Over-EXTENDing the Window for Thrombolytic Therapy in Cerebrovascular Accident

September 2019 Annals of Emergency Medicine Journal Club

Guest Contributors

Rory J. Spiegel, MD; John P. Donnelly, PhD; Ryan P. Radecki, MD, MS

0196-0644/$-see front matter

Copyright © 2019 by the American College of Emergency Physicians.

https://doi.org/10.1016/j.annemergmed.2019.07.010

Editor's Note: *You are reading the 74th installment of* Annals of Emergency Medicine *Journal Club. As the Journal Club enters its second decade of publication, the format has been revised and will focus on a monthly succinct review of high-impact articles from this journal and other premier medical journals relevant to emergency medicine. The reviews are followed by questions demonstrating principles by which readers—be they clinicians, academics, residents, or medical students—may critically appraise the literature. We are interested in receiving feedback about this feature. Please e-mail journalclub@acep.org with your comments.*

ARTICLE IN REVIEW

Ma H, Campbell BCV, Parsons MW, et al. Thrombolysis guided by perfusion imaging up to 9 hours after onset of stroke. *N Engl J Med*. 2019;380:1795-1803.

What Question Did This Investigation Aim to Answer?

In patients presenting with acute ischemic stroke between 4.5 and 9 hours after symptom onset and salvageable brain tissue on perfusion magnetic resonance imaging (MRI), did the use of intravenous alteplase improve functional outcomes at 90 days?

What Study Design Did the Authors Choose?

Design: A multicenter, randomized, placebo-controlled trial. ClinicalTrials.gov identifier: NCT00887328 and NCT01580839.

Setting: Twenty-eight centers across Australia, Asia, and Europe.

Population: Two hundred twenty-five adult patients with ischemic stroke who were treated between 4.5 and 9 hours from onset of symptoms, or adjusted onset if awakening from sleep, and feasible ischemic penumbra detected on automated cerebral perfusion imaging.

Intervention: Alteplase (0.9 mg/kg) or matching placebo.

Primary and Secondary Outcomes: The primary outcome was the proportion of patients with a modified Rankin Scale score of 0 or 1 at 90 days. Secondary outcomes included an ordinal analysis of modified Rankin Scale score status at 90 days and percentages of reperfusion. Safety outcomes included death and rate of symptomatic intracranial hemorrhage.

Sponsor: Boehringer Ingelheim provided alteplase and matching placebo.

How Did the Authors Interpret the Results?

The trial was terminated early, after enrollment of 225 of the planned 310 patients. No statistically significant difference was identified in the unadjusted primary outcome with a modified Rankin Scale score of 0 or 1 at 90 days occurring for 35.4% of patients in the alteplase group and 29.5% in the placebo group (risk ratio [RR] 1.2; 95% confidence interval [CI] 0.82 to 1.76; P=.35). In accordance with the prespecified analysis plan, modified Poisson regression was used to estimate an RR adjusted for age and National Institutes of Health Stroke Scale score. In the adjusted analysis, the association reached statistical significance favoring the alteplase group (adjusted RR 1.44; 95% CI 1.01 to 2.06; P=.04). No differences were observed in the ordinal analysis or 90-day mortality. Symptomatic intracranial hemorrhage within 36 hours occurred in 6.2% of the alteplase group and 0.9% of the control group (adjusted RR 7.22; 95% CI 0.97 to 53.54; P=.053).

Conclusion: In patients meeting eligibility criteria for the trial, alteplase administration resulted in no or minor neurologic deficits more often than use of placebo. These results should be interpreted cautiously owing to premature termination of the trial.

How Might This Study Affect Your Clinical Practice in the Emergency Department?

The results of this trial may lead to consideration of alteplase for a subset of patients with acute ischemic stroke in the 4.5- to 9-hour window. However, most patients enrolled also manifested large-vessel occlusions and would otherwise have been treated with endovascular therapy. The addition of this narrow scope for alteplase, if any, would

likely depend on available resources and individual patient features.

DISCUSSION POINTS

1. *When examining their primary outcome, the authors used an adjusted analysis. What is an adjusted analysis, and why might such a strategy have been used?*

Randomization is a tool used to control for bias resulting from confounding of a treatment-outcome association. The underlying concept relies on random distribution of measured and unmeasured confounders with respect to treatment status. Observed differences in outcomes can then reasonably be thought to be caused by the intervention tested.

