J Neurosurg (Spine 1) 93:1–7, 2000

Methylprednisolone for acute spinal cord injury: an inappropriate standard of care*

R. JOHN HURLBERT, M.D., PH.D., F.R.C.S.(C)

University of Calgary Spine Program, Foothills Hospital and Medical Centre, Calgary, Alberta, Canada

Object. Since publication in 1990, results from the National Acute Spinal Cord Injury Study II (NASCIS II) trial have changed the way patients suffering an acute spinal cord injury (SCI) are treated. More recently, recommendations from NASCIS III are being adopted by institutions around the world. The purpose of this paper is to reevaluate carefully the results and conclusions of these studies to determine the role they should play in influencing decisions about care of the acutely spinal cord–injured patient.

Methods. Published results from NASCIS II and III were reviewed in the context of the original study design, including primary outcomes compared with post-hoc comparisons. Data were retroconverted from tabular form back to raw form to allow direct inspection of changes in treatment groups. These findings were further analyzed with respect to justification of practice standards.

Although well-designed and well-executed, both NASCIS II and III failed to demonstrate improvement in primary outcome measures as a result of the administration of methylprednisolone. Post-hoc comparisons, although interesting, did not provide compelling data to establish a new standard of care in the treatment of patients with acute SCI.

Conclusions. The use of methylprednisolone administration in the treatment of acute SCI is not proven as a standard of care, nor can it be considered a recommended treatment. Evidence of the drug's efficacy and impact is weak and may only represent random events. In the strictest sense, 24-hour administration of methylprednisolone must still be considered experimental for use in clinical SCI. Forty-eight-hour therapy is not recommended. These conclusions are important to consider in the design of future trials and in the medicolegal arena.

KEY WORDS • spinal cord injury • pharmacotherapy • methylprednisolone • steroid • practice guideline • spinal cord injury trial

In 1987, high forceps were used to deliver a newborn girl in a hospital in Southern California.[9] The attending physician inadvertently rotated her head 180° during the procedure. While in the neonatal intensive care unit, the patient suffered hypotonicity and breathing difficulties. Neurological examination yielded a diagnosis of Werdnig–Hoffmann disease. After 5 days, the infant was transferred to a larger center where a diagnosis of a complete high cervical SCI was made. Several years later, the Court of California assigned more than 13 million dollars to the family of the infant, who remained quadriplegic and dependent on a ventilator for breathing. In determining negligence, 40% of the blame was assigned to the attending obstetrician, while 60% was assigned equally between two neonatal intensive care physicians. The intensivists were faulted for erring in diagnosis and failing to institute appropriate steroid treatment. Two expert witnesses testified that had steroids been administered within an 8-hour window postinjury, there would have been a 95% probability of significant improvement in the infant's ability to breathe and use her arms.

Acute SCI remains a devastating condition, and little progress has been made toward improving or curing neurological outcome. However, since the results of NASCIS II were published in 1990, the use of MP in the treatment of this disorder has become an accepted standard.[3,7] The 1-year follow-up results from NASCIS II have been reported[4] but were less publicized. Since early and 1-year results from NASCIS III have become available, acute care medical institutions are presently faced with the decision of whether to incorporate the recommended 48-hour protocol proposed in this latest study into their existing 24-hour protocols.[5,6] More recently, a Cochrane review has been authored that further extols the use of steroids for SCI.[2]

Abbreviations used in this paper: ASIA = American Spinal Cord Injury Association; FIM = Functional Independence Measure; MP = methylprednisolone; NASCIS = National Acute Spinal Cord Injury Study; SCI = spinal cord injury.

* See the Letter to the Editor and the Response in this issue in *Neurosurgical Forum*, pp 175–182.

In the physician's role to act as a patient advocate, emerging information must be carefully weighed before accepting it as a new form of treatment. In certain instances, such a decision can be made somewhat unconsciously over time, as the original information is disseminated and casually quoted. Gradually a new treatment can become a standard of care more through the strength of ignorance and tradition than through the strength of science. In allowing this to occur, however, we risk subjecting our patients to arbitrary, unproven, and possibly dangerous treatments. In addition we jeopardize future trials with proper placebo control by creating unfounded ethical dilemmas.

