Number Needed to Treat to Benefit and to Harm for Intravenous Tissue Plasminogen Activator Therapy in the 3- to 4.5-Hour Window

Joint Outcome Table Analysis of the ECASS 3 Trial

Jeffrey L. Saver, MD, FAAN, FAHA; Jeffrey Gornbein, PhD; James Grotta, MD; David Liebeskind, MD; Helmi Lutsep, MD; Lee Schwamm, MD; Phillip Scott, MD; Sidney Starkman, MD

Background and Purpose—Measures of a therapy’s effect size are important guides to clinicians, patients, and policy-makers on treatment decisions in clinical practice. The ECASS 3 trial demonstrated a statistically significant benefit of intravenous tissue plasminogen activator for acute cerebral ischemia in the 3- to 4.5-hour window, but an effect size estimate incorporating benefit and harm across all levels of poststroke disability has not previously been derived.

Methods—Joint outcome table specification was used to derive number needed to treat to benefit (NNTB) and number needed to treat to harm (NNTH) values summarizing treatment impact over the entire outcome range on the modified Rankin scale of global disability, including both expert-dependent and expert-independent (algorithmic and repeated random sampling) array generation.

Results—For the full 7-category modified Rankin scale, algorithmic analysis demonstrated that the NNTB for 1 additional patient to have a better outcome by ≥1 grades than with placebo must lie between 4.0 and 13.0. In bootstrap simulations, the mean NNTB was 7.1. Expert joint outcome table analyses indicated that the NNTB for improved final outcome was 6.1 (95% CI, 5.6–6.7) and the NNTH 37.5 (95% CI, 34.6–40.5). Benefit per 100 patients treated was 16.3 and harm per 100 was 2.7. The likelihood of help to harm ratio was 6.0.

Conclusions—Treatment with tissue plasminogen activator in the 3- to 4.5-hour window confers benefit on approximately half as many patients as treatment 3 hours, with no increase in the conferral of harm. Approximately 1 in 6 patients has a better and 1 in 35 has a worse outcome as a result of therapy. (Stroke. 2009;40:2433-2437.)

Key Words: cerebral infarction ■ clinical trial ■ fibrinolysis ■ ischemic stroke ■ thrombolysis

In human cerebral ischemia, the infarct core progressively expands over minutes to hours. In a typical ischemic stroke, nearly 2 million nerve cells are lost in each minute in which reperfusion is not established. Consequently, the benefits conferred by intravenous thrombolytic therapy for acute cerebral ischemia are highly time-dependent. The 2 NINDS tissue plasminogen activator (TPA) trials and subgroups of other trials demonstrated that treatment with intravenous TPA confers substantial net benefit when administered between 1 and 3 hours after onset. Although the 0- to 1-hour window has not directly been investigated (only 2 of the 624 patients in the NINDS trials were treated within 1 hour of onset), an even greater benefit is presumed in hyper-early patients. Recently, the ECASS 3 trial prospectively confirmed a lesser, but still statistically significant, benefit for intravenous TPA in the 3- to 4.5-hour window, a benefit previously suggested by pooled analysis of 3- to 4.5-hour window patients from 4 previous trials.

For adoption in routine care, the results of a clinical trial must be not only statistically significant but also clinically significant. As the clinical maxim observes, “a difference to be a difference must make a difference.” In deciding how best to apply the findings of a positive clinical trial, physicians, patients, family, and payors must consider the magnitude of the treatment effect demonstrated in the ECASS 3 trial.

The number needed to treat (NNT) is a statistically valid and clinically useful indicator of treatment effect magnitude. In the primary trial paper, the ECASS 3 investigators reported...
Materials and Methods

Derivation of number needed to treat values across an ordinal outcome range requires knowledge of final group distributions and the within patient variance. In the primary ECASS 3 report, the group distributions on all 7 levels of the mRS of global disability was reported and treatment with alteplase was associated with a statistically significant (P=0.02) beneficial shift in outcomes across the outcome range. Parallel group clinical trials do not provide data regarding the within-patient variance. Accordingly, in this study, equivalent information was derived using 3 methods of joint outcome table completion: (1) expert specification; (2) minimum and maximum algorithms; and (3) random sampling from all permitted joint distributions.

The expert specification followed methods previously described. Briefly, 7 emergency physician and neurologist clinicians highly experienced in acute stroke care, and with no significant TPA-related competing interests, independently populated a joint outcome table for a model population of 1000 patients matching the ECASS 3 cohort. Each expert produced an array under the injunction that individual patient responses be those that are biologically most plausible within the constraint that group outcomes match the observed trial results.

