

Can We Improve the Statistical Analysis of Stroke Trials? Statistical Reanalysis of Functional Outcomes in Stroke Trials

The Optimising Analysis of Stroke Trials (OAST) Collaboration

Background and Purpose—Most large acute stroke trials have been neutral. Functional outcome is usually analyzed using a yes or no answer, eg, death or dependency versus independence. We assessed which statistical approaches are most efficient in analyzing outcomes from stroke trials.

Methods—Individual patient data from acute, rehabilitation and stroke unit trials studying the effects of interventions which alter functional outcome were assessed. Outcomes included modified Rankin Scale, Barthel Index, and “3 questions”. Data were analyzed using a variety of approaches which compare 2 treatment groups. The results for each statistical test for each trial were then compared.

Results—Data from 55 datasets were obtained (47 trials, 54 173 patients). The test results differed substantially so that approaches which use the ordered nature of functional outcome data (ordinal logistic regression, *t* test, robust ranks test, bootstrapping the difference in mean rank) were more efficient statistically than those which collapse the data into 2 groups (χ^2 ; ANOVA, $P < 0.001$). The findings were consistent across different types and sizes of trial and for the different measures of functional outcome.

Conclusions—When analyzing functional outcome from stroke trials, statistical tests which use the original ordered data are more efficient and more likely to yield reliable results. Suitable approaches included ordinal logistic regression, *t* test, and robust ranks test. (*Stroke*. 2007;38:1911-1915.)

Key Words: stroke ■ randomised controlled trial ■ statistical analysis

The management of patients with acute or recent stroke has benefited significantly from the results of randomized controlled trials and meta-analyses of these. For example, functional outcome is improved with alteplase, aspirin, management in a Stroke Unit, and community occupational therapy.¹⁻⁷ In contrast, some studies were overtly negative finding that treatment worsened outcome, eg, DCLHb, enlimomab, selfotel, or tirilazad.⁸⁻¹¹ However, the majority of acute stroke trials were neutral in spite of positive preclinical findings. The failure of these latter studies can be attributed to multiple causes, including the relevance of laboratory findings to clinical stroke,¹² inadequate sample size,¹³ choice of primary outcome, and its statistical analysis.

Measures of functional outcome such as the modified Rankin Scale (mRS),¹⁴ Barthel Index (BI)¹⁵ and “3-questions”¹⁶ are ordinal in nature: that is, they consist of ≥ 3 categories which have a natural ordering, eg, the mRS has 7 categories ranging from no symptoms to dead. It might then be expected that statistical analysis would preserve and use the data in this ordinal form. However, most published trials have used a “yes/no” (dichotomized) analysis of functional outcome, eg, combining categories within the mRS into 2 groups, such as “dead or dependent” (eg, mRS 3 to 6) and

“independent” (mRS 0 to 2), and then comparing these between the treatment groups. Unfortunately, there is little agreement where mRS data should be divided (ie, 0,1 versus 2 to 6,¹ 0 to 2 versus 3 to 6,¹⁷ or 0 to 3 versus 4 to 6,¹⁸) and whether this matters.¹⁹ Furthermore, collapsing data in this way generally lowers statistical power and therefore reduces the chance of finding a significant treatment effect because information from many subjects are ignored. For example, patients responding to treatment and achieving a mRS of 3 rather than 4 or 0 rather than 1 are not detected in a analysis comparing mRS 0 to 2 with 3 to 6.

Inadequacies in the statistical analysis of trials in acute stroke are apparent in 2 examples. First, the ECASS II trial of alteplase showed no treatment effect for its primary outcome (when comparing mRS 0,1 with mRS 2 to 6) but was positive when reanalyzed using the data collapsed in a different place (mRS 0 to 2 versus 3 to 6)²⁰ or when analyzed using a “bootstrapping” technique (Figure 1).²¹ Second, 5 trials of tirilazad individually showed no treatment effect when analyzed using dichotomous outcomes²²⁻²⁴ although a meta-analysis found that the intervention was associated with a worse outcome²⁵; post hoc analysis then suggested that one of these trials was negative²⁴ (not neutral) when analyzed using a

Received September 27, 2006; final revision received November 2, 2006; accepted November 21, 2006.