The purpose of this paper is to reevaluate carefully the results and conclusions of NASCIS II and III with respect to current practice standards in the treatment of acute SCI. The studies are examined using a series of steps intuitively valid in appraising and accepting a new form of medical treatment. It was the author's intent to provide detailed information and independent, unbiased interpretation to help establish a perspective on standard-of-care treatment in acute spinal cord–injured victims.

Clinical Material and Methods

Overall reevaluation of both trials was performed in the context of the quality of evidence that is necessary to change a pattern of practice. The author assumed that to accept a new drug as a standard of care for the treatment of acute SCI that 1) the evidence should be obtained from a prospective randomized double-blind trial; 2) the study should be well designed and well executed; 3) the data should be compelling (face validity and internal consistency) and obtained using appropriate statistical methods; 4) the study should yield changes meaningful to the patient; and 5) the result should be reproducible.

In considering these criteria, it becomes immediately apparent that any new treatment must meet not just one but all of the aforementioned requirements. Failure to satisfy even one requirement must disqualify a treatment from becoming a gold standard. The results obtained in both NASCIS II and III were subsequently and systematically reevaluated with these requirements in mind.

Data obtained in the NASCIS II and III were converted from the tabular form in which they were reported into raw data, and they were then graphically represented. For the purposes of these analyses, it was assumed that mean baseline neurological scores within treatment groups were not significantly influenced when patients were lost to follow up, an assumption that benefits the original studies. The y-axis scales were chosen to reflect a range from minimum (no neurological function) to maximum (normal neurological function) possible scores. In instances in which preplanned comparisons were omitted, it was assumed that these data were uninteresting and hence not reported. In such cases, statistical insignificance was also assumed. Data obtained from the 6-month and 1-year NASCIS II and III publications were combined. The results pertaining to pharmaceuticals other than MP (naloxone and tirilazad) were not considered in detail.

Results

Prospective Randomized Trial: Design and Execution

The NASCIS II Trial. Three treatment arms including appropriate placebo controls were defined in this prospective randomized double-blind study, which was designed to test the effect of MP and naloxone on acute SCI. A 12-hour time limit was arbitrarily established in which to randomize patients and initiate treatment from the time of injury. Primary outcome measures were preplanned as bilateral motor scores (seven segments for each arm and leg, scored 0–5 for a possible range of 0–140 points), bilateral light touch and pinprick scores (following standard dermatomes scored 1–3 for a possible range of 29–87), and assignment to one of five motor, five light touch, and five pinprick categories (two- or four-limb involvement with complete, incomplete, or normal function [note: no patient had normal function in all three motor and sensory categories]). These latter assignments were made to determine if one treatment group changed category (for example, from quadriparetic to paraparetic) more frequently than the others. Important concerns about the validity of the motor scoring system have been raised.[10] However, at present a superior motor assessment has not been established.

Four hundred eighty-seven patients were randomized across 10 centers. Eighty percent received the drugs within the 12-hour time limit, and 90% received the drugs according to dosing protocol. Ninety-eight percent of the patients were available for 6-week follow up, 97% for 6-month follow up, and 95% for 1-year follow up. Based on these accomplishments, the authors of this study are to be highly commended on both the design and execution of the protocol.

The NASCIS III Trial. Three treatment arms were once again defined in this prospective randomized study, which was designed to test the effect of MP given over a period of 48 hours (the 48-hour MP group) and tirilazad mesylate given over 48 hours compared with the standard 24-hour MP protocol used in NASCIS II. A true placebo group was not included for "ethical" reasons. Based on results from NASCIS II, an 8-hour time limit was chosen, beyond which patients were ineligible for entry. Primary outcome measures were preplanned as bilateral motor scores (15 segments on each side of the body [an additional side was added to include all ASIA motor groups]), bilateral light touch and pinprick scores (range 29–87), bilateral deep pain and pressure, and assignment to one of five injury categories: quadriplegic; quadriparetic; paraplegic; paraparetic; and normal motor function (with impaired sensation). The FIM assessment was performed based on ASIA criteria.

Four hundred ninety-nine patients were randomized across 16 centers. Ninety-three percent received the drugs according to dosing protocol, whereas 94 to 97% of the assigned milligram dose was administered successfully in each group. Ninety-eight percent of patients were available for 6-week, 95% for 6-month, and 92% for 1-year follow-up review. Except for the lack of a proper placebo group, this study was otherwise acceptably designed and well executed.

