EWG 2014

EXPERT WORKING GROUP

ACTILYSE (ALTEPLASE) BA LANCE O F BEN EFITS AND RISKS WH EN USED IN THE TREATMENT OF ACUTE ISCHAEMIC STROKE

Title of paper: Paper 5: Discussion of individuals' concerns on specific aspects of the supporting clinical evidence.

Product: Actilyse 10, 20, 50mg	**Assessors:** Medical assessor: Dr [redacted] Scientific assessor: Dr [redacted] Statistical assessor: Dr [redacted] Epidemiological assessors: Dr [redacted] Dr [redacted]
MAHs: Boehringer Ingelheim Limited	**Previous Assessments:** CHM May 2014
Active constituents: Alteplase (rt-PA)	**Legal status:** **POM**
Therapeutic classification: Antithrombotic agent, ATC code B01AD02	

Contents

Introduction

This paper addresses specific concerns that have been raised with MHRA relating to the use of rt-PA in the treatment of acute ischaemic stroke.

We have received three separate submissions from interested individuals, and their data submissions have been circulated to the EWG. The submissions have been received from the following:

- Dr [redacted] stroke physician
- Dr Pitchaiah Mandava, Director of the Stroke Unit at the Michael E. DeBakey VA Medical Center, US (information provided to Dr [redacted] and separately to MHRA)
- Professors Daniel Fatovich and Simon Brown, Discipline of Emergency Medicine, University of Western Australia (19 September 2014)

1. Issue

MHRA was contacted by Dr [redacted] regarding the use of rt-PA in the treatment of acute ischaemic stroke and requesting an updated evaluation of the evidence. Because new data had become available since the last regulatory review the MHRA conducted a critical appraisal of these data and of Dr [redacted] s specific concerns. In May 2014, the Commission on Human Medicines (CHM) carefully considered this review, and advised that the new data and Dr [redacted] s concerns did not impact on the positive balance of benefits and risks of rt-PA in the treatment of acute ischaemic stroke. However, the CHM advised that in order to be assured that all relevant sources of evidence have been taken into consideration, an expert working group should be set up. The CHM paper has been circulated to the group.

Dr [redacted] has subsequently provided a further submission, dated 5 September and an updated version with additional data, dated October 2014, which is discussed in this paper. As the October version contains all information provided in the September version, and more, only the October version has been circulated to the group.

Since the CHM meeting in May, [redacted] MHRA has been contacted by Professors Fatovich and Brown, and Dr Mandava who have raised additional issues for consideration. This paper also discusses these issues.

2. Studies supporting the EU product licence

The application for the indication in treatment of acute ischaemic stroke was made via the mutual recognition procedure, with Germany as the lead country (or Reference Member State, RMS) in 2000. The studies that formed the basis of the assessment were NINDS part 1 and 2, ECASS I and II, and ATLANTIS. The variation to extend the time-window for treatment from 0-3 hours to 0-4.5 hours after onset of symptoms was based mainly on data from the ECASS III trial, with supportive data from the observational registry SITS-ISTR and a pooled analysis. A number of the points raised as issues of concern in the submissions considered in this paper relate to the original clinical trials. A summary of these trials is provided in paper 3.

3. Submission by Dr [redacted]

Dr [redacted] s updated submission of October has been circulated to the group and the concerns [redacted] has raised are discussed in this section. Several of these issues have been addressed previously in the paper discussed by CHM in May 2014. Therefore where relevant, sections of the CHM paper have been reproduced below. The full paper has been circulated to the group.

Extracts and summaries of Dr [redacted] s submission are included below in blue italic font, followed by our assessment of each concern.

3.1 Concerns raised by animal experiments

With reference to initial experiments in rabbits published in 1985 (Zivin et al, 1985):

"Key concerns include a doubtful clot model, the very modest numbers included in the study (15 were given alteplase), the report that the delay to treatment was only 2 minutes, the lack of clarity over blinding of those recording neurological outcome in the rabbits and the use of complex statistical modelling to produce an attractive graphic display of benefit."

The study by Zivin *et al* published in Science in 1985, was one of the first of several animal studies using rt-PA. This small, early study was intended to a) describe a new animal model for embolic stroke, b) provide proof of concept that thrombolysis may be of value in the treatment of stroke and c) provide *in vitro* data demonstrating that rt-PA produces clot lysis at concentrations comparable to those achieved *in vivo*.

It is very likely that this initial small non-clinical study was not conducted to the same standards as would be expected for a non-clinical study submitted for a present-day marketing authorisation application. The same is likely to be true for any drugs developed prior to the latest regulations.

Many studies in animals using rt-PA have subsequently been published, including several published prior to the initiation of the NINDS study in 1991, for example:

- Papadopoulos *et al* (J Neurosurg, 1987): studied the effect of rt-PA administered to rats 2 hours after middle cerebral artery embolic stroke had been caused by injection of 0.025 cc of human blood clot. 16 rats were included, 8 treated with rt-PA and 8 placebo. The rt-PA treated rats were found to have blood flow increased significantly at 30 minutes post treatment, and the same as pre-embolic levels within 60 minutes, whilst no improvement in blood flow was observed in the placebo group. There was also improvement in the EEG recordings for the treated vs. untreated animals.
- Kissel *et al* (J Neurosurg, 1987): studied the effect of rt-PA in a rabbit cerebroembolic stroke model. Fourteen animals underwent pre-embolus angiogram, blood clots were then injected and the occlusion of the internal carotid artery at the circle of Willis was documented with repeat angiogram. Animals either received rt-PA or saline and follow-up angiograms were performed every 15 minutes. The rt-PA treated animals showed progressive improvement in flow.
- Phillips *et al* (Am J Neuroradiol, 1988): studied the effect of rt-PA in a rabbit model of thromboembolic stroke. Fifteen minutes after embolisation, 8 out of 14 rabbits received 1 mg/kg rt-PA over a 30 minute period, and 6 received saline. Cerebral arteriograms obtained at 30 minute intervals for 180 minutes found partial or complete thrombus dissolution in 7 of the 8 rt-PA treated animals and none of the controls.

- Chehrazi *et al* (Neurosurgery, 1989): studied the effect of rt-PA in a rabbit model of embolic stroke (at the bifurcation of the internal carotid artery at the circle of Willis), in 17 animals. Experimental animals received rt-PA 30 minutes, 2 hours or 4 hours after clot embolisation. Control animals received saline. Digital subtraction angiograms were performed before, and every 30 minutes after treatment. In the rt-PA group, all clots dissolved and circulation was re-established within 120 minutes, whilst in control animals the clots were stable and the internal carotid artery remained occluded.
- Bednar *et al* (Stroke, 1990): studied the effect of rt-PA in a rabbit model of thromboembolic stroke. Six animals received rt-PA and 11 received control. Intracarotid embolisation reduced cerebral blood flow (cm^3/100 g/min, mean +/- SEM) from 55.2 +/-7.7 to 8.5 +/-2.5 in the control group and from 61.8 +/-14.8 to 10.0 +/-3.5 in the treated group. Cerebral blood flow recovered significantly within 60 minutes in the rt-PA group, reaching 59.6 +/-10.0 four hours after embolization, whilst in the control group blood flow reached 15.3 +/-8.9. Cerebral infarct size (% of hemisphere) was 34.4 +/-5.6 in the control group compared with 8.8 +/-5.6 in the rt-PA treated animals.

Whilst all of these studies involved small numbers of animals, the findings are supportive of a possible role for rt-PA in the treatment of acute ischaemic stroke. As discussed above, these studies would not necessarily constitute sufficient non-clinical evidence for a present-day new drug application, however it would not be reasonable to apply these standards retrospectively.

3.2 Relevance of streptokinase

"The 3 main streptokinase trials were not encouraging and a review of their 7 and 30 day mortality profiles would alert any careful physician. (table 1 and 2) The increased 30 day mortality with treatment appeared consistently over 10 percent (figure 2). (MAST-I 1995, MAST-E 1996, ASK 1996) The pharmacological effects of the various thrombolytic agents may reasonably be considered as similar, given matching pharmacological activity. Only alteplase survived the expert reviewer's, possibly conflicted, evaluation…

… In coronary heart disease, trials in tens of thousands of patients found a similar benefit with both streptokinase and alteplase. A similar adverse effect, via intracerebral haemorrhage, was also noted. Both work via the lysis of fibrin by plasmin so major differences would not be expected if the drug activity was similar."

Unlike rt-PA, streptokinase is not licensed for the treatment of acute ischaemic stroke. The effects of streptokinase cannot be directly compared with rt-PA purely on the basis that they are both thrombolytics. Differences in their properties include the following:

- Streptokinase is isolated/purified from streptococcus bacteria; rt-PA is a human protein produced by recombinant technology
- Streptokinase does not have fibrin-specificity; rt-PA has fibrin-specificity, therefore rt-PA mainly has a local action whilst streptokinase does not have this selectivity and disrupts haemostasis to a greater extent than rt-PA.
- Streptokinase half-life is longer (biphasic, ~18 minutes in association with antibodies and ~80 minutes) than rt-PA (4-5 minutes).
- Streptokinase treatment results in accumulation of fibrinogen-degradation products (which can increase bleeding risk by affecting platelet function); rt-PA does not.
- Streptokinase has high antigenicity; rt-PA does not, although it can in some cases cause allergic reactions as per any medicine.