From the Division of Stroke Medicine (P.M.W.B., L.J.G.), University of Nottingham, Nottingham, UK; and the Medical Statistics Unit (T.C., S.P.), London School of Hygiene and Tropical Medicine, London, UK.

Correspondence to Prof Philip M.W. Bath, Division of Stroke Medicine, University of Nottingham, D Floor, South Block, Queen's Medical Centre, Nottingham, NG7 2UH UK. E-mail philip.bath@nottingham.ac.uk

© 2007 American Heart Association, Inc.

Stroke is available at <http://www.strokeaha.org>

DOI: 10.1161/STROKEAHA.106.474080

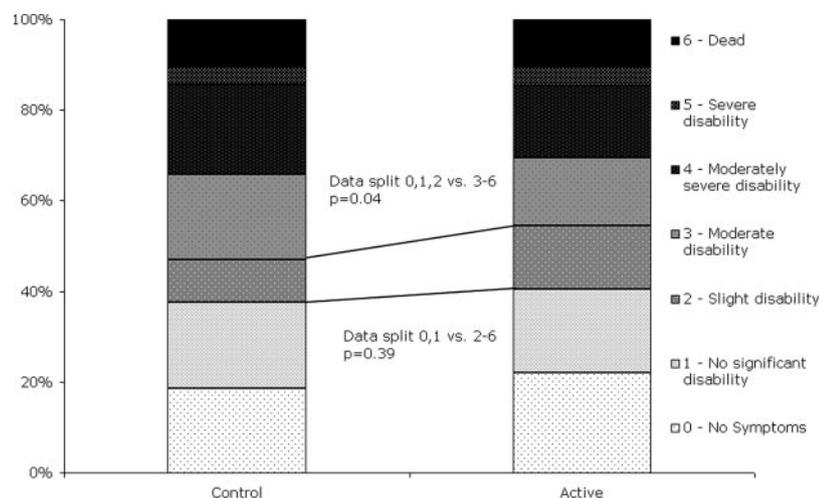


Figure 1. Distribution in mRS in the ECASS II trial showing the primary and post hoc analyses.²⁰

method which preserved the original ordered data (P.B., unpublished data, 2004).

We aimed to identify which statistical methods might optimize the analysis of data from functional outcome scales in stroke trials.

Methods

Identification of Trials

We sought individual patient data from randomized controlled trials assessing functional outcome after stroke for interventions which were either positive or negative according to the trial publication, or were included in a meta-analysis showing benefit or harm; neutral trials in a neutral meta-analysis were excluded. Published studies (full article or abstract) fulfilling these criteria were identified from electronic searches of the Cochrane Library (to end of 2005). In each case, we invited the chief investigator to join the collaboration and share their data. In some cases where individual data could not be obtained it was possible to extract it from the original publication.

Trial Data

Demographic (age, gender), trial (setting, intervention, length of follow-up, result), patient severity, and functional outcome (BI, mRS, "3 question" scale [3Q, a derivative of mRS], or another measure) data were collected for each trial. In factorial trials or those having >2 treatment groups, data were analyzed for each comparison of active therapy versus control. Where outcome data were scored at several time points (eg, 1, 3 and 6 months) the time point used for the primary outcome was included.

Statistical Tests

We compared different statistical tests for assessing treatment effect. Some of these required the data to be collapsed into groups (such as the χ^2 test), whereas others used the original ordinal data (such as Wilcoxon test and *t* test). Statistical tests which dichotomized ("yes/no") data were assessed multiple times collapsing the data in different places, eg, mRS 0,1 versus 2 to 6, 0 to 2 versus 3 to 6 and 0 to 5 versus 6. A description of the statistical tests used is given in the supplemental Appendix I, available online at <http://stroke.ahajournals.org>.

Comparison of Statistical Tests

Each data set was analyzed using each statistical test. These results were then ordered within each trial and given a rank, with the lowest rank given to the test which produced the most significant result, ie, the largest *z* score, within that trial. A 2-way analysis of variance test was then used to see on average which statistical test had produced the lowest ranks. We were then able to order the statistical tests in

terms of their efficiency in identifying treatment effects. We also assessed how many statistically significant (at 5%) results each test found.

