Steroids and Spinal Cord Injury: Revisiting the NASCIS 2 and NASCIS 3 Trials

Nesathurai, Shanker MD, FRCP(C)

Author Information
From the Boston University School of Medicine and New England Regional Spinal Cord Center, Boston Medical Center, Boston, Massachusetts.
This work was supported in part by grant H133N950014-97 from the National Institute on Disability and Rehabilitation Research.
Address for reprints: Shanker Nesathurai, MD, F511, East Newton Campus, Boston Medical Center, 1 Boston Medical Place, Boston, MA 02118-2393.

Abstract

The National Acute Spinal Cord Injury Study (NASCIS) 2 and 3 trials are often cited as evidence that high-dose methylprednisolone is an efficacious intervention in the management of acute spinal cord injury. Neither of these studies convincingly demonstrate the benefit of steroids. There are concerns about the statistical analysis, randomization, and clinical endpoints. Even if the putative gains are statistically valid, the clinical benefits are questionable. Furthermore, the benefits of this intervention may not warrant the possible risks. This paper comments on these two clinical trials.

The use of intravenous high-dose methylprednisolone in nonpenetrating acute spinal cord injury (SCI) has become the standard of care throughout the United States and Canada. The rationale for this intervention was derived from the National Acute Spinal Cord Injury Study (NASCIS) 2 and 3 trials. [1,2] For a number of reasons, however, these two trials do not convincingly demonstrate the benefits of methylprednisolone. First, there are some deficiencies in the statistical analysis. Second, the clinical endpoints are poorly defined. Third, the data are presented in an ambiguous manner that does not allow for clear interpretation. More importantly, the putative gains attributed to steroids are relatively modest and may not warrant the risks of treatment.

SUMMARY OF NASCIS 2 AND NASCIS 3 RESULTS

NASCIS 2 randomized acute SCI patients to placebo, methylprednisolone, or naloxone. Patients were treated within 12 hours of SCI. Patients enrolled in the steroid arm received a methylprednisolone bolus of 30 mg/kg, followed by a continuous infusion...
of 5.4 mg$\text{[center dot]}$kg$^{-1}$[center dot]$\text{h}^{-1}$ for 23 hours. At 1 year after injury, there were no significant differences in neurologic recovery among the treatment groups. [3] Steroids, however, were reported to be efficacious in a subgroup of patients who were treated within 8 hours of injury. It should be noted that the "positive" subgroup consisted of less that 50% of the 487 enrolled patients. In this subgroup, the average increases in motor score (maximum possible value is 70) for the steroid and placebo groups were 17.2 and 12.0 points, respectively. This 5.2-point difference ($p = 0.030$) was at 1 year after injury with an intent-to-treat analysis. If only those patients who complied with the protocol were included (i.e., ignoring the intent-to-treat paradigm), then motor scores improved by 6.6 points ($p = 0.014$) at 1 year. Improvements in pinprick sensation and light touch sensation at 1 year after injury were 2.4 points ($p = 0.251$) and 3.4 points ($p = 0.122$), respectively.

In NASCIS 2, the subgroup of patients who were treated with methylprednisolone 8 hours after injury had worse neurologic outcomes than the placebo group. This would suggest that steroid treatment may be detrimental to some SCI patients. Although the actual motor scores were not reported, this difference did not reach statistical significance ($p = 0.08$). [3]

In NASCIS 3, 499 patients were randomized into three treatment groups. There was no placebo arm. All patients received an initial bolus of intravenous methylprednisolone (20-40 mg/kg). One group then received tirilazad mesylate for 48 hours. The other two groups received 48 or 24 hours of methylprednisolone. Patients were treated within 8 hours of SCI. Clinical examinations were completed at 6 weeks and 6 months after injury. When all patients were considered with an intent-to-treat analysis, the differences between the groups that received 48 and 24 hours of steroids were 2.8 motor points ($p = 0.09$) at 6 weeks and 3.4 motor points ($p = 0.07$) at 6 months. If only patients who complied with the protocol were included, the improvements at 6 weeks and 6 months were 3.6 points ($p = 0.04$) and 3.7 points ($p = 0.06$), respectively. At best, these motor score gains are of borderline statistical significance. There was no statistically significant difference in sensory function (i.e., pinprick or light touch) at 6 months after injury.

