

The Impact of Imbalances in Baseline Stroke Severity on Outcome in the National Institute of Neurological Disorders and Stroke Recombinant Tissue Plasminogen Activator Stroke Study

Thomas Kwiatkowski, MD
Richard Libman, MD
Barbara C. Tilley, PhD
Christopher Lewandowski, MD
James C. Grotta, MD
Patrick Lyden, MD
Steven R. Levine, MD
Thomas Brott, MD
and the National Institute of
Neurological Disorders and
Stroke Recombinant Tissue
Plasminogen Activator Stroke
Study Group

From the Department of Emergency Medicine (Kwiatowski) and the Department of Neurology (Libman), Long Island Jewish Medical Center, New Hyde Park, NY; the Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, SC (Tilley); the Department of Emergency Medicine, Henry Ford Hospital, Detroit, MI (Lewandowski); the Department of Neurology, University of Texas Medical Center, Houston, TX (Grotta); the Department of Neurology, University of California–San Diego, San Diego, CA (Lyden); the Department of Neurology, Mount Sinai School of Medicine, New York, NY (Levine); and the Department of Neurology, Mayo Clinic, Jacksonville, FL (Brott).

Study objective: The National Institute of Neurological Disorders and Stroke (NINDS) Recombinant Tissue Plasminogen Activator (rtPA) Stroke Study demonstrated a clinically meaningful and statistically significant benefit of tissue plasminogen activator (tPA). Adjusting for the baseline National Institutes of Health (NIH) Stroke Scale, the benefit of tPA remained. However, other authors suggest that an imbalance in baseline stroke severity between the tPA and placebo groups confounded the results. Another issue that has been raised concerns a possible increase in early mortality for individuals given tPA. In post hoc subgroup analysis, we describe the effect of tPA across a spectrum of time from stroke onset to treatment and stroke severity subgroups. Stroke severity was based on the NIH Stroke Scale. We also compare early mortality (2-week and 30-day) in the tPA and placebo groups.

Methods: Using combined data from the 2 NINDS rtPA Stroke Study trials, we performed post hoc subgroup analyses of 3-month favorable outcome (defined by the NIH Stroke Scale, Barthel, Rankin, and Glasgow Outcome Scales). We categorized patients from the trials into onset to treatment (0 to 90 minutes, 91 to 180 minutes) by NIH Stroke Scale (≤ 5 , 6 to 20, >20) subgroups. Analyses were adjusted for all variables previously shown to be associated with favorable outcome at 3 months. We also compared early mortality within onset-to-treatment subgroups.

Results: For all the 12 specified onset-to-treatment–NIH Stroke Scale subgroups, the adjusted odds ratio for a favorable 3-month outcome was greater than 1.0 and favored tPA. We detected no difference in mortality between patients treated with rtPA and those treated with placebo by 2 weeks posttreatment (rtPA=9%, placebo=13%; $P=.49$) or by 30 days (rtPA=11%, placebo=16%; $P=.30$).

Conclusion: These are descriptive post hoc subgroup analyses. Using cut points defined in previous critiques of the NINDS trials, these analyses give results consistent with previous NINDS Study Group reports. Baseline NIH Stroke Scale imbalance does not account for the better outcome of rtPA-treated patients. [Ann Emerg Med. 2005;45:377-384.]

Editor's Capsule Summary

What is already known on this topic

The National Institute of Neurological Disorders and Stroke (NINDS) Recombinant Tissue Plasminogen Activator (rtPA) Stroke Study demonstrated clinically significant benefit for patients treated with tissue plasminogen activator (tPA) within 3 hours of stroke onset, but baseline imbalances in the severity of stroke between the tPA and placebo groups contributed to controversy about the validity of the results.

What question this study addressed

What is the benefit of tPA across the spectrum of stroke severity in the NINDS study, and is there any difference in early stroke mortality between the tPA and placebo groups?

What this study adds to our knowledge

Although any post hoc study, such as this one, has limitations, there appears to be a benefit from tPA in patients with acute stroke, even accounting for differences in baseline stroke severity. Early mortality is not different between the 2 groups.

How this might change clinical practice

tPA treatment of acute stroke gives clinically significant benefit in patients across the range of stroke severity. The treating physician should consider the severity of neurologic deficits in the individual patient when making decisions about tPA treatment rather than using any set cut points on the NIH stroke scale.

SEE EDITORIAL, P. 385.

