



# Thrombolysis with alteplase 3–4.5 hours after acute ischaemic stroke: trial reanalysis adjusted for baseline imbalances

Brian Scott Alper <sup>1,2</sup> Gary Foster,<sup>3,4</sup> Lehana Thabane,<sup>3,4</sup> Alex Rae-Grant,<sup>5</sup> Meghan Malone-Moses,<sup>2</sup> Eric Manheimer<sup>2</sup>

10.1136/bmjebm-2020-111386

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjebm-2020-111386>).

<sup>1</sup>Medical Knowledge Office, EBSCO Information Services, Ipswich, Massachusetts, USA

<sup>2</sup>Innovations and Evidence-Based Medicine Development, EBSCO Health, Ipswich, Massachusetts, USA

<sup>3</sup>Biostatistics, St Joseph's Healthcare, Hamilton, Ontario, Canada

<sup>4</sup>Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

<sup>5</sup>DynaMed, EBSCO Health, Ipswich, Massachusetts, USA

Correspondence to: Dr Brian Scott Alper, Medical Knowledge Office, EBSCO Information Services, Ipswich, MA 01938, USA; [balper@ebSCO.com](mailto:balper@ebSCO.com)



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Alper BS, Foster G, Thabane L, *et al.* *BMJ Evidence-Based Medicine* 2020;**25**:172–179.

## Abstract

**Objectives** Alteplase is commonly recommended for acute ischaemic stroke within 4.5 hours after stroke onset. The Third European Cooperative Acute Stroke Study (ECASS III) is the only trial reporting statistically significant efficacy for clinical outcomes for alteplase use 3–4.5 hours after stroke onset. However, baseline imbalances in history of prior stroke and stroke severity score may confound this apparent finding of efficacy. We reanalysed the ECASS III trial data adjusting for baseline imbalances to determine the robustness or sensitivity of the efficacy estimates. **Design** Reanalysis of randomised placebo-controlled trial. We obtained access to the ECASS III trial data and replicated the previously reported analyses to confirm our understanding of the data. We adjusted for baseline imbalances using multivariable analyses and stratified analyses and performed sensitivity analysis for missing data. **Setting** Emergency care.

**Participants** 821 adults with acute ischaemic stroke who could be treated 3–4.5 hours after symptom onset.

**Interventions** Intravenous alteplase (0.9 mg/kg of body weight) or placebo.

**Main outcome measures** The original primary efficacy outcome was modified Rankin Scale (mRS) score 0 or 1 (ie, being alive without any disability) and the original secondary efficacy outcome was a global outcome based on a composite of functional end points, both at 90 days. Adjusted analyses were only reported for the primary efficacy outcome and the original study protocol did not specify methods for adjusted analyses. Our adjusted reanalysis included these outcomes, symptom-free status (mRS 0), dependence-free status (mRS 0–2), mortality (mRS 6) and change across the mRS 0–6 spectrum at 90 days; and mortality and symptomatic intracranial haemorrhage at 7 days. **Results** We replicated previously reported unadjusted analyses but discovered they were based on a modified interpretation of the National Institutes of Health Stroke Scale (NIHSS) score. The secondary efficacy outcome was no longer significant using the original NIHSS score. Previously reported adjusted analyses could only be replicated with significant effects for the primary efficacy outcome by using statistical approaches not reported in the trial protocol or statistical analysis plan. In analyses adjusting for

## Key message

### What is already known about this subject?

- Thrombolysis with alteplase at 3–4.5 hours after onset of acute ischaemic stroke is widely reported to be effective for improving functional outcome.
- The evidence basis for efficacy is primarily derived from the Third European Cooperative Acute Stroke Study (ECASS III) trial which had substantial baseline imbalances in stroke severity and history of prior stroke.
- The influence of baseline differences on effect estimates from the ECASS III trial has not been thoroughly analysed.

### What are the new findings?

- Adjusting for baseline imbalances using multiple methods fails to find significant benefits and continues to find significant risks with alteplase 3–4.5 hours after stroke.
- Originally reported unadjusted analyses suggesting efficacy in functional end points other than the primary outcome were not reproduced when using the original National Institutes of Health Stroke Scale score.

### How might it impact on clinical practice in the foreseeable future?

- Any prior decisions or interpretations based on ECASS III results warrant reconsideration.
- Estimation of efficacy for alteplase at 3–4.5 hours after stroke should be reassessed with independent access to original trial data from all trials contributing to the assessment.
- Until thorough reassessment is done clinicians should consider any estimates of efficacy for alteplase beyond 3 hours after stroke onset to have very low certainty.

baseline imbalances, all efficacy outcomes were not significant, but increases in symptomatic intracranial haemorrhage remained significant.

**Conclusions** Reanalysis of the ECASS III trial data with multiple approaches adjusting for baseline imbalances does not support any significant benefits and continues to support harms for the use of alteplase 3–4.5 hours after stroke onset. Clinicians, patients and policymakers should reconsider interpretations and decisions regarding management of acute ischaemic stroke that were based on ECASS III results.