In addition, there were differences in the design of the clinical trials with streptokinase compared with rt-PA, for example in the use of anticoagulants and/or aspirin in the MAST studies:

MAST-I (enrolment up to 6 hours post-symptom onset) (MAST-I group, 1995):

- 622 patients randomised in a 2x2 factorial manner to 15MU streptokinase (n=157), 300 mg/day buffered aspirin for 10 days (n=153), both active treatments (n=156), or none (n=156).
- Other medications were permitted, however thrombolytics, heparin, oral anticoagulants and antiplatelet treatments were avoided for the first 10 days. Subcutaneous heparin was allowed at no higher than 15,000 U daily.

MAST-E (enrolment up to 6 hours post-symptom onset) (Europe study group, 1996):

- 310 patients randomised to streptokinase (n=156) or placebo (n=154)
- 65% of streptokinase group, 75% of placebo group received concomitant heparin (31% and 12% were within 12 hours of randomisation respectively)
- 21 patients in each group received aspirin within 48 hours of randomisation.

In contrast, in the NINDS trial (NINDS stroke study group, 1995), anticoagulants and anti-platelet agents were not allowed during the first 24 hours, after this time they could only be used once CT scan at 24 hours had confirmed the absence of haemorrhage. Patients who were taking anticoagulants or received heparin in the preceding 48 hours and with an elevated partial thromboplastin time, or prothrombin time >15s, or platelet count below 100,000/mm^3.Similarly in the ECASS III trial (Hacke *et al*, 2008), treatment with i.v. heparin, oral anticoagulants, aspirin or volume expanders during the first 24 hours was not permitted. Subcutaneous heparin (≤10,000IU) or equivalent low-molecular-weight-heparin was permitted for VTE prophylaxis.

For these reasons it is not considered appropriate to consider the results obtained with these two agents to be interchangeable. The scope of the issues under evaluation includes only rt-PA, and the balance of benefits and risks of streptokinase in the unlicensed indication of acute ischaemic stroke will not be considered further.

In addition to streptokinase, there are a number of other thrombolytic medicines although none are licensed for the treatment of acute ischaemic stroke. Some of these may provide more appropriate comparisons with rt-PA.

Desmoteplase is structurally related to rt-PA and may be even more fibrin specific than rt-PA. As yet, clinical trial data for desmoteplase in the treatment of acute stroke have not proven to be very promising, with the phase III study DIAS-3 failing to meet its primary endpoint (mRS=0-2) with 51.3% in the desmoteplase group and 49.8% in the placebo group (Press Release, Lundbeck, 27 June 2014). However, there may be several reasons for the disappointing results, for example the time-window for enrolment of patients following stroke onset was 3-9 hours.

Tenecteplase is similar to rt-PA and has been suggested to be more effective than rt-PA in a phase IIb study involving 75 patients randomised in three groups up to 6 hours post-symptom onset (mean 2.9 hours). The co-primary endpoint was a measure of reperfusion and improvement on the NIHSS (Parsons *et al*, 2012). The higher dose of tenecteplase was superior to the lower dose and to rt-PA for all efficacy outcomes, including absence of serious disability at 90 days (72% vs 40% rt-PA).

Reteplase is a third thrombolytic agent that has similarities with rt-PA. Data on use of reteplase in acute ischaemic stroke is limited. A phase I dose-ranging study using 4 doses of intra-arterial reteplase with intravenous abciximab in 20 patients with acute ischaemic stroke treated between 3-6 hours following symptom onset demonstrated

partial or complete recanalization in 13 out of 20 patients, and 13 patients had early neurological improvement (Qureshi *et al*, 2006). A further study treating patients with reteplase and abciximab between 3 and 24 hours of stroke onset (ReoPro Retavase Reperfusion of Stroke Safety Study – Imaging Evaluation (ROSIE)) presented results in the form of an abstract for 34 patients who received increasing doses of reteplase. Reperfusion rates increased with increasing dose of reteplase.

3.3 30 day mortality

"30 day mortality is increased in those treated with alteplase in all studies bar NINDS (see figure 2). 30 day mortality is not examined in Cochrane Reviews or pooled analyses. 30 day mortality is very widely used outside stroke (eg by MHRA in coronary thrombolysis evidence). [redacted] *"*

It is generally recognised that the longer the duration of follow-up in a clinical trial the better, particularly where the disease under evaluation is chronic, may have long-term consequences or the performance of the patient is expected to change substantially in the weeks/months after the event. Realistically duration of follow-up in a clinical trial will be influenced by financial considerations, with a balance struck between longer follow-up and increasing financial burden. It is considered that the use of 90 day mortality in the assessment of thrombolysis for ischaemic stroke in preference to 30 day mortality likely reflects this situation. In general the rate of recovery from stroke is usually highest in the first few weeks after the event. Functional improvement may continue for many months and up to several years for some patients, albeit at a slower rate. The rate and completeness of recovery varies greatly, and the pattern of recovery reflects initially the recovery of ischaemic neurones in the penumbra, with neuroplasticity/adaptive changes being the most important later. On average, it is thought that around the first 3 months or so after stroke is most important in terms of a patient's recovery. Disability has been found to remain stable between 6-9 months and 5 years after stroke (Carod-Artal and Egido, 2009).

Cramer *et al* (2007) provide examples from a number of different studies into stroke recovery of general recovery rates for different neurological functions after stroke, for example, motor function was shown to have the most dramatic improvements in the first 30 days, gains in constructional apraxia made for up to 6 months after stroke and language deficits having gains for months to years. Gait was found to improve over a 14 week period, but can continue to improve in some patients, whilst resolution of urinary incontinence has been observed to occur beyond 20 weeks and cognitive function may continue to improve for many months.

These examples of recovery rates suggest that an assessment made at 90 days would likely provide a better indication of a patient's probable outcome from their stroke than an assessment at 30 days, and therefore may explain why day 90 and not day 30 mortality has been analysed within the Cochrane Reviews and the pooled-analyses. The Committee for Proprietary Medicinal Products (CPMP) Points to Consider guidelines (2001) recommend a study duration of three months for pivotal trials of medications for acute stroke.

3.4 The NINDS trial

3.4.1 Design of the NINDS trials:

- *Primary outcome changed between part 1 and part 2*
- *Outcome assessment potentially not blinded*
- *Some on-site nurses employed by manufacturer*
- *Randomisation local, not central*
- *Modest size with 624 participants*

- *Minor, not major disability was primary outcome*
- [redacted]

Each of the above concerns will be considered in turn.

- *Primary outcome changed between part 1 and 2*

The FDA report of the NINDS study [http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm080832.pdf] states that Part 1 of the NINDS study was designed as a study of early activity of rt-PA in stroke. The planned size of the study was 280 patients and there were no plans for an immediate continuation into an additional study. Part 1 of the study had a primary outcome of an improvement by 4 points in NIHSS or resolution of the deficit within 24 hours of the onset of stroke. The FDA reports that "by late in the conduct of this study the investigator group had determined that the outcome at 90 days was more informative of clinical benefit, and for a variety of reasons wished to proceed directly into a phase 3 efficacy study. Thus it was decided to add a second study, of approximately 300 patients, which would commence immediately upon ending the enrolment in the Part 1 study. Interim analyses of the Part 1 study results were used in the selection of the Part 2 study's primary endpoint. CBER [FDA] was involved with the discussions leading to this plan of immediate initiation of the Part 2 study, and the analytical differences between the studies….Although the analytic plans regard the two studies as completely independent, they were not entirely so with regards to randomization. The two studies in effect used a single randomization list. The only apparent interaction between the studies with regard to this is that the randomization lists were blocked (within each clinical center), and there will exist a transition block at each clinical center where the earlier patients were enrolled into the Part 1 study, and the remainder of the block were enrolled into the Part 2 study." The FDA comments that this is unlikely to have had a significant effect upon the studies.

Part 2 of NINDS, described as the pivotal study, had a primary outcome of proportion of patients who recovered with minimal or no deficit three months after treatment (using 4 different stroke scales). The FDA states that "Selection of the Global Statistic as the primary endpoint [used in Part 2], and the uncertainty of interpreting it for clinical meaningfulness caused some concern for its suitability for regulatory purposes. Consequently Genentech began in 1994 to discuss alternative analytic plans with CBER for the NINDS Stroke Studies to use in potential support of a licensure application…..The final Genentech analytical plan for these studies retained the concept of the Part 1 study designed with the objective of a 24 hour assessment activity endpoint, and the Part 2 Study objective of the 3 month efficacy outcome."

Changing the primary outcome of a trial part-way through a single study would normally be considered to be problematic. However, using a different primary outcome for a second trial based on the interim results from a first trial does not raise issues. In this situation the second trial must initially be analysed as an independent trial and this trial must be positive alone. This is because the second trial is an unbiased assessment of the new primary endpoint, while in the first trial that endpoint is a retrospective choice and so including those data could introduce bias. The second trial was analysed independently from the first in the assessment report for the licensing application and NINDS part 2 was considered be a positive pivotal trial. Subsequently analysis of the results from the two studies together is reasonable given that the study designs are otherwise the same and provided the data from the two studies seem reasonably consistent.