INTRODUCTION

Background

The National Institute of Neurological Disorders and Stroke (NINDS) Recombinant Tissue Plasminogen Activator (rtPA) Stroke Study demonstrated clinically and statistically significant benefit over placebo for patients treated with rtPA within 3 hours of onset of ischemic stroke.¹ More severely affected patients were enrolled more often in the 0- to 90-minute arm of the study than in the 91- to 180-minute arm. In addition, in the 91- to 180-minute cohort, the tissue plasminogen activator (tPA) group included more patients with less severe deficits and fewer patients with severe deficits. After adjustment for baseline National Institutes of Health (NIH) Stroke Scale, a greater benefit was demonstrated for patients in the 0- to 90-minute group (odds ratio [OR] 2.11; 95% confidence interval [CI] 1.33 to 3.35) than in the 91- to 180-minute group (OR 1.69; 95% CI 1.09 to 2.62), although the rtPA effect remains significant in the 91- to 180-minute stratum.²

However, because it has been shown that baseline stroke severity is a powerful predictor of outcome after stroke,³ others

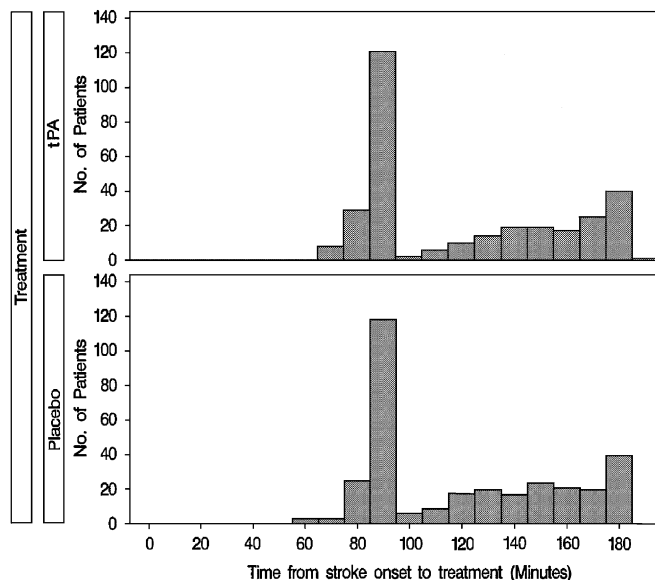


Figure 1. Histogram of number of patients per interval of time from stroke onset to start of treatment by treatment group (tPA, N=312, or placebo, N=312).

have continued to raise a question about whether the demonstrated benefits of rtPA, given within the 91- to 180-minute window, might be attributable to the baseline imbalance in randomization, rather than to a true biological effect of the drug.⁴

As previously reported, we did not detect a difference in mortality between the tPA and placebo groups at 3 months¹ or 1 year,⁵ despite the increased risk of intracerebral hemorrhage. Yet questions continue to be raised about a “tradeoff” between early mortality after tPA use and long-term improvement in functional outcome.

Importance

Given these criticisms about the use of rtPA for stroke, it is essential to show that the data in question do in fact support a beneficial effect of the drug, despite baseline imbalances in stroke severity.

Goals of This Investigation

We performed additional analyses on specific subgroups to explore outcome after tPA treatment to ascertain whether imbalances in these subgroups biased the results of the initial trial. Taking into account the limitations of subgroup analysis, our goal was to better understand the benefits of tPA across the spectrum of stroke severity in our study population. In addition, we explored whether early mortality differed in tPA versus placebo patients in the 0- to 90- and 91- to 180-minute strata.

METHODS

Study Design, Methods of Measurement, and Data

Collection and Processing

The NINDS rtPA Stroke Study consisted of 2 randomized clinical trials for patients with acute ischemic stroke randomized within 180 minutes from stroke onset.¹ The methodology and