**Trial registration number** NCT00153036.

## Introduction

The Third European Cooperative Acute Stroke Study (ECASS III) was a randomised controlled trial comparing alteplase (a thrombolytic medication) with placebo between 3 and 4 hours and 30 min (3–4.5 hours) after stroke onset in patients with acute ischaemic hemispheric stroke.<sup>1</sup> In 2008, the ECASS III trial results were reported with a conclusion that alteplase increased the chance of being alive with minimal symptoms 3 months after stroke.<sup>1</sup> Primarily based on or heavily influenced by the ECASS III trial, most clinical practice guidelines for the management of acute ischaemic stroke currently recommend extending the use of alteplase up to 4.5 hours after stroke onset.<sup>2–12</sup> To date, ECASS III is the only trial to have reported benefit from use of alteplase 3–4.5 hours after stroke onset.

The ECASS III trial had substantial baseline imbalances in two variables that are prognostic for efficacy outcomes: National Institutes of Health Stroke Scale (NIHSS) score and history of prior stroke.<sup>1 13–15</sup> The NIHSS score at baseline was lower (ie, better) in the alteplase group (mean 10.7, median 9) than in the placebo group (mean 11.6, median 10) ( $p=0.03$ ).<sup>1</sup> Prior stroke was reported in 7.7% alteplase patients and 14.1% of placebo patients ( $p=0.003$ ).<sup>1</sup>

The primary efficacy outcome in ECASS III (being alive with minimal symptoms 3 months after stroke) was reported to be significant in some analyses adjusting for baseline imbalances. In the original report, an analysis adjusted for NIHSS score, time to treatment, smoking and hypertension (but not history of stroke) reported an OR 1.42 (95% CI 1.02 to 1.98) based on analysis of 785 patients.<sup>1</sup> A subsequently reported 'full model' analysis which adjusted for NIHSS score, history of prior stroke and other prognostic variables that did not have baseline imbalances reported an OR 1.43 (95% CI 1.02 to 2.00) based on analysis of 784 patients.<sup>16</sup> However, an analysis limited to 732 of 821 patients (89%) in ECASS III who did not have a prior stroke reported no significant difference in the primary efficacy outcome.<sup>13 16</sup> Aside from the primary outcome, the other clinically relevant outcomes have not been reported with analyses adjusted for the baseline imbalances.<sup>13 16</sup>

It is unclear how much of the alteplase-attributed effect in the ECASS III trial 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 (such that baseline imbalances lead to a false signal of benefit). We reanalysed the ECASS III trial data to provide a view of clinically relevant patient-important outcomes adjusted for baseline imbalances, including sensitivity analyses and different analytical approaches to assess robustness of reported findings. Consistency across the results of the primary analysis, reanalysis

and sensitivity analyses would provide reassurance about the credibility of the primary findings.<sup>17–19</sup>

## Methods

We obtained access to the original data from the ECASS III trial with a study protocol posted at ClinicalStudyDataRequest.com (proposal number 1619) and included in online supplementary appendix.

We first attempted to reproduce results from table 2 through table 5 from the original publication<sup>1</sup> to ensure that we had the same database and were using the same variables as those used for the publication by the original study team. We also attempted to reproduce the adjusted analyses for the primary outcome reported in the original publication<sup>1</sup> and in a subsequently reported 'full model'.<sup>16</sup>

We specified seven outcomes for our reanalysis:

1. Symptom-free status (modified Rankin Scale (mRS) score 0) at 90 days.
2. Disability-free status (mRS 0–1) at 90 days.
3. Dependence-free status (mRS 0–2) at 90 days.
4. Mortality at 7 days.
5. Mortality at 90 days.
6. Symptomatic intracranial haemorrhage (by ECASS III and National Institute of Neurological Disorders and Stroke (NINDS) study definitions) at 7 days.
7. Change across mRS 0–6 spectrum at 90 days (ordinal shift analysis).

The first six outcomes are dichotomous and the seventh outcome will have discrete values of 0–6.

For dichotomous outcomes we estimated the effect of alteplase on the probability of having the outcomes. For the change across mRS 0–6 spectrum at 90 days (ordinal shift analysis), we estimated the effect of alteplase on the aggregated probability of having an mRS score less than  $k$  where  $k$  varies between 1 and 6.

For all seven outcomes, the descriptive statistics were reported using the sample proportions. Inferential statistics were reported using relative risks and absolute risk differences for the six dichotomous outcomes and ORs for the ordinal shift analysis.

We assessed for baseline imbalances between the two treatment groups and identified baseline variables that were statistically significant ( $\alpha=0.05$ ) between the two groups. The statistically significant baseline imbalances identified in the original publication were NIHSS score and history of prior stroke.<sup>1</sup>

To test for robustness of results according to the statistical analysis applied, we planned to conduct three analytical approaches (multivariable modelling, matching and stratified analysis) to dichotomous outcomes and multivariable modelling to the ordinal shift analysis to apply different valid approaches to adjusting for potential confounders.

For the multivariable modelling, the independent variables were treatment assignment, NIHSS score, history of prior stroke and possibly other covariates if identified as significant baseline imbalances. For the dichotomous outcomes, we used log link, and binomial or Poisson distribution to obtain the estimated adjusted relative risk for alteplase compared with placebo.<sup>20–22</sup> For the ordinal shift analysis, we did not categorise the mRS score into categories, but rather left the mRS score as 0–6 and analysed the data using ordinal logistic regression under the assumption of proportional odds. If the proportional odds assumption did not hold well, we would use multinomial logistic regression.