- *Outcome assessment potentially not blinded*

Based upon other communications from Dr [redacted] it is anticipated that this point either refers to:

a) the composition of the placebo and the possibility that it did not froth as would be expected for a protein or
b) study investigators being unblinded to treatment as a result of visible bleeding events in patients treated with rt-PA

The composition of the placebo used in the NINDS study was one of the points raised with the MAH in the list of questions provided on 1 August 2014. The MAH has responded stating that the source data of NINDS was requested from Genentech however Genentech has informed the MAH that the documentation for this study has been transferred to an external archiving company. The documents are therefore not readily accessible, however as soon as Genentech provides the data the MAH will submit it to the MHRA.

On the subject of visible bleeding, this issue was considered in the May 2014 CHM paper as follows (see section 4.2.4 of CHM paper):

"The Cochrane review states that in NINDS, follow-up at all stages was to be by a doctor (blinded) who had not been involved in the randomisation or care of the patient in the first 24 hours. The issue of visible external bleeding would not be avoidable. The NINDS publication reports a small number of serious systemic bleeds (5 in the rt-PA group, none on placebo), and a higher number of minor external bleeds in the first 10 days (23% in the rt-PA group vs. 3% in the placebo group)."

It is not possible to determine what effect, if any, this would have had on the results. The endpoints employed were fairly objective and therefore doctors carrying out the assessments should not have had their ratings affected by information on bleeding. Further, as there were some bleeding events recorded in the placebo group, the presence of bleeding alone would not conclusively confirm any individual's treatment assignment.

- *Some on-site nurses employed by manufacturer*

This issue was also discussed in the May CHM paper (section 4.2.4), as follows:

"The NINDS re-analysis states that Genentech nurses determined which specific medications recorded on forms were considered antihypertensive therapies, and in some cases these nurses performed pharmacologic monitoring (relating to blood pressure). It is not possible to determine what effect, if any, this would have on the results."

- *Randomisation local, not central*

Dr [redacted] s concerns regarding the local randomisation protocol are that this protocol broke down and delivered unbalanced arms to the trial. This was also discussed in the May CHM paper (section 4.2.4), as follows:

"The Cochrane review explains that randomisation was by selection of a sealed, sequentially numbered, pre-pack (of active drug or identical appearing placebo), followed within 2 hours with a telephone call to the co-ordinating office to notify them of the patient and number of the drug pack. This system was designed to reduce delays in treatment. An error led to 'out of order' treatment allocations in between 13 and 31 patients which affected every patient until the error was detected, and led to patients appearing to cross between treatment allocations.

The details of the 31 patients, the effects on their treatment and their outcome are described in the FDA clinical review at

http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm080832.pdf

The [FDA's] conclusion regarding these patients was that the error in the process appears not to have contributed to any bias in overestimating the treatment effect, and the errors do not seem to have altered the overall outcome of the studies."

Whilst centralised randomisation provides the most reliable method of randomisation, when the time to treatment must be minimised as in this case, it is perhaps understandable why local randomisation was considered more appropriate, particularly for a study carried out in the early 1990's when methods of communication were less efficient. The FDA analysis of the results for the individual patients affected by randomisation errors determined that there would have been no impact on the study results due to these errors.

- *Modest size with 624 participants*

The NINDS publication (NINDS stroke study group, 1995) explains the sample size calculation, which for part 1 was designed to have a power of 0.90 to detect an absolute difference of 24 percentage points in outcome given a rate of 16% in the placebo group. For part 2, the power was 0.95 to detect a difference of 20 percentage points between groups in a single outcome measure (i.e. mRS, BI, GOS or NIHSS). Retrospective assessment of sample size is not usually a profitable exercise. As the study was positive it was by definition acceptably powered for the primary endpoint.

The NINDS study required patients to be enrolled and treated within a maximum of 3 hours following onset of stroke symptoms, and was conducted at a time when treatment of stroke was not viewed as a medical emergency and so there would not have been the infrastructure that is in place today, to deal with stroke patients. This requirement, together with other exclusion criteria, resulted in the need to screen a large number of patients (n=17,991) in order to achieve the 624 participants[1]. The practicalities of the study and the novelty of treating stroke, let alone as a medical emergency, may have rendered a larger study unfeasible and of unacceptably long duration – as it is, the NINDS trials were conducted over a four year period.

- *Minor, not major disability was primary outcome*

The primary endpoint in part 2 of the NINDS trial was the proportion of patients with minimal or no deficit at day 90 in the rt-PA group vs. the placebo group. Four outcome scales were used in the assessment, the Barthel Index (BI), the modified Rankin Scale (mRS), Glasgow outcome scale (GOS) and the NIH stroke scale (NIHSS).

There is no objection in principal to the choice of endpoint. An increase in the proportion of patients with only minor disability is considered to be clinically useful. In addition, a full assessment of the data is not restricted to the primary endpoints alone and the full scales are also examined when considering the overall effect of treatment.

[1] Of the 17,367 excluded patients, 8,708 were excluded based on the time from onset of symptoms being too long [FDA clinical review].

A separate discussion on the choice of endpoints in trials in acute ischaemic stroke, in particular the use of dichotomised vs. ordinal analyses, and the different outcome scales is provided in section 5.

3.4.2 Conduct of the NINDS trials:

- *Problems over randomisation process*
- *Highly distorted spread of onset to treatment time*
- *Blood pressure evaluation poor*
- *A member of trial design team asked to resign*

Each concern is considered separately below:

- *Problems over randomisation process*

The randomisation process used in NINDS is discussed in the previous comments box.

- *Highly distorted spread of onset to treatment time*

This concern, that 50% of all patients who were treated in the 0-90 minute time window were reportedly treated between 89-90 minutes, was evaluated in the re-analysis of the NINDS trial in 2004 and discussed in the May CHM paper (section 4.2.3.1) as follows:

"The distribution of the time from onset of symptoms is indeed strange, but it is not agreed that this raises concerns about the results or study conduct. As noted in the review *[published re-analysis of the NINDS trial, 2004 (O'Fallon et al, 2004)]* "Considering the questionable precision with which many patients' 'time of onset' must have been estimated and the intense setting of an emergency department the precision of these OTT values and their accumulation just before 90 minutes is questionable." It should also be remembered that this variable was used to stratify the randomisation based upon only two categories, whether time from onset was ≤ 90 minutes or > 90 minutes, and this dichotomised variable was to be used as a covariate. In this setting it seems possible that investigators just focussed on capturing the correct categorisation in relation to the 90 minute threshold and were entering values of 89 and 90 minutes to capture this and not bothering with the precise time. It is also possible that the data are genuine and investigators were targeting a treatment time of just before 90 minutes, or rushing in some way to get into the early strata having fulfilled their allocation into the later strata. Given this distribution, whether genuine or an artefact of the planned dichotomisation, as also concluded by the review authors there is little value in analyses using the time to randomisation as a continuous variable.

The review also notes the poor performance of placebo in the 91-133 minute window and the imbalance with a larger number of placebo patients than expected falling into this group. The treatment difference is indeed largest in this group [in favour of rt-PA] and there is an imbalance – however it must be remembered that when looking at

sub-group analyses, retrospectively looking for the largest difference will often lead to an extreme result. The important thing is to be reassured that there is not a group of patients eligible for treatment that the data suggest receive no benefit.

If it seemed that the overall treatment difference was entirely driven by the 91-133 minute subgroup, and was consequently magnified by the imbalance, then there would be concern. However this does not seem to be the case. Statistical significance is reached for all 4 endpoints when the 0-90 minute subgroup is taken alone. The trend is positive in all subgroups for all 4 endpoints, and on many occasions point estimates from the later 3 groups exceed that for 0-90 minutes. There is no clear ordering of the later 3 groups across endpoints, and there is considerable overlap between the confidence intervals for the different groups.

Therefore there seems no reason based upon these data to be concerned that within the 3 hour window there is a group of patients defined by time to randomisation who seem to receive no benefit from treatment."

- *Blood pressure evaluation poor*

There were issues identified in the published re-analysis of the NINDS study (O'Fallon *et al*, 2004) regarding the measurement of blood pressure in NINDS, these are reproduced in the May CHM paper, section 4.3.2.1. As described in the paper, the statistical assessment of these issues is that:

"While it seems that the collection and monitoring of blood pressure data and treatment could have been better and this has meant that conclusions regarding the impact of blood pressure management measures cannot be drawn, in line with the [NINDS re-analysis] committee it is not considered that this calls into question the primary conclusions of the study regarding the efficacy of t-PA."

- *A member of trial design team asked to resign*

[redacted]

3.4.3 Results of the NINDS trials:

- *Primary outcome in Part 1 not significant at p<0.05*
- *Asymmetrical 'funnel plot' of outcome by centre*
- *Randomisation yielded fitter patients in treatment arm*

- *Minimal improvement in individual Stroke Score*
- *Unconvincing spread of modified Rankin scores*
- *Surprisingly low 7 day mortality with treatment*
- *Increase in fatal and non-fatal cerebral haemorrhage*

Each concern is considered separately below:

- *Primary outcome in Part 1 not significant at p<0.05*

As this was the first ever trial of rt-PA in the treatment of acute ischaemic stroke there was no precedent in terms of the optimal outcomes/endpoints. Based on the results observed in animal models the primary outcome in part 1 was defined as an improvement of 4 points over baseline values in the NIHSS score, or the resolution of the neurologic deficit, within 24 hours of the onset of stroke. The day 90 evaluations of the mRS, BI, GOS and NIHSS were secondary endpoints in the NINDS part 1 trial, and the primary outcome for part 2.