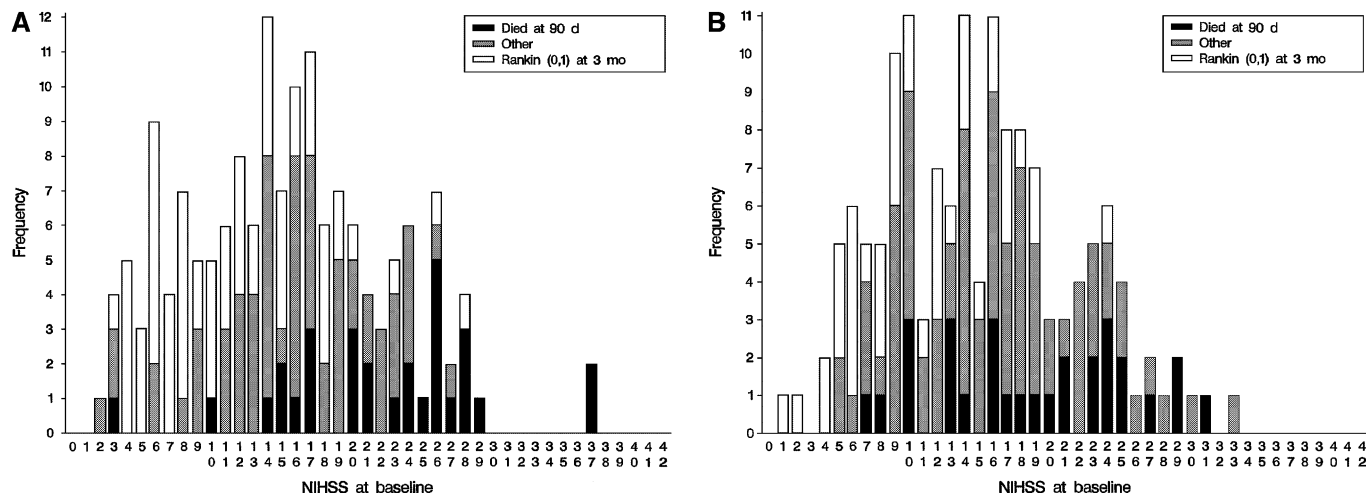


Figure 2. A, Distribution of 3-month favorable outcome defined as a Rankin Scale score of 0 or 1 and death at 90 days by baseline NIH Stroke Scale score intervals for patients treated with tPA and with time from stroke onset to treatment equal to 0 to 90 minutes. B, Distribution of 3-month favorable outcome defined as a Rankin Scale score of 0 or 1 and death at 90 days by baseline NIH Stroke Scale score intervals for patients treated with placebo and with time from stroke onset to treatment equal to 0 to 90 minutes. *NIHSS*, NIH Stroke Scale.

results of the NINDS rtPA Stroke Study have been published for 3-month¹ and 1-year outcomes.⁵ Randomization was stratified according to the time since onset of stroke (about half the patients were enrolled within 90 minutes after onset of symptoms and the remainder between 91 and 180 minutes). The protocols for the trials were approved by the institutional review board at each of the participating sites, and informed consent was obtained for all patients.

Setting

The study was conducted in 40 emergency departments across the country, 30 located in community hospitals.

Selection of Participants

Each trial enrolled patients with acute ischemic stroke, a measurable neurologic deficit on the NIH Stroke Scale, who