NASCIS 3 then subdivided the participants into those who received steroids within 3 hours of injury and those who received steroids between 3 and 8 hours after injury. In those patients treated within 3 hours of injury, there was no difference in the motor scores among the treatment arms. In the subgroup of patients treated within 3 to 8 hours after SCI, however, there was a reported advantage to 48 hours of methylprednisolone compared with 24 hours of methylprednisolone. In this subgroup, patients who were treated with 48 hours of steroids had an improvement of 6.4 motor points ($p = 0.01$) at 6 months after injury. If only compliers were analyzed, a 7.2-point gain ($p = 0.008$) was reported.

NASCIS 3 demonstrated that treatment with 48 hours of steroids resulted in a small but statistically significant improvement in the self-care and sphincter-control domains of the Functional Independence Measure compared with 24 hours of steroids. This gain was noted when all enrolled patients were included (i.e., no stratification analysis). The
self-care subset had a maximum value of 42 points. Treatment with 48 hours of steroids resulted in a further 2.4-point gain ($p = 0.03$). The sphincter-control domain had a maximum value of 14 points, and 48 hours of steroids resulted in a 1.1-point gain ($p = 0.01$).

In NASCIS 2, patients with disorders "of nerve root or cauda equina only" were excluded. The conus medularis ends at approximately L1. It would appear that injuries below the L1 segment were not randomized. Consequently, the results of NASCIS 2 are not applicable to patients with lesions below the conus medularis (i.e., paraplegia secondary to an L3 burst fracture). The NASCIS 3 trial did not specifically exclude patients with isolated nerve-root or cauda equina disorders.

(* In NASCIS 3, patients were classified as quadriplegic, quadriparetic, paraplegic, paraparetic, or normal. A total of 499 patients were enrolled. Assume 166 patients were in each of the steroid treatment arms: 41 patients (24.7%) were classified as normal and 125 patients were classified as abnormal in the 24-hour steroid group. In the 48-hour steroid group, 23 patients (13.9%) were classified as normal and 143 patients were classified as abnormal. As such, two-by-two contingency tables can be constructed to complete the chi squared analysis.

**RANDOMIZATION AND STATISTICAL ANALYSIS**

In NASCIS 3, there were difficulties with the randomization. The authors have conceded that patients in the tirilazad treatment arm had significantly lower baseline motor scores and diastolic blood pressures compared with patients in the other two treatment arms. It also appears that the two steroid arms were not equivalent. In the group that received 24 hours of steroids, 24.7% of patients had normal motor function. In the 48-hour steroid group, 13.9% of patients had normal motor function. This difference is statistically significant ($p = 0.012$). As the treatment arms appear unbalanced, the small gains attributed to treatment with 48 hours of methylprednisolone should be interpreted with caution.

* In NASCIS 3, patients were classified as quadriplegic, quadriparetic, paraplegic, paraparetic, or normal. A total of 499 patients were enrolled. Assume 166 patients were in each of the steroid treatment arms: 41 patients (24.7%) were classified as normal and 125 patients were classified as abnormal in the 24-hour steroid group. In the 48-hour steroid group, 23 patients (13.9%) were classified as normal and 143 patients were classified as abnormal. As such, two-by-two contingency tables can be constructed to complete the [chi squared] analysis.