For matching, we planned to use the optimal matching procedure<sup>23</sup> to obtain a 1 to 1 match. The match factors would have been age, sex, NIHSS score, history of prior stroke, time from stroke onset

to treatment initiation and possibly other covariates if identified as significant baseline imbalances or potential confounders.

For the stratified analysis, we stratified the sample by NIHSS score and history of prior stroke. NIHSS score was trichotomised into lower (0–9), intermediate (10–19) and higher (20–42) groups reflecting three strata of stroke severity. We used the Cochran-Mantel-Haenszel test to obtain weighted relative risk from the pooled data of the six strata.

The reported finding of the treatment effect by the study authors assumed no interaction effects between randomised assignment and other covariates such as history of prior stroke.<sup>1</sup> We assessed if the treatment effect varied differently across levels of covariates. We tested for the significance of the interaction term between randomised assignment and history of prior stroke and between randomised assignment and NIHSS score. If significant interactions were found, we would report the treatment effects across the levels of these covariates, but we did not find such interactions.

Where the missing data for outcomes or covariates were less than 2%, we did not attempt imputation of missing data and treated missing data as missing at random. If missing data occurred for more than 2% and less than 20%, we produced intention-to-treat analyses using best-case and worst-case assumptions for missing data. If missing data occurred for more than 20%, we did not use the variables.

We planned to use the SAS software V.9.4<sup>24</sup> for all data management and statistical analyses.

## Results

### Unadjusted analyses

We reproduced tables 2–5 from the original publication<sup>1</sup> with only minor discrepancies. The baseline characteristics data matched exactly except for age in the alteplase group (which was reported as mean of 64.9 with SD of 12.2 in the original report and was calculated as mean of 64.7 with SD of 12.1 in our analysis) (online supplementary appendix table S1). This discrepancy was explained by anonymisation rules mandated by the European Union General Data Protection Regulation.

The unadjusted analyses for efficacy outcomes matched initially with the exception of finding 209 patients in the alteplase group (rather than 210 patients) having the NIHSS score of 0 or 1 (online supplementary appendix table S2). This discrepancy resulted in the outcomes at 90 days of NIHSS score of 0 or 1 and the global outcome becoming no longer statistically significant. We identified the patient with discrepant data and clarified with the study sponsor. The NIHSS total score at day 90 (possible range 0–46 based on total of 17 individual item scores) was reported as 2 in the original data rather than calculated as the sum of 17 individual item scores. The 17 individual item scores were each recorded as 0. Using a calculated NIHSS total score of 0 for this patient replicated the originally reported analysis.

We reproduced the results in unadjusted analyses reported for symptomatic intracranial haemorrhage by ECASS II and NINDS definitions but did not find data reporting symptomatic intracranial haemorrhage by ECASS III definition (online supplementary appendix table S2). The ECASS III definition was the same as the ECASS II definition plus the requirement that ‘the hemorrhage must have been identified as the predominant cause of the neurologic deterioration’. The only other difference from the published results was that we found 11 cases of vascular serious adverse events in the placebo group rather than 10 cases originally reported.

All analyses reported for ECASS III were based on ‘NIHSS (inclusive of distal motor function left/right)’, resulting in an NIHSS total score with a range of 0–46 calculated from 17 component scores (the 15 components of the NIHSS score plus two additional distal motor function components). This approach using ‘NIHSS (inclusive of distal motor function)’ had been prespecified in the ECASS III protocol and statistical analysis plan. However, the original publication<sup>1</sup> erroneously stated NIHSS scores ranged from 0 to 42 and this is based on 15 components for calculation of NIHSS scores.<sup>25</sup> Our reanalysis protocol specifies to analyse NIHSS scores based on the standard 0–42 range. In our reanalysis we found all non-primary efficacy outcomes were not significant when using the NIHSS score as reported in the original publication (table 1).

### Replication of previously reported adjusted analyses

We attempted to reproduce the adjusted analysis in the original publication<sup>1</sup> and a subsequently reported ‘full model’ analysis.<sup>16</sup> Our initial attempts failed to replicate the previously reported adjusted analyses. We clarified with the study sponsor and were able to reproduce both the previously reported adjusted analyses with three conditions:

1. Exclusion of patients who had incomplete or missing baseline NIHSS scores.
2. Treatment of the baseline NIHSS score as a categorical variable with five categories (0–5, 6–10, 11–15, 16–20, >20).
3. Treatment of the time from symptom onset to treatment as a categorical variable with seven categories (15 min windows).

The ECASS III protocol and statistical analysis plan provided detailed prespecification of unadjusted analyses for all end points but only predefined the role of adjusted analyses as for exploratory purpose and to identify predictive/confounding variables. The three conditions necessary to replicate the adjusted analyses were thus not prespecified. They were subsequently identified in the clinical trial report or in communication with the study sponsor (table 2). The previously reported adjusted analyses were statistically significant only under all three of these conditions (table 3).

### Analyses adjusted for baseline imbalances

To follow our originally intended protocol we calculated NIHSS scores based on the 15 component scores in the original NIHSS score definition. If a baseline NIHSS component score had a missing value, it was imputed with a 0 and if a 90-day NIHSS component score had a missing value it was imputed with the maximum value for the item. The baseline and 90-day NIHSS scores were then computed.

The only two variables that were statistically significant as baseline imbalances were the NIHSS score and history of prior stroke. We were unable to conduct the matching analysis because the necessary tools were not present in the SAS software version available through the portal used for study data access.

