Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Alper BS, Foster G, Thabane L, Rae-Grant A, Malone-Moses M, Manheimer E. Thrombolysis with Alteplase 3 to 4.5 Hours after Acute Ischemic Stroke: Trial Reanalysis Adjusted for Baseline Imbalances.

SUPPLEMENTARY APPENDIX

Table of Contents

Protocol for ECASS III Reanalysis

Here is the Research Plan from the Data Sharing Agreement used to access the data for the reanalysis of the European Cooperative Acute Stroke Study III (ECASS III) trial. This Research Plan is reproduced nearly verbatim with minor adjustments for placement and style of reference citations.

SECTION A: RESEARCH PLAN

A.1 Title of the Proposed Research

Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke: reanalysis of ECASS III trial to assess robustness of findings across statistical models

A.2 Lay Summary

For many patients the only available treatment for reversing the effects of a stroke is alteplase, a drug that breaks up blood clots. When a stroke is caused by blood clots in the arteries leading to the brain, using alteplase soon after stroke onset may reverse the process and minimize the brain damage that occurs which in turn leads to more patients surviving the stroke with less functional deficit. However, a stroke involves damaged brain tissue and a risk or side effect of using a clot-busting drug can be bleeding at the site of this damaged brain tissue which results in a serious worsening of function and can cause death.

Alteplase was initially used up to 3 hours following stroke onset based on a trial reported in 1995 showing such use would increase the chance of being alive with minimal symptoms three months after a stroke. A trial reported in 1999 that tested use of alteplase 3-5 hours after stroke onset did not find this benefit but found an increase in the chance of dying from bleeding. For many years the standard of stroke treatment was the use of alteplase within 3 hours of stroke onset and not after 3 hours.

In 2008 the ECASS III trial reported that use of alteplase 3-4.5 hours after stroke increased the chance of being alive with minimal symptoms 3 months after stroke. However it is unclear how much of this difference is related to the effect of the drug (a true benefit) and how much is related to differences between the drug group and the placebo group that were occurring at the time of trial entry (called baseline difference, leading to a false signal of benefit). There are multiple valid methods for adjusting for baseline differences. Analyses using these different methods would further support the finding of "true benefit" if the results are consistent (considered robustness of results), whereas finding different results would mean that we are less certain the "benefit" is related to use of alteplase and more concerned the results are influenced by baseline differences.

We will evaluate the ECASS III trial data for three functional outcome measures at 90 days after stroke (being alive with no symptoms, being alive with minimal symptoms, and being alive without dependence) and three adverse effect outcome measures (symptomatic bleeding in the head at 7 days, death at 7 days, and death at 90 days) with analyses adjusting for baseline differences in different ways.

When completed, the ability to view the spectrum of outcomes (and not just a selection of outcomes) and the robustness of these results will allow clinical decision-makers and policy-makers a greater view to expected results with the alteplase decision in this timeframe. These results will be reported publicly through a major medical journal.

The results will help patients (or their proxy) further understand and balance the risks and benefits while deciding to consent for alteplase therapy.

A.3 Study Design

The ECASS III trial was a randomized controlled trial comparing alteplase with placebo between 3 hours and 4 hours 30 minutes after stroke onset in patients with acute ischemic hemispheric stroke.[1] The ECASS III trial had baseline differences between study arms and only some of the clinically relevant outcomes were reported with analyses fully adjusted for the baseline differences.[2]

The study design is a reanalysis to provide a fully adjusted view of clinically relevant patient-important outcomes, including different analytic approaches to assess robustness of reported findings. The outcomes to be reported include:

- Symptom-free status (mRS 0) at 90 days
- Disability-free status (mRS 0 or 1) at 90 days
- Dependence-free status (mRS 0 or 1 or 2) at 90 days
- Mortality at 7 days
- Mortality at 90 days
- Symptomatic intracranial hemorrhage (by ECASS III and by NINDS definitions) at 7 days
- Ordinal shift analysis (change across mRS 0-6 spectrum) at 90 days

A.4 Studies Selected and Study Populations

BI-135.312

ECASS III — European Cooperative Acute Stroke Study III: A Placebo Controlled Trial of Alteplase (RtPA) in Acute Ischemic Hemispheric Stroke Where Thrombolysis is Initiated Between 3 and 4 Hours 30 Minutes After Stroke Onset

Medicine: alteplase, Condition: Acute Ischemic Hemispheric Stroke, Phase: 3, Clinical Study ID: 135.312, Sponsor: Boehringer Ingelheim

A.5 Primary and Secondary Endpoints for the Study

Primary Endpoints:

- Symptom-free status (mRS 0) at 90 days
- Disability-free status (mRS 0 or 1) at 90 days
- Dependence-free status (mRS 0 or 1 or 2) at 90 days

Secondary Endpoints:

- Mortality at 7 days
- Mortality at 90 days
- Symptomatic intracranial hemorrhage (by ECASS III and by NINDS definitions) at 7 days
- Change across mRS 0-6 spectrum at 90 days (ordinal shift analysis)

With the exception of the ordinal shift analysis where the outcome will have discrete values of 0 to 6, the other 6 outcomes will be dichotomous outcomes with binary variables with “Yes” for having the outcome and “No” for not having the outcome.