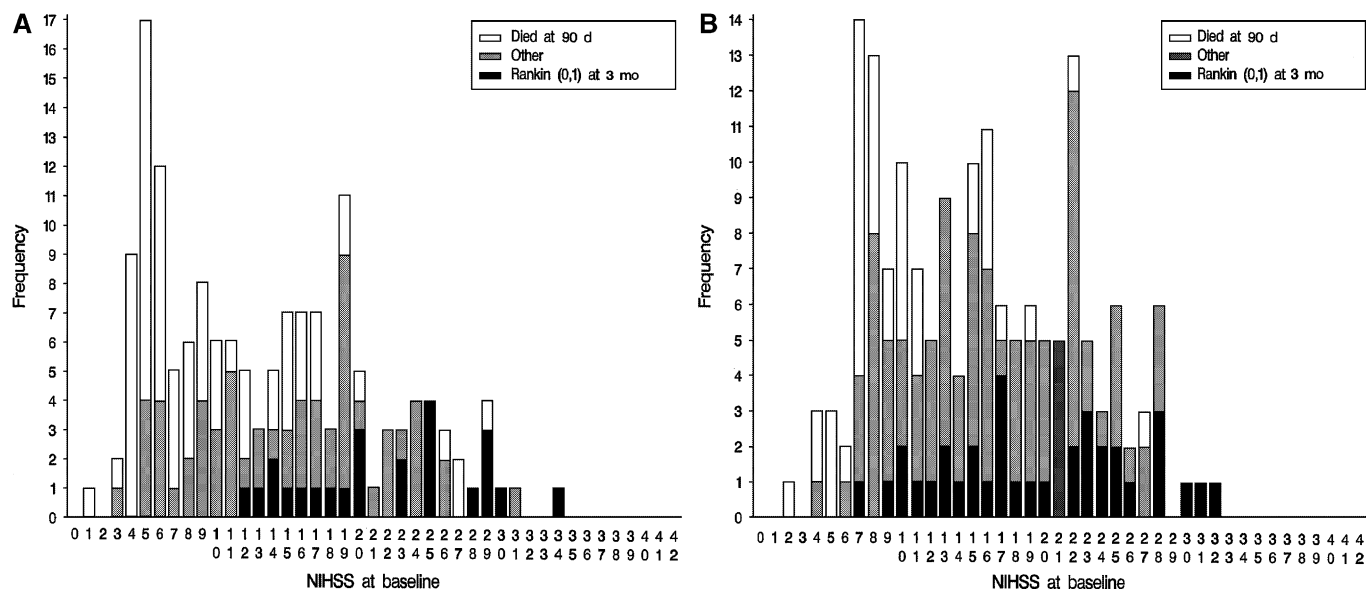


Figure 3. A, Distribution of 3-month favorable outcome defined as a Rankin Scale score of 0 or 1 and death at 90 days by baseline NIH Stroke Scale score intervals for patients treated with tPA and with time from stroke onset to treatment equal to 91 to 180 minutes. B, Distribution of 3-month favorable outcome defined as a Rankin Scale score of 0 or 1 and death at 90 days by baseline NIH Stroke Scale score intervals for patients treated with placebo and with time from stroke onset to treatment equal to 91 to 180 minutes.

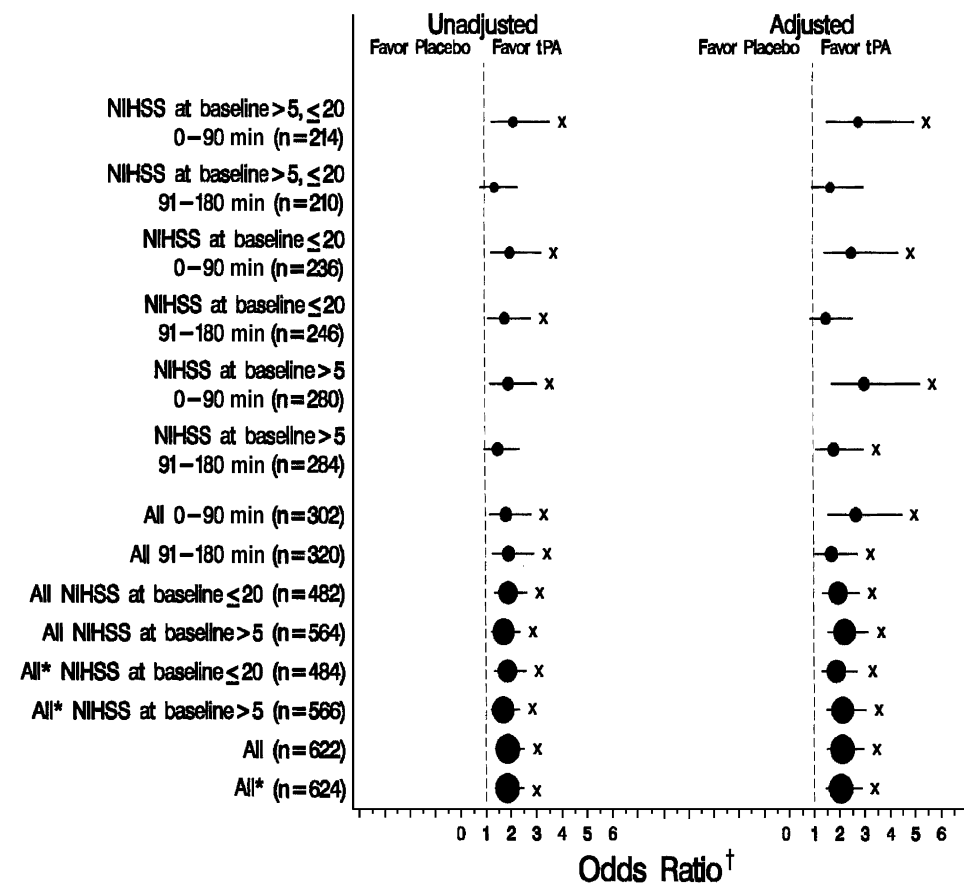


Figure 4. ORs for favorable 3-month outcome (global statistical test) for total population and selected subgroups in the NINDS rtPA Stroke Study. The sizes of the symbols (dots) are in proportion to the sizes of the populations of interest. × Indicates 95% CIs that do not include 1.0. *Included 2 patients who were randomized after 180 minutes from stroke onset. †ORs are adjusted for age, baseline NIH Stroke Scale score, admission mean arterial pressure, diabetes, early computed tomographic findings (edema, hypodensity, or intravascular thrombus), age by NIH Stroke Scale, age by admission mean arterial pressure, and center.

could be treated within 3 hours of stroke onset. More detailed eligibility criteria are given in the original article on the trials.¹

Selection of Participants for This Analysis

To increase our power to detect treatment differences in this analysis, we combined the data from the 2 NINDS trials for the present post hoc analysis. When outcomes within categories of time from stroke onset were assessed, 2 patients randomized after 180 minutes from stroke onset (protocol violations) were excluded.