For all efficacy outcomes at 90 days, multivariable analyses and stratified analyses adjusting for baseline imbalances found no significant treatment effect (table 4). This occurred for every dichotomous classification of mRS, for all other dichotomous efficacy outcome measures and for the ordinal shift analysis. For the ordinal shift analysis, the assumption of proportional odds was violated for our multivariable modelling ( $p=0.0004$ ). Therefore, we used a multinomial logistic regression model and found no significant difference for this outcome ( $p=0.440$ ). Increases in symptomatic intracranial haemorrhage with alteplase remained statistically significant in five of six adjusted analyses (table 4).



**Table 1** Unadjusted analyses of efficacy outcomes

Outcome at day 90	Scale description	Relative risk (95% CI) P value	
		Modified NIHSS 0–46 model*	Original NIHSS 0–42 model†
Modified Rankin Scale 0 or 1 vs >1	Scores range from 0 (no symptoms at all) to 6 (death). A score of 0 or 1 indicates no disability.	<b>1.16 (1.01 to 1.34)</b> <b>0.039</b>	
Barthel Index ≥95 vs <95	Assesses the ability to perform activities of daily living on a scale that ranges from 0 (complete dependence) to 100 (independence).	1.08 (0.97 to 1.21) 0.157	
Modified NIHSS 0–46 0 or 1 vs >1	Original NIHSS score modified with addition of distal motor function with scores ranging from 0 to 46.	<b>1.16 (1.00 to 1.35)</b> <b>0.044</b>	
Original NIHSS 0–42 0 or 1 vs >1	Scores on the NIHSS range from 0 to 42, with higher values reflecting more severe neurological impairment.		1.15 (0.99 to 1.32) 0.062
Glasgow Outcome Scale 1 vs >1	A 5-point scale on which 1 indicates independence, 3 severe disability and 5 death.	1.12 (0.97 to 1.29) 0.113	
Global outcome‡	A multidimensional calculation of a favourable outcome, defined by several individual outcome scales and entered into a statistical algorithm.	<b>1.28 (1.00 to 1.64)</b> <b>0.049</b>	1.27 (0.995 to 1.63) 0.055

Intention-to-treat population (alteplase (n=418) vs placebo (n=403)).

Significant differences are in bold.

\*Modified NIHSS 0–46 model uses the original NIHSS score with 15 items and addition of distal motor function left/right as two items, both with scores 0–2, and total score range 0–46.

†Original NIHSS 0–42 model uses original NIHSS score with 15 items and total score range 0–42.

‡The statistical approach is a global OR test based on a linear logistic regression model. The global OR is the probability of a favourable outcome with alteplase as compared with placebo.

NIHSS, National Institutes of Health Stroke Scale.

Descriptive statistics for our seven prespecified outcomes of interest are reported in online supplementary appendix table S2. There were no missing data for four safety outcomes. For the three prespecified efficacy outcomes, a best-case sensitivity analysis found significant effects for the outcome of mRS 0 or 1 across most methods of analysis, and variable results for mRS 0 and for mRS 0–2 based on method of analysis (online supplementary appendix table S3). In the worst-case sensitivity analysis (using either baseline NIHSS score in 0–42 range or in 0–46 range), none of the efficacy end points were significant in any unadjusted

analysis or any analysis adjusted for baseline imbalances (online supplementary appendix table S4).

## Discussion

Reanalysis of the ECASS III trial data with multiple approaches adjusting for baseline imbalances does not support any statistically significant benefits that were previously reported and continues to support statistically significant harms for the use of alteplase 3–4.5 hours after stroke onset.

Strengths of this reanalysis include prespecified outcomes of clinical importance and prespecified analytical methods to avoid selective analysis and selective reporting. Limitations of this reanalysis include limitations to the trial data access such that matching analysis could not be performed. However, it is implausible for a matching analysis to find statistically significant results because the high likelihood of mismatched patients would be expected to reduce the sample sizes and resulting statistical power. Another limitation of reanalysis, or any method for adjusting for non-randomised factors influencing the effect estimates from a randomised trial, is such analyses cannot confidently produce new conclusions (neither a claim of efficacy nor a claim of absence of efficacy). The role of reanalysis and adjusted analyses is limited to increasing or decreasing the certainty in the unadjusted analysis of the randomised trial. In this case the reanalysis does not negate the original findings, but it greatly reduces the certainty for those findings.

In addition to the adjusted analyses conducted as our primary reanalysis per our predefined protocol, we discovered two unique differences between our expectations and the published findings when attempting to reproduce the originally reported results. First, the original authors used a modified NIHSS score, consistent with their prespecified protocol but inconsistent with the results reported in the trial publication. Analysis using the original NIHSS score (without additional scores from distal motor function) results in loss of significance in non-primary efficacy outcomes previously reported as significant in unadjusted analysis. Second,

**Table 2** Methodological approaches to previously reported adjusted analyses

Concept	Approach used for clinical trial report*	Most informative model†
Handling of patients who had incomplete data for baseline NIHSS score	Exclusion‡	Inclusion using imputed scores§
Approach to baseline NIHSS score variable for adjustment	Categorical with five categories (0–5, 6–10, 11–15, 16–20, >20)	Continuous
Approach to time from symptom onset variable for adjustment	Categorical with seven categories (3–4.5 hours in 15 min increments, >4.5 hours)	Continuous

\*As described in clinical trial report (dated 15 December 2008).

