- 20 Matsumoto T, Tamaki T, Kawakami M, Yoshida M, Ando M, Yamada H. Early complications of high-dose methylprednisolone sodium succinate treatment in the follow-up of acute cervical spinal cord injury. *Spine* 2001; **26**: 426–430.
- 21 Qian T, Guo X, Levi AD, Vanni S, Sherbert RT, Sipski ML. High-dose methylprednisolone may cause myopathy in acute spinal cord injury patients. *Spinal Cord* 2005; **43**: 199–203.
- 22 Bracken MB, Shepard MJ, Holford TR, Leo-Summers L, Aldrich EF, Fazl M *et al.* Methylprednisolone or tirilazad mesylate administration after acute spinal cord injury: 1-year follow up. *J Neurosurg* 1998; **89**: 699–706.
- 23 Committee on Trauma, American College of Surgeons. Spine and spinal cord trauma: *Advanced Trauma Life Support Student Course Manual*. American College of Surgeons: Chicago, 2008, p 170.
- 24 Liu JC, Patel A, Vaccaro AR, Lammertse DP. Methylprednisolone after traumatic spinal cord injury: yes or no? *PMR* 2009; **1**: 669–673.
- 25 Geisler FH, Dorsey FC, Coleman WP. Recovery of motor function after spinal-cord injury—a randomized, placebo-controlled trial with GM-1 ganglioside. *N Engl J Med* 1991; **324**: 1829–1838.
- 26 Geisler FH, Coleman WP, Grieco G, Poonian D. The Sygen multicenter acute spinal injury study. *Spine* 2001; **26**: S87–S98.
- 27 Harrop JS, Maltenfort MG, Geisler FH, Coleman W, Jones LAT, Wirth E *et al.* Traumatic thoracic ASIA A examinations and potential for clinical trials. *Spine* 2009; **34**: 2525–2529.
- 28 Fawcett JW, Curt A, Steeves JD, Coleman WP, Tuszynski MH, Lammertse D *et al.* Guidelines for the conduct of clinical trials in 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* 2007; **45**: 190–205.
- 29 Steeves JD, Kramer JK, Fawcett JW, Cragg J, Lammertse DP, Blight AR *et al.* Extent of spontaneous motor recovery after traumatic cervical sensorimotor complete spinal cord injury. *Spinal Cord* 2011; **49**: 257–265.
- 30 Hansabout RR, Blight AR, Fawcett S, Reddy K. 4-aminopyridine in chronic spinal cord injury: a controlled, double-blind, crossover study in eight patients. *J Neurotrauma* 1993; **10**: 1–18.
- 31 Hayes KC, Potter PJ, Wolfe DL, Hsieh JTC, Delaney GA, Blight AR. 4-aminopyridine-sensitive neurologic deficits in patients with spinal cord injury. *J Neurotrauma* 1994; **11**: 433–446.
- 32 Cardenas DD, Ditunno J, Graziani V, Jackson AB, Lammertse D, Potter P *et al.* Phase 2 trial of sustained-release fampridine in chronic spinal cord injury. *Spinal Cord* 2007; **45**: 158–168.
- 33 Knoller N, Auerbach G, Fulga V, Zelig G, Attias J, Bakimer R *et al.* Clinical experience using incubated autologous macrophages as a treatment for complete spinal cord injury: phase I study results. *J Neurosurg Spine* 2005; **3**: 173–181.
- 34 Jones LAT, Lammertse DP, Charlifue SB, Kirshblum SC, Apple DF, Ragnarsson KT *et al.* A phase 2 autologous cellular therapy in patients with acute, complete spinal cord injury: pragmatics, recruitment, and demographics. *Spinal Cord* 2010; **48**: 798–807.
- 35 Lammertse DP, Jones LAT, Charlifue SB, Kirshblum SC, Apple DF, Ragnarsson K *et al.* Autologous incubated macrophage therapy in acute, complete spinal cord injury: results of the Phase 2 randomized controlled multicenter trial. *Spinal Cord* 2012; **50**: 661–671.
- 36 Rossignol S, Chau C, Brustein E, Belanger M, Barbeau H, Trevor D. Locomotor capacities after complete and partial lesions of the spinal cord. *Acta Neurobiol Exp* 1996; **56**: 449–463.
- 37 Wernig A, Muller S. Laufband locomotion with body weight support improved walking in persons with severe spinal cord injuries. *Paraplegia* 1992; **30**: 229–238.
- 38 Behrman AL, Harkema SJ. Locomotor training after human spinal cord injury: a series of case studies. *Phys Ther* 2000; **80**: 688–700.
- 39 Dobkin BH, Apple D, Barbeau H, Basso M, Berhman A, Deforge D *et al.* Weight-supported treadmill vs over-ground training for walking after acute incomplete SCI. *Neurology* 2006; **66**: 484–493.

- 40 Wolpaw J. Treadmill training after spinal cord injury: good but not better. *Neurology* 2006; **66**: 466–467.
- 41 Wernig A. Weight-supported treadmill vs over-ground training for walking after acute incomplete SCI. *Neurology* 2006; **67**: 1900.
- 42 Barbeau H, Basso M, Behrman A, Harkema S. Treadmill training after spinal cord injury: good but not better. *Neurology* 2006; **67**: 1900–1901.
- 43 Dobkin BH, Duncan PW. Should body weight-supported treadmill training and robotic-assistive steppers for locomotor training trot back to the starting gate? *Neurorehabil Neural Repair* 2012; **26**: 308–317.
- 44 Blight A, Curt A, Ditunno JF, Dobkin B, Ellaway P, Fawcett J *et al*. Position statement on the sale of unproven cellular therapies for spinal cord injury. *Spinal Cord* 2009; **47**: 713–714.
- 45 Curt A. The translational dialogue in spinal cord injury research. *Spinal Cord* 2012; **50**: 352–357.
- 46 Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010; **152**: 1–7.
- 47 Dobkin BH. What matters in cellular transplantation for spinal cord injury: the cells, the rehabilitation, or the best mix? *Neurorehabil Neural Repair* 2010; **24**: 7–9.