The minimum and maximum algorithm specifications were also conducted according to previously described methods. The distribution of harm conferred by therapy was set at the median estimate of the expert panel. The maximum possible NNTB compatible with the data were then derived by completing the joint outcome table following the rule that every patient who benefits from therapy does so by improving only the minimum possible number of levels compatible with the final trial group outcome distributions. The minimum possible NNTB compatible with the data were derived by completing the table following the rule that every patient who benefits from therapy does so by improving the maximum possible number of grades compatible with the group data. The central value within the possible range was obtained by calculating the geometric mean of the minimum and maximum NNTB.

The automated random sampling technique followed previously described principles. Within the 7×7 joint distribution matrix with fixed marginal frequencies adding to 1000, there are 26 unknown frequencies subject to 12 constraints (6 columns, 6 rows). The vector “x” is a 26-dimension vector that satisfies the 12 constraints. This is a linear inequality system of the form Ex ≤ f, where x ≥ 0 and the 12×26 E matrix and 12×1 f vector are obtained from the constraints.

Table 1. Number Needed to Treat Values for ECASS 3 Trial

<table>
<thead>
<tr>
<th>Category</th>
<th>NNTB</th>
<th>NNTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>For transitions across all 7 levels of the mRS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologically most plausible</td>
<td>6.1</td>
<td>37.5</td>
</tr>
<tr>
<td>Lowest possible</td>
<td>4.0</td>
<td>37.5</td>
</tr>
<tr>
<td>Highest possible</td>
<td>13.0</td>
<td>37.5</td>
</tr>
</tbody>
</table>

Net NNTB

For individual dichotomizations of the mRS

<table>
<thead>
<tr>
<th>Category</th>
<th>NNTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 vs 1–6</td>
<td>17.5</td>
</tr>
<tr>
<td>0–1 vs 2–6</td>
<td>13.7</td>
</tr>
<tr>
<td>0–2 vs 3–6</td>
<td>20</td>
</tr>
<tr>
<td>0–3 vs 4–6</td>
<td>34.5</td>
</tr>
<tr>
<td>0–4 vs 5–6</td>
<td>−66.7</td>
</tr>
<tr>
<td>0–5 vs 6</td>
<td>71.4</td>
</tr>
</tbody>
</table>

Because computing all possible values of x is not practical, we used the R function “xsample” in the R program library “limSolve” to randomly sample values of the vector “x” from all possible joint outcome table solutions. Using the mirror algorithm, 3000 samples were taken from the large population of all possible solutions, without replacement. The mean and range NNTB values of these random samples from all possible NNTB values under the constraints were analyzed.

In the initial analysis, NNT for change by ≥1 grades across all 7 levels of the mRS were derived. However, nearly half of individuals at risk for stroke do not consider a severely disabled outcome such as mRS rank 5 (severe disability, bedridden, incontinent, and requiring constant nursing care and attention) to be a better outcome than an mRS of 6 (dead). Accordingly, NNT were also derived from the algorithmic and expert joint outcome tables across 6 levels of the mRS, collapsing mRS levels 5 and 6 into 1 worst outcome category.

Benefit per 100 patients treated was derived in standard fashion by the formula, benefit per 100 patients treated = 100/NNTB, and harm per 100 patients treated can be derived similarly by the formula, harm per 100 patients treated = 100/NNTH. Likelihood of help to harm ratios were derived in standard fashion by the formula, likelihood of help to harm ratio = benefit per 100 patients treated/ harm per 100 patients treated.17

Results

In the ECASS 3 trial, the mean mRS score in the TPA group was 1.99 (SD, 1.9) and in the placebo group 2.20 (SD, 1.9), yielding a mean difference of 0.21. Net number needed to treat values for each of the 6 possible dichotomizations of the mRS are shown in Table 1, and range from −66.7 to 13.7. For the full 7-category mRS, algorithmic analysis found the lowest possible NNTB consistent with the group outcomes of the ECASS 3 trial for 1 additional patient to have a better outcome by ≥1 grades with TPA than with placebo was 4.0 and the highest possible 13.0. The geometric mean of this range was an NNTB of 7.2. The independent expert joint outcome table analyses indicated that the biologically most plausible NNTB for improved final outcome on the mRS was 6.1 (95% CI, 5.6–6.7; range, 5.7–7.4) and the NNTH 37.5 (95% CI, 34.6–40.5; range, 34.0–41.7). The resulting expert likelihood of help to harm ratio was 6.0. The results of the
bootstrap simulations are shown in the Figure. The mean NNTB was 7.08 (SD, 0.90).

For the 6-category mRS, algorithmic joint outcome table analysis of the group outcomes of the ECASS 3 trial indicated that the lowest possible for 1 additional patient to have a better outcome by ≥1 grades with TPA than with placebo was 4.2 and the highest possible 16.1. The independent expert joint outcome table analyses indicated that the biologically most plausible NNTB for improved final outcome on the mRS was 6.9 (95% CI, 6.3–7.5; range, 6.3–8.3) and the NNTH 39.9 (95% CI, 36.9–43.1; range, 34.5–45.5). The resulting expert likelihood of help to harm ratio was 5.8.

