Thrombolysis Is Associated With Consistent Functional Improvement Across Baseline Stroke Severity

A Comparison of Outcomes in Patients From the Virtual International Stroke Trials Archive (VISTA)

Nishant K. Mishra, MBBS; Patrick Lyden, MD; James C. Grotta, MD; Kennedy R. Lees, MD, FRCP; for the VISTA Collaborators

Background and Purpose—Baseline stroke severity predicts outcomes among thrombolysed patients. The baseline National Institutes of Health Stroke Scale (NIHSS) thresholds are sometimes used to select patients for thrombolysis, clinical trial enrollment, or both. Using data lodged with Virtual International Stroke Trials Archive, we compared adjusted outcomes between thrombolysed and nonthrombolysed patients enrolled in neuroprotection trials (1998–2007) to assess the influence of various levels of baseline NIHSS.

Method—We assessed the association of treatment with outcome, measured across the modified Rankin scale score distribution, in patients categorized by baseline NIHSS in increments of 4. We used an age and baseline NIHSS adjusted Cochran-Mantel-Haenszel test followed by proportional odds logistic regression analysis. We report the Cochran-Mantel-Haenszel $P$ values and estimated odds ratios (OR) for improved modified Rankin scale score distribution with treatment for patients within each baseline NIHSS category.

Results—Data were available for 5817 patients (1585 thrombolysed and 4232 nonthrombolysed). Baseline severity was greater among thrombolysed than nonthrombolysed (median baseline NIHSS, 14 vs 13; $P<0.05$). An association of treatment with outcome was seen independently and was of similar magnitude within each of the baseline NIHSS categories 5 to 8 ($P=0.04$; OR, 1.25; 95% confidence interval [CI], 1.0–1.6; N=278/934 thrombolysed/nonthrombolysed), 9 to 12 ($P=0.01$; OR, 1.3; 95% CI, 1.1–1.6; N=404/942), 13 to 16 ($P<0.05$; OR, 1.6; 95% CI, 1.3–2.1; N=342/814), 17 to 20 ($P<0.05$; OR, 1.7; 95% CI, 1.3–2.1; N=311/736), and 21 to 24 ($P<0.05$; OR, 1.6; 95% CI, 1.1–2.1; N=178/466). No association was observed within baseline NIHSS categories 1 to 4 ($P=0.8$; OR, 1.1; 95% CI, 0.3–4.4; N=8/161) or ≥25 ($P=0.08$; OR, 1.1; 95% CI, 0.7–1.9; N=64/179).

Conclusions—In this nonrandomized comparison, outcomes after thrombolysis were significantly better than in untreated comparators across baseline NIHSS 5 to 24. The significant association was lost only at extremes of baseline NIHSS when sample sizes were small and confidence limits were wide. (Stroke. 2010;41:2612-2617.)

Key Words: baseline • functional • severity • stroke • thrombolysis

Intravenous thrombolysis with alteplase is a proven therapy for acute ischemic stroke patients presenting before 4.5 hours of symptom onset.1 However, some patients are denied therapy for fear of poor outcomes.2 European guidelines recommend that patients with baseline stroke severity, National Institutes of Health Stroke Scale (NIHSS) ≥25, and minor/rapidly improving strokes should not be thrombolysed,2,3 because it is believed that many patients who show rapid improvement/have minor strokes would not display residual deficit, and treatment with thrombolytic therapy would expose them to risk of complications, such as cerebral hemorrhage. Similarly, those patients who present with baseline NIHSS ≥25 are also supposed to have poorer outcomes because of excess symptomatic hemorrhages.2–3 Baseline stroke severity (baseline NIHSS) is known to affect outcomes among thrombolysed patients4 and therefore was incorporated for patient selection in the ECASS III trial.1 Although the regulatory authorities have recommended withholding thrombolytic therapy among patients with minor/rapidly improving strokes and for those with severe stroke at baseline, poorer response to therapy in these subgroups has never been demonstrated in randomized, controlled trials.5 Post hoc analyses of the NINDS and ECASS-III trials suggest equal efficacy across severity range, although power to examine
subgroups is inevitably lower than chosen for the primary analyses, and patients at extremes of severity were under-represented.\textsuperscript{6–8} The logistical challenges involved in generating randomized trial evidence for these limited subgroups militate against any prospect for producing a definitive answer in the foreseeable future. Therefore, we must turn to alternative sources of evidence.

The Virtual International Stroke Trials Archive (VISTA) is a repository of data from many rigorously controlled clinical trials.\textsuperscript{9} Although most of these trials examined putative neuroprotectant agents, use of recombinant tissue plasminogen activator was generally recorded. We planned to use data from VISTA, hypothesizing that clinical practice over the past decade would have been sufficiently diverse to allow analysis of existing rigorously collected clinical data lodged in VISTA to examine the influence of baseline stroke severity on outcomes after thrombolytic therapy.