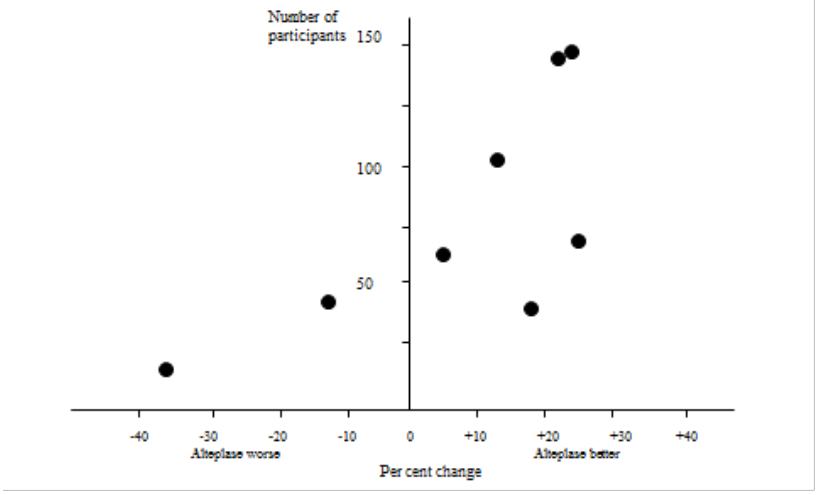
In Part 1, the 24 hours primary outcome showed a trend to early improvement with rt-PA treatment, but did not reach statistical significance (p=0.21 for 0-3 hours). Meanwhile, the secondary endpoints evaluated at 90 days post-stroke were found to demonstrate benefit of rt-PA treatment (p=0.001 for Global test). The interim analyses of Part 1 were used to inform the selection of the primary endpoint of Part 2. In addition outcome at 90 days following stroke was considered to be of more clinical relevance than outcome at 24 hours.

The fact that the primary endpoint for Part 1 did not achieve statistical significance is not in itself of concern, provided that the outcome at day 90 following a stroke is of more clinical relevance than the outcome at 24 hours post-stroke – the reason for the change in outcome between the two studies – and the positive findings are replicated in the second part of the trial. The conclusion on clinical relevance seems reasonable and part 2 of the trial was positive.

- *Asymmetrical 'funnel plot' of outcome by centre*

The issue that the apparent success of rt-PA across the eight NINDS clinical trial centres differed and that smaller centres did not underpin the results from larger centres was discussed in the May CHM paper. The following figure is the funnel plot referred to by [redacted]:

Figure 1



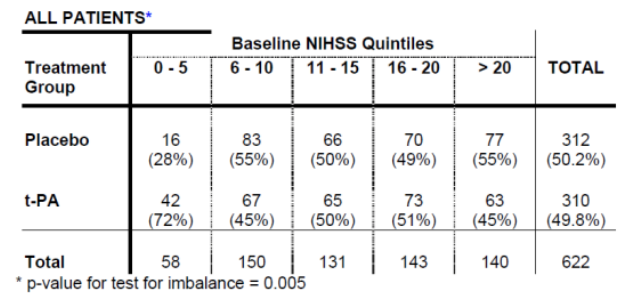
The results for all four of the outcome measures per centre are provided in the May CHM paper, in section 4.2.3.1. The statistical assessment of this issue, as provided in the paper, is:

"Interaction tests generally lack power, so it would be unwise to conclude on the basis of negative interaction tests that there were no important differences between centres. However the subgroup analyses by centre do not present concern. It does not seem that the results from the largest centres always give the largest treatment differences; centre 5 [the largest] ranks 3rd, 2nd, 3rd and 5th of the 8 centres across the 4 endpoints, while centre 4 [the second largest] is 1st, 3rd, 2nd and 4th. The two smallest centres (6&9 and 7) were consistently the two worst; however centre 1, a similar size to the 6&9 grouping saw good results (2nd, 4th, 1st, and 3rd). Statistically, the expectation would be that smaller centres are generally the furthest away from the true result because of the higher standard error associated with estimates based on a small number of patients. Simply plotting the point estimates of the treatment effect seen in each centre without considering the variability of those estimates is likely to be misleading. Looking at the confidence intervals from each centre there is considerable overlap, even between the best and the worst results. There is nothing in the data to suggest the positive conclusions are entirely driven by a few large centres or anything to suggest the performance truly differed between centres and that small centres were truly worse."

- *Randomisation yielded fitter patients in treatment arm*

The baseline imbalance in stroke severity in the arms of the NINDS trial, and the suggestion that this may have been the driver for the benefit observed with rt-PA treatment, is an issue that has been highlighted frequently over the years since the trial was published. Addressing this issue was one of the main purposes of the committee set up to re-analyse the NINDS data in 2004, and it was discussed in the May CHM paper, section 4.2.3:

"The committee re-analysis confirmed the existence of an imbalance, with more rt-PA patients with NIHSS scores 0-5, the patients with a better prognosis (O'Fallon *et al*, 2004):



ALL PATIENTS*

Treatment Group	Baseline NIHSS Quintiles 0 - 5	6 - 10	11 - 15	16 - 20	> 20	TOTAL
Placebo	16 (28%)	83 (55%)	66 (50%)	70 (49%)	77 (55%)	312 (50.2%)
t-PA	42 (72%)	67 (45%)	65 (50%)	73 (51%)	63 (45%)	310 (49.8%)
Total	58	150	131	143	140	622

* p-value for test for imbalance = 0.005

The committee presented sub-group analyses broken down across the baseline NIHSS quintiles. Results across the sub-groups were as would be expected, with more favourable outcomes in the milder groups and very few in the severe groups; the impact of treatment was seen in the much steeper decline in favourable outcomes with severity for the placebo arm:

Barthel index by baseline NIHSS quintiles

Baseline NIHSS	n/N (%) Favourable outcome			
	rt-PA	Placebo	Difference	95% CI
0-5	35/42 (83%)	15/16 (94%)	-10.4%	(-27.1, 6.3)
6-10	53/67 (79%)	46/83 (55%)	23.7%	(9.1, 38.3)
11-15	34/65 (52%)	27/66 (41%)	11.4%	(-5.7, 28.5)
16-20	26/73 (36%)	18/70 (26%)	9.9%	(-5.2, 25.0)
>20	14/63 (22%)	13/77 (17%)	5.3%	(-8.0, 18.7)

Modified Rankin score by baseline NIHSS quintiles

Baseline NIHSS	n/N (%) Favourable outcome			
	rt-PA	Placebo	Difference	95% CI
0-5	33/42 (79%)	13/16 (81%)	-2.7%	(-26.0, 20.6)
6-10	46/67 (69%)	38/83 (46%)	22.9%	(7.3, 38.4)
11-15	27/65 (42%)	15/66 (23%)	18.8%	(3.0, 34.6)
16-20	21/73 (29%)	14/70 (20%)	8.8%	(-5.3, 22.9)
>20	6/63 (10%)	3/77 (4%)	5.6%	(-2.9, 14.1)

Glasgow outcome scale by baseline NIHSS quintiles

Baseline NIHSS	n/N (%) Favourable outcome			
	rt-PA	Placebo	Difference	95% CI
0-5	34/42 (81%)	14/16 (88%)	-6.5%	(-27.1, 14.0)
6-10	48/67 (72%)	44/83 (53%)	18.6%	(3.3, 34.0)
11-15	30/65 (46%)	18/66 (27%)	18.9%	(2.5, 35.2)
16-20	22/73 (30%)	15/70 (21%)	8.7%	(-5.7, 23.1)
>20	7/63 (11%)	6/77 (8%)	3.3%	(-6.6, 13.2)

NIHSS by baseline NIHSS quintiles

Baseline NIHSS	n/N (%) Favourable outcome			
	rt-PA	Placebo	Difference	95% CI
0-5	29/42 (69%)	10/16 (63%)	6.5%	(-21.6, 34.7)
6-10	35/67 (52%)	29/83 (35%)	17.3%	(1.4, 33.2)
11-15	22/65 (34%)	13/66 (20%)	14.1%	(-1.0, 29.3)
16-20	16/73 (22%)	10/70 (14%)	7.6%	(-5.0, 20.3)
>20	4/63 (6%)	2/77 (3%)	3.8%	(-3.3, 10.8)

Favourable trends were seen in favour of rt-PA for all 4 endpoints in all of the sub-groups and statistical significance in favour of rt-PA was achieved for the 6-10 sub-

group taken alone for all four endpoints. The exception was the 0-5 group where the trend favoured placebo for three of the four endpoints, though the confidence intervals were wide and the percentage with a favourable outcome high in both groups.

These sub-group analyses make it clear that the overall benefit of rt-PA was not seen only because of the baseline imbalance.