When the results of the entire study population do not demonstrate a difference between the treatment and control groups, extreme caution must be used in the interpretation of positive results in some subgroups. In this situation, subgroup analysis may be particularly misleading. The positive outcomes of any one subgroup are not as scientifically compelling as the negative results from the entire study population.
NASCIS 2, benefit with steroids was reported only among patients treated within 8 hours of injury. There was no benefit when all study participants were included (i.e., treatment within 12 hours of injury). The authors contend that they had an a priori hypothesis that time to treatment would affect outcome and that stratification of outcomes on the basis of time to treatment was permissible. If treatment within 8 hours was such a strong a priori concern, then the patients should have been randomized in a stratified manner (i.e., up to 8 hours and between 8 and 12 hours). It would be interesting to know if other subgroups at other time demarcations were also analyzed (i.e., less than 6 hours or between 2 and 9 hours). There is a temptation to present results from the subgroup with the greatest difference between treatment and placebo. A similar stratification analysis was completed in NASCIS 3. There was no substantial statistically significant difference among the treatment groups when all patients were reviewed. A statistically significant benefit was attributed to 48 hours of methylprednisolone, however, when patients received the drug between 3 and 8 hours.

Regardless of what the subgroups may demonstrate, these two trials do not convincingly document the benefit of high-dose methylprednisolone. In an ideal prospective, randomized, controlled trial, the treatment (steroid) group and the control group should differ only in the nature of the experimental intervention. The subgroups in a randomized trial, however, are not necessarily clinically equivalent. In the subset of NASCIS 2 patients treated within 8 hours of injury, Rosner [5,6] has argued cogently that the treatment and placebo groups were dissimilar. With the exception of the baseline motor scores, the demographic characteristics of the subset of NASCIS 2 patients treated within 8 hours has not been published. The investigators have not convincingly demonstrated that the subgroups with positive findings in the NASCIS trials were equivalent. Even if these subgroups had similar baseline characteristics (i.e., age, gender, site of treatment, motor scores, etc.), there is no assurance that some unidentified confounding factor (other than steroids) was not responsible for the improvement in motor scores. After all, NASCIS 2 and NASCIS 3 were designed to randomize SCI patients to treatment within 12 and 8 hours, respectively. These are the study populations that are most likely to be clinically equivalent.

Stratification by time to treatment allows for a multitude of potential subgroups. In NASCIS 2 and NASCIS 3, patients were randomized to treatment within 12 and 8 hours of injury, respectively. If one were to determine the number of possible subgroups available for analysis at discrete 1-hour intervals (i.e., 0-3 hours, 2-7 hours, etc.), there would be 78 potential subgroups in NASCIS 2. Similarly, there would be 36 potential subgroups available for analysis in NASCIS 3. With 20 subgroups, by chance alone, 1 subgroup would be positive for a p value of 0.05.

In the NASCIS trials, patients are assigned to one of five motor groups: quadriplegic, paraparetic, paraplegic, paraparetic, and normal. Likewise, there are five potential sensory groups: analgesic and anesthetic at or above T1, analgesic and anesthetic below T1, hypalgesic and hypoesthetic at or above T1, hypalgesic and hypoesthetic below T1, and normal. In total, there are five motor subgroups and five sensory subgroups; therefore, 25 possible subgroups exist for each arm of the NASCIS 2 and 3 trials. The authors have commented on outcomes in various combinations of these subgroups, such as "plegic patients with total sensory loss below their level of injury." [3] The more
subgroups that are available, the more likely that by chance alone some subgroups will demonstrate statistically significant differences. [7]

CLINICAL END POINTS

Even if the statistical analysis is valid, there are serious questions regarding the interpretation of the results. The baseline data were obtained in an inexact manner, and the results were presented ambiguously. More importantly, the reported gains are relatively small and may not warrant the associated risks.

One of the key clinical end points for the NASCIS 2 and 3 trials was the aggregate motor score. The motor score was determined by examining 14 muscle groups bilaterally in NASCIS 2. In NASCIS 3, 1 extra muscle was added, so that 15 muscle groups were examined. This extra muscle group, however, was excluded from calculations of the motor score. Although both sides of the body were examined, the investigators reported only the results from the right side in both NASCIS trials. They have stated that the results from the left side were similar.