**Patients and Methods**

**Data Source and Patients**

We collated the demographics, clinical data, and measures of functional outcome from neuroprotection trials conducted in the period 1998 to 2007, held within VISTA (www.vista.gla.ac.uk).\textsuperscript{10} All trials held necessary review board and regulatory approvals, and patients consented to participation; only anonymous data are held by VISTA. We sought VISTA data derived from trials in which the investigational neuroprotection agent was not vasoactive and did not interfere with clotting or from placebo groups. We excluded any patient who had cerebral hemorrhage or stroke of undetermined etiology. To avoid dual publication, we excluded patients who may have been enrolled in SITS-MOST; we determined this from their country and date of enrollment. Finally, we excluded patients lacking our chosen outcome measure, 90-day modified Rankin scale (mRS) score, or secondary outcome, 90-day NIHSS score. Patients who died within 90 days were attributed the mRS score of 6 and categorized separately for NIHSS analysis.

**Statistical Analysis**

We compared outcome between patients who received thrombolysis and patients who did not receive thrombolysis (controls) among the categories of baseline NIHSS scores (<4, 5–8, 9–12, 13–16, 17–20, 21–24, and $\geq 25$). Note that the reason for withholding thrombolysis in each case was not recorded but will include absence of marketing approval in the region at that time, clinical uncertainty over the use of thrombolysis for stroke generally, absence of treatment facilities for thrombolysis in the hospital at that time, and contraindications to thrombolysis for the individual patient. For each contrast, we compared the overall distribution of all 7 categories of day 90 mRS scores of the 2 groups, ie, from 0 (asymptomatic) through 5 (bed-bound and completely dependent) to 6 (dead). The European Medicines Evaluation Agency Points to Consider for reporting trials allow for use of the full distribution of the mRS but suggest that this may be supported by a secondary analysis of a second outcome measure, such as NIHSS.\textsuperscript{11} For analysis of our supporting end point, NIHSS, we grouped adjacent scores into categories: 0 (no measurable deficit), 1 to 4, 5 to 8, 9 to 12, 13 to 16, 17 to 20, 21 to 24, and $\geq 25$ (most severe neurological deficit) or dead. The distribution of patients across these categories was then compared between the groups as it was for mRS.

To test for a significant association of outcome distribution with thrombolysis exposure, we used the Cochran-Mantel-Haenszel statistic, adjusting for both age and baseline NIHSS as continuous variables.\textsuperscript{12,13} This nonparametric approach avoids invoking an assumption of proportional odds in which there should be a common odds ratio across all cut points on the ordinal outcome scale, and we consider that the Cochran-Mantel-Haenszel test provides the most conservative estimate of statistical significance. However, it does not express the extent of the association. For this, we applied logistic regression analysis, also adjusted for age and baseline NIHSS, to estimate the odds ratio under the assumption of proportional odds and its associated 95% confidence interval.

Our choice of baseline factors for adjustment was based on 2 influences. First, age and baseline severity are the 2 most powerful prognostic factors for stroke and are usually included in outcome distribution analyses.\textsuperscript{15–17} Second, age and NIHSS data were available for our entire sample, whereas other factors of potential interest were incomplete. However, we also undertook a sensitivity analysis in which we adjusted for ECASS III variables diabetes and previous stroke. In addition, we also undertook an adjusted analysis by combining the variables that differed significantly at baseline; however, if this resulted in excessive diminution of our sample, we reported the limitations.

Our objective was mainly to undertake ordinal distribution or “shift” analysis, which is an efficient end point analytic technique recommended by European Medicines Evaluation Agency.\textsuperscript{11,18–20} Shift analysis is considered better than dichotomization of end point measures, although there are differences of opinion.\textsuperscript{21,22} Dichotomization is criticized for the statistical information it discards, ie, loss of power, and shift analysis is especially useful when the treatment effect is mild and/or uniform across all Rankin categories. Odds ratios in our analysis express the common odds of an improved distribution of outcome in association with alteplase treatment.

Cochran-Mantel-Haenszel and logistic regression analyses were undertaken using SAS 9.2 software (SAS Software Limited, United Kingdom) and other analyses by Stats Direct software (StatsDirect Limited, United Kingdom).

Reliable information on symptomatic hemorrhage was not available because post-treatment imaging was not routinely applied in neuroprotection trials to patients who had not been treated with alteplase.

**Results**

**Patient Sample**

We collated data on 9665 patients, of whom 5342 (59%) were enrolled from non-European sites. To avoid dual publication with SITS-MOST, we excluded 2789 patients (28%) enrolled from European sites between 2002 and 2006, and 177 patients for whom we lacked information on country. Complete data were available for analysis of mRS for 5817 patients and data on NIHSS were available for 5715 (Figure 1).