†As considered in alternative models applied to replication of previously reported analyses (see table 3).

‡According to correspondence with the study sponsor.

§Imputed values for baseline NIHSS scores were very close to the likely actual values for 10 of 11 patients requiring imputation. These 10 patients had values for 15 or 16 of the 17 component scores used to calculate the baseline NIHSS scores.

NIHSS, National Institutes of Health Stroke Scale.

**Table 3** Adjusted analyses of primary efficacy outcome with previously used and modified approaches

Model*	n†	OR (95% CI)	P value
mRS score of 0 or 1 at 90 days, analysis adjusted for baseline NIHSS score, time from onset to treatment (OTT), smoking and prior hypertension as reported in Hacke <i>et al</i> <sup>1</sup> and Bluhmki <i>et al</i> <sup>16</sup>			
Previously reported adjusted analysis <sup>1,16</sup>	785	<b>1.42 (1.02 to 1.98)</b>	<b>0.037</b>
Missing—impute, NIHSS score—continuous, OTT—continuous	794	1.25 (0.91 to 1.73)	0.173
Missing—impute, NIHSS score—continuous, OTT—categorical	794	1.33 (0.96 to 1.85)	0.086
Missing—impute, NIHSS score—categorical, OTT—continuous	794	1.29 (0.94 to 1.78)	0.121
Missing—impute, NIHSS score—categorical, OTT—categorical	794	1.38 (0.99 to 1.92)	0.054
Missing—exclude, NIHSS score—continuous, OTT—continuous	785	1.30 (0.94 to 1.80)	0.119
Missing—exclude, NIHSS score—continuous, OTT—categorical	785	1.38 (0.99 to 1.93)	0.056
Missing—exclude, NIHSS score—categorical, OTT—continuous	785	1.33 (0.96 to 1.84)	0.087
Missing—exclude, NIHSS score—categorical, OTT—categorical	785	<b>1.42 (1.02 to 1.98)</b>	<b>0.037</b>
mRS score of 0 or 1 at 90 days, analysis adjusted for history of stroke, baseline NIHSS score, OTT, smoking, prior hypertension and nine other variables as reported in Bluhmki <i>et al</i> <sup>16</sup> ('full model')			
Previously reported 'full model' <sup>16</sup>	784	<b>1.43 (1.02 to 2.00)</b>	<b>0.040</b>
Missing—impute, NIHSS score—continuous, OTT—continuous	791	1.27 (0.91 to 1.76)	0.156
Missing—impute, NIHSS score—continuous, OTT—categorical	791	1.34 (0.96 to 1.87)	0.086
Missing—impute, NIHSS score—categorical, OTT—continuous	791	1.31 (0.94 to 1.81)	0.112
Missing—impute, NIHSS score—categorical, OTT—categorical	791	1.38 (0.99 to 1.93)	0.059
Missing—exclude, NIHSS score—continuous, OTT—continuous	782	1.32 (0.95 to 1.84)	0.102
Missing—exclude, NIHSS score—continuous, OTT—categorical	782	1.39 (0.99 to 1.95)	0.055
Missing—exclude, NIHSS score—categorical, OTT—continuous	782	1.35 (0.97 to 1.88)	0.076
Missing—exclude, NIHSS score—categorical, OTT—categorical	782	<b>1.43 (1.02 to 2.00)</b>	<b>0.039</b>

Statistically significant differences are in bold.

\*Missing—impute means patients with missing components of baseline NIHSS score had a value of 0 imputed for the missing component(s) and then baseline NIHSS score was computed with range 0–46. Missing—exclude means patients with missing components of baseline NIHSS score were excluded from analysis. NIHSS score—continuous means baseline NIHSS score treated as a continuous variable. NIHSS score—categorical means baseline NIHSS score treated as a five-category (0–5, 6–10, 11–15, 16–20, 20+) variable. OTT—continuous means OTT treated as a continuous variable. OTT—categorical means OTT treated as a seven-category variable (15 min windows).

†n=794 for analyses excluding 27 patients with missing OTT data. n=785 same as for n=794 for analyses also excluding nine further patients with missing baseline NIHSS data. n=791 same as for n=794 analysis, also excluding one person missing dose information and two missing age. n=782 same as for n=791 analyses also excluding nine patients with missing baseline NIHSS data.

mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; OTT, time from symptom onset to treatment.

the original authors reported adjusted analyses which supported their original findings, but the precise methods and assumptions used for these adjusted analyses were not prespecified. Previously reported adjusted analyses suggesting efficacy show statistical significance only under multiple conditions that do not represent the most informative use of available data. Seven other 'less selective' approaches to these adjusted analyses fail to replicate significant effects.