Interventions

Patients in the NINDS rtPA Stroke Study received either rtPA or an identical placebo.

Methods of Measurement

Methods for measuring the 3-month outcomes are described in the original report on the trial.¹ Time of stroke onset was determined by interviewing patients and any available observers present when the stroke was first noticed.² Physicians sought

corroborating evidence (such as ambulance reports) and carefully screened for the possibility of onset during sleep. If the patient awoke with stroke symptoms, the time of onset was taken as the last time the patient was known to be awake and without any symptoms of stroke. If the onset time could not be established with confidence, the patient was not randomized. After randomization, the onset time was rereviewed using the emergency medical system logs and notes in the medical record. If discrepancies were found, corrections were made. It was through this review that 2 patients were identified who were randomized outside the 180-minute window. Audits of source records for all patients were performed to confirm consistency of stroke onset time and treatment start time determinations. The admission blood pressure was a single measurement made at patient triage, admission to the emergency department, admission to the hospital, or admission to the ICU.

The NIH Stroke Scale used to assess stroke severity has been shown to be a valid and reliable scale for this purpose.⁶ An NIH Stroke Scale score less than 5 was used to indicate the group with the least severe strokes, and a cut point of greater than 20 was used

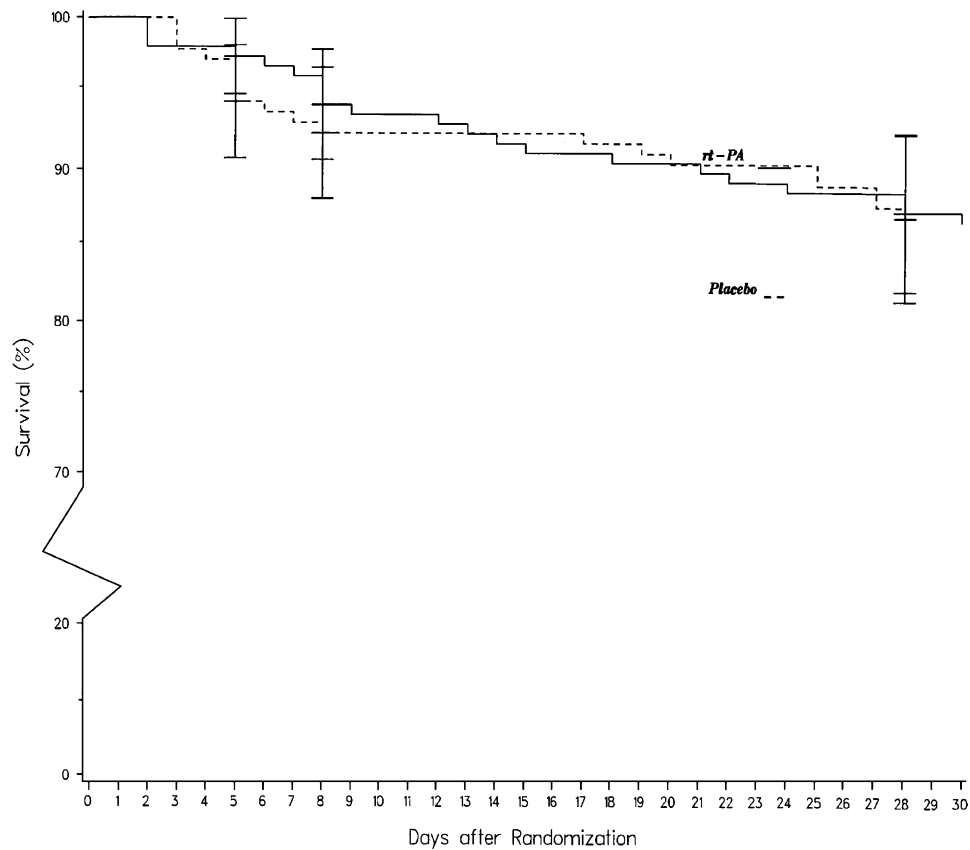


Figure 5. Kaplan-Meier survival curves during the first 30 days after randomization by treatment group (rtPA and placebo) for group randomized with time from stroke onset to treatment equal to 0 to 90 minutes. CIs are calculated at the time an event occurs. Those CIs are presented for events occurring closest to the 2-week, 30-day time points of interest in the treatment group.

for the most severe strokes. All study investigators were certified in proper use of the NIH Stroke Scale and recertified periodically.