All stroke patients were treated as per institutional practice and stroke guidelines acceptable at the point of trial conduct. Monitoring for protocol compliance was undertaken on behalf of sponsors for these trials. This implies that when thrombolysis was administered, this was in accordance with marketing authorization for the relevant country, ie, that treatment commenced within 3 hours of stroke onset; however, the onset to treatment delay is not recorded for thrombolysis in these trials. Our data derived mainly from North American (60%), European (16%), and Australasian (13%) centers. Baseline characteristics are shown in the Table. Of the 5817 patients with mRS outcome data, 1585 (27.2%) received thrombolysis.

**Does Baseline Stroke Severity Influence Stroke Outcomes?**

In an ordinal logistic regression analysis, we found that baseline severity ($P<0.0001$), use of recombinant tissue plasminogen activator, and age were significant predictors of outcomes. Then, in an age-adjusted ordinal logistic regression analysis, we found that baseline stroke severity ($P<0.0001$) and the interaction between severity and use of alteplase
were associated with outcome of stroke, but we did not see an independent effect of alteplase (P=0.65).

Supported by this interaction test, we classified the baseline stroke severity into 7 baseline NIHSS score categories: 1 to 4, 5 to 8, 9 to 12, 13 to 16, 17 to 20, 21 to 24, and ≥25, and undertook tests of association for thrombolysis with outcomes in each of these categories.

Are There Improved Outcomes Across All Baseline Stroke Severity Categories?

Findings from age and baseline NIHSS-adjusted analysis of functional outcomes are shown in Figure 2. This essentially shows a significant association of better outcomes with use of alteplase for patients presenting with baseline NIHSS 5 to 24. The patients with NIHSS <4 at baseline had a mixed distribution of outcomes at 90 days, some Rankin categories appeared to have improved, and others worsened. Patients with baseline NIHSS >24 showed generally improved Rankin distribution with alteplase; however, proportionality of the treatment effect was maintained.

Findings were consistent for the neurological outcomes (by NIHSS on day 90) and also for the sensitivity analyses (ie, unadjusted analysis and analysis adjusting for age, baseline NIHSS, diabetes, and previous stroke). We could not adjust for onset to treatment time because time to initiation of thrombolytic therapy was not recorded within our source neuroprotection trials. Fifty-nine percent of records lacked coding for the variable “antithrombotic” (N=3432), 4.8% lacked coding for atrial fibrillation (N=278), and 3.2% lacked coding for patients with previous strokes (N=186). This limitation to our sample precluded a reliable analysis that was adjusted for all variables that differed at baseline (age, baseline NIHSS, previous use of antithrombotic drugs, previous stroke, and atrial fibrillation).

<table>
<thead>
<tr>
<th>Table. Baseline Characteristics of the Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age, median (range)</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Baseline NIHSS, median (range)</td>
</tr>
<tr>
<td>Previous antiplatelet use</td>
</tr>
<tr>
<td>Previous anticoagulation use</td>
</tr>
<tr>
<td>Previous stroke</td>
</tr>
<tr>
<td>Congestive heart failure</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>Hypertension</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
</tr>
<tr>
<td>Myocardial infarction</td>
</tr>
</tbody>
</table>

NIHSS, National Institutes of Health Stroke Scale.
Discussion

Patients with mild and severe strokes are under-represented in randomized trials and post-marketing analyses. As a result, the European Medicines Evaluation Agency marketing authorization for alteplase in acute ischemic stroke lists minor neurological deficit or symptoms rapidly improving before start of infusion and severe stroke as assessed clinically (eg, NIHSS >25) and/or by appropriate imaging techniques as contraindications.\(^2\)\(^{23}\) Such patients do present to hospital services, however, and this places the physician in a dilemma of whether to offer treatment. Some experienced physicians treat such patients. For example, 12% of patients in the SITS-ISTR thrombolysis registry had a baseline NIHSS score in the range of 0 to 4, and 4% had severe stroke as assessed clinically (eg, NIHSS >25)\(^7\)\(^8\)\(^9\)\(^{20}\)\(^{21}\) report similar findings. Randomized trials to establish the existence or extent of benefit at extremes of baseline severity may be difficult to conduct and delayed in execution. Other sources of evidence must be examined, and high-quality registry data are the obvious choice.

In our present nonrandomized comparison of data held in VISTA, outcomes after thrombolysis were significantly better than in untreated comparators across baseline NIHSS scores 5 to 24. This significant association was lost only at extremes of baseline NIHSS (ie, 1–4 and ≥25). Although the point estimates for both adjusted and unadjusted odds ratios remain favorable in the extreme groups, they are lower than those observed at other levels of stroke severity.