A.6 Statistical Analysis Plan

When we have full access to the study data, we will first attempt to reproduce results from Table 2 through Table 5 from the publication.[1] This exercise will assure us that we have the same database and are using the same variables as those used for the publication by the original study team.

For all 6 dichotomous outcomes specified above, we are interested in estimating the effect of alteplase on the probability of "Yes" for these outcomes. For the Change across mRS 0-6 spectrum at 90 days (ordinal shift analysis), we are interested in estimating the effect of alteplase on the aggregated probability of having an mRS score less than k where k is allowed to vary between 1 and 6.

The variable GROUP is binary, and is equal to 1 for alteplase, and 0 for placebo.

For all 7 outcomes, the descriptive statistics will be reported using the sample proportions. Inferential statistics will be reported using relative risks and absolute risk differences for the 6 dichotomous outcomes, and odds ratios for the ordinal shift analysis.

The point estimate, 95% confidence interval, and p-value associated with variable GROUP are the quantities of interest in the statistical models for all 7 outcomes.

We will assess for baseline differences between the two treatment groups (GROUP=1, GROUP=0), and identify baseline variables that are statistically significant (type-I error of 0.05) between the two groups. In the original publication, the study authors have identified NIHSS score and history of stroke as having statistically significant differences (at the 0.05 threshold) between the 2 treatment groups. We will check to make sure that these two variables are the only variables that are statistically different between the 2 groups.

To test for robustness of results according to the statistical analysis applied, we will conduct the following analyses for each of the 7 outcomes to apply different valid approaches to adjusting for potential confounders:

1) Multivariable modeling
2) Matching
3) Stratified analysis

These different methods of controlling for confounding are used to ascertain whether or not the effect estimates of GROUP are dependent on the statistical method of estimation.

For the multivariable modeling, the independent variables are GROUP, NIHSS score, history of stroke, and possibly other covariates if identified as significant baseline differences. For the 6 dichotomous outcomes, we will use log link, and binomial or poisson errors[3-5] to obtain the estimated adjusted relative risk for GROUP. For the ordinal shift analysis, we will not categorize the mRS score into categories, but rather will leave the mRS score as 0 to 6 and will analyze the data using ordinal logistic regression under the assumption of proportional odds. If the proportional odds assumption does not hold well, we will use multinomial logistic regression.

For matching, we will use the optimal matching procedure[6] to obtain a 1 to 1 match. The match factors will be age, sex, NIHSS score, history of stroke, time from stroke onset to treatment initiation, and possibly other covariates if identified as significant baseline differences or potential confounders. We will use a caliper (difference between those with outcome, and those without) that is about half the

standard deviation of 6 years for age, 3 for NIHSS score, identical value for sex, and history of stroke. The caliper used for other covariates (if any) will be based on the available data. After the matching procedure, we will obtain the relative risk estimate and 95% confidence interval of GROUP via conditional logistic regression which takes into account the match design. Matching will not be applied to the ordinal shift analysis.

For the stratified analysis, we will stratify the sample by NIHSS score and history of stroke. NIHSS score will be trichotomized into lower (0-9), intermediate (10-19), and higher (20-42) groups reflecting three strata of stroke severity. We will use the (Cochran-Mantel-Haenszel (CMH) test to obtain weighted relative risk from the pooled data of the 6 strata. Stratified analysis will not be applied to the ordinal shift analysis.

The reported finding of the treatment effect by the study authors assumes no interaction effects between GROUP and other covariates such as history of stroke. We will assess if the GROUP effect varies differently across levels of covariates. We will test for the significance of the interaction term between GROUP and history of stroke, and GROUP and NHISS score, and if significant interactions are found we will report the effects of GROUP across the levels of these covariates.

If the missing data for outcomes or covariates are less than 2%, we will attempt no imputation of missing data, and will treat missing data as missing at random. If missing data occurs for more than 2% and less than 20%, we will produce intention-to-treat analyses using best-case and worst-case assumptions for missing data. If missing data occurs for more than 20%, we will not use the variables.