Although control of confounding can be achieved in randomized studies with limited sample sizes, deviations from random sampling in small trials can lead to imprecise estimates, inefficiency, and exaggerated effect sizes.[1] This type of error is best ameliorated by increasing the size of the sample in question; thus, smaller trials may not be definitive. For additional information on the relationship between sample size and random error, see the May 2008 Journal Club.[2] The CI surrounding the unadjusted RR for improved function at 90 days (RR 1.2; 95% CI 0.82 to 1.76; P=.35) observed in the Extending the Time for Thrombolysis in Emergency Neurological Deficits (EXTEND) trial may reflect this problem. An adjusted analysis, such as the strategy used by the EXTEND authors, attempts to improve estimation of the treatment effect. A number of simulation studies have demonstrated that adjusting for prespecified factors strongly associated with the study outcome can result in more precise and potentially less biased estimates compared with unadjusted analyses.[3-8] Adjusted analyses carried out in this manner are now recommended by experts and the most recent Consolidated Standards of Reporting Trials guidelines.[9] For adjusted analyses to be valid, several conditions must hold: the authors should explain their rationale for choosing to adjust for baseline variables; they should prespecify variables for adjustment according to known influence on the primary outcome; and variables should not be selected retrospectively according to baseline differences noted to be statistically significant.

In this case, the authors adjusted for age and National Institutes of Health Stroke Score at baseline because both are strong prognostic variables for patients with ischemic stroke.

2. *The trial planned to enroll 310 patients but was stopped prematurely because of what the authors describe as "loss of equipoise." How might this early stoppage affect the trial's results?*

As trial enrollment proceeds and samples enlarge, observed variations in estimates may decrease in amplitude, eventually approaching the true treatment effect. When trials are stopped before reaching their intended sample size, the observed results may not reflect the true effect. For example, trials stopped early for benefit are likely to overestimate the effect size of the treatment in question. In fact, the smaller the sample size at closure, the greater the potential for amplification.[10-12]

The authors initially estimated 400 participants would be required for the trial. However, based on an interim assessment of the amount of missing data and loss to follow-up, the final intended enrollment was revised to 310. In terms of power and sample size, the authors stated this number "would provide the trial 80% power to detect a between-group difference of 15 percentage points in the primary outcome (36% in the alteplase group and 21% in placebo group) at a two-sided significance level of p=0.05." In May 2018, after enrollment of 225 patients, the results of the Efficacy and Safety of MRI-Based Thrombolysis in Wake-Up Stroke (WAKE-UP) trial were published.[13] The results suggested the potential of benefit for an MRI-guided approach to the use of alteplase. The data and safety monitoring board recommended the discontinuation of the EXTEND trial because of loss of equipoise.

At stoppage for the EXTEND trial, a total of 113 participants were included in the alteplase group and there were 112 in the placebo group. This implies 42 and 43 participants were to be enrolled in the alteplase and placebo groups, respectively. If this trial were run to its original attending sample size, what would be the probability the effect size would reach the designated threshold for significance?

Because the trial had no loss to follow-up and there were no planned formal interim analyses, the remaining enrollment can be modeled with a series of random binomial variates with assumed outcome distributions. To evaluate the sensitivity of the results, we show the distribution of possible effect sizes as absolute differences for 100,000 simulated trials, assuming 3 different underlying outcome proportions for the alteplase and placebo groups. Our specifications of outcome proportions for the remaining participants were the current effect size observed in the 225 patients enrolled in the EXTEND cohort (–alteplase=0.36 and placebo=0.29); the expected effect size proposed by the authors in their original power calculation (–alteplase=0.36 and placebo=0.21); and an even greater effect size (–alteplase=0.42 and placebo=0.15) (Figure 1). Histograms for P values from a test for differences in proportions are shown in Figure 2. For conditions consistent with the results at stoppage, the