Each muscle tested was assigned a point value from 0 to 5. A score of 0 indicated no contraction; 1 indicated reduced contraction; 2 indicated active movement without antigravity; 3 indicated active movement with antigravity; 4 indicated reduced function but active movement against resistance; and 5 indicated normal motor function. The sum of the motor scores for all 14 muscles was calculated to obtain a total motor score. Motor scores could range from 0 to 70 on each side.

This motor scoring system equates the improvement from 0 to 1 or 4 to 5 (i.e., a 1-point gain) as equally important as that from 2 to 3 or 3 to 4. This is a fundamental flaw. The functional improvement from 0 to 1 or 4 to 5 in any particular muscle group is far less than that from 2 to 3 or 3 to 4. Naturally, patients would like to maintain as much motor strength as possible. Yet, spinal cord injury patients gain very little function when an individual muscle improves from no contraction (0 grade) to a reduced contraction (1 grade, a flicker of movement). Conversely, for most functional activities, grade-4 strength is satisfactory. There are few functional tasks that require a minimum of grade-5 strength in any particular muscle group.

Moreover, the location of the motor function gains is essential to meaningful functional gains. A 5-motor-point gain in a quadriplegic patient distributed over five dispersed myotomes would not translate into any functional improvement. Conversely, a 5-point gain concentrated at one of two motor levels could result in significant functional benefit for the patient.

The functional capabilities of SCI patients is closely related to motor level. [8-10] The American Spinal Injury Association (ASIA) scoring protocol [11,12] is the most commonly accepted scoring system for spinal cord patients. There are 10 key muscles
that correspond to 10 motor levels: C5 (biceps), C6 (extensor carpi radialis longus and brevis), C7 (triceps), C8 (flexor digitorum profundus), T1 (abductor digiti minimi), L2 (iliopsoas), L3 (quadriceps), L4 (tibialis anterior), L5 (extensor hallucis longus), and S1 (gastrocnemius). Each of these muscles can be reasonably examined in the supine position. In NASCIS 2, 9 of the 10 ASIA key muscles were examined.  In NASCIS 3, all of the ASIA key muscles were included. In the ASIA scoring system, [11, 12] the motor level is determined by identifying the most caudal key muscle group with at least grade-3 strength. All remaining cephalad key muscles must have at least grade-4[dagger] strength. The one exception to this rule is the C5 motor level, because there is no C4 index muscle.

[dagger] In NASCIS 2, the only ASIA key muscle that was not included was the flexor digitorum profundus. The extensor digitorum, which was examined, would be a reasonable substitute. Both muscles are partly supplied by the C8 root. Therefore, it would be relatively easy to obtain an ASIA motor level for NASCIS 2 patients.

[double dagger] Under the new ASIA scoring system, [12] the motor level is determined by identifying the most caudal key muscle with at least grade-3 power. The remaining cephalad key muscles groups, however, must have grade-5 strength. This new system makes it more difficult to gain a motor level, and in my judgment, it does not correlate as well with functional capabilities. As such, I prefer the old system. [11] This distinction, however, does not materially change the arguments based on the hypothetical case (Table 1).

<table>
<thead>
<tr>
<th>Movements</th>
<th>Index Muscle (Right Side)</th>
<th>Column I: No Benefits</th>
<th>Column II: Steroids (Least Scenario)</th>
<th>Column III: Steroids (Worst Case)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5</td>
<td>Biceps</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>C6</td>
<td>Extensor carpi radialis</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>C7</td>
<td>Triceps</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>C8</td>
<td>Flexor digitorum profundus</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>T1</td>
<td>Abductor digiti quinti</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>L2</td>
<td>Iliopsoas</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>L3</td>
<td>Quadriceps</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>L4</td>
<td>Tibialis anterior</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>L5</td>
<td>Extensor hallucis longus</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S1</td>
<td>Gastrocnemius</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aggregate Motor Score</td>
<td></td>
<td>6</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>ASIA Motor Level</td>
<td></td>
<td>C5</td>
<td>C7</td>
<td>C5</td>
</tr>
</tbody>
</table>

Table 1. Three possible neurologic outcomes for a hypothetical spinal cord injury patient 1 year after injury.