To assess the validity and reliability of the results, a number of supplementary analyses were carried out. First, the comparison of statistical tests was repeated within subgroups of trials sharing similar characteristics; second, the statistical assumptions of the tests were assessed; and last, the sensitivity of the tests was explored to make sure treatment effects were only detected when they truly existed (the type I error rate). Technical details of these supplementary analyses can be found in the supplemental Appendix II, available online at <http://stroke.ahajournals.org>.

Analyses were carried out in SAS (version 8.2) and Stata (version 7) and significance was taken at $P < 0.05$.

Results

Trials Characteristics

A total of 55 comparisons of active versus control treatment (54 173 patients) were included, these comprising individual patient data from 38 trials and summary data extracted from the publications of a further 9 studies; 6 trials had 2 active treatment groups, and 1 had 3 active groups so a further 8 comparisons were available (Figure 2). The data related to 34 acute stroke trials, 7 trials of rehabilitation (1164 patients) and 6 trials of stroke units (1399 patients). BI was used to measure functional outcome in 22 trials, 18 used the mRS, 3 used the 3Q scale, 1 used the Rivermead scale, 2 related trials

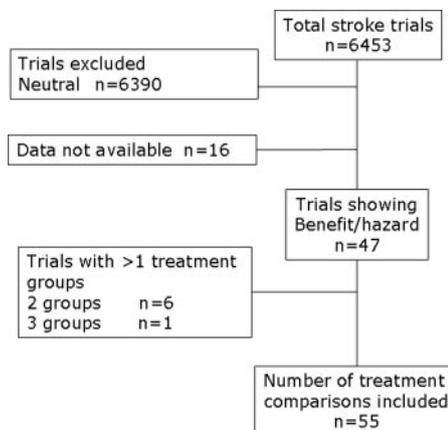


Figure 2. Identification of included trials.

used the Nottingham ADL scale, and 1 trial used its own ordinal measure.²⁶ Included trials studied the following interventions: abciximab (AbESTT); alteplase (ATLANTIS A & B, ECASS II, NINDS); aspirin (CAST, IST); atenolol (BEST); citicoline; DCLHb; ebselen; edaravone; enlimomab (EAST); factor VIIa; feeding (FOOD 3); nadroparin (FISS, FISS-TRIS); nimodipine (INWEST); occupational therapy (Corr, Gilbertson, Logan, TOTAL, Walker); physiotherapy (Young); pro-urokinase (PROACT II); selfotel (ASSIST); streptokinase (ASK, MAST-E, MAST-I); stroke unit (Dover, Helsinki, Kuopio, Nottingham, Orpington, Newcastle); and tirilazad (RANTTAS I & II, STIPAS, TESS I & II). Data relating to 16 trials or interventions which fulfilled the inclusion criteria were not made available.

The method of analyzing functional outcome used in the original trial publication varied considerably, see supplemental Appendix III, available online at <http://stroke.ahajournals.org>. Twenty-three (48.9%) trials assessed the treatment effect using a method which required the data to be collapsed into groups, eg, χ^2 test; 17 (36.2%) used a test based on comparing medians and 4 (8.5%) used a test which compared means; the remaining trials were unpublished so the method of analysis is not known.

Comparison of Statistical Tests

The statistical tests assessed differed significantly in the results they gave for each trial (2-way ANOVA, $P < 0.0001$). The ordering of the tests showed that those which analyze the original ordinal data generally perform better than those which collapse the data into ≥ 2 groups. The most efficient tests included ordinal logistic regression, t test, robust rank test and bootstrapping the difference in mean rank (Table). The subgroup analysis showed the same ordering of tests irrespective of type of intervention (acute, rehabilitation, stroke unit), trial size, time between randomization and onset, patient age, baseline severity, outcome measure, length of follow-up, and trial result (supplemental Appendix IV, available online at <http://stroke.ahajournals.org>).

When assessed by how many trials were statistically significant, those tests which did not collapse the data into groups again out-performed the other approaches; for example, ordinal logistic regression (using raw data) gave a statistically significant result in 25.9% of trials, whereas the 2×2 χ^2 test comparing death or poor outcome to an excellent outcome only gave a significant result in 9.3% of the trials (Figure 3).

