Clinical Impact of NXY-059 Demonstrated in the SAINT I Trial
Derivation of Number Needed to Treat for Benefit Over Entire Range of Functional Disability

Jeffrey L. Saver, MD

Background and Purpose—The SAINT I trial demonstrated that the neuroprotective agent NXY-059 improves the distribution of acute stroke patient outcomes on the modified Rankin Scale (mRS) of global disability. Standard dichotomized number-needed-to-treat (NNT) analyses for the magnitude of treatment benefit range widely, from 22.7 to infinity, and each capture only a portion of the observed beneficial shift in outcomes. Derivation of an NNT value reflecting the treatment’s benefit over the entire range of the mRS is required to describe the clinical import of the trial results.

Methods—The minimum and maximum possible NNTs for benefit over the range of mRS global disability outcomes were calculated by completing a joint distribution outcome table for a model population of 1000 patients, shifting responders by the greatest and smallest possible increments, respectively. The biologically most plausible NNT within this range was derived by having 10 neurologist and emergency-physician acute stroke–care experts independently specify the joint distribution of outcomes in model samples of 1000 patients assigned to placebo and active therapy.

Results—The minimum possible NNT for benefit incorporating all mRS state transitions is 7.9 and the maximum is 16.7. The biologically most plausible NNT for 1 additional patient to have a better outcome by 1 or more grades on the mRS outcomes is 9.8 (95% CI, 8.7 to 10.9).

Conclusions—Considering improvements in global disability over the entire outcome range, the SAINT I trial results indicate that for every 100 patients treated, 10 will benefit and none be harmed as a result of treatment. (Stroke. 2007; 38:1515-1518.)

Key Words: acute stroke ■ cerebrovascular disease ■ clinical trials ■ neuroprotective agents ■ stroke ■ therapy

Strokes produce a wide range of functional impairments, and treatments for acute stroke may improve patient outcomes anywhere over the entire continuum of functional disability. Because stroke outcomes are distributed over a functional spectrum, rather than binary, dichotomizing end points in stroke clinical trial analysis reduces outcome information and will often lead to reduced study power and underestimation of clinically important treatment effects.1–5 Leading investigative groups and regulatory agencies have called for stroke trials to analyze transitions in health states over the full distribution of potential outcomes.5–8

A challenge of trials that test whether a treatment produces a shift in outcomes over a functional range is translating the treatment effect into terms that are meaningful to patients and clinicians. As the first phase 3 stroke trial to use “shift analysis” as its primary prespecified end point–analysis technique and the first neuroprotective phase 3 trial to have a positive outcome requiring clinical translation, the Stroke Acute Ischemic NXY

Treatement (SAINT) I trial provides an important instance of this challenge.9

Expert generation of joint outcome tables is a recently described method permitting derivation of clinically useful number-needed-to-treat (NNT) estimates over the entire range of functional outcomes.4 The goal of this study was to derive from the SAINT I data clinically useful NNT estimates of the benefit of NXY-059 in acute stroke across the entire range of functional outcomes.

Methods

Ten neurologist and emergency-physician experts in acute stroke care independently specified the joint distribution of outcomes of a model sample of patients assigned to placebo and NXY-059 therapy, with group outcomes constrained to fit the observed outcome distributions in the SAINT I trial. The joint outcome table specification method described previously was followed,4 except that the model population consisted of 1000 patients rather than 100 patients to increase precision and permit capture of the smaller treatment effects observed with NXY-059 than tissue plasminogen activator.

Received and accepted December 22, 2006.
From the Stroke Center and Department of Neurology, David Geffen School of Medicine at the University of California, Los Angeles.
Correspondence to Jeffrey L. Saver, MD, University of California, Los Angeles (UCLA) Stroke Center, 710 Westwood Plaza, Los Angeles, CA 90095.
E-mail jsaver@ucla.edu

© 2007 American Heart Association, Inc.

Stroke is available at http://www.strokeaha.org DOI: 10.1161/01.STR.0000263135.47779.9e

1515
Each expert was presented with an Excel spreadsheet. The spreadsheet included: (1) definitions of each mRS outcome category, (2) the distribution of mRS outcomes in the placebo and NXY-059 treatment groups in the SAINT I study, and (3) the rates of symptomatic intracerebral hemorrhage in the placebo and NXY-059 treatment groups in the SAINT I study. As in the primary prespecified SAINT I trial analysis, mRS categories 5 (severe disability) and 6 (death) were collapsed into a single worst outcome category, yielding a 6 level mRS scale. In the center of the spreadsheet was a joint-distribution table of outcomes, initially with all 1000 model patients arrayed in placebo distribution cells. The panel member redistributed patients iteratively to complete the joint-distribution table, under the instruction to specify the joint distribution most likely to occur among a typical group of 1000 patients who are treated with NXY-059 and who match the SAINT I study population. The expert first filled cells via improved outcomes to achieve the observed SAINT I study distribution, then filled cells to capture worsened outcomes if they felt any occurred, and then added to cells via improved outcomes to reach the target observed SAINT I study distribution. The expert was then asked to globally re-examine all cells and readjust the joint distribution, if needed, to achieve maximum biological plausibility, constrained by maintaining the observed trial group outcomes (the marginal distributions). Each expert’s estimate of the proportion of patients per 1000 experiencing benefit or harm from NXY-059 treatment compared with placebo across the entire mRS was calculated by adding all off-diagonal cells in the appropriate direction from the expert-specified joint-distribution table.