The committee also conducted many covariate adjusted analyses and concluded that "*After a thorough evaluation of this issue, we found no evidence that the imbalance in the distribution of baseline NIHSS between the treatment groups had either a statistically or clinically significant effect on the study results. We have determined that the original models using both Age and BsNIHSS [baseline NIHSS] as continuous variables properly adjust for the complex roles played by these two variables, both so strongly (negatively) related to the likelihood of a favorable outcome. There was a strong interaction between age and baseline NIHSS in the Global analysis and in the analyses of each of the four outcome measures. The likelihood of a favorable outcome was particularly low in patients older than 70 who had a baseline NIHSS more than 20. However, there was no evidence of any Age by BsNIHSS subgroup responding significantly differently to t-PA treatment than the study group at large.*""

- *Minimal improvement in individual Stroke Score*

No elaboration of this point has been provided by Dr [redacted] in his submission or in previous communications but it may be referring to the change from baseline in NIHSS score analysis, this is discussed in section 7.2.

- *Unconvincing spread of modified Rankin scores*

A similar point was raised by Dr [redacted] previously and discussed in the May CHM paper:

> *"…the patterns seen when the pooled outcome data* [from a pooled analysis of clinical trials by Lees *et al*] *are examined graphically raise concern. A plausible spread of modified Rankin scores in observational stroke cohorts, contrasts with an uneven pattern in the alteplase trialist's analyses. The data on patients randomised within 180 min are particularly uneven and are predominantly from the NINDS trial. The authors of the pooled analyses made no reference to the Cochrane review, which indicated this pooling and modelling of data 'may be incorrect'"*

The concern would appear to relate mostly to the smaller proportion of patients classified as mRS 2 and 3 in RCTs compared with the observational cohorts at all time points, which results in these curves appearing uneven in comparison with the observational cohorts. The discussion and graphical representation of these data are provided in the May CHM paper, in section 4.3.5:

"The figures do not show the variability around these curves, which is an important consideration when assessing if the distributions are truly different. Notably figure a) is based on 10,231 patients [from the SITS-ISTR registry] and so would be expected to be smooth, while figure e) is based on only about 150 for each curve [randomised trials of rt-PA, 3 month outcome for patients treated within 90 minutes], divided across the 7 points. Curves based on such a small number of data points would be expected to be uneven. Figure g) has the largest number of patients of any of the graphs representing the randomised trials at about 800 per curve [3 month outcome, treated between 181-270 minutes], and here it seems that the rt-PA curve is actually very similar to the curves in figures a) and b) which also represent rt-PA treatment [data from SITS-ISTR].

Furthermore there is no *a priori* expectation that the distribution of scores in a randomised trial should mirror those from an observational cohort, as the two situations have many differences, including inclusion/exclusion criteria and the level of monitoring and follow-up etc. The focus in a randomised trial is the comparison between the randomised groups rather than the distribution in an individual group.

For both these reasons it is not considered that there is anything here which should lead to concern regarding the outcomes of the placebo controlled t-PA trials."

- *Surprisingly low 7 day mortality with treatment*

Mortality is an unambiguous endpoint, and therefore it can be assumed that the data obtained for percentage mortality at 7 days is accurate. The data on 7 day mortality as presented in Dr [redacted] s submission for the treated and untreated arms of the NINDS trial do not fall outside of the ranges observed for the other randomised controlled trials of rt-PA:

Trial	Participants (n)	Percent mortality at 7 days		
		rt-PA	Control or placebo	Improvement with rt-PA
NINDS (i)	624	4.3	8.1	3.8
ECASS (ii)	620	11.8	8.5	-3.3
ECASS 2 (iii)	800	6.1	5.1	-1.0
ECASS 3	821	2.9	3.2	0.3
ATLANTIS (iv)	755	2.6	0.3	-2.3
EPITHET (v)	101	11.5	2.0	-9.5
IST-3	3035	10.8	7.0	-3.8

i Data read from Kaplan-Meier curve, absent from Cochrane
ii '7 or 10 day' data from Cochrane 2009
iii '7 or 10 day' data from Cochrane 2009
iv Data from table on fatal intracerebral haemorrhage, absent from Cochrane
v Estimated from 5 day mortality imputed from main paper and Cochrane review 2009

Table reproduced from Dr [redacted] s submission. Note that these data refer to all-cause mortality, except for the information on ATLANTIS, which are data on fatal intracranial haemorrhage for ATLANTIS B.

- *Increase in fatal and non-fatal cerebral haemorrhage*

Intracranial haemorrhage (ICH) is a well-recognised risk of treatment with rt-PA and constitutes the main safety concern associated with treatment.

Different definitions of ICH were used in different trials, and defining when an ICH is clinically significant and related to treatment is not straightforward. For example haemorrhagic transformation of an ischaemic stroke can occur as a natural event with bleeding into an already infarcted area of brain – this may have no clinical

consequences in itself. In NINDS symptomatic ICH was defined as "any CT-documented haemorrhage that was temporally related to deterioration in the patient's clinical condition in the judgement of the clinical investigator", irrespective of the size of the haemorrhage on the CT scan. In addition symptomatic ICH attributable to study medication was defined as "symptomatic haemorrhage that occurred within 36 hours from treatment onset".

Further discussion of the most appropriate definitions of symptomatic ICH and the frequencies with which this occurs after treatment with rt-PA is provided in section 4.

3.4.4 Generalisability of the NINDS results:

- *Only highly selected patients enrolled*
- *Few over 80 years of age*
- *Within 3 hours*

Each specific concern is considered below:

- *Only highly selected patients enrolled*

The population included in NINDS was highly selected, as previously discussed, a total of 17,991 patients were screened in order to enrol 624 participants – in other words, only 3.6% of screened patients were suitable to be included in the trial. The FDA clinical review provides details of the main reasons that patients were excluded during screening, with 51.6% (8,708) on the basis of the time from onset of symptoms being too long. The other reasons for exclusion were:

Symptoms rapidly improving	1749 (10.4%)
Intracranial haemorrhage	1306 (7.8%)
Symptoms too minor	1106 (6.6%)
Outside age range[a]	1021 (6.1%)
Other serious illness	490 (2.9%)
Seizure at stroke onset	391 (2.3%)
Stroke not present	373 (2.2%)
Time from onset 90-180 min[b]	267 (1.6 %)
Recent prior stroke	219 (1.3%)
Oral anticoagulants	210 (1.2%)
Subarachnoid haemorrhage	169 (1.0%)

[a] inclusion criterion: age ≥18 years

[b] site enrolment unbalanced at the time the specific patient was screened. Individual sites were required to maintain the number enrolled in the two time strata equal within 3 hours

The most frequent reason for exclusion, in more than half of screened patients, was the requirement for treatment to occur within 3 hours of symptom onset. As discussed above, at the time of the NINDS trial, treatment of stroke had not been viewed as a medical emergency, either by doctors or by patients and in addition there was very poor public awareness of the symptoms of stroke. Therefore it is not perhaps surprising that a large proportion of screened individuals were excluded on the basis of time from onset of symptoms being too long. The other reasons for exclusion make up a smaller proportion of screened patients, and are generally either related to the appropriateness of treatment (haemorrhagic stroke) or to safety precautions (taking oral anticoagulants) or are populations commonly excluded in trials (patients under 18 years of age).

Generalisability of results from any randomised controlled trial will always be an issue to some degree, because trials are designed to determine the level of treatment effect and therefore require a degree of uniformity in the included population and

because often some patients will be excluded on safety grounds. As a result it will often be necessary to strictly define the patient population for whom the treatment is eventually licensed – as was the case for rt-PA, which was initially licensed for treatment up to 3 hours post-symptom onset - and impose a number of other restrictions (which reflect the exclusions) in the form of contraindications.

- *Few over 80 years of age*

Patients aged over 80 years of age were not excluded from NINDS, although the numbers enrolled were small [19 in the placebo group and 25 in the rt-PA group (Longstreth *et al*, 2010)]. The low number of patients over 80 is presumably due to either smaller numbers of patients in this age category presenting within 3 hours of symptom onset, or patients in this age category having a higher frequency of exclusion criteria, or both. The other main rt-PA trials excluded patients >80 years of age, and therefore only a very small number were included in total. When rt-PA was licensed in the UK, patients aged over 80 years were contraindicated. The more recent IST-3 trial (IST-3 collaborative group, 2012) included a large number of patients aged over 80 years – this study is discussed in more detail in section 3.6 and in paper 3. It will also be considered in terms of off-label use in papers to be considered in January.

- *Within 3 hours*

This inclusion criterion resulted in over half of screened patients being ineligible for inclusion in the trial, and clearly therefore restricts the generalisability of the results. However, this criterion was based on the data available at the time, and the understanding that the quicker treatment can be delivered, the greater the likelihood of a favourable outcome. The EU rt-PA licence was initially also restricted to treatment within 3 hours of symptom onset and, even with the advances in stroke care since NINDS, the majority of patients who present with a suspected stroke are not eligible for treatment with rt-PA. From that perspective the trial is reflective of use within the terms of the licence.