Compelling Data. Two irregularities in data reporting serve to undermine credibility and are common to both

NASCIS II and III: timing of therapy, and motor assessment. Although in the results section of the NASCIS II study, the authors assured that an a priori hypothesis was for earlier treatment to produce greater improvement, no a priori intuition is provided concerning an otherwise arbitrary 8-hour window for 24-hour MP treatment. Hence, it is reasonable to expect that the data were sequentially grouped into all possible categories, hunting for differences (in NASCIS II, $<$ 1 hour as compared with $\geq$ 1 hour, $<$ 2 hours as compared with $\geq$ 2 hours, and so on). In NASCIS III it is also likely that all possible comparisons between 0 and 8 hours were made ($<$ 1 hour and 1–8 hours, $<$ 2 hours and 2–8 hours, $>$ 3 hours and 3–8 hours, and so on. With certainty, only the most interesting data were reported. It is intuitive that a time-to-treatment factor may affect outcome, but it is highly likely that such an effect would be graduated. One might expect a progressively diminished effect of 24-hour MP group dependent on the time from injury to administration. However, "all-or-none" cutoffs of 8 hours (NASCIS II) and 3 hours (NASCIS III) are neither intuitive nor likely physiological. To establish a time-dependent treatment effect, data are best displayed as a function of time and subjected to mathematical examination for the degree of correlation. In the absence of published data, the reader has no choice but to assume such correlation does not exist. Thus, although the "time windows" for treatment in both studies are interesting, the strength of the data within these windows is seriously weakened because the time windows are unplanned and apparently arbitrary in nature. Because of the multiple post-hoc comparisons required to discover these differences, quite possibly the observations reflect random chance alone.

In both NASCIS II and III motor scores were reported to improve significantly more in treatment groups at 6 weeks, 6 months, and 1 year posttreatment. Only the right-sided motor scores were reported, purportedly to simplify the presentation of the results. "Essentially identical" scores for the left side were not reported. It would seem highly unusual and counterproductive to omit from the report statistically significant differences that serve to strengthen conclusions within a clinical study. Especially considering that bilateral sensory scores were reported, it is possible to speculate that left-sided motor data did not reflect the same magnitude of change as the right-sided data, and that statistical significance was not achieved. More important, because only selected (right-sided) motor scores were reported, it must be inferred that the data as a whole (combined right- and left-sided motor scores) were uninteresting. These inconsistencies in data reporting argue against a robust nature to the results; observed differences between treatment groups may simply represent random events.

Primary Outcome Measures: NASCIS II. At 1-year follow up no significant difference was found in motor scores among the three groups of patients treated within 12 hours of injury. Although pinprick and light touch scores obtained at 6 months showed marginally more improvement in patients treated with the 24-hour MP protocol compared with those receiving placebo, this finding was not substantiated at 6 weeks or 1 year follow up. There was no reliably increased propensity to change overall severity of injury among any of the motor or sensory categories.

Post-Hoc Analyses: NASCIS II. The majority of the patients received their treatment after the 8-hour window had passed and were therefore excluded from further analyses. All subsequent noteworthy results and conclusions were then no longer based on a study population of 487 but, rather, on only 66 patients receiving MP and 69 patients receiving placebo.

Raw right-sided motor scores obtained in patients receiving treatment within 8 hours over the 1-year follow-up period are depicted in Fig. 1 *left*. Although statistical significance was reported, the data do not appear compelling. Bilateral sensory scores were provided for both light touch and pinprick in patients receiving treatment within 8 hours of injury. Marginal improvements in patients treated with the 24-hour MP protocol were reported to be statistically significant at 6 weeks and 6 months as compared with patients receiving placebo. However, the effect was lost for both sensory modalities at 1 year post-treatment (Fig. 1 *center* and *right*).

Subgroup analyses were performed for both groups according to injury severity. In contrast to the preplanned 15 categories of motor and sensory severity, results were arbitrarily compiled into three simplified categories: motor/sensory complete injury; motor complete and sensory incomplete injury; and motor/sensory incomplete injury. No lasting differences in pinprick and light touch sensations were found. However, motor scores at 6 months and 1 year posttreatment were reported to improve more in the motor/sensory complete group and in the motor/sensory incomplete group when the patients received the 24-hour MP treatment compared with those receiving placebo. Interestingly, patients with motor complete and sensory incomplete injuries seemingly fared better when given placebo than when given 24-hour MP protocol. Because of the small numbers of patients in each group, statistical inference was not drawn. The overall difference in recovery between the two treatment groups is small, not uniformly observed, not within preplanned comparisons, and therefore suspicious for random events.