In these extreme groups, the small sample size seriously undermines the power of the statistical tests and, with wide confidence intervals, the true point estimate is not reliably indicated. There is a second statistical issue to consider relating to the outcome measure that we used. By examining the full distribution of the mRS, we have used a test that is less dependent on case-mix than dichotomization. We are able to use the same test for patients with mild stroke as is used severe stroke and may still detect benefit. Even so, at the extremes of baseline severity, outcomes are generally so good or so poor that only a few mRS categories are well-represented in the control groups. Both the Cochran-Mantel-Haenszel test and the proportional odds estimations will be compromised if some categories are not contributing to the analysis. Effectively, the test of treatment effect will be diluted by the noncontributing groups. For Cochran-Mantel-Haenszel, this means that it becomes more difficult to reach statistical significance; however, but for the proportional odds tests, the basic assumption has been breached and the effect is not proportional. There is no easy solution to this problem. If case-mix is altered to deliver a significant result, then patients with mild or severe stroke must be excluded, which is the solution used by the trials. Conversely, if the outcome measure is varied according to the sample case-mix (the sliding dichotomy approach discussed by Murray et al\(^24\)), then interpretation is rendered difficult. Is an odds ratio for achieving mRS 0 vs 1 to 6 equivalent to an odds ratio for achieving mRS 0 to 5 vs 6, ie, is survival free from symptoms equivalent to survival at any cost?

Here, we have chosen to present 1 analytic approach for all severities of stroke, but we also illustrate the range of outcomes at extremes of severity. From these, although the summary statistics show only a nonsignificant but favorable trend, we can draw further conclusions. Among patients with severe stroke, there are evident trends toward benefit across almost all boundaries of mRS. Among patients with mild stroke, all boundaries except 0 to 1 vs 2 to 6 show benefit, but 4 of the mRS categories are entirely unrepresented. Our data show no reason to withhold treatment from either group of patients but are not in themselves sufficient evidence to justify treatment.

Our findings draw validity from the fact that our source clinical trials rigorously reported concomitant treatments and outcomes and had strict on-site data verification procedures. However, the nonrandom allocation to treatment vs control groups is a significant weakness of our design. We could not determine the degree and cause of exclusion of patients from...
our database. We can only consider factors known to be associated with prognosis.

We have adjusted statistically for factors that have a large influence on outcome. We can also ‘anchor’ our findings by comparison of treatment associations for patients with moderate stroke severity in our study against known treatment effects in comparable patients from randomized trials. For example, we find an odds ratio for favorable outcome of 1.3 to 1.6 for patients with baseline NIHSS 9 to 12 and 13 to 16; the comparable estimate from treatment within 3 hours of stroke onset in a randomized control trial would be 1.64 and for 3 to 4.5 hours would be 1.34.25 Our estimates are comparable and perhaps conservative.

The decay of benefit across later onset to treatment times raises a second issue. We do not have information on the onset to treatment delay for alteplase in our current analysis. Because the patients were permitted only 1 investigational drug in the participating VISTA trials, with alteplase being used as standard of care, and because these trials were closely monitored by their sponsors, we assume that patients were largely treated within 3 hours of stroke onset. We also assume that the onset-to-treatment time is comparable to those from the CASES and SITS-MOST registries (155 [130–175] minutes and 140 [115–165] minutes, respectively; n=6483). Unfortunately, the latency between stroke onset and recording of initial severity differed between our treatment group (3.7 hours) and controls (5.1 hours). Severity is associated with onset to hospital arrival time: patients with more severe stroke present earlier.26 We adjusted our analyses for stroke severity, but it is conceivable that residual bias persists. Such a bias would cause underestimation of true initial severity among our controls and through the baseline adjustment would lead us to overestimate treatment effect. It will influence all patients across our severity range but may be less evident at extremes of severity: the NIHSS criterion will be responsible for discouraging use of alteplase, and so the proportion of patients who are treated with alteplase will have extremes of NIHSS.

With these caveats, it would be desirable to replicate our findings. Supporting evidence could come from a comparison of SITS ISTR data against VISTA controls, a collaborative analysis that is underway. We lack data on symptomatic hemorrhages because patients who are not treated with thrombolysis generally do not undergo follow-up cerebral imaging for routine detection of hemorrhagic transformation. However, the outcome measure that we use takes into account effects of hemorrhage or other adverse events on function.

We adjusted for age and baseline severity because these are the established most important variables known to influence outcomes.27 We could not adjust for all age, baseline NIHSS, previous use of antithrombotic drugs, previous stroke, and atrial fibrillation data together because 1 of the contributing trial programs did not record pretreatment medications. However, we were able to undertake an adjusted analysis for the variables that were found significant in ECASS III, namely diabetes and previous stroke, and our estimates remained consistent.

Some of the patients in our study received an investigational medicinal product. Each contributing trial has already tested for, and excluded, a significant interaction of that product with alteplase, both in vitro and in vivo.

Conclusion
In conclusion, our findings imply that patients with extremes of NIHSS scores recorded at baseline may still benefit from treatment but the supporting evidence remains weak.