3.4.5 Review committee which conducted re-analysis of NINDS in 2004 (O'Fallon *et al*, 2004):

- *Chosen through collaboration with trial organisers*
- *Narrow terms of reference*
- *Key information placed on temporary website*
- *Website only reinstated in 2014*
- *Declined to investigate conflict of interest*
- *Accepted 'prototype' status of trial*
- *Highlighted 'onset to treatment time' gradient unproven*

Addressing each concern in turn:

- *Chosen through collaboration with trial organisers*

The re-analysis of the NINDS trial was commissioned by NINDS as a result of public debate around the study findings and discussion about the appropriate use of rt-PA in the treatment of stroke. As such, NINDS appointed the chair of the committee, Professor W Michael O'Fallon (professor of biostatistics), and the chair chose the members of the committee. It is stated in the committee report that NINDS did not participate in the selection of any member of the committee other than the chair and neither NINDS staff nor original NINDS rt-PA Stroke Study Group investigators participated in any of the committee's meetings, though there are anecdotal reports

that the project leader of NINDS may have provided a list of emergency physicians for consideration (Lenzer 2002). It is also stated that NINDS staff/investigators communicated only rarely with the committee/chair and then only at the invitation of the committee. The committee members were paid as hired consultants to an independent contractor to NINDS. None of the committee members had a connection with the previous published study or with the manufacturer of rt-PA.

- *Narrow terms of reference*

The main charge given to the committee was:

"to address whether there is concern that eligible stroke patients may not benefit from rt-PA given according to the protocol used in the trials and, whether the subgroup imbalance (in baseline stroke severity) invalidates the entire trial as claimed by some of the critics."

This charge reflects the extensive debate around the results of the NINDS trials at the time, with the main issue of concern being the imbalance in baseline stroke severity. However, in addition to examining this issue, the committee also ensured they could reproduce the tables in the original analyses from the original dataset as well as reviewing a number of other specific issues including blood pressure assessment/management, intracerebral haemorrhage, onset to treatment time, centre effect, stroke subtype, pre-existing disability and diabetes mellitus.

Dr [redacted] s submission of 5 September/October elaborates on this concern [Narrow terms of reference] in the text to state that the review did not examine the results in the context of other trial data, particularly referring to day 7 and day 30 mortality data and comments that the NINDS data for these data points tends to outlie other trial data. As discussed above, while the difference between the % mortality in the rt-PA treated group vs. the placebo group is indeed outside of the range of the other trials quoted, the actual % mortality in each group is within that observed in other trials.

Although the main aim of the NINDS re-analysis was to ascertain the impact of the imbalance in stroke severity at baseline, it also reviewed a number of important additional aspects of the data. As this committee was specifically convened to examine the issues with the NINDS data only, and not with the aim to review all clinical trials of rt-PA in stroke, the terms of reference are considered to be appropriate.

- *Key information placed on temporary website*
- *Website only reinstated in 2014*

In the publication in Stroke summarising the committee's findings the link in the references to the full committee report is no longer active, and the full report is now provided at a new website address (http://stroke.nih.gov/resources/t-pa-review-committee.htm). It is not clear how long the report was unavailable or the reason why it was unavailable however, it is possible this was an administrative or data-migration error when the new website was set up.

- *Declined to investigate conflict of interest*

The charge to the committee also stated:

"The issue of whether pharmaceutical company participation biased the results of the trial is an important, but secondary issue for the group."

However, the committee declined to consider this secondary issue on the grounds that "it was in no position to assess whether financial arrangements biased any of the parties involved in the study, approval and endorsement of rt-PA." In general most

drugs that are approved for use are based on trials conducted with some form of company sponsorship or involvement. Dr [redacted] goes on to state of the industry trials of rt-PA (ATLANTIS A and B, ECASS I, II, III) that "These trials seemed of mainly sound design and conduct."

- *Accepted 'prototype' status of trial*

The committee report (O'Fallon *et al*, 2004)describes the NINDS trial as a prototype study, in recognition of the fact that it was the first randomised controlled trial of rt-PA and was designed to demonstrate differences in the entire group of patients and not in subgroups. It would not be expected or be able to address all possible questions around treatment, and the committee concluded that there were a number of questions arising from the exploratory analyses of the trial that required addressing:

- Is there a subgroup of patients for whom the risk of ICH is so high that the group as a whole has no benefit from rt-PA treatment?
- Is there a subgroup of patients with mild symptoms in whom rt-PA provides no net benefit?
- Within the time-frame of NINDS (up to 3 hours) is there evidence for a differential rt-PA treatment effect related to time from symptom onset to treatment?
- What is the impact on outcome of elevated blood pressure and its management, before and after rt-PA treatment?
- Can data from other trials be used to validate the cut-off for rt-PA treatment used by the NINDS investigators (≤185/110)?
- Can the exploratory analysis finding that stroke patients with diabetes do not benefit from rt-PA be confirmed?

- *Highlighted 'onset to treatment time' gradient unproven*

The pattern of onset to treatment times is discussed above in relation to the conduct of the study. As a result of the pattern of onset to treatment times, with many patients enrolled between 89-90 minutes following onset of stroke symptoms, analyses cannot consider onset to treatment time as a continuous variable. Therefore these data cannot support any conclusions about the relationship between onset to treatment time and efficacy of rt-PA treatment. As noted by the committee, it cannot be concluded from these data that a relationship does not exist.

3.4.6 Legacy of the NINDS trials:

- *The study has not been replicated*
- *Repeatedly added to pooled and meta-analysed studies*
- *Recurrent use in complex statistical models*
- *Years of controversy polarising stroke care providers*
- *Ongoing potential harm to stroke patients*

Each concern is addressed below:

- *The study has not been replicated*

The NINDS trial has not been replicated, with subsequent studies involving a different dose (ECASS I), a longer time-window for enrolment (ECASS I, II, II, ATLANTIS A&B), with different exclusion criteria (e.g. patients ≥80 years of age were excluded from trials other than NINDS) and different endpoints. The results of the NINDS study were widely considered to demonstrate efficacy of treatment with rt-PA within 0-3 hours of symptom onset, within the constraints of the exclusion/inclusion

criteria, and therefore randomised, placebo-controlled trials in the licensed population were not considered to be feasible/ethical. The ATLANTIS study, for example, was modified during the study to a 3-5 hour treatment time-window post-symptom onset (from 0-5 hours) in light of the NINDS trial results. At the time of this modification, only 31 patients had been enrolled from 0-3 hours.

Similarly, the arbitration procedure that took place in the EU during the assessment of the initial application for the indication in acute ischaemic stroke in 2002 (see May CHM paper, section 3.1.2 and http://www.ema.europa.eu/docs/en_GB/document_library/Referrals_document/Actilyse_29/WC500010327.pdf) concluded after advice from an EMA ad-hoc expert meeting that a placebo-controlled trial in patients treated up to 3 hours post-symptom onset would likely face recruitment issues and may not be feasible. Instead, the company was obliged to conduct a randomised placebo controlled trial to assess efficacy and safety within 3-4 hours (later modified to 4.5 hours) of symptom onset (ECASS III).

- *Repeatedly added to pooled and meta-analysed studies*

This concern is based on the view that the NINDS trial data are unreliable and therefore should not be considered within pooled or meta-analyses. In general no trial is ever free from limitations and so meta-analysis can be useful because, by including trials with different limitations, a summary result is provided, and this is not too biased towards any one trial, unless it is substantially larger than all the other trials and substantially skews the findings.

Based on the above discussions the concerns raised regarding the NINDS trial may not be so important as to render the results too unreliable to include in a pooled analysis/meta-analysis.

- *Recurrent use in complex statistical models*

No elaboration of this point has been provided by Dr [redacted] in his submission or in previous communications.

- *Years of controversy polarising stroke care providers*
- *Ongoing potential harm to stroke patients*

It is agreed that there have been years of controversy, with polarised views from different stroke care providers. The NINDS trial could either be considered to be responsible for ongoing potential harm to stroke patients or to have provided important benefits for stroke therapy, depending on your viewpoint.

3.5 Industry trials of rt-PA (ATLANTIS A and B, ECASS I, II, III)

"These trials seemed of mainly sound design and conduct. They did not show a clearcut and consistent benefit (table 4). The variable small benefit in reported disability seen in some trials may be due either to biased assessment due to problems with blinding, uneven randomisation or even chance. (Cochrane 2009, ECASS 3 2008) Mortality tended to be worse with treatment. Computed tomogram (CT) head scan outcomes revealed the expected significant rise in cerebral haemorrhage but no clear benefit with infarction. (ATLANTIS 1999) These trials were in lower risk individuals and recruited slowly. Generalisability seems questionable, even if the very modest benefit was considered plausible."

Table: The 4 largest industry sponsored trials of rt-PA

Trial (participants)	*Outcome (%)*			
	*Good**	*Independent+*	*30 day mortality*	*Primary outcome (p<0.05)*
Atlantis (755)	*Not stated*	*Worse (- 3.6)*	*Worse (i) (- 5.3)*	*Not significant Trial halted*
ECASS (620)	*Better (+6.4)*	*Better (+ 5.6)*	*Worse (- 5.2)*	*Not significant*
ECASS II (800)	*Better (+ 3.7)*	*Better (+ 8.2)*	*Worse (ii) (- 0.3)*	*Not significant*
ECASS III (821)	*Better (+ 7.2)*	*Better (+ 5.0)*	*Worse (iii) (- 0.1)*	*Positive p<0.05*

**Good = modified Rankin Score of 0 or 1,*
+Independent = modified Rankin Score of 0 - 2
(i) sum of data presented in trial papers for part A and B
(ii) data read from Kaplan-Meier curve
(iii) sum of data provided in text

Some clarifying remarks on the table above which has been taken from Dr [redacted] submission: the good/independent data quoted are outcomes at the end of follow-up. In the Cochrane review (Wardlaw *et al*, 2014), data on death or dependency (defined as mRS 2-6), indicates that the result for ATLANTIS for the 'good' outcome is -1.7%.