Summary: NASCIS II. All primary outcome measures of NASCIS II were negative. Post-hoc comparisons proved interesting only if more than 70% of the patients were excluded from the analyses (conforming to an arbitrary 8-hour therapeutic window); if bilateral pinprick scores were ignored; if bilateral light touch scores were ignored; and if left-sided or bilateral motor scores were ignored. Reexamination of the raw data in graphic form lacks face validity. Internal consistency is absent within the various reported (and unreported) outcome measures. Therefore, the data from NASCIS II must be classified as both weak and noncompelling.

Primary Outcome Measures: NASCIS III. At 6-week, 6-month, or 1-year follow-up examinations no significant difference was found in motor scores among the three groups of patients treated within 8 hours of injury (Fig. 2 *left*). Similarly there was no difference in light touch sensation (Fig. 2 *center*), pinprick sensation (Fig. 2 *right*), deep pain, or pressure (data not provided). The study drug did not affect the patients' different propensities to change injury severity category. Total FIM scores were not appre-

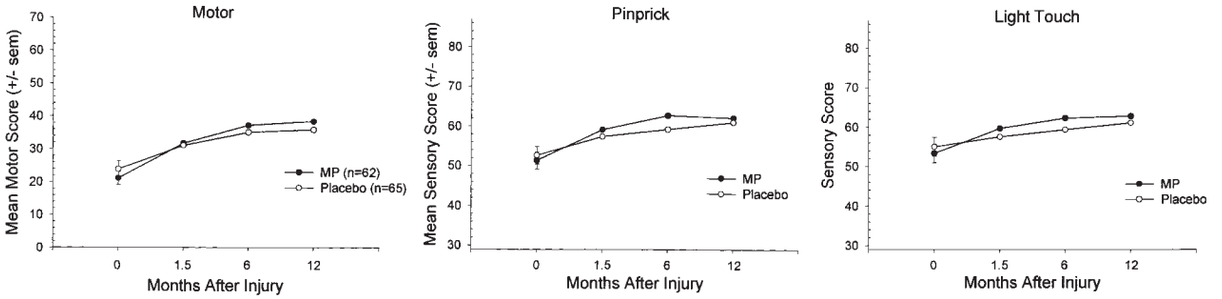

FIG. 1. Graphs depicting mean neurological scores for NASCIS II patients receiving treatment within 8 hours of SCI. *Left:* Improved motor function in patients receiving MP compared with those receiving placebo was believed by the investigators to be significant at 6 weeks, 6 months, and 1 year after injury. *Center:* Pinprick scores were reported to be improved in steroid-treated patients at 6 weeks and 6 months posttreatment, but the effect is lost at 1 year. *Right:* Light touch data showed an essentially identical trend as that obtained on pinprick sensation. When absolute scores from these three outcomes are compared in their entirety, the difference in recovery between the two treatment groups is not compelling and can easily be explained by random chance. Note that of 487 patients entered into the study, this post-hoc analysis applies to only 127 and represents only half of the available motor data. sem = standard error of the mean.

ciably different between the groups at 6 weeks, 6 months, and at 1 year posttreatment (Fig. 3). Although subgroup analyses are typically reserved for cases in which a significant difference has been discovered between total scores, FIM scores for self care and sphincter control were reported to be statistically improved by two points and one point, respectively, in the group receiving the 48- as compared with the 24-hour MP protocol at 6 months. These differences were not present at 6 weeks and were lost again at 1 year posttreatment.

Post-Hoc Analyses: NASCIS III. In addition to omitting left-sided motor scores, the preplanned 15th ASIA motor segment was also eliminated from the reported results. Therefore, inferentially, left-sided motor scores, combined motor scores, and right-sided ASIA motor scores were uninteresting. Data were presented in two additional categories: intent-to-treat and compliers. Intent-to-treat analyses are the primary analyses upon which clinical trials are built.[1] Exclusion of patients for protocol violations ("noncompliers") introduces bias. Violations such as these are best controlled through the randomization process.