Table 2 compares the benefit and harm per 100 patients treated values for expert joint outcome table and dichotomized analyses of the ECASS 3 trial (3- to 4.5-hour window) and the 2 trials constituting the NINDS TPA Study (1- to 3-hour window). Across all 7 levels of the mRS, in both time windows, ≈3 per 100 patients are harmed by therapy, whereas in the 1- to 3-hour window ≈2 per 100 patients benefit and in the 3- to 4.5-hour window 16 per 100 patients benefit.

### Discussion

Patients and physicians consider as desirable several health state transitions across the spectrum of poststroke disability. Well-informed treatment decisions will take all valued transitions into account. Using previously described expert-based and expert-independent methods, we derived NNTB and NNTH values for outcome shifts over the entire range of the mRS when patients with acute cerebral ischemia are treated with intravenous TPA in the 3- to 4.5-hour time window. The biologically most likely NNTB of 6.1 and NNTH of 37.5 indicate that, for every 100 patients treated in the 3- to 4.5-hour window, 16.4 will have a better and 2.7 a worse outcome by ≥1 levels on the mRS of global disability. Clinicians can use these values in the acute setting to inform

patients and family members of the benefits and risks associated with intravenous TPA administered >3 and <4.5 hours after onset.

It is instructive to contrast the magnitude of the treatment effect of IV TPA in the 3- to 4.5-hour window with that in the 1- to 3-hour window. In biologically most plausible analyses, only half as many patients in the later time window benefit as in the earlier, whereas the risk of harm is approximately the same. These results emphasize the fundamental importance of treating patients quickly and continuous quality improvement to reduce time to therapy at every level of regional stroke systems of care. Nonetheless, the benefit to risk ratio in the 3- to 4.5-hour window of 6.0 suggests that, for those patients who have unavoidably missed the 3-hour window but are treatable within 4.5 hours, therapy will be worth pursuing.

The expert-independent methods used in this study provide support for the expert-based best estimates. The algorithmic technique demonstrates that the NNTB over the entire outcome range must lie between 4.0 and 13.0, which are all values lower than any of the dichotomized NNTB. The geometric mean within this possible range is 7.2, fairly close to the expert derived value of 6.1, lending it plausibility. This plausibility is further enhanced by the results of the sampling simulations, which similarly yielded a mean NNTB of 7.1.

The small difference between the expert-based and the expert-independent NNTB values provides insight into the experts’ views of the likely pattern of treatment effects yielded by TPA therapy in the 3- to 4.5-hour window. In parallel design clinical trials, the group outcome distribution data circumscribe a total amount of benefit of a treatment at a population level that can be allocated variously to individual patients. When individual patients benefit a lot, fewer
patients benefit at all, and the NNTB is high. When individual patients benefit a little, more patients benefit at all, and the NNTB is low. The bootstrap simulation (and the geometric mean of the algorithmic-determined potential range) project that the degree of benefit experienced by individual patients is the average of the possible range. In contrast, experts project the degree of benefit experienced by individual patients based on physiological knowledge and clinical experience. Their projections of individual benefit may be higher or lower than the average. In the current study of ECASS 3, that the expert-specified NNTB is smaller than the mean of all possible NNTB determined by bootstrap simulations indicates that the experts project that individual patients who respond to TPA in a later time window have their outcomes improved only to a restricted degree. As a result, the total group benefit indicated by the trial outcome distribution is allocated to numerous, rather than only a few, individual patients, lowering the expert-derived NNTB estimate to below the expert-independent NNTB estimates.

The number needed to harm values were similar in the 1- to 3-hour and 3- to 4.5-hour window, reflecting the similar rates of symptomatic intracerebral hemorrhage (SICH) in the 2 NINDS TPA trials and in the ECASS 3 trial. Any minor or major symptomatic hemorrhage (NINDS protocol definition of SICH) occurred 5.8% more often among TPA than placebo patients in the 1- to 3-hour NINDS trials, vs 4.4% more often among TPA than placebo patients in ECASS 3, supporting the expert assigned slightly lower harm per 100 treated values in the 3- to 4.5-hour window, and likely reflecting the milder strokes at entry in ECASS 3. For worsening of final outcome, the NNTH values derived by the experts more closely matched the difference in the rate of major SICH (2.2%, ECASS 3 protocol definition of SICH) than the difference in the rate of any SICH, in accord with prognostic models indicating that patients experiencing minor SICH generally do not have their final outcomes altered as a result. Comparison of the results for intravenous fibrinolysis with IV TPA in the 3- to 4.5-hour window with those for intra-arterial fibrinolysis in the overlapping 3- to 6-hour window is also instructive. Based on the PROACT 2 trial results, across the entire mRS, for every 100 patients treated with intra-arterial lysis in the 3- to 6-hour window, 20.8 benefit and 3.5 are harmed (likelihood of help to harm, 6.0), vs 16.4 and 2.7 for IV TPA in the 3- to 4.5-hour window. Differences in trial populations between the PROACT 2 trial and the ECASS 3 trial include important prognostic features such as age, pretreatment stroke severity, and confirmed presence of middle cerebral artery occlusion, making direct comparisons tentative. Nonetheless, the concordant findings in these 2 trials of a reduced frequency of benefit in the >3-hour window, compared with treatment in the first 3 hours, supports a biological constraint on the degree of benefit that can be achieved with recanalization therapy once several hours have passed since onset.