- *"These trials seemed of mainly sound design and conduct. They did not show a clearcut and consistent benefit (table 4)."*

The implication of this comment appears to be that, unlike the NINDS study, these studies were well-conducted and thus the lack of consistent benefit observed is more reliable than the benefit observed in NINDS.

- *"The variable small benefit in reported disability seen in some trials may be due either to biased assessment due to problems with blinding, uneven randomisation or even chance. (Cochrane 2009, ECASS 3 2008)"*
- *"These trials were in lower risk individuals and recruited slowly."*

This comment implies that its author believes the studies were negative and that the only reason for any suggestion of benefit was through error or chance.

These industry sponsored studies were assessed in detail at a European level either as part of the initial licensing procedure when the indication for treatment of acute ischaemic stroke was granted (ECASS I, ECASS II and ATLANTIS), or during the procedure to increase the time-window for treatment up to 4.5 hours (ECASS III).

The designs of these studies varied and differed from the NINDS trial, in particular with regards to the time-window for treatment (as discussed above), the dose (in ECASS I), endpoints, and baseline stroke severity, all of which should be considered when drawing comparisons between the trial results. A brief summary of these trials is provided in the May CHM paper, sections 3.1 and 3.2. A substantial number of concerns were raised, including by the MHRA, about the findings of the studies and the interpretation thereof, both at the time of authorisation and during the procedure to extend the time from onset of symptoms to treatment from 3 to 4.5 hours. However, these concerns were addressed through further analysis and presentation of additional data by the MAH.

- *"Mortality tended to be worse with treatment."*

This statement refers to mortality at day 30, as opposed to day 90/end of trial. The use of day 30 endpoints is discussed earlier (section 3.3).

- *Computed tomogram (CT) head scan outcomes revealed the expected significant rise in cerebral haemorrhage but no clear benefit with infarction. (ATLANTIS 1999)"*

Radiology outcomes are discussed later (section 3.7.2)

- *"Generalisability seems questionable, even if the very modest benefit was considered plausible."*

Generalisability of study results is discussed in more detail in sections 3.4.4 and 3.7.6.

3.5.1 Limitations of the ECASS III trial

3.5.1.1 Design of ECASS III trial:

- *An industry funded and managed trial, run by enthusiasts.*
- *Primary outcome based only on non-disabling symptoms (mR 0-1).*
- *Final place of residence not assessed or reported.*
- *30 day CT head data not assessed or reported.*
- *Outcome assessment potentially not blinded.*
- *Potential assessment bias was either not examined or not presented.*
- *Aged 18 to 80 years.*

Addressing each concern in turn:

- *An industry funded and managed trial, run by enthusiasts.*

The ECASS III trial (Hacke *et al*, 2008) was a licensing commitment following the approval of the indication in treatment of acute ischaemic stroke. As a commitment of the MAH, the study was by definition industry funded and managed. Whilst this may not be ideal, this should not, of itself, be considered to invalidate the results of the study. It is noted that the ECASS III trial (and the ECASS I, II and ATLANTIS trials are described as being "*of mainly sound design and conduct"* (see 3.1.5 above).

- *Primary outcome based only on non-disabling symptoms (mR 0-1).*

A discussion on the most appropriate endpoints for trials in acute ischaemic stroke is provided in section 5. However, this endpoint was selected prior to the study start in conjunction with CHMP whose reason for proposing it was for comparability with previous studies – NINDS, ECASS I and II. The use of the mRS as a single outcome scale (as opposed to a composite global endpoint using several scales) was

recommended because when the individual scales are dichotomised, the results for each of the different scales are highly correlated.

- *Final place of residence not assessed or reported.*

This parameter was not included as an endpoint in the ECASS III study.

Whilst the utility of this parameter is clear, it is also not without limitations. It is considered to be a fairly crude measurement, as there will be a whole spectrum of care e.g. in patients who returned to their own home some will have required no help at all whilst others will have required full time care but chose to remain in their own home. In addition, there may be many reasons why any individual chooses to move into a residential home following a stroke, which may be unrelated to the outcome of their stroke/functional ability. For example an individual may feel less confident living alone after a stroke, despite an excellent recovery, and prefer to move into sheltered living accommodation; alternatively injuries, e.g. from falling due to the stroke, could lead to patients not being able to return to their own homes, despite good recovery from the stroke itself. Bond *et al* (2000) found that factors including marital status, living arrangements and mental health status at admission were all significantly related to place of discharge for stroke and hip fracture patients. Schlegel *et al* (2004) found that increasing age was a predictor of nursing home use after stroke, and the authors considered this to be likely related to psychosocial and clinical factors.

The appropriateness of endpoints used in the clinical trials is discussed in section 5.

- *30 day CT head data not assessed or reported.*

This parameter was not included as an endpoint in the ECASS III study. CT or MRI scans were performed before treatment and 22-36 hours after treatment and then at the discretion of the investigators.

CT scans reliably exclude ICH and are therefore used prior to treatment with rt-PA. CT scans are widely available at all times of day, they are quick to perform and are suitable for patients with metallic implants and those with claustrophobia. As most ICHs related to thrombolysis treatment occur in the first 36 hours, there would be no good reason to repeat the CT scan at 30 days post stroke if this is not included as an outcome measure. At 1 month plus, infarcts will be undergoing atrophy/cystic cavitation and so measures of ischaemic lesion volume would not be helpful.

- *Outcome assessment potentially not blinded.*

The reason for this comment is not clear. The ECASS III publication (Hacke *et al*, 2008) specifies that patients were assessed by an examiner who was unaware of treatment assignment, at the time of enrolment, at 1, 2 and 24 hours after administration of the drug and on days 7, 30, 90 post treatment.

Members of the safety outcome adjudication committee who were unaware of the treatment assignments reviewed all CT or MRI scans.

It is possible that visible bleeding reactions may have helped physicians to guess which treatment a patient has been assigned to simply because more bleeds occurred in the rt-PA group. However some placebo patients did also have visible bleeding which would introduce an element of uncertainty.

- *Potential assessment bias was either not examined or not presented.*

It is not clear to what this limitation is referring.

As was discussed in the original assessment report at the time of licensing, the safety outcome adjudication committee for ECASS III, blinded to treatment allocation,

reviewed all CT and MRI scans. For stroke-related and neurological deaths, the committee adjudicated whether each death case was likely to be due to intracranial haemorrhage or to other brain pathology, or neither. Cases of intracranial haemorrhage were centrally adjudicated by the committee, including classification of a symptomatic intracranial haemorrhage according to the study protocol definition. The committee also adjudicated cases of symptomatic oedema (defined as brain oedema with mass effect as the predominant cause of clinical deterioration).

Review of all scans/cases of ICH/oedema/death by a blinded committee should have the advantage of reducing inter-investigator variability.

- *Aged 18 to 80 years.*

The inclusion/exclusion criteria mirrored the EU SmPC apart from the treatment time-window. This was because the trial was intended to provide further reassurance of the positive balance of benefits and risks of the EU marketing authorisation. Even though many stroke patients are aged over 80 years, this is a contraindication for use and so these patients should not receive rt-PA.
Although NICE does not say anything definitive about use of rt-PA in these populations the latest technology appraisal (TA264, Sept 2012) does state that *"patients outside the licensed indication for alteplase (under 18 years and over 80 years of age may have the potential to benefit from treatment with the technology. However, consistent with NICE methods, the Committee was aware that it can only make recommendations based on the current marketing authorisation for alteplase"*

3.5.1.2 Conduct of the ECASS III trial:

- *Patients highly selected – 1.5 recruited per centre per year.*

The small number of patients recruited per centre each year is not ideal, however this is likely to be mainly related to the narrow treatment time-window specified for the trial, of initially 3-4 hours, later widened to 4.5 hours. This time-window was selected because the NINDS trial was considered to provide evidence of a positive balance of benefits and risks from 0-3 hours post symptom onset and therefore enrolling patients in a placebo controlled trial in this time-window would be unfeasible/unethical. The other restrictions on enrolment (as per the EU SmPC) will also have reduced the number of eligible patients,

As with every clinical trial these restrictions are in place to ensure that the patients selected for the trial are those in whom the balance of benefits and risks is most likely to be maximal.

Although the average number of patients recruited per centre per year is low, these centres will presumably have been treating other stroke patients with rt-PA and who presented within 3 hours of symptom onset, and therefore the centres overall are unlikely to be treating such low numbers of patients.

- *Slow recruitment led to change of inclusion criteria mid trial.*

The time from onset of symptoms to treatment inclusion criterion was changed to allow recruitment of patients within a time-window of 3-4.5 hours post-symptom onset (initially the time-window was set at 3-4 hours). This change was permitted based on the publication of a pooled analysis of the ATLANTIS A and B, ECASS I and II, and NINDS 1 and 2 studies (Hacke *et al*, 2004).

In general, where recruitment into a trial is very slow, a change in the inclusion criteria or increase in the number of recruiting centres is common. Provided the population defined by the changed inclusion criteria is still a clinically relevant

population and the changes are not based upon unblinded data from the ongoing trial such changes are not considered to raise a concern.