Exclusion of noncompliers in NASCIS III is invalid for three reasons. First, only eight noncompliers were identified of 76 (24-hour MP group) and 84 (48-hour MP group) patients at 6 weeks. Excluding these patients from the analyses changed the level of statistical significance for motor scores by nearly an order of magnitude ($p = 0.04$ intent-to-treat; $p = 0.008$ compliers). For only eight of 160 patients to have such an effect suggests the data were somewhat unstable. Second, the patients labeled as noncompliers did not complete their dosing protocol for a reason. Although it is possible that this reason was simply technical, alternatively the medication may have produced an adverse reaction, that, in turn, may have resulted in neurological deterioration. Hence, omitting the noncompliers from analysis may misleadingly inflate the benefits of the 48-hour MP protocol. Finally (and most importantly), in examining the data it is apparent that elimination of noncompliers from the group treated with the 48-hour MP protocol improved motor scores, which initially seems intuitive. However, omitting noncompliers from the analysis of the 24-hour MP protocol group worsened the residual mean scores. Therefore, the noncompliers in the 24-hour MP group fared better than those who actually received the medication. This is not intuitive. Because both treatment groups received active drugs (no true pla-

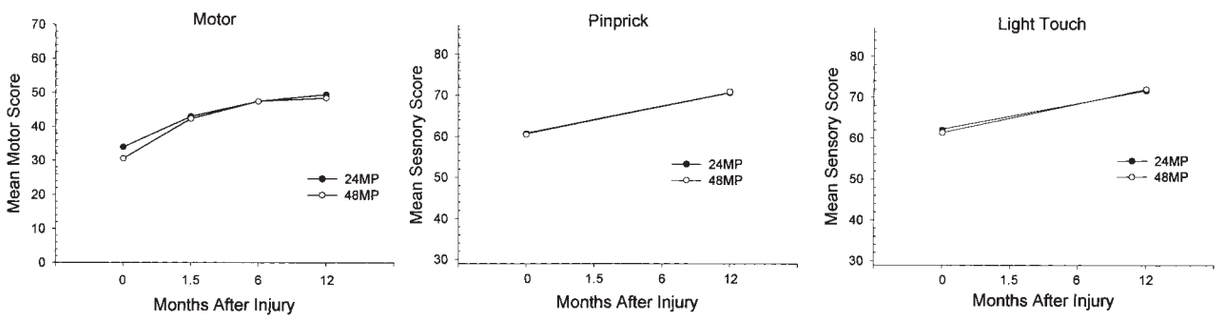

FIG. 2. Graphs showing the primary (preplanned) neurological scores obtained in patients enrolled in the NASCIS III. All patients were treated within 8 hours of SCI. *Left:* Motor scores. *Center:* Pinprick sensation scores. *Right:* Light touch sensation scores. Missing data points in sensory scores were not reported. The standard error cannot be determined from published data. Nonetheless, the results unambiguously fail to demonstrate a difference between the 24-hour MB (24MB) and 48-hour MB (48MB) treatment groups.

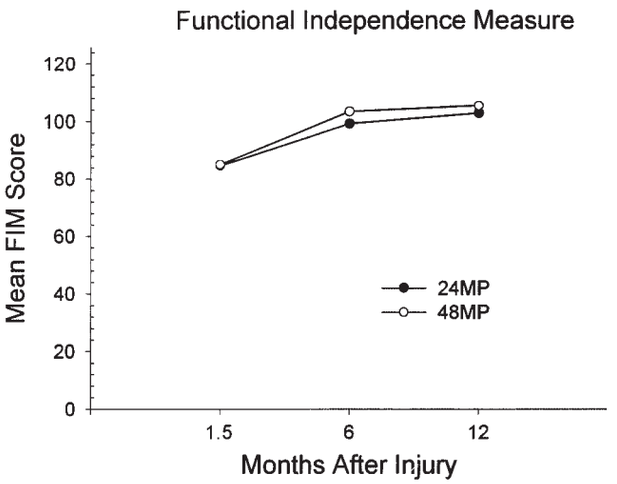

FIG. 3. Graph depicting the mean total FIM scores in NASCIS III patients treated with the 24- and 48-hour MB protocol at 6 weeks, 6 months, and 1 year postinjury (intent to treat). No compelling (or statistically significant) difference in FIM scores was demonstrated as a result of treatment.

cebo), dropping noncompliers does not clarify but rather confuses subsequent comparisons. Therefore, analysis of compliance in this experiment is undesirable, invalid, and should not be considered.