NNT and number-needed-to-harm (NNH) values were obtained from each of the 10 experts. The geometric mean and corresponding 95% CI for NNT and NNH across the 10 experts were calculated, using the sample standard deviation. In addition to the expert judgments, the minimum and maximum possible NNTs consistent with the SAINT I results were also obtained. For both derivations, it was assumed that no patients were harmed by treatment with NXY-059, which coincided with the experts’ judgment. To calculate the minimum NNT, the joint distribution table was populated by following the rule that all patients who improve as a result of therapy improve by only one step on the mRS. To calculate the maximum NNT, the joint-distribution table was populated by following the rule that each patient who improves does so by the maximum amount permitted by the residual distribution resulting from the previous patient move. This rule was applied first to patients in the worst mRS outcome category (mRS 5), then next worse, and continued through ordered levels of lessened severity (supplemental Figure I, available online at http://stroke.ahajournals.org). NNT values were also calculated for each dichotomized breakpoint of the mRS in standard fashion by taking the inverse of the absolute risk difference. For dichotomized NNT calculations, it was assumed that no patients were harmed by treatment with NXY-059, which coincided with the experts’ judgment.

Results

The distribution of mRS outcomes in the placebo and treatment groups of the SAINT I trial is shown in the Figure. The mean (SD) mRS scores in the placebo group was 2.84 (1.75) and in the NXY-059 group 2.71 (1.81). The mean (SEM) difference in the mRS score was 0.13 (0.09).

Discussion

The Table shows NNTs for (1) all dichotomized breakpoints of the mRS, (2) trial specified minimum and maximum possible NNTs for benefit by 1 or more steps over the entire range of the 6-level mRS, and (3) the expert-derived biologically most plausible NNT for benefit by 1 or more steps over the entire range of the 6-level mRS. No expert judged that NXY-059 produced harm, so NNH estimates were not calculated.

### NNT Values for Different 6-Level Rankin Transitions

<table>
<thead>
<tr>
<th>Rankin Transition</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dichotomized</strong></td>
<td></td>
</tr>
<tr>
<td>0 vs 1–5</td>
<td>22.7</td>
</tr>
<tr>
<td>0–1 vs 2–5</td>
<td>41.7</td>
</tr>
<tr>
<td>0–2 vs 3–5</td>
<td>47.6</td>
</tr>
<tr>
<td>0–3 vs 4–5</td>
<td>27.8</td>
</tr>
<tr>
<td>0–4 vs 5</td>
<td>Infinity</td>
</tr>
<tr>
<td><strong>Across all transitions</strong></td>
<td></td>
</tr>
<tr>
<td>Minimum possible</td>
<td>7.9</td>
</tr>
<tr>
<td>Maximum possible</td>
<td>16.7</td>
</tr>
<tr>
<td>Biologically most plausible (mean, 95% CI)</td>
<td>9.8 (8.7–10.9)</td>
</tr>
</tbody>
</table>

The Table shows NNTs for (1) all dichotomized breakpoints of the mRS, (2) trial specified minimum and maximum possible NNTs for benefit by 1 or more steps over the entire range of the 6-level mRS, and (3) the expert-derived biologically most plausible NNT for benefit by 1 or more steps over the entire range of the 6-level mRS. No expert judged that NXY-059 produced harm, so NNH estimates were not calculated.

The SAINT I trial showed that NXY-059 changes the distribution of outcomes on the mRS of global disability in a statistically significant manner. However, small shifts in outcome can be statistically significant without being clinically important. Consequently, it is critical to translate clinical trial results into terms that are meaningful to patients and clinicians.

Standard approaches to assessing the clinical relevance of the treatment effect observed in the SAINT I trial include stating the mean outcome difference and providing dichotomized NNTs. Each of these approaches provides only an incomplete index of treatment impact. It is difficult for clinicians and patients to intuit the clinical worth of the mean improvement of 0.13 on the Rankin Scale that treatment with NXY-059 on average yields. Dichotomizing the Rankin Scale permits conventional derivation of NNT estimates, but these range widely depending on the breakpoint used (22.7 to infinity). Proponents of a therapy may highlight only the most favorable NNTs, as in the SAINT I primary report, whereas skeptics may highlight the least favorable. More importantly, neither of these NNTs reflects the treatment effect of most relevance to decision-makers, which is benefit over the full range of outcomes rather than just at a single, arbitrarily selected, health state transition.