To demonstrate the importance of the distribution of motor point gains, Table 1 has been constructed. This Table depicts three possible neurologic examination outcomes for a hypothetical SCI patient at 1 year after injury. For clearer understanding, assume that the motor scores results are the same bilaterally. Column I depicts the motor scores of a patient who received placebo. This untreated hypothetical patient has an aggregate motor score of 6. He would have a C5 motor level. He would be a quadriplegic who would require assistance with self-care and transfers. [8, 9] This patient would require a power wheelchair. Bladder management would probably require
a Foley catheter or a suprapubic catheter. He would require some degree of daily nursing care (probably a minimum of 3 hours per day). Column II depicts the scores of a patient who received steroids. The 5-point gain (as per NASCIS 2) in the motor score is concentrated in a few myotomes. The patient, with 11 motor points, can now be reclassified as a C7 quadriplegic. This individual would, with proper training and assuming that the gains on both sides were similar, be independent in self-care and transfers. He could propel a manual wheelchair and possibly drive a modified vehicle. This patient may manage his bladder independently with intermittent catheterization. The final column (column III) depicts the scores of a patient who also had a 5-point gain. Although this patient has 11 motor points, the patient would remain at a C5 motor level. The motor improvements are dispersed over a number of myotomes, with no meaningful functional improvement.

Granted, this is only one example of a multitude of scoring possibilities. Other scenarios could be constructed. The C5** motor level, however, was the most common level (27.9%) of injury in NASCIS 2. [13] This case also demonstrates that the distribution of motor gains is paramount. This example assumes a 5-point gain in motor scores over the 10 ASIA key muscles. In NASCIS 2, the approximately 5-point motor gain is distributed over 14 muscle groups. As such, the ideal scenario (column II) probably overstates an improvement in motor level. NASCIS 3 reported that in patients treated with 24 hours of steroids, 85.6% of all motor function gained was below the level of the lesion. This would suggest that the worst-case scenario (column III) is more likely than the ideal scenario (column II).

** The motor level described here is not an ASIA motor level, but a motor level derived from a model devised by Holford and Bracken. [15] This is an elegant, but relatively complex, statistical model. It is uncertain if this model correlates with ASIA motor levels or conventional anatomic levels. For example, in NASCIS 2, between 33 and 37% of the patients in each treatment arm were classified as paraplegic or paraparetic. Paraplegic and paraparetic patients should have intact upper extremity function and, by definition, a motor level below C8-T1. In a later paper published by Bracken and Holford [13] that discusses long-tract improvement in NASCIS 2, however, less than 10 of the approximately 500 patients (2%) had a NASCIS motor level below C8-T1.

RELIABILITY OF DATA COLLECTION

It has been reported that approved personnel completed the exhaustive neurologic examination in NASCIS 2. In NASCIS 3, these examinations were completed by physicians and nurses. Among the muscle groups tested were the hamstring and opponens muscles. The hamstring muscles are the primary flexors of the knee. To accurately grade the hamstring muscle between as 1 grade (reduced contraction) and 2 grade (active movement without antigravity), the patient must be positioned on his or her side to eliminate the effects of gravity. For 3 grade (active movement with antigravity), the patient should placed in the prone position. [14] It is unclear if acute spinal cord injury patients, with potentially unstable injuries, were positioned in this manner. The opponens pollicis muscle was also scored. Once again, it is unclear how an examiner could accurately score this muscle. The opponens has no pure plane of movement that
would allow a clinician to distinguish its function from that of nearby surrounding muscles. For example, a flicker of movement of the thumb (presumably grade-1 opponents function) could also be the result of contraction of the flexor pollicis longus, flexor pollicis brevis, extensor pollicis longus, extensor pollicis brevis, abductor pollicis longus, or abductor pollicis brevis. To distinguish between grade-2 and grade-3 strength in the opponents pollicis, the muscle must be placed in a position to eliminate gravity. There is no position in which the hand can be placed to reliably test the opponents in a gravity-eliminated position.