Results of post-hoc analyses suggested that change in motor score obtained at 6-week and 6-month follow up was significantly higher in those patients receiving MP over 48 hours when it was administered between 3 hours and 8 hours after injury as compared with receiving MP over 24 hours. By the authors' own strict statistical standards this effect was reduced to only a trend at 1 year ($p = 0.053$). These results are portrayed graphically in Fig. 4. There were no appreciable differences between the groups with respect to deep pain, pinprick, and light touch. That results of pressure testing were not reported causes one to acknowledge the possibility that they may have favored the 24-hour MP group. An increased propensity for patients receiving MP over 48 hours (when administered within 3–8 hours) to improve their motor severity category compared with patients receiving MP over 24 hours was found to be statistically significant at 6 months, but significance was lost at 1 year. The FIM scores were not reported for the patients receiving the drug with 3 to 8 hours of injury, also implying uninteresting data.

Summary: NASCIS III. All primary outcome measures of NASCIS III were negative. Post-hoc analyses proved interesting only when the data were arbitrarily restricted to patients treated within 3 to 8 hours of injury, excluding almost 70% of the original study population, and when bilateral light touch sensation, bilateral pinprick sensation, bilateral deep pain, bilateral pressure, left-sided and bilateral motor scores, and right-sided ASIA motor scores were ignored. Even then, the potential treatment effect was lost at 1 year. Analysis of compliance is especially invalid in this protocol because all groups received active medications. Reexamination of the "interesting" raw data lacks face validity. Internal consistency is absent among the various outcome measures. In addition, internal consistency is lacking in that FIM scores were reported to be improved at 6 months in the 48-hour MP group for all entry times (0–8 hours), whereas motor scores were only improved for entry times restricted to 3 to 8 hours. Therefore, the data from NASCIS III must be classified as both weak and noncompelling.

Statistical Significance With Appropriate Tests

In NASCIS II, ignoring naloxone data, there were more than 66 statistical comparisons performed. In NASCIS III, excluding data obtained in patients who received tirilazad mesylate over 48 hours, there were over 100 comparisons performed. Based on an alpha of 0.05, the chance of a type I error (concluding the treatment has an effect when indeed it does not) is one in 20. This problem of multiple comparisons would be expected to create at least three positive statistical tests from random chance alone in NASCIS II. At least five such falsely positive tests would be expected in NASCIS III. Various techniques are available to help correct for this problem. For example, a Bonferroni correction would stipulate a minimum p value of less than or equal to 0.0007 in NASCIS II and less than or equal to 0.0005 in NASCIS III to conclude significance.[1] This can be criticized as an overly aggressive correction. Nonetheless, additional techniques lend themselves to help reduce the problem of multiple comparisons particularly in this type of experimental design, such as two-way or repeated-measures analysis of variance.[12] The statistical methods used in NASCIS II and III were not corrected and are therefore inappropriate. Because of internal inconsistencies within the data (see previous section), the positive statistical comparisons reported in both studies are at least equally as likely to represent random events.

Appreciable Impact on Issues Important to Patients

The NASCIS II has been widely criticized for its failure to include outcomes important to the patient.[8,13] It serves no purpose to do so further at this time. To correct this oversight, the NASCIS III protocol included the ASIA FIM assessment. As discussed previously, no differences in FIM scores were demonstrated between treatment groups at any of the follow-up points. Therefore, neither NASCIS II nor NASCIS III has demonstrated a benefit of importance to the spinal cord–injured patient.

Alternatively, both studies reported potential adverse effects due to steroid administration that have serious negative ramifications for patients. In NASCIS II there was a 1.5-fold higher incidence of gastrointestinal hemorrhage, twofold higher incidence of wound infection, and threefold higher incidence of pulmonary embolism in the MP group as compared with controls. Similarly in NASCIS III there was a twofold higher incidence of severe pneumonia and a fourfold higher incidence of severe sepsis in the 48-hour MP group as compared with the 24-hour MP group. Although these differences did not reach statistical significance, one cannot conclude that steroids are not harmful. Sample size calculations based on NASCIS II data indicate that to prove statistically that there was no difference in the rate of wound infection between both groups would require over 1400 patients in each group (sample size for