We plan to use the SAS software version 9.4[7] for all data management and statistical analyses proposed in this document.

A.7 Publication Plan

The results of this reanalysis and new analysis of ECASS III data on clinically relevant outcomes adjusted for all baseline differences will be summarized in a manner able to rapidly interpret the full spectrum of effects.

The manuscript will be submitted to the New England journal of Medicine as the first choice for submission because NEIM was where the ECASS III data was originally reported. If not accepted by the NEJM, the British Medical journal will be the next journal of choice for submission as the BMJ was where the call for reanalysis of such data was published.

References

1. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med 2008 Sep 25;**359**:1317-29. doi: 10.1056/NEJMoa0804656.

2. Bluhmki E, Chamorro A, Dávalos A, et al. Stroke treatment with alteplase given 3.0-4.5 h after onset of acute ischaemic stroke (ECASS III): additional outcomes and subgroup analysis of a randomized controlled trial. Lancet Neurol 2009 Dec;**8**:1095-102. doi: 10.1016/S1474-4422(09)70264-9

3. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. Am J Epidemiol 1986 Jan 1;**123**:174-84.

4. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am J Epidemiol 2004 Aug 15;**160**:301-5.

5. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol 2005 Aug 1;**162**:199-200.

6. Rosenbaum PR. Optimal matching for observational studies. J Am Stat Assoc 1989;**84**:1024-32.

7. SAS Institute Inc., SAS/STAT version 9, Cary, NC: SAS Institute Inc.

Table S1. Replication of Summary of Baseline Characteristics

Baseline Variable	Alteplase (n=418)	Placebo (n=403)	p-value*
Age (yr)	**64.7 +/- 12.1†**	65.6 +/- 11.0	**0.28†**
Male sex (%)	63.2	57.3	0.10
Weight (kg)	78.5 +/- 15.0	78.0 +/- 15.7	0.62
NIHSS score‡			
Mean	10.7 +/- 5.6	11.6 +/- 6.0	**0. 037‡**
Median	9	10	**0.062‡**
Systolic Pressure (mm Hg)	152.6 +/- 19.2	153.3 +/- 22.1	0.63
Diastolic Pressure (mm Hg)	84.4 +/- 13.5	83.9 +/- 13.6	0.58
Diabetes (%)	14.8	16.6	0.47
Previous use of aspirin or antiplatelets (%)	31.1	32.5	0.65
Hypertension (%)	62.4	62.8	0.88
Atrial flutter or fibrillation (%)	12.7	13.6	0.67
History of stroke (%)	7.7	14.1	0.003
Smoking status (%)			**0.38§**
Never smoked	48.6	46.2	
Ex-smoker	20.6	24.6	
Current smoker	30.6	28.8	
Time to treatment initiation			
Median	3h 59m	3h 58m	0.49
By 0.5 hr period‖			**0.46¶**
≥3.0 to ≤ 3.5 hr (%)	9.6	10.4	
>3.5 to ≤ 4.0 hr (%)	45.7	47.9	
> 4.0 to ≤ 4.5 hr (%)	41.6	36.7	

This table shows replication of the original Table 2 from the 2008 N Engl J Med publication. Differences from the original report are bolded with explanations in the footnotes.

*p-value is based on t-statistic for continuous variables when testing the means, Wilcoxon 2 sample test for continuous variables when testing the medians, and chi-square statistic for categorical variables.

†Minor and inconsequential differences from the original report for age with mean 64.9 and standard deviation 12.2 in the alteplase group and p-value for difference 0.36 in the original report. Explained as due to anonymization rules mandated by the European Union General Data Protection Regulation.

‡National Institutes of Health Stroke Scale, range 0-42 with higher values indicating more severe neurologic impairment (< 5 indicates mild impairment; > 25 indicates very severe impairment). It was not clear in Table 2 of the 2008 publication whether the p-value reported comparing NIHSS scores at baseline (p = 0.03) was referring to a comparison of means or medians. We report both p-values in this table.

§Original reported p-value for smoking status as p = 0.93. That result appears to match the Mantel-Haenszel chi-square test result.

‖Percentages do not add up to 100 because no exact time to treatment initiation available for 27 patients (12 in alteplase group and 15 in placebo group) and treatment was initiated > 4.5 hours after stroke onset for 6 patients (1 in alteplase group and 5 in placebo group).

¶Minor and inconsequential difference from the p-value reported in 2008 publication (ie, p = 0.44 instead of p = 0.46), but the proportions in each 0.5-hour period matched exactly.