Test Assumptions and Sensitivity

The statistical assumptions of the t test were not met for the majority of trials, and the assumptions of the ordinal logistic regression analysis failed for 8 of the 55 data sets; in contrast, the assumptions for the other tests were maintained. The sensitivity analysis showed that the top performing statistical tests were not overly sensitive, and statistically significant treatment effects were only found where they truly existed; see supplemental Appendix V, available online at <http://stroke.ahajournals.org>, for detailed results.

Comparison of Rank Scores for 16 Statistical Tests

| Test | Mean | No. of | Banding |
|---|-------|----------|---------|
| | Rank | Datasets | |
| Ordinal logistic regression | 6.11 | 54 | |
| t test | 6.51 | 55 | |
| RRT | 6.53 | 55 | |
| Bootstrap difference in mean rank | 6.85 | 55 | |
| Wilcoxon test | 7.31 | 55 | |
| Cochran-Armitage trend test (4 groups) | 7.36 | 50 | |
| Ordinal logistic regression (4 groups) | 7.50 | 50 | |
| Ordinal logistic regression (3 groups) | 7.92 | 51 | |
| Cochran-Armitage trend test (3 groups) | 8.27 | 51 | |
| χ^2 - death or poor outcome vs good | 8.87 | 55 | |
| χ^2 - death or poor outcome vs excellent | 9.24 | 54 | |
| Median test | 9.47 | 55 | |
| χ^2 - 2×3 test | 9.96 | 51 | |
| χ^2 - death vs alive | 9.98 | 51 | |
| χ^2 - 2×4 test | 10.02 | 50 | |
| Kolmogorov-Smirnov test | 11.29 | 55 | |

Comparison of rank scores for 16 statistical tests; lower ranks imply the test is more efficient. Analysis by two-way ANOVA and Duncan's multiple comparison procedure; tests joined by the same band are not significantly different from each other at $P < 0.05$.

RRT indicates robust rank test.

Discussion

These results show that statistical approaches which analyze the original ordinal data for functional outcome are more efficient than those which work on preprocessed data which has been collapsed into ≥ 2 groups. Interestingly, this point was originally demonstrated mathematically by Shannon in 1948.²⁷ In particular, ordinal logistic regression, t test, robust ranks test, and bootstrapping (the difference in mean rank) performed well and appear to be useful irrespective of the type of stroke trial, patient or intervention. Although individual tests based on dichotomized data using χ^2 analysis (eg, "dead/dependent" versus "independent") were effective for some data sets, they performed poorly in many and therefore cannot be recommended as general solutions for analyzing stroke trials. From an historical perspective, it is quite possible that trials which collapsed mRS or BI in 2 groups may have used a suboptimal analysis, and this may have contributed to false neutral findings in some cases in the past. For example, MAST-E²⁸ and STIPAS²⁴ were neutral as reported using dichotomous analysis but negative when assessed with ordinal approaches.

Several comments can be made about this study. First, it aimed to include data from all stroke trials assessing a beneficial or harmful intervention. Unfortunately, data were not made available for all identified trials; where possible, we created individual data from publications which provided patient numbers by outcome score. Data were missing for a variety of trial types (acute/rehabilitation/stroke unit) and

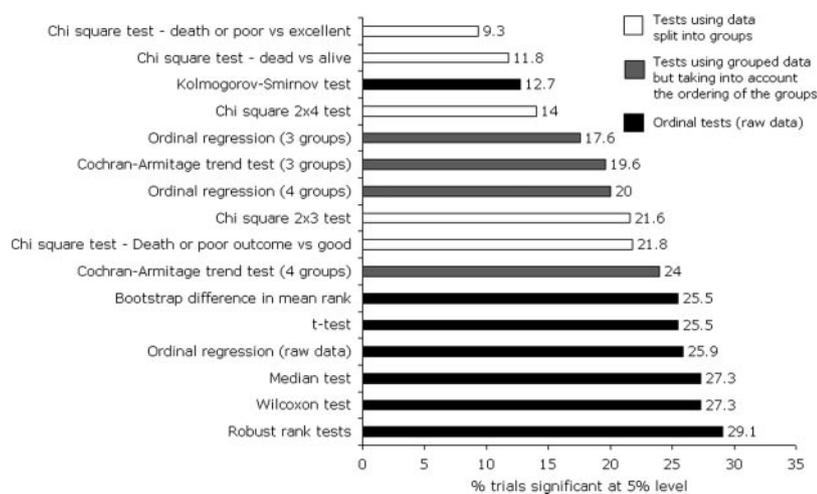


Figure 3. Percentage of trials significant at the 5% level with each statistical test.