Acknowledgments

Disclosures
The analyses reported in this article were based on a research proposal approved by the VISTA Steering Committee and were undertaken by N.K.M. at University of Glasgow, UK. N.K.M. is supported by a British ORS Scholarship and the University of Glasgow scholarship. VISTA has received financial support from the European Stroke Organisation in the form of an unrestricted grant and contributions toward data extraction and capacity building from the Universities of Glasgow, California San Diego, Nottingham, Edmonton, Calgary, Texas, and Massachusetts; from commercial groups, including Brainsgate, Novartis, Boehringer Ingelheim, and the Vertical Group; and from grant agencies and charities, including the UK Stroke Association, N.K.M. and K.R.L. designed and interpreted the analyses and drafted the manuscript; both had access to the VISTA data. P.L. and J.G. contributed data, reviewed the outline proposal, commented on the manuscript, and approved the final version. All authors take full responsibility for the content. The manuscript was reviewed and approved by the VISTA steering committee (www.vista.gla.ac.uk). No commercial organization was involved in the origination, execution, or reporting of this work. P.L. has received research support from the American Heart Association, the National Institutes of Health, and the Veteran’s Affairs Medical Research Department; consulting fees from Mitsubishi, Co-Axia, Benechill, and Photothera; and research contracts from AstraZeneca, Bayer, and Innercool. J.G. has received consulting compensation from Lundbeck. K.L. has received honoraria (modest) from Boehringer Ingelheim, Lundbeck, Thrombogenics, and Talecirs.

References


Thrombolysis Is Associated With Consistent Functional Improvement Across Baseline Stroke Severity: A Comparison of Outcomes in Patients From the Virtual International Stroke Trials Archive (VISTA)
Nishant K. Mishra, Patrick Lyden, James C. Grotta and Kennedy R. Lees
for the VISTA Collaborators

Stroke. 2010;41:2612-2617; originally published online October 14, 2010;
doi: 10.1161/STROKEAHA.110.589317

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/41/11/2612

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2012/04/02/STROKEAHA.110.589317.DC1

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/
Thrombolysis Is Associated With Consistent Functional Improvement Across Baseline Stroke Severity

A Comparison of Outcomes in Patients From the Virtual International Stroke Trials Archive (VISTA)

Nishant K. Mishra, MBBS; Patrick Lyden, MD; James C. Grotta, MD; Kennedy R. Lees, MD, FRCP; for the VISTA Collaborators

Background and Objectives: Stroke baseline score severity predicts the outcomes of thrombolysis. The National Institutes of Health Stroke Scale (NIHSS) baseline threshold is sometimes used to select patients for thrombolysis, to enroll in clinical trials, or to choose stroke patients for thrombolysis treatment.

We performed a comparative study of outcomes, data acquired and stored in the Virtual International Stroke Trials Archive (VISTA), mainly to assess the influence of NIHSS baseline scores on outcomes in patients included in neuroprotective trials (1998-2007).

Methods: Patients were divided into groups based on baseline NIHSS score increments of 4 points, and the relationship between treatment and outcome was assessed, where outcome was determined by the modified Rankin score. We applied age and baseline NIHSS stratified Cochran-Mantel-Haenszel tests and proportional-odds logistic regression analysis. We reported Cochran-Mantel-Haenszel-stratified test P values and the odds ratio for different baseline NIHSS score groups of patients treated with thrombolysis compared to non-thrombolysis patients.

Results: Data included 5817 patients (1585 for thrombolysis, 4232 for non-thrombolysis). Thrombolysis patients had a higher baseline severity compared to non-thrombolysis patients (median baseline NIHSS score, 14 vs. 13; P<0.05). Treatment outcome relationships showed independence in baseline NIHSS score at 5-8 points (P=0.04; OR, 1.25; 95% CI, 1.0-1.6; N=278/934), 9-12 points (P=0.01; OR, 1.3; 95% CI, 1.1-1.6; N=404/942), 13-16 points (P<0.05; OR, 1.6; 95% CI, 1.3-2.1; N=342/814), 17-20 points (P<0.05; OR, 1.7; 95% CI, 1.3-2.1; N=178/466), and ≥25 points (P<0.08; OR, 1.1; 95% CI, 0.7-1.9; N=64/179); no such relationship was found for baseline NIHSS scores of 1-4 points (P=0.8; OR, 1.1; 95% CI, 0.3-4.4; N=8/161) or >25 points (P=0.08; OR, 1.1; 95% CI, 0.7-1.9; N=64/179).

Conclusion: This study, through non-random comparisons, compared outcomes of patients with NIHSS baseline scores of 5-24 points, and thrombolysis patients showed significantly better results than non-thrombolysis patients. Only when the sample size was small and confidence intervals were wide, and in the case of baseline NIHSS scores at 1-4 or >25 points, was there no significant relationship between thrombolysis and outcome.

Keywords: baseline, functional, severity, stroke, thrombolysis

(Stroke. 2010;41:2612-2617. 吉林大学第一医院神经内科 金涛 译 吴江 董铭 校)
some improvements or rapid improvement patients and those with severe stroke at baseline for thrombolysis, but in these subgroups, no RCTs have confirmed a worse response to treatment. The NINDS and ECASS III trials showed that thrombolysis was effective for different stroke severities, even though the sample size for each group was lower than the first analysis, and the results were usually at the extremes of stroke severity. Therefore, we needed to seek alternative evidence sources.