This study has limitations. The NNTB and NNTH values are derived from a randomized trial and do not necessarily translate to an observational setting unless only patients are considered that would have participated in ECASS 3. The analysis treats only transitions in health state from 1 mRS level to the next as valuable, and does not give value to improvements or worsening in health that occur entirely within the bounds of a scale level, which may lead to underestimation of the full impact of treatment. The expert-independent analyses are independent with regard to derivation of NNTB, but not NNTH. Completely expert-independent models cannot be used because they are vulnerable to random person statistical effects that lead to overestimation of degree of both benefit and harm.

The positive results of the ECASS 3 trial are a major advance in stroke science that now require translation into everyday clinical practice, including hospital treatment policies and consent processes. Joint outcome table analysis indicates that among individuals matching the ECASS 3 cohort, as a result of treatment with TPA in the 3- to 4.5-hour window, in 6 patients has a better and 1 in 35 a worse disability final outcome. Clinicians, patients, family members, and policy-makers should consider these effect size indices when making consent and policy decisions about lytic therapy in later time windows.

**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. has served as a scientific consultant regarding trial design and conduct to CoAxia, Concentric Medical, Talecris, and Cygnis (all modest); received lecture honoraria from Ferrer and Boehringer Ingelheim (modest); received devices for use in an NIH multicenter clinical trial from Concentric Medical (modest); was an unpaid investigator in a multicenter prevention trial sponsored by Boehringer Ingelheim; has declined consulting/honoraria monies from Genentech since 2002; and is an employee of the University of California, which holds a patent on retriever devices for stroke. J. Gornbein is an employee of the University of California, which holds a patent on retriever devices for stroke. J. Grotta has served as a scientific consultant regarding trial design and conduct to Lundbeck (modest), and has patent on caffeine in stroke licensed by InnerCool (significant). D.L. has served as a scientific consultant regarding trial design and conduct to CoAxia, Concentric Medical, Brainsgate (all modest); was an unpaid investigator in a multicenter prevention trial sponsored by Boehringer Ingelheim; and is an employee of the University of California, which holds a patent on retriever devices for stroke. H.L. has served as a scientific consultant regarding trial design and conduct to CoAxia, Concentric Medical, and Talecris (all modest), and has declined consulting/honoraria monies from Genentech since 2002; and is an employee of the University of California, which holds a patent on retriever devices for stroke. L.S. has served as a scientific consultant regarding trial design and conduct to CoAxia, Concentric Medical, and Talecris (all modest), and is on prevention Speaker’s Bureau for Boehringer Ingelheim (modest). L.S. has served as a scientific consultant regarding trial design and conduct to CoAxia, Concentric Medical, Phressia, RTI Health Solutions, and CryoCath (all modest), and was a site investigator in a multicenter trial run by Forest for which the MGH Neurology Service received payments based on the clinical trial contracts for the number of subjects enrolled (modest). S.S. has served as a scientific consultant regarding trial design and conduct to Astra Zeneca; was a site investigator in a multicenter trial run by Vernalis for which the UC Regents received payments based on the clinical trial contracts for the number of subjects enrolled; received devices or study agent for use in NIH multicenter clinical trials from Concentric Medical, Genentech, Eks Medical (all modest); and is an employee of the University of California, which holds a patent on retriever devices for stroke.
References

Number Needed to Treat to Benefit and to Harm for Intravenous Tissue Plasminogen Activator Therapy in the 3- to 4.5-Hour Window: Joint Outcome Table Analysis of the ECASS 3 Trial

Jeffrey L. Saver, Jeffrey Gornbein, James Grotta, David Liebeskind, Helmi Lutsep, Lee Schwamm, Phillip Scott and Sidney Starkman

Stroke. 2009;40:2433-2437; originally published online June 4, 2009;
doi: 10.1161/STROKEAHA.108.543561

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

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/40/7/2433

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/