sizes, and functional outcome measure (mRS/BI), so it is unlikely that a systematic bias was introduced into the findings; however, the precision of the results may have been attenuated by the missing trials. Second, we did not exhaustively search for all possible statistical tests relevant to the problem of analyzing ordered categorical data; instead, we focused on those approaches which are available in standard statistical textbooks and computer packages. Additionally, we could not include some tests used in recent trials, eg, patient specific outcomes²⁹ and Cochran Mantel-Haenszel test,³⁰ because these require access to individual data for both baseline and outcome variables, and these data were not available uniformly. Third, some of the statistical assumptions underlying the more efficient tests were not met in all trials; for example, the *t* test assumes data are normally distributed, whereas ordinal logistic regression assumes that any treatment effect is similar across outcome levels (“proportionality of odds”, ie, the odds of moving a treated patient from mRS 2 to 1 is similar to that for moving them from 5 to 4). Nevertheless, the robustness of these tests to deviations from their underlying assumptions means that they remain relevant for analyzing functional outcome data from stroke trials.

If alternative approaches to analyzing functional outcome data are to be used in the future, it is pertinent to ask how sample size should be calculated at the trial design stage. Historically, most calculations assumed that functional outcome would be dichotomized and analyzed using a χ^2 test approach.¹³ Although future trials could continue to calculate sample size in the same way (and then gain extra power by analyzing their data using an ordinal approach), specific sample size calculations are available when data are to be analyzed using ordinal logistic regression³¹ or the *t* test. Ideally, the extra power gained by using an ordinal statistical approach should not be used to reduce sample size; stroke trials have been too small in the past, as shown in a recent meta-analysis,¹³ and this may also have contributed to the failure of some of them.

A further issue with using a statistical test which analyses ordered categorical data are how to report the results to patients, carers, clinicians, and health-policy makers. The results of dichotomous tests may be summarized easily as the

proportion of patients who benefit (or suffer) with a treatment, ie, alteplase reduced absolute death or dependency (mRS >1) by 13% in the NINDS part 2 trial.¹ In contrast, ordinal tests will need to be presented as the average absolute improvement in outcome, eg, alteplase improved the mRS by 1 (of 7) point and BI by 22.5 (of 100) points. Alternatively, the combined odds ratio and its confidence intervals would be reported if ordinal logistic regression was used. In this respect, health consumers will need to decide what differences in mRS and BI are worthwhile, both clinically and in terms of health economics. In reality, it is reasonable to present the effect on functional outcome using both absolute percentage change and mean or median change in functional outcome score, and show this data graphically (as in Figure 1).

In summary, we suggest that ongoing and future trials should consider using statistical approaches which use the original ordered categorical data in the primary analysis of functional outcome measures. Such ordinal tests include ordinal logistic regression, and the robust ranks test; the *t* test may also be used although its assumptions were not meant in the majority of trials.

Acknowledgments

Secretariat and writing committee: Philip M.W. Bath (chief investigator, Nottingham, UK); Laura J. Gray (lead statistician, Nottingham, UK); Timothy Collier (statistical advisor, London, UK); Stuart Pocock (statistical advisor, London, UK). Statistical advisor: James Carpenter (Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK).

The following contributors provided individual patient data from their trial, and commented on the draft manuscript: Abciximab: H Adams (USA), E. Barnathan (USA); W Hacke (Germany); ASK: G. Donnan (Australia); ASSIST 07 & 10: S. Davis (Australia); ATLANTIS A & B: G. Albers, S. Hamilton (USA); BEST Pilot & Main: D. Barer (UK); Citicoline 1, 7, 10, 18: A. Davalos (Spain); Corr: S. Corr (UK); Dover Stroke Unit: P. Langhorne (UK); DCLHb: P. Koudstaal, R. Saxena (Netherlands); Ebselen: T. Yamaguchi (Japan); ECASS II: W. Hacke, E. Bluhmki (Germany); Factor VII: S. Mayer (USA), K. Begtrup (Denmark); FISS: R. Kay (Hong Kong); FOOD 3: M. Dennis (UK); Gilbertson: L. Gilbertson (UK); INWEST: N.-G. Wahlgren, N. Ahmed (Sweden); IST: P. Sandercock (UK); Kuopio Stroke Unit: J. Sivenius (Finland); Logan: P. Logan (UK); MAST-I: L. Candelise (Italy), J. Wardlaw (UK); Newcastle Stroke Unit: H. Rodgers (UK); NINDS: J. Marler (USA);