NASCIS 2 collected sensory responses for light touch and pinprick from 29 sites (C2 to S5) bilaterally. In addition, in NASCIS 3, deep pain and pressure were evaluated in six sites bilaterally. Participants were asked if the sensory perception was absent, dysfunctional, or normal for each sensory test. In total, 116 (128 in NASCIS 3) sensory tests were performed in addition to testing the strength of 28 (30 in NASCIS 3) muscle groups. If one were to allocate 20 seconds (probably a conservative estimate) for each sensory and motor test, an expert examiner would require approximately 1 hour of the patient’s undivided attention to accurately complete the neurologic examination. Many acutely ill spinal cord patients would not be able to tolerate such a detailed clinical examination because of pain, fractures, distraction, anxiety, etc. In NASCIS 2, between 18.5 and 19.9% of enrolled patients had associated musculoskeletal disorders. In NASCIS 3, between 15.7 and 23.3% of the patients had concomitant musculoskeletal injuries. Incomplete clinical examinations appear a possibility. In NASCIS 2, the authors do not clearly state what percentage of assessments were incomplete and how the statistical analysis was completed with incomplete data. In NASCIS 3, the authors report that incomplete data “was never more than 2.5 percent of the examinations.” In such cases, the scores from the other side of the body were used in the statistical analysis.

**INEXACT DEFINITIONS OF FUNCTIONAL MOTOR LEVELS**

The investigators in both trials distinguished between four levels of motor injury: quadriplegic, quadriparetic, paraplegic, and paraparetic. A patient was considered quadriplegic (presumably complete motor injury) when the most cephalad muscle with no contraction (i.e., grade-0 strength) was the first dorsal interosseous or higher. A patient was considered quadriparetic (presumably incomplete motor injury) if the most cephalad muscle with a trace of contraction or active movement without gravity (i.e., grade 1 or 2, respectively) was the first dorsal interosseous or higher. For example, a patient with a grade-5 biceps and no motor function in any caudal myotomes would be considered quadriplegic. A patient with a grade-5 biceps and a grade-1 (i.e., reduced contraction, flicker of movement) extensor carpi radialis would be considered quadriparetic. By gaining 1 grade of motor function in one caudal myotome, a patient can be deemed to have improved from being quadriplegic (complete injury) to quadriparetic (incomplete injury). The distinction between “plegic” and “paretic” injury, as presented in the NASCIS papers, is inexact and artificial. Use of the terms quadriparesis and paraparesis is discouraged by spinal cord clinicians because it “describes incomplete lesions imprecisely.” [12]

The investigators have commented that patients have “improved” in neurologic level
from complete motor injury (quadriplegic) to incomplete motor injury (quadriparetic). With the previously stated system for defining quadriplegic and quadriparetic, however, a patient can have a clinical decline in function and yet be considered to have experienced an improvement in neurologic recovery. Consider a patient with initial neurologic results of grade-5 biceps, grade-5 extensor carpi radialis, and grade-0 strength in all caudal myotomes. He would be considered to be quadriplegic (complete motor injury), with a motor score of 10. At a later examination, he could have grade-5 biceps and grade-2 extensor carpi radialis. This patient would then have a motor score of 7 and would clearly have declined in motor function. Yet, he would now be reclassified as quadriparetic (incomplete motor injury). By the investigators’ definitions, this patient would have improved in neurologic level (i.e., complete to incomplete motor injury), although he had deteriorated clinically. Given the imprecise application of these terms, reported improvements in neurologic level must be interpreted with caution.

The deficiencies in data collection and the imprecise definition of terms are relatively minor points. In theory, both the treatment and placebo group should be equally affected by these inadequacies. The selection and presentation of clinical end points, however, provide some insight into the overall quality of a clinical trial.