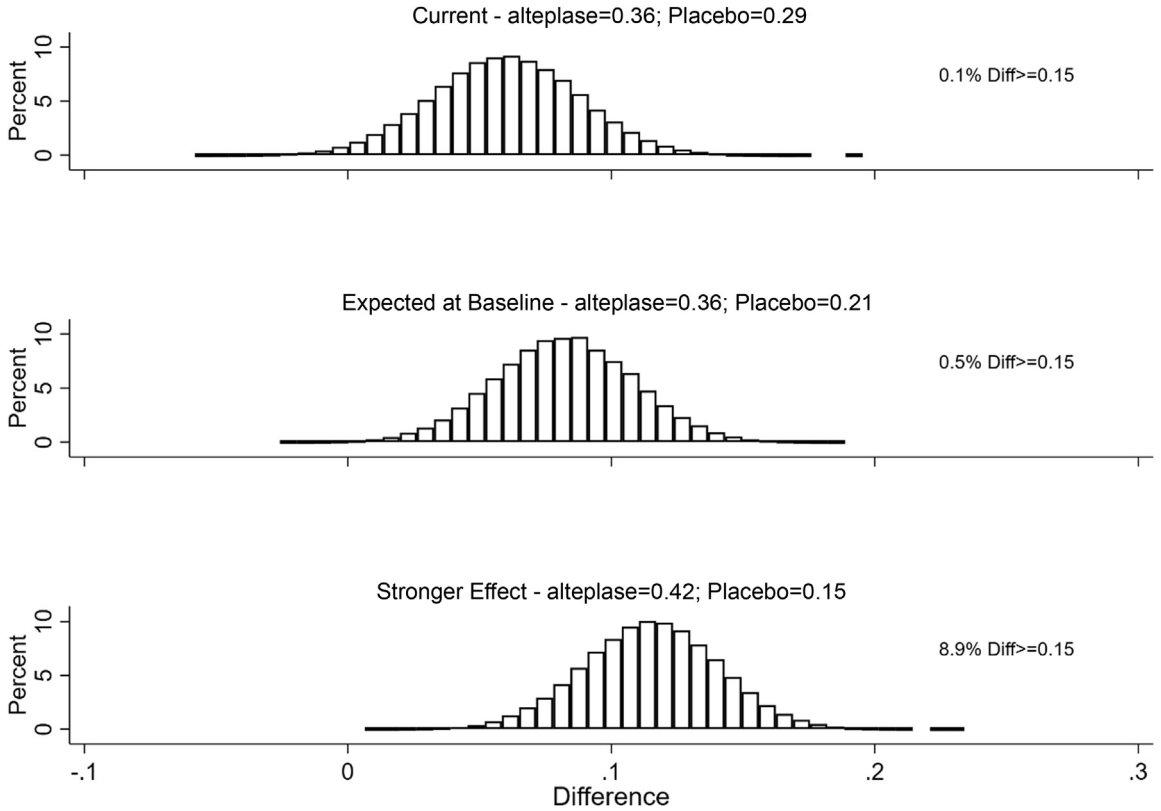

Figure 1. Distribution of absolute differences at trial end by assumed outcome proportions for remaining enrollment.

median absolute difference at the end of the trial was 6% (25th to 75th percentile 5% to 8%), with only 0.1% of simulated trials having an effect size at or above the clinically meaningful difference described at baseline (15% absolute risk reduction). Only 6.1% had a statistically significant *P* value at the .05 level. Under the conditions expected at baseline, the median absolute difference at the end of the trial was 8% (25th to 75th percentile 6% to 10%), with only 0.5% of simulated trials having an effect size of at least 15%. Only 23.8% had a statistically significant *P* value at the .05 level. For the condition with a stronger effect for the remaining participants, the median effect was 12% (25th to 75th percentile 10% to 14%), with 8.9% having an effect size of 15% or greater; 71.7% had a statistically significant *P* value at the .05 level.

3. *How should this early stoppage and the use of an adjusted analysis affect our interpretation of the trial results?*

EXTEND was a modest-sized trial enrolling a total of 225 patients. This trial was stopped early because of loss of equipoise after the publication of the WAKE-UP trial,[13] which was also stopped early and found a signal of benefit in favor of intravenous alteplase. As discussed in question 2, there may be substantial variability in treatment effects for small trials, and in some cases this can result in exaggerated estimates. The effect size observed in EXTEND between the alteplase and control groups at stoppage was a small, 5.9% absolute difference. This difference was not statistically significant when an unadjusted analysis was used. However, the simulations above show that if enrollment had continued to the specified sample size of 310, there could have been some variability in the results. As an example, for the situation in which the effect size among the remaining participants was similar to that observed for the 225 enrolled patients, only 0.1% of simulated trials had an unadjusted effect size greater than 0.15 and only 6.1% demonstrated a $P<.05$.