Parker: C. Parker (UK); Nottingham Stroke Unit: N. Lincoln, P. Berman (UK); RANNTAS I & II, STIPAS, TESS I & II: P. Bath (UK); Walker 1 & 2: M. Walker (UK); Young: J. Young, A. Forster (UK).

We thank the patients who took part in these studies, and the trialists who shared their data. The study was conceived, initiated, managed, analyzed, and interpreted independently of any pharmaceutical company. Each collaborator listed above commented on the draft manuscript.

Sources of Funding

L.J.G. is funded, in part, by BUPA Foundation and The Stroke Association (UK). P.M.W.B. is Stroke Association Professor of Stroke Medicine. The funding sources had no involvement in this project.

Disclosures

None.

References

1. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute stroke. *N Engl J Med*. 1995;333:1581–1587.
2. CAST (Chinese Acute Stroke Trial) Collaborative Group. CAST: randomised placebo-controlled trial of early aspirin use in 20 000 patients with acute ischaemic stroke. *Lancet*. 1997;349:1641–1649.
3. Walker M, Gladman J, Lincoln N, Siemonsma P, Whiteley T. Occupational therapy for stroke patients not admitted to hospital: a randomised controlled trial. *The Lancet*. 1999;354:278–280.
4. Chen ZM, Sandercock P, Pan HC, Counsell C, Collins R, Liu LS, Xie JX, Warlow C, Peto R; on behalf of the CAST and IST Collaborative Groups. Indications for early aspirin use in acute ischemic stroke: a combined analysis of 40 000 randomized patients from the chinese acute stroke trial and the international stroke trial. *Stroke*. 2000;31:1240–1249.
5. Stroke Unit Trialists' Collaboration. Organised inpatient (stroke unit) care for stroke. *The Cochrane Library*. Oxford: Update Software; 2002.
6. The ATLANTIS ECASS and NINDS rt-PA Study Group Investigators. Association of outcome with early stroke treatment: pooled analysis of ATLANTIS, ECASS, and NINDS rt-PA Stroke trials. *The Lancet*. 2004;363:768–813.
7. Walker MF, Leonardi-Bee J, Bath P, Langhorne P, Corr S, Drummond A, Gilbertson L, Gladman JRF, Jongbloed L, Parker C. An individual patient data meta-analysis of randomised controlled trials of community occupational therapy for stroke patients. *Stroke*. 2004;35:2226–2232.
8. Saxena R, Wijnhoud AD, Carton H, Hacke W, Kaste M, Przybelski RJ, Stern KN, Koudstaal PJ. Controlled safety study of a hemoglobin-based oxygen carrier, DCLHB, in acute ischemic stroke. *Stroke*. 1999;30:993–996.
9. Davis SM, Lees KR, Albers GW, Diener HC, Markabi S, Karlsson G, Norris J; for the ASSIST Investigators. Selfotel in acute ischemic stroke: possible neurotoxic effects of an NMDA antagonist. *Stroke*. 2000;31:347–354.
10. Enlimomab acute stroke trial investigators. Use of anti-ICAM-1 therapy in ischemic stroke: results of the enlimomab acute stroke trial. *Neurology*. 2001;57:1428–1434.
11. Bath PMW, Bleic S, Bogousslavsky J, Boysen G, Davis S, Diez-Tejedor E, Ferro JM, Gommans J, Hacke W, Indredavik B, Norrving B, Orgogozo JM, Ringelstein EB, Sacchetti ML, Iddenden R, Bath FJ, Musch BC, Brosse DM, Naberhuis-Stehouwer SA. Tirilazad mesylate in acute ischemic stroke: a systematic review. *Stroke*. 2000;31:2257–2265.
12. Grotta J. Why do all drugs work in animals but none in stroke patients? 2 neuroprotective therapy. *J Intern Med*. 1995;237:89–94.
13. Weaver CS, Leonardi-Bee J, Bath-Hexall FJ, Bath PMW. Sample size calculations in acute stroke trials: a systematic review of their reporting, characteristics, and relationship with outcome. *Stroke*. 2004;35:1216–1224.