Table S2. Descriptive Statistics for Efficacy and Safety Outcomes*

Outcome	Intention-to-Treat Re-analysis		Per Protocol Re-analysis	
	Alteplase (n=418)	**Placebo (n=403)**	**Alteplase (n=375)**	**Placebo (n=355)**
Modified Rankin Scale 0 (Symptom free)	115 (27.5)	88 (21.8)	109 (29.1)	79 (22.3)
Modified Rankin Scale 1 (Symptoms but no disability)	104 (24.9)	94 (23.3)	97 (25.9)	82 (23.1)
Modified Rankin Scale 2 (Slight disability)	59 (14.1)	66 (16.4)	54 (14.4)	60 (16.9)
Modified Rankin Scale 3 (Moderate disability)	39 (9.3)	46 (11.4)	38 (10.1)	42 (11.8)
Modified Rankin Scale 4 (Moderate to severe disability)	39 (9.3)	55 (16.7)	33 (8.8)	53 (14.9)
Modified Rankin Scale 5 (Severe disability)	23 (5.5)	16 (4.0)	18 (4.8)	15 (4.2)
Modified Rankin Scale 0 or 1 (Disability free)	219 (52.4)†	182 (45.2)†	206 (54.9)†	161 (45.4)†
Modified Rankin Scale 0, 1 or 2 (Dependence free)	278 (66.5)	248 (61.5)	260 (69.3)	221 (62.3)
Modified Rankin Scale 0, 1, 2 or 3	317 (75.8)	294 (73.0)	298 (79.5)	263 (74.1)
Modified Rankin Scale 0, 1, 2, 3 or 4	356 (85.2)	349 (86.6)	331 (88.3)	316 (89.0)
Modified Rankin Scale 0, 1, 2, 3, 4 or 5	379 (90.7)	365 (90.6)	349 (93.1)	331 (93.2)
Mortality by day 7 Yes	12 (2.9)†	13 (3.2)†	9 (2.4)	10 (2.8)
Mortality by day 90 Yes	28 (6.7)†	31 (7.7)†	23 (6.1)	22 (6.2)
Symptomatic ICH (ECASS II) Yes	22 (5.3)†	9 (2.2)†	19 (5.1)	9 (2.5)
Symptomatic ICH (NINDS) Yes	33 (7.9)†	14 (3.5)†	28 (7.5)	14 (3.9)
Barthel Index ≥ 95	265 (63.4)†	236 (58.6)†	248 (66.1)†	211 (59.4)†
NIHSS 0 or 1	**209 (50.0)#**	174 (43.2)†	197 (52.5)†	155 (43.7)†
Glasgow Outcome Scale 1	213 (51.0)†	183 (45.4)†	200 (53.3)†	165 (46.5)†

* All entries are n (%). All outcomes are at 90 days unless otherwise specified. NIHSS denotes National Institutes of Health Stroke Scale

If a patient had a missing value for any of Modified Rankin Scale, Barthel Index, NIHSS Score, or Glasgow Outcome Scale, a value was entered using the Last Observation Carried Forward approach if an earlier observation was available. If an earlier observation was not available or the patient died by Day 90 the worst possible score was entered.

† Re-analyzed results are identical to those reported in original publication
# Re-analyzed results are not identical to those reported in original publication [original publication reported 210 (50.2%)]

Table S3. Best-case Sensitivity Analysis of Three Efficacy Outcomes

Intention-to-Treat Population (Alteplase (n=418) vs Placebo (n=403))

Multivariable and Stratified Analyses Adjusted for Baseline NIHSS score and Prior Stroke Status

Table S3a. Baseline NIHSS scores based on original NIHSS score (range 0 to 42)

Outcome at 90 days	Unadjusted RR (95% CI) p-value	Multivariable RR (95% CI) p-value	Stratified Common RR (95% CI) p-value	Unadjusted RD (95% CI) p-value	Multivariable RD (95% CI) p-value
Modified Rankin Scale 0 vs > 0 (Symptom free)	**1.40 (1.10, 1.77) 0.006**	1.24 (0.99, 1.56) 0.064	**1.33 (1.06, 1.67) 0.014**	**0.09 (0.03, 0.15) 0.005**	**0.06 (0.00, 0.12) 0.050**
Modified Rankin Scale 0, 1 vs > 1 (Disability free)	**1.23 (1.07, 1.42) 0.003**	**1.11 (1.02, 1.21) 0.017**	**1.17 (1.03, 1.33) 0.014**	**0.10 (0.04, 0.17) 0.003**	DNC
Modified Rankin Scale 0, 1, 2 vs > 2 (Dependence free)	**1.13 (1.02, 1.25) 0.016**	1.04 (0.99, 1.10) 0.108	**1.10 (1.00, 1.20) 0.045**	**0.08 (0.02, 0.15) 0.015**	0.05 (-0.004, 0.11) 0.068