VISTA (Virtual International Stroke Trials Archive) is a database containing data from well-controlled RCTs [9]. Despite most trials validating the supposed neuroprotective agents, tissue plasminogen activator (t-PA) was widely used. We planned to use the data from VISTA, assuming that there were sufficient and well-controlled clinical trials between 1998 and 2007 on different types of patients, which have been stored in VISTA and could be used to test the effect of different stroke baseline severity on thrombolysis outcomes.

**Patients and Methods**

Data Source and Patients

We collected and collated data from 1998 to 2007 RCTs on neuroprotective agents and stored in VISTA, including the demographic data, clinical data, and functional outcomes (www.vista.gla.ac.uk) [10].

In all trials, the review committee, registration approval, and patient consent were required. The data from VISTA were anonymous. We needed the VISTA data from trials where the neuroprotective agent was not active in the vessel and not affected by placebo or inhibitor. We excluded all hemorrhagic stroke and those with unknown etiology. We excluded patients enrolled in the SITS-MOST trial. Finally, we excluded patients lacking our chosen treatment results or mRS at 90 days or NIHSS scores at 90 days.

**Statistical Analysis**

We compared the NIHSS scores in different groups (<4, 5-8, 9-12, 13-16, 17-20, 21-24, >25) between thrombolysis and non-thrombolysis patients. We noted that the reason for non-selection of thrombolysis was not recorded, which may have included that thrombolysis was not available in those areas, concerns about clotting or safety, and lack of some facilities or personal contraindications. In each comparison, we applied the 90-day mRS score comparison analysis to all thrombolysis and non-thrombolysis patients. The European Medicines Agency recommended that the results of RCTs should be reported using the full distribution of mRS, but this should be built on a second-level result (like using NIHSS) for reanalysis [11]. To analyze our results, we grouped the adjacent NIHSS scores into several groups: 0 (no neurological deficit), 1-4, 5-8, 9-12, 13-16, 17-20, 21-24, and >25 (severe neurological deficit) or death. We compared the distribution of patients in different groups using the stratified Cochran-Mantel-Haenszel test and stratified on age and baseline NIHSS scores as continuous variables [12,13]. This non-parametric method avoids the assumption of a proportion advantage, which is common at the boundary point of all continuous scoring results. We used logistic regression analysis to assess the proportion advantage (OR) and 95% confidence interval (CI). Our main goal was to use the continuous distribution of shift analysis, which is recommended by the European Medicines Agency as a more effective endpoint analysis tool [11,18-20]. Although different views exist, shift analysis is considered better than dichotomy [21,22]. Dichotomy is not preferred because it loses some statistical power, especially when the Rankin scores are consistent.

We used the SAS 9.2 software (SAS Software Limited, United Kingdom) for stratified Cochran-Mantel-Haenszel analysis and logistic regression analysis.
神经保护剂试验的 VISTA 数据 (1998-2007) N=9665

颅内出血 N=571

缺血性卒中患者 N=9058

病因不明的卒中 N=36

与 SITS-MOST 试验潜在重叠的患者 N=2966

VISTA 分析 N=6092

排除缺失 m-RS 数据，5817 例患者用于功能结果的分析

排除缺失 NIHSS 数据，5715 例患者用于神经功能结果分析

所有卒中患者接受的治疗措施所依从的每个治疗中心的治疗常规及卒中指南均在试验指导可接受的范围内。试验的赞助商负责监督试验流程图的依从性，这意味着患者接受的溶栓治疗与其所属国家的推销许可相一致。也就是说，这一治疗在卒中发生 3 小时内展开；但是，在试验中治疗延迟的将不做记录。我们的试验数据 60% 来源于北美试验中心，欧洲 16%，澳大利亚 13%。基线特征列于表中。5817 名患者中 mRS 数据的患者中，有 1585 名 (27.2%) 接受了溶栓治疗。