The joint distribution outcome table completion method permits derivation of NNTs for the SAINT I trial that parallel the “shift analysis” design of the trial, capture a broad range of health state transitions of interest to the patient, and are easily understandable by clinicians, patients, family members, and payors. Because each level of the mRS is better than the next lower level in a clinically meaningful way (once levels 5 and 6 are collapsed), improvements by 1 or more levels are widely recognized as clinically desirable.
When a treatment does not produce harmful effects, the minimum and maximum possible values for the NNT for benefit over all levels of the mRS are fully specified by the trial data. Assuming that every patient who benefits from treatment improves by only 1 mRS category allows the lowest possible NNT compatible with the data to be derived. This approach was taken by the SAINT I investigators and yields a NNT just under 8.9. However, because assuming that all patients who benefit do so by only 1 mRS level and none by 2 or more is biologically implausible for many treatments, this approach is best viewed as providing a lower bound for the NNT rather than the likely true NNT. Assuming that all patients who benefit from treatment improve by the maximum amount possible permits rule-based specification of the joint outcome table and derivation of the upper bound for the range of possible NNTs compatible with observed data. The SAINT trial results indicate that the NNT for 1 patient to benefit by 1 or more steps on the mRS must fall between 7.9 to 16.7, indicating that all the dichotomized NNTs substantially underestimate the impact of this therapy.

Within these extreme patterns of possible response (many patients benefit each by a little bit versus few patients benefit each by a lot) lies the actual pattern (typically that specific proportions of patients benefit by minimal, moderate, and maximal degrees). The expert-derived estimate is that the biologically most-plausible NNT consistent with the SAINT I data is 9.8. Accordingly, patients and clinicians can approach treatment decisions with the knowledge that SAINT I data is 9.8. The experts independently concluded from the SAINT I trial results that for every 10 patients treated with NXY-059, one will have their functional disability outcome improved in a clinically important manner. Although subjective, the expert approach is not “arbitrary”. The experts drew on their knowledge of disease mechanism and natural history, drug mechanism, preclinical data, and clinical experience in making the multiple “shift” judgments needed to populate the joint outcome table. Most experts, for example, projected greater individual benefits for responders to the recanalization agent tissue plasminogen activator (tPA) than for responders to the neuroprotective agent NXY-059.

For the expert-derived NNT, no assumptions regarding the NNH value for NXY-059 were made. Every one of the 10 experts independently concluded from the SAINT I trial results that the NNH was 0. The experts arrived at this conclusion from review of preclinical and actual SAINT I trial data. SAINT I indicated that NXY-059 is a very safe treatment for stroke.4 Neuroprotective therapies in general are likely to yield lesser degrees of benefit than recanalization treatments. In the case of NXY-059, the shifts in mRS outcome distribution in SAINT I are substantially less than those observed in the National Institute of Neurological Disorders and Stroke (NINDS) trials I and II. Accordingly, the expert-derived NNT over the 6-level range of mRS outcomes is higher for NXY-059 than for tPA, 9.8 versus 3.3.

For the end point of global disability levels, according to the NINDS tPA trials I and II and the SAINT I trial, for every 100 patients treated with IV tPA, 32 will benefit and 3 will be harmed; for every 100 patients treated with NXY-059, 10 will benefit and none will be harmed in clinically important ways. These contrasting profiles of treatment impact should inform individual treatment decisions and stroke healthcare policies. Although the current best estimate of the NNT for NXY is about 10, this may well change in the light of the results of SAINT II and the planned combined analysis of the 2 trials which will yield a more robust and precise estimate of effect than SAINT I alone.

Appendix

Subsequent to the acceptance of this article, the SAINT II trialists announced that SAINT II failed to confirm the findings of SAINT I. The SAINT II results do not alter the validity of the methods used in the current article to derive NNT values from shift analysis trials, nor of the current article’s derivation of NNHs that index the strength of the treatment effect shown in SAINT I. However, because SAINT II failed to confirm the treatment effect shown in SAINT I, the true clinical impact of NXY-059, if any, is likely to be of lesser magnitude than indicated by the SAINT I trial results alone and by NNHs derived only from SAINT I.

Acknowledgments

The author thanks the members of the expert panel: Greg Albers, MD; Stanley Cohen, MD; James Grotta, MD; Steven Levine, MD;
David Liebeskind, MD; Helmi Lutsep, MD; Phil Scott, MD; Sidney Starkman, MD; Janet Wilterdink, MD.

Sources of Funding
This work was supported in part by NIH-NINDS Awards NIH-NINDS U01 NS 44364 and NIH-NINDS P50 NS044378.

Disclosures
J.L.S. was a site subinvestigator in the SAINT II trial.

References
Clinical Impact of NXY-059 Demonstrated in the SAINT I Trial: Derivation of Number Needed to Treat for Benefit Over Entire Range of Functional Disability

Jeffrey L. Saver

Stroke. 2007;38:1515-1518; originally published online March 29, 2007; doi: 10.1161/01.STR.0000263135.47779.9e

The online version of this article, along with updated information and services, is located on the World Wide Web at: http://stroke.ahajournals.org/content/38/5/1515

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at: http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at: http://stroke.ahajournals.org/subscriptions/