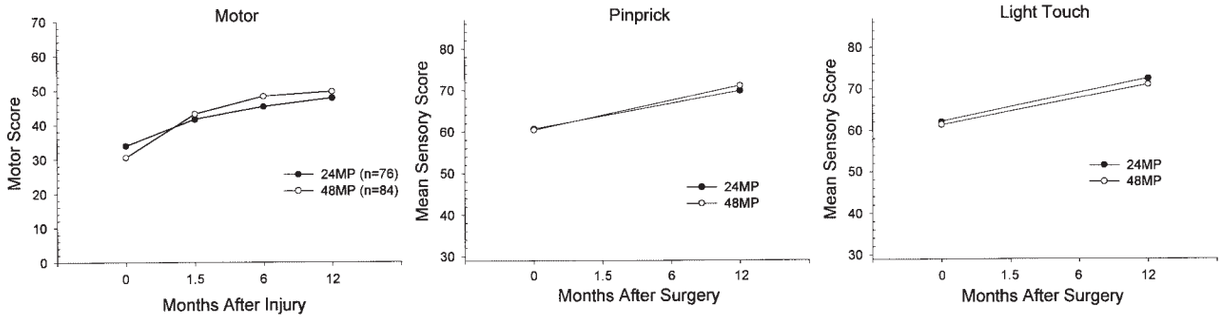

FIG. 4. Graphs showing the mean neurological scores for patients receiving treatment between 3 to 8 hours of SCI in the NASCIS III. *Left:* Improved motor function in patients receiving MP compared with those receiving placebo was reported to be statistically significant at 6 weeks, 6 months, and 1 year after injury. However, mean scores viewed in their entirety, across all 12 months, do not appear to differ substantially between groups. *Center:* No difference in admission and 1-year pinprick scores is demonstrated between patients receiving the 24- and 48-hour MB protocol. *Right:* Light touch data are similarly uninteresting.

binomial proportion, $\beta = 0.8$). Of more concern is the observation in NASCIS III of a sixfold higher incidence of death due to respiratory complications in the 48-hour MP group compared with the 24-hour MP group ($p = 0.056$), suggesting higher mortality rate associated with a 48-hour protocol.

Reproducibility of Results.

Although NASCIS II is purported to have been replicated independently in Japan, the results have not been formally translated into English, are not available through the National Library of Medicine, and have not been subjected to appropriate peer review.[11] In addition, the study appears to suffer from several critical design flaws.[10] To date, NASCIS III has not been reproduced; it is unlikely that this study will ever be undertaken again because of its negative outcomes.

Discussion

In accordance with the criteria proposed at the beginning of this paper to help guide clinical acceptance of a new treatment strategy, both NASCIS II and NASCIS III were well-designed and well-executed trials. The work of the clinical investigators in acquiring and randomizing patients, ensuring protocol compliance, obtaining data, and providing appropriate follow up was monumental. However, independent detailed reexamination of available data and statistical methods reveals several flaws that critically undermine the credibility of these studies. Despite qualifying as Class I experimental trials, both NASCIS II and III fail to meet four of six requirements intuitively important in validating a new treatment (Table 1). In addition, evidence of an unacceptably high mortality rate resulting from respiratory complications is found in the NASCIS III. It is interesting to note that despite the overwhelmingly negative (and potentially harmful) 1-year follow-up results of NASCIS III the authors were led to conclude that "Patients starting therapy 3 to 8 hours after injury should be maintained on the regimen for 48 hours unless there are complicating medical factors." [6] Clearly, from an unbiased viewpoint, the results of the study can be interpreted to indicate quite the opposite.

Although our present ability to improve functional status following acute SCI is frustratingly hindered, as patient advocates it is undesirable for the medical community to propagate unproven or poorly proven therapeutic regimens. Especially in today's evidence-based environment, perpetual vigilance is necessary to guard against the pressures attendant with publication. In reporting findings from complicated clinical trials, it is highly desirable to present all negative results fairly, especially preplanned or primary outcome measures.

There remains room to speculate about the potential beneficial effect of MP in the treatment of acute SCI. Marginally positive p values, especially in NASCIS II (if reproducible with more appropriate statistical techniques), might simply reflect inadequacies of sample size. However, until data obtained from larger populations of patients are published, such hypothetical beneficial effects must be treated as speculation only and cannot be considered in determining standard of care.