Table S3b. Baseline NIHSS scores based on modified NIHSS score (range 0 to 46)

Outcome at 90 days	Unadjusted RR (95% CI) p-value	Multivariable RR (95% CI) p-value	Stratified Common RR (95% CI) p-value	Unadjusted RD (95% CI) p-value	Multivariable RD (95% CI) p-value
Modified Rankin Scale 0 vs > 0 (Symptom free)	**1.40 (1.10, 1.77) 0.006**	1.24 (0.99, 1.56) 0.060	**1.33 (1.06, 1.66) 0.014**	**0.09 (0.03, 0.15) 0.005**	0.06 (-0.002, 0.12) 0.056
Modified Rankin Scale 0, 1 vs > 1 (Disability free)	**1.23 (1.07, 1.42) 0.003**	**1.10 (1.00, 1.20) 0.040**	**1.17 (1.03, 1.32) 0.015**	**0.10 (0.04, 0.17) 0.003**	0.06 (-0.01, 0.12) 0.081
Modified Rankin Scale 0, 1, 2 vs > 2 (Dependence free)	**1.13 (1.02, 1.25) 0.016**	1.04 (0.99, 1.09) 0.155	1.09 (1.00, 1.19) 0.051	**0.08 (0.02, 0.15) 0.015**	0.05 (-0.01, 0.11) 0.103

Table S4. Worst-case Sensitivity Analysis of Three Efficacy Outcomes

Intention-to-Treat Population (Alteplase (n=418) vs Placebo (n=403))

Multivariable and Stratified Analyses Adjusted for Baseline NIHSS score and Prior Stroke Status

Table S4a. Baseline NIHSS scores based on original NIHSS score (range 0 to 42)

Outcome at 90 days	Unadjusted RR (95% CI) p-value	Multivariable RR (95% CI) p-value	Stratified Common RR (95% CI) p-value	Unadjusted RD (95% CI) p-value	Multivariable RD (95% CI) p-value
Modified Rankin Scale 0 vs > 0 (Symptom free)	1.11 (0.88, 1.40) 0.374	1.01 (0.81, 1.26) 0.942	1.06 (0.85, 1.33) 0.617	0.03 (-0.03, 0.09) 0.373	0.01 (-0.05, 0.06) 0.853
Modified Rankin Scale 0, 1 vs > 1 (Disability free)	1.10 (0.95, 1.26) 0.196	1.03 (0.93, 1.14) 0.607	1.04 (0.92, 1.18) 0.569	0.05 (-0.02, 0.11) 0.195	-0.01 (-0.07, 0.06) 0.779
Modified Rankin Scale 0, 1, 2 vs > 2 (Dependence free)	1.04 (0.94, 1.15) 0.499	0.99 (0.93, 1.05) 0.770	1.00 (0.92, 1.09) 0.999	0.02 (-0.04, 0.09) 0.498	-0.02 (-0.09, 0.04) 0.529

Table S4b. Baseline NIHSS scores based on modified NIHSS score (range 0 to 46)

Outcome at 90 days	Unadjusted RR (95% CI) p-value	Multivariable RR (95% CI) p-value	Stratified Common RR (95% CI) p-value	Unadjusted RD (95% CI) p-value	Multivariable RD (95% CI) p-value
Modified Rankin Scale 0 vs > 0 (Symptom free)	1.11 (0.88, 1.40) 0.374	1.01 (0.81, 1.26) 0.950	1.05 (0.84, 1.32) 0.670	0.03 (-0.03, 0.09) 0.373	0.003 (-0.05, 0.06) 0.922
Modified Rankin Scale 0, 1 vs > 1 (Disability free)	1.10 (0.95, 1.26) 0.196	1.03 (0.93, 1.14) 0.621	1.03 (0.91, 1.17) 0.661	0.05 (-0.02, 0.11) 0.195	-0.01 (-0.08, 0.05) 0.706
Modified Rankin Scale 0, 1, 2 vs > 2 (Dependence free)	1.04 (0.94, 1.15) 0.499	0.99 (0.93, 1.05) 0.738	0.99 (0.91, 1.08) 0.851	0.02 (-0.04, 0.09) 0.498	-0.02 (-0.09, 0.04) 0.463