We had a priori concerns for risk of bias due to known baseline imbalances in prognostic variables. Our reanalysis (intended to evaluate the robustness of ECASS III findings to multiple standard analyses adjusted for these baseline imbalances) resulted in numerous reasons to reduce our certainty in the primary results. First, all our prespecified adjusted analyses were inconsistent with the primary analysis, so we have substantial uncertainty for efficacy estimates. Second, the previously reported adjusted analysis appears to be selective analysis and reporting (not necessarily intentional and potentially just an oversight in reporting scientific findings with a lot of complexity). Rather than viewing consistent findings in one highly selected adjusted analysis we find inconsistent findings in seven of eight adjusted analyses for this method of adjustment, further reducing our certainty. Finally, the absence of significant effects in any non-primary efficacy outcome (with use of the original NIHSS score for unadjusted analyses) adds reasons for less certainty in the primary outcome unrelated to the risk of bias from baseline imbalances. All other

methods of assessing functional outcomes being inconsistent with the primary outcome make it less likely that the difference in the primary outcome represents a true effect.

Overall decision-making should not be determined based on a single trial but rather careful synthesis of a body of evidence. We previously reported that the most current comprehensive meta-analyses supported the use of alteplase 3–4.5 hours after stroke despite the evidence directly comparing alteplase to no alteplase suggesting a 2% absolute increase in mortality and no clear benefit.<sup>26</sup> The individual patient data meta-analysis reported an improvement in functional outcome with alteplase 3–4.5 hours after stroke (adjusted OR 1.26, 95%CI 1.05 to 1.51)<sup>27</sup> despite combining data from a meta-analysis of the same trials except the Third International Stroke Trial (IST-3) (adjusted OR 1.34 with 1620 patients)<sup>28</sup> and the IST-3 (adjusted OR 0.73 with 1177 patients).<sup>29</sup> It is possible that differences might be explained by selective outcome reporting or analytical approach but the data are not easily available for independent and comprehensive analysis. Our findings with the ECASS III trial data provide further evidence of a need for comprehensive independent reanalysis using all available data informing use of alteplase for acute ischaemic stroke.

Concerns for certainty in reported analyses are not limited to use of alteplase 3–4.5 hours after stroke. Among the multiple randomised trials assessing alteplase for thrombolysis after acute stroke only two have been reported to support significant claims

Table 4 Unadjusted analyses and analyses adjusted for baseline NIHSS score and prior stroke status using reanalysis protocol

Outcome (day 90 unless stated otherwise)	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)	1.26 (0.99 to 1.60) 0.061	1.11 (0.88 to 1.39) 0.395	1.19 (0.94 to 1.50) 0.140	0.06 (−0.002 to 0.12) 0.059	0.03 (−0.03 to 0.09) 0.320
Modified Rankin Scale* 0, 1 vs >1 (disability free)	<b>1.16 (1.01 to 1.34)</b> <b>0.039</b>	1.05 (0.95 to 1.16) 0.324	1.09 (0.96 to 1.24) 0.163	<b>0.07 (0.004 to 0.14)</b> <b>0.038</b>	DNC
Modified Rankin Scale* 0, 1, 2 vs >2 (dependence free)	1.08 (0.98 to 1.20) 0.139	1.01 (0.95 to 1.06) 0.848	1.04 (0.95 to 1.14) 0.397	0.05 (−0.02 to 0.12) 0.138	0.02 (−0.05 to 0.08) 0.625
Modified Rankin Scale* 0, 1, 2, 3 vs >3	1.04 (0.96 to 1.13) 0.344	1.00 (0.95 to 1.05) 0.973	1.01 (0.94 to 1.08) 0.775	0.03 (−0.03 to 0.09) 0.344	−0.00 (−0.07 to 0.07) 0.971
Modified Rankin Scale* 0, 1, 2, 3, 4 vs >4	0.98 (0.93 to 1.04) 0.555	0.98 (0.93 to 1.04) 0.513	0.97 (0.92 to 1.02) 0.222	−0.01 (−0.06 to 0.03) 0.555	−0.03 (−0.09 to 0.04) 0.446
Modified Rankin Scale* 0, 1, 2, 3, 4, 5 vs >6	1.01 (0.98 to 1.05) 0.495	1.00 (0.95 to 1.05) 0.984	1.00 (0.97 to 1.04) 0.834	0.01 (−0.02 to 0.05) 0.495	−0.00 (−0.06 to 0.06) 0.987
Barthel Index ≥95 vs <95	1.08 (0.97 to 1.21) 0.157	1.00 (0.95 to 1.06) 0.949	1.04 (0.94 to 1.14) 0.463	0.05 (−0.02 to 0.12) 0.155	0.03 (−0.03 to 0.09) 0.362
NIHSS 0 or 1 vs >1	1.15 (0.99 to 1.32) 0.062	1.05 (0.95 to 1.16) 0.375	1.09 (0.95 to 1.23) 0.214	0.07 (−0.003 to 0.13) 0.061	DNC
Glasgow Outcome Scale 1 vs >1	1.12 (0.97 to 1.29) 0.113	1.03 (0.93 to 1.15) 0.568	1.07 (0.94 to 1.21) 0.310	0.06 (−0.01 to 0.12) 0.111	DNC
Global outcome	1.27 (1.00 to 1.63) 0.055	1.14 (0.87 to 1.50) 0.335	NA	NA	NA
Mortality by day 7 Yes vs No	0.89 (0.41 to 1.93) 0.767	1.06 (0.50 to 2.28)† 0.877	1.05 (0.49 to 2.22) 0.902	−0.004 (−0.03 to 0.02) 0.768	0.005 (−0.06 to 0.07) 0.889
Mortality Yes vs No	0.87 (0.53 to 1.42) 0.582	1.05 (0.65 to 1.69)† 0.846	0.99 (0.61 to 1.60) 0.953	−0.01 (−0.05 to 0.03) 0.582	0.003 (−0.06 to 0.06) 0.913
Symptomatic ICH (ECASS II) Yes vs No	0.82 (0.50 to 1.36) 0.420	2.33 (1.08 to 5.01) 0.030	2.17 (1.01 to 4.64) 0.041	0.03 (0.004 to 0.06) 0.021	0.025 (−0.01 to 0.06) 0.118
Symptomatic ICH (NINDS) Yes vs No	2.27 (1.23 to 4.18) 0.008	2.37 (1.31 to 4.29)† 0.004	2.32 (1.13 to 4.39) 0.007	0.04 (0.01 to 0.08) 0.006	0.04 (0.00 to 0.08) 0.049