Data Collection and Processing

These processes are described in the original article on the trial.¹

Outcome Measures

In this analysis, we used the primary outcome from the second NINDS rtPA Stroke Study trial (ie, favorable outcome at 90 days).¹ Favorable outcome at 3 months was defined as minimal or no disability according to combined information from the dichotomized NIH Stroke Scale, Barthel Index, modified Rankin Scale, and Glasgow Outcome Scale. Minimal or no disability was defined separately for each scale as a 0 or 1 on the NIH Stroke Scale, 95 or 100 on the Barthel Index, 0 or 1 on the Rankin Scale, and 1 on the Glasgow Outcome Scale. We also used 2-week, 30-day, and 90-day mortality as outcomes for some analyses.

Primary Data Analysis

We used histograms to describe the distribution of time from stroke onset to treatment within onset-to-treatment subgroups

of 0 to 90 minutes and 91 to 180 minutes. We used stacked histograms to describe the distribution of baseline NIH Stroke Scale by onset-to-treatment subgroups along with favorable 3-month outcome on the Rankin Scale (0, 1) and 90-day mortality. The global statistical test, used to make comparisons within the onset-to-treatment/NIH Stroke Scale subgroups, could not be used in constructing the histograms because it does not classify individuals as successes or failures.

Outcomes at 3 months were analyzed within subgroups according to a cross-classification of baseline stroke severity as measured by the NIH Stroke Scale at baseline (<5, 6 to 20, and >20) and onset to treatment (0 to 90 minutes, 91 to 180 minutes). The chosen cut points had been proposed by others who were questioning the results of the trial.⁴ These cut points represent the areas of most imbalances between the 2 groups within a time stratum.

To combine information across the 4 measures of outcome, we used a global test statistic. From the global statistical approach, we computed an OR for a favorable outcome at 3 months to assess treatment differences in favorable outcome within each subgroup. The global statistical test (a Wald test) is derived from a generalized linear model with logit-link function, computed with the use of generalized estimating equations.^{7,8} Computations for

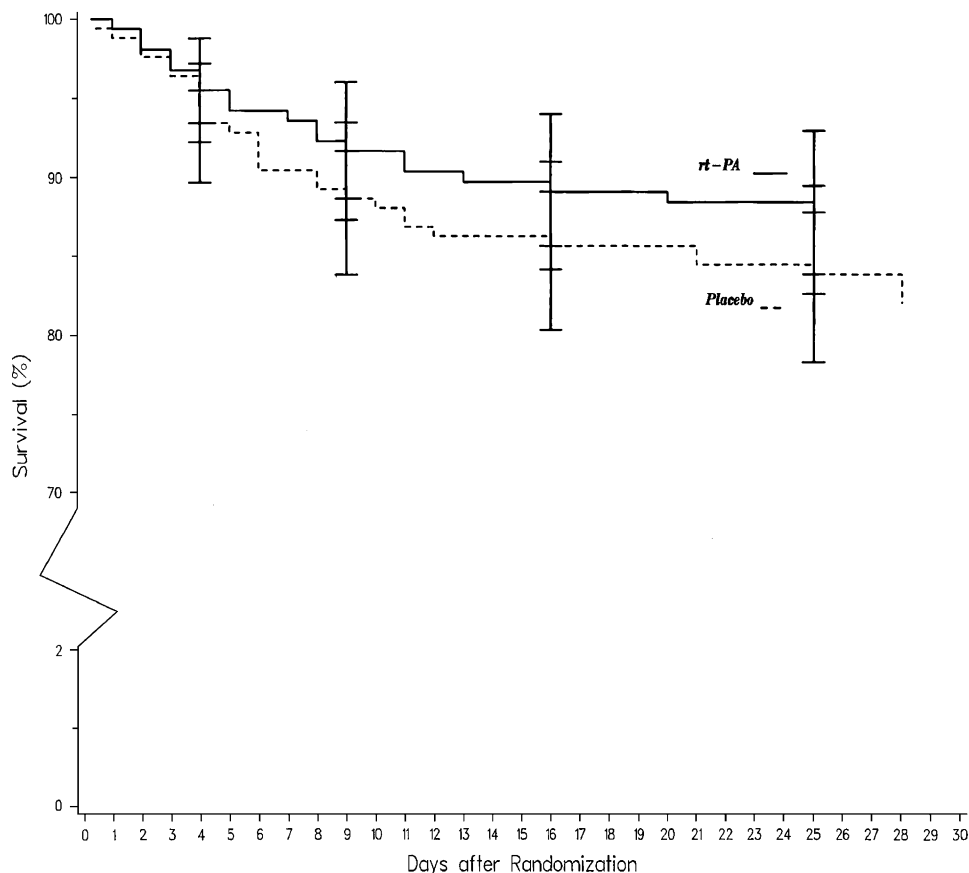


Figure 6. Kaplan-Meier survival curves during the first 30 days after randomization by treatment group (rtPA and placebo) for group randomized with time from stroke onset to treatment equal to 91 to 180 minutes. CIs are calculated at the time an event occurs. Those CIs are presented for events occurring closest to the 2-week, 30-day time points of interest in the treatment group.