Conclusions

In summary, most experts would agree that the currently accepted management for acute SCI consists of protection of airway, breathing, and circulatory status, as well as immobilization, oxygenation, and blood pressure maintenance through volume. A critical reevaluation of the clinical efficacy of steroid administration in acute SCI demonstrates that, despite a Class I trial and general widespread use, the evidence for 24-hour MP therapy in humans is negligible or weak at best. As such, it can be regarded as no more than an experimental treatment at this time. Due to the lack of compelling, objectively reported and properly analyzed evidence, steroid administration cannot be considered a standard of care, a recommended treatment, or even a proven treatment option. There is no evidence to support treatment with the 48-hour MP protocol in patients treated within 3 to 8 hours of acute injury. More important,

TABLE 1

Adherence of NASCIS II and III to requirements of a new standard of care

Requirement	NASCIS II	NASCIS III
well designed	yes	yes*
well executed	yes	yes
compelling data	no	no
statistical significance w/ appropriate tests	no	no
appreciable impact on issues important to the patient	no	no
reproducible results	no	no

* Ignoring the absence of a proper control group.

unless further mortality statistics become available, the 48-hour MP therapy should be regarded as potentially harmful and possibly lethal. The findings of this critical analysis underscore the need for clinical investigators to recognize the limitations of their studies and for peer review to provide detailed, unbiased scrutiny. In cases in which clinical ramifications are potentially large, uninterested third-party analyses of entire datasets may be desirable.

Acknowledgment

The author thanks the American Association of Neurological Surgeons/Congress of Neurological Surgeons practice guidelines committee for their encouragement of this review.

Disclosure

Dr. Hurlbert sits on the American Association of Neurological Surgeons/Congress of Neurological Surgeons practice guidelines subcommittee for the pharmacological treatment of SCI. The views expressed in this paper do not represent the final consensus of the committee.

References

1. Altman DG: **Practical Statistics for Medical Research.** Chapman & Hall: London, 1995, pp 440–476
2. Bracken MB: Pharmacological interventions for acute spinal cord injury. **Cochrane Database of Systematic Reviews (Issue 1).** 1999
3. Bracken MB, Shepard MJ, Collins, WF, et al: A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury. Results of the Second National Acute Spinal Cord Injury Study. **N Engl J Med 322:**1405–1411, 1990
4. Bracken MB, Shepard MJ, Collins WF, et al: Methylprednisolone or naloxone treatment after acute spinal cord injury: 1-year follow-up data. Results of the second National Acute Spinal Cord Injury Study. **J Neurosurg 76:**23–31, 1992
5. Bracken MB, Shepard MJ, Holford TR, et al: Administration of methylprednisolone for 24 or 48 hours or tirilazad mesylate for 48 hours in the treatment of acute spinal cord injury. Results of the third National Acute Spinal Cord Injury Randomized Controlled Trial. National Acute Spinal Cord Injury Study. **JAMA 277:**1597–1604, 1997
6. Bracken MB, Shepard MJ, Holford TR, et al: Methylprednisolone or tirilazad mesylate administration after acute spinal cord injury: 1-year follow-up. Results of the third National Acute Spinal Cord Injury Randomized Controlled Trial. **J Neurosurg 89:**699–706, 1998
7. Chiles BW III, Cooper PR: Acute spinal injury. **N Engl J Med 334:**514–520, 1996
8. Ducker TB, Zeidman SM: Spinal cord injury. Role of steroid therapy. **Spine 19:**2281–2287, 1994
9. **Mosier Versus Southern California Physicians Insurance Exchange.** 63 California Appelate 4th 1022, 74 California Rptr. 2 d 550, 1998
10. Nesathurai S: Steroids and spinal cord injury: revisiting the NASCIS 2 and NASCIS 3 trials. **J Trauma 45:**1088–1093, 1998
11. Otani K, Abe H, Kadoya S, et al: Beneficial effect of methylprednisolone sodium succinate in the treatment of acute spinal cord injury. **Sekitsui Sekizui J 7:**633–647, 1994
12. Rosner B: **Fundamentals of Biostatistics.** Boston: Duxbury Press, 1986, pp 472–480
13. Walker MD: Acute spinal-cord injury. **N Engl J Med 324:** 1885–1887, 1991

Manuscript received October 14, 1999.
Accepted in final form March 6, 2000.

Address reprint requests to: R. John Hurlbert, M.D., University of Calgary Spine Program, Foothills Hospital and Medical Centre, 1403 - 29th St. N.W., Calgary, Alberta T2N 2T9, Canada. email: jhurlber@ucalgary.ca.