卒中基线的严重性影响患者的转归吗？

在顺序的逻辑回归分析中，我们发现基线的严重程度 (P<0.0001)、使用 rt-PA 和年龄是判断预

表 患者基线特征

<table>
<thead>
<tr>
<th></th>
<th>溶栓组</th>
<th>非溶栓对照组</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>年龄，中位数 (范围)</td>
<td>71 (21–98)</td>
<td>N=1585</td>
<td>72 (21–101)</td>
</tr>
<tr>
<td>男性</td>
<td>880/1585</td>
<td>N=55.2%</td>
<td>2226/4232</td>
</tr>
<tr>
<td>基线 NIHSS，中位数 (范围)</td>
<td>14 (2–32)</td>
<td>N=1585</td>
<td>13 (2–37)</td>
</tr>
<tr>
<td>抗血小板药物史</td>
<td>429/1078</td>
<td>39.8%</td>
<td>446/1306</td>
</tr>
<tr>
<td>抗凝药物应用史</td>
<td>67/1078</td>
<td>6.2%</td>
<td>198/1306</td>
</tr>
<tr>
<td>既往卒中病史</td>
<td>319/1555</td>
<td>20.5%</td>
<td>1579/4076</td>
</tr>
<tr>
<td>充血性心衰</td>
<td>151/1262</td>
<td>12%</td>
<td>164/1409</td>
</tr>
<tr>
<td>糖尿病</td>
<td>342/1548</td>
<td>22.1%</td>
<td>992/3991</td>
</tr>
<tr>
<td>高血压</td>
<td>1030/1548</td>
<td>66.5%</td>
<td>2827/3991</td>
</tr>
<tr>
<td>房颤</td>
<td>398/1548</td>
<td>25.7%</td>
<td>1274/3991</td>
</tr>
<tr>
<td>心肌梗死</td>
<td>278/1548</td>
<td>18%</td>
<td>691/3991</td>
</tr>
</tbody>
</table>
后的重要预测因素。然后,在行年龄校正后的顺序逻辑回归分析中,我们发现卒中基线的严重程度 (P<0.0001) 及卒中病情的严重性与阿替普酶之间的相互作用与卒中预后相关,但是我们没有看到阿替普酶的独立作用 (P=0.65)。

依据这一相互作用试验,我们把卒中基线的严重程度分作7个NIHSS评分等级:1-4分,5-8分,9-12分,13-16分,17-20分,21-24分及≥25分,并着手验证在每一分级中溶栓治疗和预后的关联。

所有卒中基线严重程度分类的预后均有改善吗?

图2显示了年龄和基线 NIHSS校正分析后的功能结果。这实质上显示了基线 NIHSS评分在5到24分的患者使用阿替普酶后预后好转比较显著。基线期 NIHSS评分<4分的患者在90天内有着不同的预后,某些患者的 Rankin 分级显示有所好转,而有些则有所恶化。NIHSS 评分≥24分的患者显示使用阿替普酶后 Rankin 分级逐渐改善;但是,总体治疗效果不明显。

结果与神经病学的预后判断 (通过90天时进行的 NIHSS评分) 及灵敏度分析 (包括未校正的数据分析和对年龄、基线期 NIHSS评分、糖尿病及既往卒中病史校正后的分析) 相一致。由于在我们的神经保护治疗试验中,溶栓治疗的开始时间没有记录,所以我们不能校正起病到治疗开始的时间。数据记录中有59%缺乏抗血栓治疗的变化性数据 (N=3432),4.8%缺乏房颤病史的数据 (N=278),还有3.2%缺乏既往卒中病史的数据 (N=186)。本研究样本有这样局限性,不能在基线校正所有不同的变量 (年龄,基线 NIHSS 评分,既往抗血栓药物的使用,卒中史和房颤病史),影响了数据的可靠性。

讨论

轻中度卒中患者在药品随机试验及上市后分析中较少被关注。因此,欧洲药品评价局针对急性缺血性卒中,把轻度神经功能受损、在开始静脉用药前症状迅速恢复、以及经临床或适当的成像技术评估为严重卒中的患者都列为阿替普酶使用的禁忌症[2,23]。这类患者在去医院就医时,是否给其应用阿替普酶常常使医生处于进退两难之地。某些有经验的医师会选择溶栓来治疗这类患者。例如,在卒中治疗安全实施与国际卒中溶栓治疗登记 (SITS-ISTR) 记录的溶栓患者中,12%患者的基线 NIHSS评分0-4分,有4%的患者卒中较严重,NIHSS 评分≥25分。大多数这类患者采取了非溶栓治疗。加拿大的一系列试验发现有31%的患者被认为症状太轻或者症状恢复太快而不适宜溶栓治疗[19],而一份美国报告显示,NIHSS评分<8分患者仅有1/5选择溶栓[6]。现有结果表明此种做法不尽合理。而事后观察,被认为症状太轻而不需行溶栓治疗的患者中,有32%的患者在90天后或者死亡,或者致残[19]。其他研究也有类似报道[20,21]。为了明确卒中基线评分最轻与最重时溶栓的益处,开展随机试验可能会比较困难或者是执行起来比较拖延时间。其他的试验数据还有待验证,最好能有高质量的登记数据。
评分在1-4分与＞25分时，溶栓与否与结果的相关性不具统计学意义。尽管在两极的患者中，校正及非校正的优势比溶栓组倾向于比非溶栓组好，但低于其他卒中严重程度评分患者的改善程度。