the global statistical test take the correlations among outcomes into account.⁹ In computing the global test statistic, we also adjusted for all covariates previously shown to be associated with favorable outcome at 3 months¹⁰ or covariates used as strata in the randomization for the original trials. Covariates included age, baseline NIH Stroke Scale, admission mean arterial blood pressure, diabetes, early computed tomographic findings (edema, hypodensity, or intravascular thrombus), age by NIH Stroke Scale, age by admission mean arterial blood pressure, and center. Onset to treatment, another stratifying variable in the randomization, was not included as a covariate, because we were analyzing onset-to-treatment subgroups. We used a Kaplan-Meier approach to computing survival curves and CIs and a log-rank test to compare survival curves between treatment groups. The software package used for all analyses was SAS version 8.0 (SAS Institute, Inc., Cary, NC).

RESULTS

The distribution of onset to treatment is shown in Figure 1. Figures 2A, 2B, 3A, and 3B give the distribution of 3-month favorable outcome (according to a Rankin Scale score of 0 or 1) and mortality at 90 days by onset-to-treatment/NIH Stroke

Scale subgroups. From Figures 3A and 3B, when an arbitrary cut point of NIH Stroke Scale score less than or equal to 5 is chosen, an imbalance is noted in the proportion of patients with these lower NIH Stroke Scale scores who were randomly assigned to tPA versus placebo (tPA=19%; placebo=4.2%). When other cut points, still reflecting the least severe stroke severity, are chosen, the degree of imbalance is not consistent. For example, for NIH Stroke Scale score less than or equal to 4, tPA equals 3.8% and placebo equals 1.3%. For NIH Stroke Scale score less than or equal to 7, tPA equals 14.4% and placebo equals 7.2%.

After adjustment for potential confounding variables for favorable outcome at 3 months, the ORs for a favorable outcome were greater than 1.0 for all 12 of the onset-to-treatment/NIH Stroke Scale subgroups (Figure 4), favoring rtPA in all subgroups.

There was no detectable difference in mortality between tPA-treated patients and patients given placebo during the first 2 weeks after tPA treatment (tPA=9%, placebo=13%; $P=.49$) or at 30 days (tPA=11%, placebo=16%; $P=.30$; Figures 5 and 6).

LIMITATIONS

This is a descriptive post hoc analysis of data from a randomized clinical trial, and we have made many comparisons.

Some comparisons could be significant by chance alone, although there is a consistent benefit for tPA across all the subgroups analyzed (Figure 4).

We also have limited power to detect differences in the subgroups, and the study was not powered to detect differences in mortality. Thus, when we failed to find differences, low power might have been the cause. The cut point of less than or equal to 5 to designate the least severe stroke severity is arbitrary and was based on a previous critique.⁴ As shown in Figures 2A, 2B, 3A, and 3B, any imbalance in the proportion of tPA versus placebo patients varies considerably, depending on the cut point chosen. This effect demonstrates the hazards of choosing arbitrary subsets of the study population post hoc and inferring that imbalances in these subgroups affect the main findings of the study. Previous analyses adjusting for the baseline NIH Stroke Scale as a continuous variable gave results consistent with the results reported here,² as shown in Figure 4.