The fact that the adjusted analysis found a statistically significant *P* value with a higher-magnitude treatment effect compared with the unadjusted analysis suggests that some amount of imprecision and inefficiency may be present. In the absence of randomization failures, when adjustment variables are prespecified and there is strong theoretic support for associations with outcomes, effect estimates from adjusted analyses in clinical trials are likely more trustworthy than unadjusted estimates. However, this does not absolve the clinician of the need for careful interpretation of results when considering incorporation

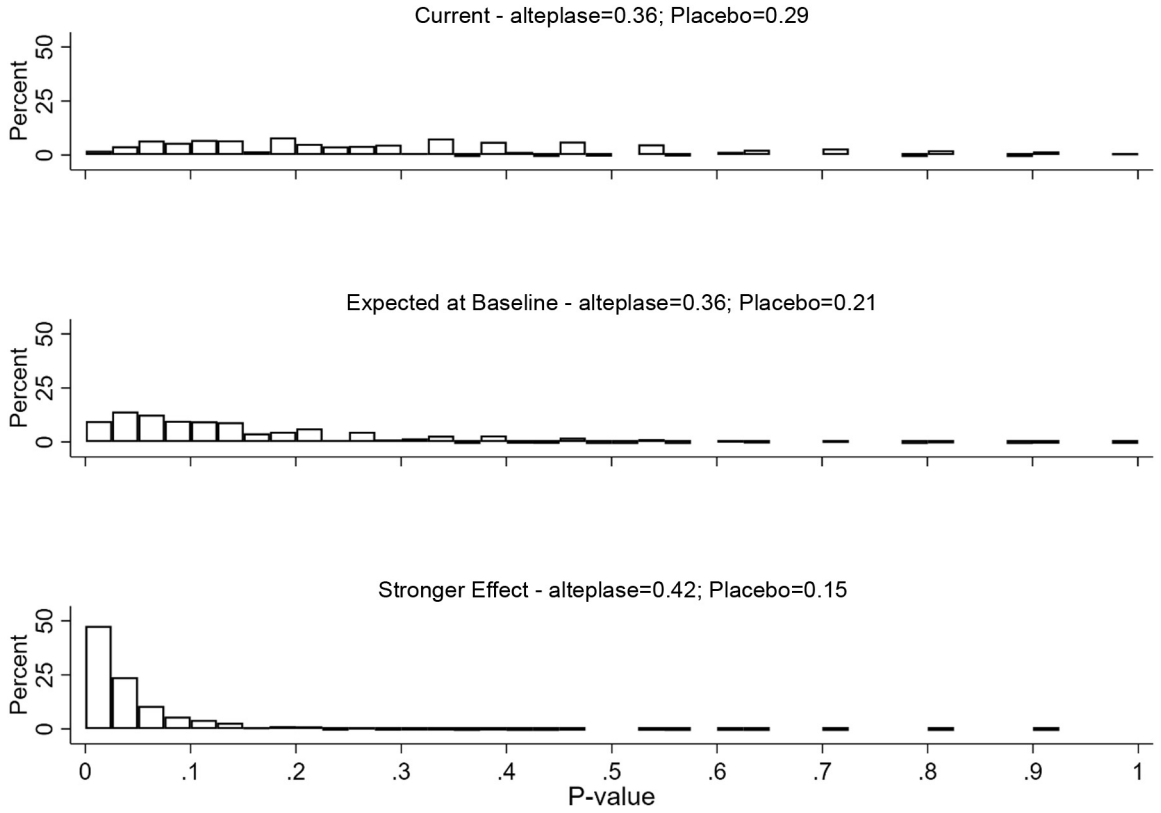

Figure 2. Distribution of *P* values for differences in proportions at trial end by assumed outcome proportions for remaining enrollment.

into everyday practice. There are numerous examples of clinical trials demonstrating benefit when an adjusted analysis was used, only to have their results reversed when larger, more robust trials were performed.[14-17] It may be inaccurate to indict adjusted analyses as the main issue for such reversals, but it does bring to light how important it is to be skeptical about all randomized interventional studies and interpret trial results in the appropriate broader context.

For the EXTEND trial, little previous evidence exists to support the plausibility of the observed effect size. Whether one examines the adjusted or unadjusted analysis, the interpretation should be the same. Although interesting, these results require further validation before being incorporated into clinical practice.