处于两极的患者，因样本量小而严重降低了统计学检验效力，较宽的可信区间则严重影响了真实值评估的可信度。关于我们用的测量结果，有另外的统计方法可以考虑。在研究mRS的全方位分布时，我们应用检测法比起二分法，更少依赖于案例组合(case-mix)。对于卒中病情较轻或者较重的患者，我们可以应用相同的检验，并且可能仍会检测到溶栓治疗的益处。虽然如此，在两极的患者，测评结果通常不是太好就是太差，因此在对照组中，仅有小部分mRS分组具代表性。如果分类不能归纳分析，无论Coehran-Mantel-Haenszel检验，还是优势比率评估的检验效率都会打折扣。那些没有贡献的组群将会显著地影响治疗效果的检测。对于分层Coehran-Mantel-Haenszel检验而言，意味着将很难达到统计学上的显著性差异。尽管如此，对于优势比率检验来说，违背了最初的假定，因而产生的影响不成比例。解决问题并非易事。如果改变case-mix来提供一个有统计学意义的结果，那么轻微的或严重的卒中患者将被排除在外，这是一些试验的解决办法。相反的，如果根据样本case-mix的变化，而相应的测量结果也变化(Murray[24]等描述的二分法(the sliding dichotomy))，那解释结果将比较困难。比较mRS评分为0与1-6的优势比和mRS评分0-5与6的优势比，能是一样的吗？完全无卒中症状的患者和不惜任何代价抢救的幸存者能一样吗?

在这里，我们选择一种分析方法来分析卒中的严重性，但是我们也阐明了严重性两个极端的卒中患者的结果。为此，尽管总结式的统计信息并未显示出有统计学意义，而仅有趋向，我们仍可进一步得出结论。在严重卒中患者中，几乎所有mRS的边界点，都有朝着良好结果发展的趋势。在轻度卒中患者中，除去0-1与2-6，所有的边界点都显示出有利的一面，但是mRS评分为4的组无整体代表性。我们的数据显示，仅因为他们没有足够的证据来证明溶栓治疗的可取性，对于任一组患者，没有理由不选择溶栓治疗。

我们发现依赖于严格记录的临床试验数据及伴随的治疗及结果，并有精确的实地的数据校正程序，因而真实有效。尽管如此，对我们的试验设计来说，溶栓治疗与对照组的非随机分配，是一个明显弱点。我们不能从数据库中判定排除患者的程度及原因。我们只能考虑所知道的因素与预后的相关性。

我们校正了统计学上对结果影响大的某些因素。我们做了下面的比较：一方为卒中基线中等严重程度患者的溶栓治疗相关性，另一方为在随机试验中可比性患者中的治疗效果。我们发现基线NIHSS为9-12和13-16的患者溶栓有明显效果的治疗优势比为1.31-1.63。在评估一个随机对照试验时，3小时内溶栓的优势比为1.64，在3-4.5小时内优势比为1.34[25]。这说明我们的结果是可以比较的，可能还比较恒定。另外一个问题就是，延迟溶栓治疗的时间将会削减疗效。在现有的分析中，我们没有明确的信息证明应用阿替普酶溶栓治疗开始时间是否有延迟。因为参与VISTA的患者在试验中仅允许应用一种观察药物，阿替普酶的应用被设为标准治疗，并且这些试验均有生产商严密监控，因而我们假定大部分患者开始溶栓治疗都在发病3小时内。我们还假定从起病到溶栓治疗的时间间隔可以与CASES及SITS-MOST记录的数据相比(分别为155[130-175]分钟和140[115-165]分钟；n=6483)。遗憾的是，由卒中起病至记录起病时间之间的间隔，在溶栓组(3.7小时)及对照组(5.1小时)有所不同。卒中的严重程度与发病后到达医院的时间相关：越重的病人到达的时间越早[26]。尽管在分析中我们校正了卒中的严重程度，但是可以想到的其他偏差还会存在。这种偏差将会低估对照组中疾病初期的严重程度，并且使我们在基线校正时高估溶栓效果。这将影响不同卒中基线水平的所有患者，但对卒中基线严重程度的两极的影响可能不太明显：应用阿替普酶的大部患者NIHSS评分有严格的限制，NIHSS评分可能是不采用阿替普酶的原因。

有了这些警示，重复我们的发现将是值得的。一个即将进行的来自SITS-ISTR数据与VISTA对照组的对比分析可能成为有力证据。我们缺乏症状性出血患者的数据，因为通常不经历溶栓治疗的患者，在随访中不会常规行影像学检查以检测出血的转归。尽管如此，我们应用的检测指标考虑到了出血或者其他不利于神经功能事件的影响。

我们校正了年龄及基线的严重程度，因为这些是比较明确的影响结果的重要变量[27]。我们不能对所有年龄、基线NIHSS、先前应用抗血栓药物史、卒中病史及房颤等数据综合校正，主要是因为其中一个试验项目并没有记录治疗前用药。尽管如
此，我们仍开展了一项多变量校正分析，如校正在ECASS III试验中发现的很重要的变量：糖尿病及卒中病史，并且我们的评估保持前后一致。

在本研究中，某些患者仅接受一种调查研究的药物。每个有贡献的试验已经验证或排除了该药品与阿替普酶的重要相互作用，包括体内和体外试验。

结论
总之，我们的发现表明，尽管支持证据尚不够充分，对于NIHSS评分基线值处于两个极端的患者来说，溶栓治疗仍能使他们受益。

参考文献