DISCUSSION

The baseline imbalance in stroke severity in the 91- to 180-minute arm of the NINDS rtPA Stroke Study does not appear to account for the better outcome of tPA-treated patients. When the difference in baseline stroke severity in the 91- to 180-minute group is controlled for, a greater odds of a favorable outcome for tPA-treated patients persists.² In addition, using all NINDS patients and after excluding patients with either the least severe or the most severe strokes, as defined on the NIH Stroke Scale, there is still a greater odds of a favorable outcome for tPA-treated patients (Figure 4). Grotta¹¹ has published further information on the numbers of patients and their outcomes in smaller NIH Stroke Scale intervals for the Rankin Scale score alone in an editorial on behalf of the NINDS rtPA Stroke Study Group. We did not perform adjusted global statistical tests within these small subsets, because the number of patients within a subset was too small to do statistical adjustments, given that such a large number of covariates was used in making the adjustment. However, there was no subgroup based on smaller, categorized NIH Stroke Scale cut points in which there was any suggestion that outcome was worse with tPA. Occasionally, patients can have disabling symptoms (eg, aphasia, hemianopia) and low NIH Stroke Scale scores. In addition, our endpoint of “minimal or no disability” may overlook, or make it difficult to document, benefits to patients at the lower end of the stroke scale. Therefore, the treating physician must consider the possible seriousness of the patient’s disability, rather than relying on an absolute NIH Stroke Scale score or cut point when making a decision to use tPA. As Yusuf et al¹² have cautioned, the overall “average” result of a randomized clinical trial is usually a more reliable estimate of treatment effect in the various subgroups examined than are the observed effects in individual subgroups.

We have previously reported no detectable difference in long-term mortality between rtPA and placebo patients.⁵ Here we also report no detectable difference in early mortality between the 2 groups. We wish to emphasize, therefore, that

there appears to be no identified “tradeoff” between long-term improvement in functional outcome and early mortality despite an increased risk of intracerebral hemorrhage among tPA patients. We detected no significant increase in mortality at any point from the initiation of treatment to 1 year.^{1,5}

As previously published, the benefit of rtPA appears to decline with time from stroke onset to treatment during the 3-hour window, emphasizing the importance of treating patients as early as possible,² although even patients treated later within the 3-hour window stand to benefit from treatment. We recommend, therefore, that in hospitals that have the appropriate resources to adhere to the NINDS protocol, patients receive treatment as early as possible to be given the opportunity for a better outcome with tPA treatment.

Author contributions: TK, RL, BCT, CL, JCG, PL, SRL, and TB were all involved with the design, analysis, and preparation of the manuscript. TK takes responsibility for the paper as a whole.

Funding and support: Supported in part by NIH/NINDS contracts NO1-NS-23373, NO1-NS-02374, NO1-NS-02377, NO1-NS-02381, NO1-NS-02379, NO1-NS-02378, NO1-NS-02376, NO1-NS-02382, and NO1-NS-02380.

Publication dates: Received for publication June 18, 2003. Revisions received December 10, 2003; April 20, 2004; and June 29, 2004. Accepted for publication June 30, 2004. Available online January 11, 2005.

Address for reprints: Thomas Kwiatkowski, MD, Department of Emergency Medicine, Long Island Jewish Medical Center, 270-05 76th Avenue, New Hyde Park, NY 11040; 718-470-7501, fax 718-470-9113; E-mail kwiatkow@lij.edu.

REFERENCES

1. The National Institute of Neurologic Disorders and Stroke rtPA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med.* 1995;333:1581-1587.
2. Marler JR, Tilley BC, Lu M, et al, for the NINDS rtPA Stroke Study Group. Early stroke treatment associated with better stroke outcome: the NINDS rtPA Stroke Study. *Neurology.* 2000;55:1649-1655.
3. Adams HP, Davis PH, Leira EC, et al. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: a report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). *Neurology.* 1999;53:126-131.
4. Mann J. Truths about the NINDS study: setting the record straight. *West J Med.* 2002;176:192-194.
5. Kwiatkowski TG, Libman RB, Frankel M, et al. Effects of tissue plasminogen activator for acute ischemic stroke at one year: National Institute of Neurological Disorders and Stroke Recombinant Tissue Plasminogen Activator Stroke Study Group. *N Engl J Med.* 1999;340:1781-1787.
6. Lyden P, Brott T, Tilley B, et al. Improved reliability of the NIH Stroke Scale using video training. *Stroke.* 1994;25:2220-2226.
7. Lefkopoulou M, Moore D, Ryan L. The analysis of multiple correlated binary outcomes: application to rodent teratology experiments. *J Am Stat Assoc.* 1989;84:810-815.
8. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42:121-130.
9. Tilley BC, Marler J, Geller NC, et al. Use of a global test for multiple outcomes in stroke trials with application to the National

- Institute of Neurological Disorders and Stroke rtPA Trial. *Stroke*. 1996;27:2136-2142.
10. The National Institute of Neurological Disorders and Stroke rtPA Stroke Study Group. Generalized efficacy of tPA for acute stroke: subgroup analysis of the NINDS tPA Stroke Trial. *Stroke*. 1997; 28:2119-2125.
 11. Grotta C. The NINDS stroke study group response. *J Stroke Cerebrovasc Dis*. 2002;11:121-124.
 12. Yusuf S, Wittes J, Probstfield J, et al. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA*. 1991; 226:93-98.