Section editors: Tyler W. Barrett, MD, MSCI; Ryan P. Radecki, MD, MS; Rory J. Spiegel, MD

Author affiliations: From the Department of Emergency Medicine, Washington Hospital Center, Washington, DC (Spiegel); the Department of Learning Health Sciences, Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI (Donnelly); and the Department of Emergency Medicine, Northwest Permanente, Portland, OR, and The University of Texas Health Science Center at Houston, Houston, TX (Radecki).

Authorship: All authors attest to meeting the four ICMJE.org authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). Dr. Donnelly was supported by the National Heart, Lung, and Blood Institute (no. K12-HL138039).

REFERENCES

1. Altman DG, Bland JM. Uncertainty and sampling error. *BMJ*. 2014;349:g7064.
2. Barrett TW, Schriger DL. *Annals of Emergency Medicine* Journal Club. Acutely decompensated heart failure in a county emergency department: a double-blind randomized controlled comparison of nesiritide versus placebo treatment. Answers to May 2008 Journal Club questions. *Ann Emerg Med*. 2008;52:458-472.
3. Neuhaus JM. Estimation efficiency with omitted covariates in generalized linear models. *J Am Stat Assoc*. 1998;93:1124-1129.

4. Hauck WW, Anderson S, Marcus SM. Should we adjust for covariates in nonlinear regression analyses of randomized trials? *Control Clin Trials*. 1998;19:249-256.
5. Steyerberg EW, Bossuyt PMM, Lee KL. Clinical trials in acute myocardial infarction: should we adjust for baseline characteristics? *Am Heart J*. 2000;139:745-751.
6. Hernandez AV, Steyerberg EW, Habbema JDF. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol*. 2004;57:454-460.
7. Gray LJ, Bath P, Collier T. Should stroke trials adjust for functional outcome for baseline prognostic factors? *Stroke*. 2009;40:888-894.
8. Kent DM, Trikalinos TA, Hill MD. Are unadjusted analyses of clinical trials inappropriately biased toward the null? *Stroke*. 2009;40:672-673.
9. Moher D, Hopewell S, Schulz K, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
10. Pocock SJ, Hughes MD. Practical problems in interim analyses, with particular regard to estimation. *Control Clin Trials*. 1989;10(suppl 4):209-221S.
11. Guyatt GH, Briel M, Glasziou P, et al. Problems of stopping trials early. *BMJ*. 2012;344:e3863.
12. Bassler D, Briel M, Montori VM, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303:1180-1187.
13. Thomalla G, Simonsen CZ, Boutitie F, et al. MRI-guided thrombolysis for stroke with unknown time of onset. *N Engl J Med*. 2018;379:611-622.
14. Papazian L, Forel JM, Gacouin A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med*. 2010;363:1107-1116.
15. Moss M, Huang DT, Brower RG, et al. Early neuromuscular blockade in the acute respiratory distress syndrome. *N Engl J Med*. 2019;380:1997-2008.
16. Annane D, Sébille V, Charpentier C, et al. Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic shock. *JAMA*. 2002;288:862-871.
17. Venkatesh B, Finfer S, Cohen J, et al. Adjunctive glucocorticoid therapy in patients with septic shock. *N Engl J Med*. 2018;378:797-808.

IMAGES IN EMERGENCY MEDICINE

(continued from p. 454)

DIAGNOSIS:

Left atrial tamponade caused by rupture of the descending thoracic aortic dissection. The left atrium was externally compressed by the hematoma propagating from the dissection, a rare but typically fatal complication of aortic dissection.[1-3]

Despite planned emergency surgical intervention, the patient experienced cardiac arrest and was unable to be resuscitated.

Author affiliations: From the Department of Emergency Medicine (C.-H. Lin, Li, Tsai, Yang) and Division of Cardiovascular Surgery, Department of Surgery (C.-Y. Lin), Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan.

REFERENCES

1. Bonser RS, Mohammed Bashir M, Ranasinghe AM. Left atrial compression secondary to contained rupture of type A aortic dissection. *Circulation*. 2011;123:e592-e593.
2. Gandhi NM, Greaves M, Brooks NH. Rare case of heart failure caused by compression of the left atrium by a thoracic aortic aneurysm. *Heart*. 2004;90:e9.
3. Corré J, Chassaing C, Blanchard D. An unusual cause of cardiogenic shock: left atrial compression by a spontaneous rupture of an aneurysm of the descending thoracic aorta. *J Cardiol Cases*. 2017;15:153-154.