Continued

Table 4 Continued		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	
Outcome (day 90 unless stated otherwise)											
Intention-to-treat population (alteplase (n=418) vs placebo (n=403)).											
Multivariable analysis adjusted for baseline NIHSS score (continuous) and prior stroke status. Baseline NIHSS scores with missing component values had value of 0 imputed for the missing component(s) and then baseline NIHSS score was computed with range 0–42. NIHSS scores at 90 days had missing component(s) imputed as maximum possible value and then NIHSS score at 90 days was computed with range 0–42.											
Stratified analyses based on Cochran-Mantel-Haenszel method (baseline NIHSS stratified into three categories: 0–9, 10–19, 20–42). RR and RD analyses based on main effects log-binomial model, or log-Poisson regression if log-binomial model did not converge. Statistically significant differences are in bold. Meaning of functional outcome measures as described in table 1 with NIHSS score interpreted as originally defined with range 0–42.											
*Scores on the modified Rankin Scale range from 0 (no symptoms) to 6 (death). A score of 1 indicates some symptoms but no disability.											
†Poisson regression model with robust variance (did not converge for log binomial model); all others are log binomial.											
DNC, did not converge for log-binomial or log-Poisson regression model; ECASS II, European Cooperative Acute Stroke Study 2; ICH, intracranial haemorrhage; NA, not applicable; NIHSS, National Institutes of Health Stroke Scale; NINDS, National Institute of Neurological Disorders and Stroke; RD, risk difference; RR, relative risk.											

of efficacy for primary functional outcomes: ECASS III and NINDS.<sup>1 30</sup> The NINDS trial has also been criticised with attention to baseline imbalances (including stroke severity),<sup>31 32</sup> and the full influence of baseline confounding may not have been thoroughly and independently analysed and reported.<sup>33</sup>

This reanalysis confirms reports of concern that the baseline imbalances introduce such a risk of bias that conclusions of efficacy based on ECASS III data cannot be considered reliable.<sup>13 26 34</sup> Clinicians, patients, policymakers, systematic review authors, clinical practice guideline developers and drug regulators should reconsider interpretations and decisions regarding management of acute ischaemic stroke that were based on ECASS III results.

Twitter Brian Scott Alper @BrianAlperMD

**Acknowledgements** The authors thank Boehringer Ingelheim for access to the study data, for multiple communications to clarify the data handling and protocols used and to the commitment to support transparency and free exchange of scientific discussion.

**Contributors** BSA, MMM and EM were part of the original team which published a call for independent analysis of the trial data for alteplase. BSA, LT, ARG, MMM and EM contributed to the protocol for ECASS III trial reanalysis. BSA negotiated the Data Sharing Agreement. GF conducted the statistical analysis. LT supervised the statistical analysis. All authors edited the manuscript. BSA is the guarantor of the study.

**Funding** GF and LT were supported in part by a charitable contribution from EBSCO Information Services.

**Disclaimer** EBSCO Information Services had no role in the design, conduct, reporting or decision to publish for this research.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data may be obtained from a third party and are not publicly available. The authors are unable to share the data as it was obtained through a Data Sharing Agreement with the study sponsor through ClinicalStudyDataRequest.com.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**ORCID iD**  
Brian Scott Alper <http://orcid.org/0000-0003-4300-4928>

**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;359:1317–29.
- 2 Powers WJ, Rabinstein AA, Ackerson T, *et al*. 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American heart Association/American stroke association. *Stroke* 2018;49:e46–110.
- 3 Lansberg MG, O'Donnell MJ, Khatri P, *et al*. Antithrombotic and thrombolytic therapy for ischemic stroke: antithrombotic therapy and prevention of thrombosis, 9th ED: American College of chest physicians evidence-based clinical practice guidelines. *Chest* 2012;141:e601S–36.