14. Rankin J. Cerebral vascular accidents in patients over the age of 60. 2. Prognosis. *Scott Med J*. 1957;2:200–215.
15. Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index. *Maryland State Medical Journal*. 1965;61–65.
16. Lindley RI, Waddell F, Livingstone M, Sandercock P, Dennis MS, Slattery J, Smith B, Warlow C. Can simple questions assess outcome after stroke? *Cerebrovascular Diseases*. 1994;4:314–324.
17. Bath P, Lindenstrom E, Boysen G, De Deyn P, Friis P, Leys D, Marttila R, Olsson J-E, O'Neill D, Orgogozo J-M, Ringelstein B, van der Sande J-J, Turpie AGG. Tinzaparin in acute ischaemic stroke (TAIST): a randomised aspirin-controlled trial. *Lancet*. 2001;358:702–710.
18. Horn J, deHaan R, Vermeulen M, Limburg M. Very early nimodipine use in stroke (VENUS): a randomised, double-blind, placebo-controlled trial. *Stroke*. 2001;32:461–465.
19. Wardlaw J, Sandercock P, Warlow C, Lindley RI. Trials of thrombolysis in acute ischemic stroke: does the choice of primary outcome measure really matter? *Stroke*. 2000;31:1133–1135.
20. Hacke W, Markku K, Fieschi C, von Kummer R, Davalos A, Meier D, Larue V, Bluhmki E, Davis S, Donnan G, Schneider D, Diez-Tejedor E, Trouillas P. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). *Lancet*. 1998;352:1245–1251.
21. Stingele R, Bluhmki E, Hacke W. Bootstrap statistics of ECASS II data: just another post hoc analysis of a negative stroke trial? *Cerebrovascular Diseases*. 2001;11:30–33.
22. Haley EC. High-dose tirilazad for acute stroke (RANTTAS II). *Stroke*. 1998;29:1256–1257.
23. The RANTTAS Investigators. A randomized trial of tirilazad mesylate in patients with acute stroke (RANTTAS). *Stroke*. 1996;27:1453–1458.
24. The STIPAS Investigators. Safety study of tirilazad mesylate in patients with acute ischemic stroke (STIPAS). *Stroke*. 1994;25:418–423.
25. The Tirilazad International Steering Committee. Tirilazad for acute ischaemic stroke (cochrane review). Oxford: Update Software; 2002.
26. Barer DH, Cruickshank JM, Ebrahim SB, Mitchell JR. Low dose beta blockade in acute stroke ("Best" Trial): an evaluation. *BMJ*. 1988;296:737–741.
27. Shannon CE. A mathematical theory of communication. *The Bell System Technical Journal*. 1948;27:379–423.
28. The Multicenter Acute Stroke Trial; Europe Study Group. Thrombolytic therapy with Streptokinase in acute ischemic stroke. *N Engl J Med*. 1996;335:145–149.
29. Adams HP, Leclerc JR, Bluhmki E, Clarke WR, Hansen MD, Hacke W. Measuring outcomes as a function of baseline severity of ischemic stroke. *Cerebrovascular Diseases*. 2004;18:124–129.
30. Lees KR, Zivin JA, Ashwood T, Davalos A, Davis SM, Diener H-C, Grotta J, Lyden P, Shuaib A, Hardernark H-G, Wasiewski WW; for the Stroke-Acute Ischemic NXY Treatment (SAINT I) Trial Investigators. NXY-059 for acute ischemic stroke. *N Engl J Med*. 2006;354:588–600.
31. Whitehead J. Sample-size calculations for ordered categorical-data. *Statistics in Medicine*. 1993;12:2257–2271.

Can We Improve the Statistical Analysis of Stroke Trials?: Statistical Reanalysis of Functional Outcomes in Stroke Trials The Optimising Analysis of Stroke Trials (OAST) Collaboration

Stroke. 2007;38:1911-1915; originally published online April 26, 2007;
doi: 10.1161/STROKEAHA.106.474080
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 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/38/6/1911>

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/>