In a randomized trial, each arm of the study should be equally affected by the imprecise data collection. This is not necessarily true, however, for any particular subgroup. One could speculate that the positive outcomes in any particular subgroup may relate to the differences in the accuracy of clinical examinations, and not to steroid treatment.

HAS NASCIS 2 BEEN REPLICATED?

In the NASCIS 3 publication, it was asserted that NASCIS 2 was replicated in Japan. This trial was published in the Japanese-language journal Sekisui Sekizui [16] (roughly translated as 'The Spine and Its Contents'). This publication is not indexed on MEDLINE. Upon review of an English-language translation, many serious deficiencies were found in this article. For example, a double-blind, placebo study design was not used. The treating physicians were cognizant of which patients were allocated to the control group. Patients in the control group were not given a placebo. In violation of the study protocol, approximately one-fifth of the control group patients received methylprednisolone at some point in their treatment. Also, there was a statistically significant difference in baseline motor scores for the control and treatment groups (p = 0.04). Furthermore, there was a statistically significant difference in the baseline functional (Frankel) levels of the patients in the two arms of the Japanese trial (p = 0.05). Recent retrospective studies have not been able to confirm the benefits of steroids in SCI. [17,18]

A copy of an English-language translation can be obtained from the author.
REPORTED ADVERSE EFFECTS OF STEROIDS

Treatment with high-dose steroids is not without risk. In NASCIS 3, patients treated with 48 hours of steroids had a higher incidence of severe sepsis (p = 0.07) and severe pneumonia (p = 0.02) compared with the patients who received 24 hours of steroids. Pneumonia and sepsis can lead to intubation, which involves some degree of neck manipulation. Clearly, this should be avoided in patients with cervical cord lesions.

CONCLUSION

The loss of function secondary to a spinal cord injury can be devastating. It is uncertain, however, whether the NASCIS 2 and 3 trials convincingly demonstrate the benefits of high-dose steroids. Even if one were to accept the statistical analysis as valid, the clinical and functional significance of these putative gains remains unclear.

This is a topic of great significance to the spinal cord injury community. If high-dose methylprednisolone is truly efficacious, then all eligible patients should receive this medication. If steroids are not of clinical value, however, then this intervention should not be admitted into medical practice. At present, many clinicians who do not believe that steroids are beneficial are placed in the unenviable position of prescribing this treatment because of medicolegal concerns. To help resolve this dilemma, an impartial blue-ribbon panel of clinicians and statisticians should be appointed to reanalyze the NASCIS 2 and 3 primary data. Their report may provide guidance to physicians in their clinical practices. At a minimum, the NASCIS 2 results should be reanalyzed to determine if reported gains in aggregate motor score translate into improvements in ASIA motor level. In addition, the baseline descriptive characteristics (i.e., demographic data, site of treatment, types of surgical intervention, etc.) of the subgroup of patients treated within 8 hours of injury should be released.

To truly ascertain if high-dose methylprednisolone is efficacious, a new prospective, randomized, placebo-controlled trial that enrolls patients within 8 hours of injury could be organized. In the current political climate, however, it would be extremely difficult to obtain approval from any institutional review board.

Ultimately, this debate is about excellence in patient care. To achieve this goal, physicians must incorporate the results of clinical studies into their medical management. This paradigm, however, requires practitioners to critically review the scientific data before altering accepted clinical practice. The evidence, as presented in the NASCIS 2 and 3 trials, does not unequivocally demonstrate the clinical efficacy of high-dose methylprednisolone.

Acknowledgments
The author acknowledges the thoughtful criticisms of Allan Meyers, PhD, Herbert Kayne, PhD, Mark Kaplan, MD, and Howard Cabral, MPH.

ADDITIONAL READINGS


Taylor TFK, Ryan M. Methylprednisolone in the management of acute spinal cord

REFERENCES


---

**IMAGE GALLERY**

Select All

Export Selected to PowerPoint

---

Table 1