- 4 Brown MD, Burton JH, Nazarian DJ, *et al.* American College of emergency physicians (ACEP) clinical policies Subcommittee (writing Committee) on use of intravenous tPA for ischemic stroke. Clinical policy: use of intravenous tissue plasminogen activator for the management of acute ischemic stroke in the emergency department. *Ann Emerg Med* 2015;66:322–33.
- 5 National Institute for Health and Care Excellence. Alteplase for treating acute ischemic stroke. technology appraisal guidance. No. TA264, 2012. Available: <https://www.nice.org.uk/guidance/ta264/resources/alteplase-for-treating-acute-ischaemic-stroke-review-of-technology-appraisal-guidance-122-82600551992005>
- 6 AD X, Wang YJ, Wang DZ. Consensus statement on the use of intravenous recombinant tissue plasminogen activator to treat acute ischemic stroke by the Chinese stroke therapy expert panel. *CNS Neurosci Ther* 2013;19:543–8.
- 7 European Stroke Organisation (ESO) Executive Committee and the ESO Writing Committee. . Chapter 9. Ischaemic stroke and transient ischaemic attack. In: Gilhus NE, Barnes MP, Bainin M, eds. *European Handbook of neurological management*. . 2nd ed., 2011: 1. 101–58. [https://www.eaneurology.org/fileadmin/user\\_upload/guideline\\_papers/EFNS\\_guideline\\_2011\\_Ischaemic\\_stroke\\_and\\_transient\\_ischaemic\\_attack.pdf](https://www.eaneurology.org/fileadmin/user_upload/guideline_papers/EFNS_guideline_2011_Ischaemic_stroke_and_transient_ischaemic_attack.pdf)
- 8 Santé HAd. Stroke: early management (alert, prehospital phase, initial Hospital phase, indications for thrombolysis). clinical practice guideline, 2009. Available: [https://www.has-sante.fr/portail/upload/docs/application/pdf/2010-03/stroke\\_early\\_management\\_-\\_guidelines\\_-\\_english\\_version.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/2010-03/stroke_early_management_-_guidelines_-_english_version.pdf)
- 9 Minematsu K, Toyoda K, Hirano T, *et al.* Guidelines for the intravenous application of recombinant tissue-type plasminogen activator (alteplase), the second edition, October 2012: a guideline from the Japan stroke Society. *Journal of Stroke and Cerebrovascular Diseases* 2013;22:571–600.
- 10 National Stroke Foundation (Australia). Clinical guidelines for stroke management, 2017. Available: <https://informme.org.au/en/Guidelines/Clinical-Guidelines-for-Stroke-Management-2017>
- 11 Bryer A, Connor M, Haug P, *et al.* South African guideline for management of ischaemic stroke and transient ischaemic attack 2010: a guideline from the South African stroke Society (SASS) and the SASS writing Committee. *S Afr Med J* 2010;100:747–78.
- 12 Casaubon LK, Boulanger JM, Blacquiere D, *et al.* Heart and stroke Foundation of Canada Canadian stroke best practices Advisory Committee. Canadian stroke best practice recommendations (CSBPR): hyperacute stroke care guidelines, update 2015. *Int J Stroke* 2015;10:924–40.
- 13 Shy BD. Implications of ECASS III error on emergency department treatment of ischemic stroke. *J Emerg Med* 2014;46:385–6.
- 14 Hill MD, Yiannakoulis N, Jeerakathil T, *et al.* The high risk of stroke immediately after transient ischemic attack: a population-based study. *Neurology* 2015;2004:62.
- 15 Hier DB, Foulkes MA, Swiontoniowski M, *et al.* Stroke recurrence within 2 years after ischemic infarction. *Stroke* 1991;22:155–61.
- 16 Bluhmki E, Chamorro Ángel, 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 randomised controlled trial. *The Lancet Neurology* 2009;8:1095–102.
- 17 Thabane L, Mbuagbaw L, Zhang S, *et al.* A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol* 2013;13:92.
- 18 Morris TP, Kahan BC, White IR. Choosing sensitivity analyses for randomised trials: principles. *BMC Med Res Methodol* 2014;14:11.
- 19 de Souza RJ, Eisen RB, Perera S, *et al.* Best (but oft-forgotten) practices: sensitivity analyses in randomized controlled trials. *Am J Clin Nutr* 2016;103:5–17.
- 20 Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk DIFFERENCES1. *Am J Epidemiol* 1986;123:174–84.
- 21 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;160:301–5.
- 22 Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol* 2005;162:199–200.
- 23 Rosenbaum PR. Optimal matching for observational studies. *J Am Stat Assoc* 1989;84:1024–32.
- 24 SAS Institute Inc. SAS/STAT version 9. Cary, NC: SAS Institute Inc, 2013.
- 25 Lyden P. Using the National Institutes of health stroke scale: a cautionary tale. *Stroke* 2017;48:513–9.
- 26 Alper BS, Malone-Moses M, McLellan JS, *et al.* Thrombolysis in acute ischaemic stroke: time for a rethink? *BMJ* 2015;350:h1795.
- 27 Emberson J, Lees KR, Lyden P, *et al.* Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. *The Lancet* 2014;384:1929–35.
- 28 Lees KR, Bluhmki E, von Kummer R, *et al.* Time to treatment with intravenous alteplase and outcome in stroke: an updated pooled analysis of ECASS, Atlantis, NINDS, and epithet trials. *The Lancet* 2010;375:1695–703.
- 29 Sandercock P, Wardlaw JM, IST-3 Collaborative Group, *et al.* The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 hours of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial. *Lancet* 2012;379:2352–63.
- 30 National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1995;333:1581–8.
- 31 Friedman HS. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1996;334:1405–6.
- 32 Koroshetz WJ. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1996;334:1405–6.
- 33 Lenzer J. Alteplase for stroke: money and optimistic claims buttress the “brain attack” campaign. *BMJ* 2002;324:723–9.
- 34 Alper BS, Brown CB. Expanding recombinant tissue plasminogen activator time window is premature. *Stroke* 2009;40:e632.