3.5.1.3 Results of the ECASS III trial:

- *Modest imbalance of stroke severity in favour of alteplase.*
- *Treatment arm had 7.7% previous stroke and control arm 14.1% (p<0.003)*
- *Disability (mR 0-2) outcome not statistically significant (p<0.05).*
- *Severe disability (mR 5) was increased by 2.9 per cent with alteplase*
- *The study population had very low 30 day mortality – 5 per cent.*
- *This study was unusual in not showing increased early mortality.*

Taking each concern separately:

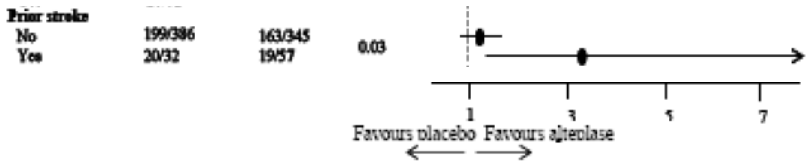
- *Modest imbalance of stroke severity in favour of alteplase.*

The mean baseline NIHSS score in the rt-PA group was 10.7 ±5.6, compared with 11.6 ±5.9 in the placebo group (p=0.03, unadjusted for multiple comparisons).

The unadjusted OR for mRS 0-1 was 1.34 95% CI [1.02-1.76], p=0.04; whilst the adjusted OR was 1.42 95% CI [1.02-1.98], p=0.04 (adjusted for baseline NIHSS and time to start of treatment).

- *Treatment arm had 7.7% previous stroke and control arm 14.1% (p<0.003).*

The subgroup analyses of patients with and without a prior history of stroke found a statistically significant subgroup by treatment interaction for prior stroke (p=0.03), the treatment effect being larger in patients with a history of prior stroke (although the results were also in the direction of favour of rt-PA in patients without a history of prior stroke):



The response rate (mRS 0-1) on rt-PA was 20/32 (62.5%) for those with a history of prior stroke compared with 19/57 (33.3%) on placebo. For those without a history of prior stroke the results were 199/386 (51.6%) on rt-PA compared with 163/345 (47.2%). There is no clear pattern of worse response rates for those with prior stroke, and the sub-group with prior stroke is very small.

Given these results, it is not considered that this imbalance in baseline history of stroke has biased the result of ECASS III in favour of rt-PA.

- *Disability (mR 0-2) outcome not statistically significant (p<0.05).*

The primary outcome measure was mRS 0-1, which demonstrated a statistically significant result in favour of rt-PA treatment in intention to treat analyses at day 90 (rt-PA: 52%, placebo: 45%; OR: 1.34, 95% CI 1.02-1.76; RR: 1.16, 95% CI 1.01-1.34; p=0.04). The odds ratio adjusted for NIHSS score at baseline and the time to start of treatment was 1.42, 95% CI 1.02-1.98, p=0.04. Similar results were observed at day 30.

The outcome mRS 0-2 was a further functional endpoint evaluated based on predefined cut-off points. Although not statistically significant at the 95% level there was a numerical imbalance in favour of rt-PA at day 90 (rt-PA 66.5%, placebo:

61.5%; OR 1.30 [0.95-1.78], p=0.11). The per protocol mRS 0-2 outcome at day 90 just achieved statistical significance (OR 1.41 [1.01-1.96], p=0.04).

The appropriateness of endpoints used in the clinical trials is discussed in section 5.

- *Severe disability (mR 5) was increased by 2.9 per cent with alteplase*

Although this finding was noted during the original regulatory assessment of the ECASS III trial for the extension of the time-window to 4.5 hours, it was similarly noted that the percentage of patients who died was slightly lower (1.5%) in the rt-PA group compared with placebo and the group with a favourable outcome (mRS 0-1) was 7.3% higher in the rt-PA group compared with placebo:

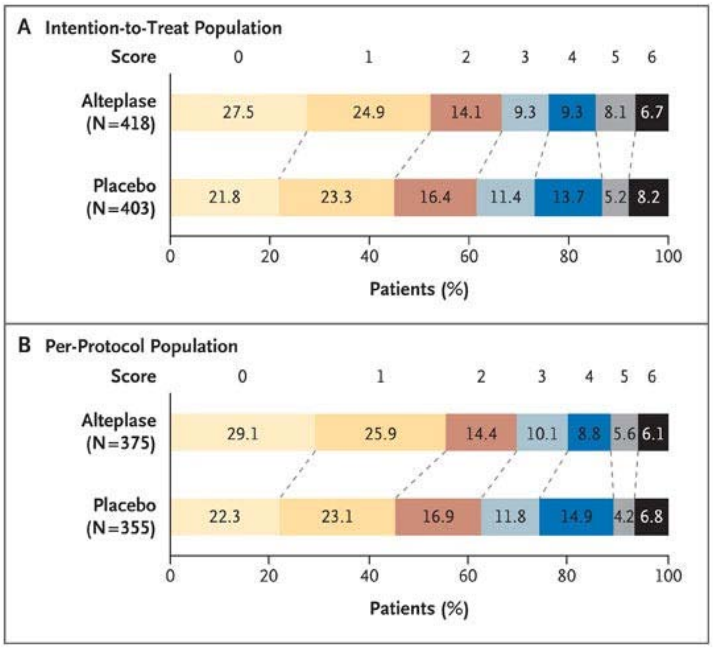
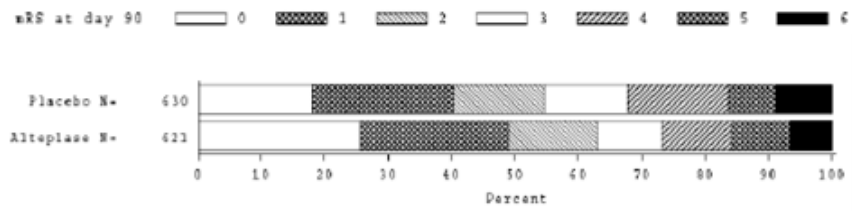


Figure: Distribution of scores on the mRS at day 90 [taken from Hacke et al, 2008]

In the overall evaluation the proportion of patients in category 5 is only one aspect to consider when looking at the efficacy profile for the treatment. The apparent increase in this category and consequent possible slight increase in patients in the 5-6 range would have to be considered in the context of improvements elsewhere on the scale.

In the original regulatory assessment, the MAH provided a pooled analysis of data from ECASS III, ECASS II, ATLANTIS A and B, and NINDS 1 and 2, for patients treated between 3-4.5 hours post symptom onset and treated according to the SmPC (n=1251):

Figure 2: mRS distribution at day 90 among patients treated in the 3-4.5 hour time window according to current SmPC (Source appendix 3, figure 2.1 and table 2.2)



Footnote to figure 2: Values for the for the mRS category 6 are slightly different to the ones presented in the Summary of Clinical Efficacy (SCE) (U09-0115-01) on page 22, since in the presentation of the SCE five death cases after day 90 were included in the calculation of the mRS 6 value. Differences between the two presentations are depicted in table 1 below.

Table 1: mRS 6 in the pooled data set, 3-4.5 h time window, patients treated according to current SmPC: Differences of the mRS 6 category in the documentation of the SCE compared to the presentation in fig. 2

	Summary of Clinical Efficacy (Fig. 2.7.3.2.3.1:1), page 22 Death (mRS 6) at day 90 and beyond	Figure 2 (see above) Death (mRS 6) at day 90
Alteplase	7.3 % (N=45)	6.6 % (N=41)
Placebo	9.1% (N=57)	8.9 % (N=56)

Note: placebo and rt-PA are shown the opposite way round to the previous figure.

The 'worst outcome (mRS 5-6)' in the pooled data occurred in a similar percentage for placebo and for rt-PA (16.4% and 15.8% respectively), with a net numerical shift from mRS 6 to mRS 5 in the rt-PA group.

The NIHSS score is a validated and widely used tool for measuring stroke outcome but is not an ideal measure in isolation, because the score is not directly and specifically associated with an individual's ability to functionally compensate for a neurological deficit, which is better assessed using scales that provide a measure of global disability such as the mRS. In ECASS III, NIHSS was measured at both baseline and at day 90, and therefore the NIHSS results for patients with a day 90 mRS score of 5 were presented. In both placebo and rt-PA groups, most of the patients with day 90 mRS=5 had unchanged or improved NIHSS scores compared with baseline. However, a greater percentage of rt-PA treated patients with day 90 mRS of 5 deteriorated in NIHSS score compared with their baseline than placebo treated patients. A worsening of >4 points in the NIHSS occurred in 14.7% of rt-PA compared with 9.5% of placebo treated patients with day 90 mRS=5.

In the original assessment within Europe, this finding was considered to be acceptable because the majority of patients in both treatment arms with mRS=5 at day 90 had improved with respect to NIHSS compared with baseline, and a higher percentage of patients in the placebo arm had a day 90 outcome of mRS=6 (death) and mRS=5 + 6 *[in the pooled analysis]*, and that the overall net effect of rt-PA was positive.

- *The study population had very low 30 day mortality – 5 per cent.*

The lower mortality rate observed at day 30 in ECASS III compared with some of the other randomised controlled trials (which have mortality rates in the range 4.2-17.9%) may be related to the baseline stroke severity of the study population. In ECASS III, the baseline median NIHSS score was 9 (range 1-24) for the rt-PA group, whilst in NINDS it was 14 (range 1-37). This may reflect the difference in time-window for treatment in the two studies (NINDS 0-3 hours, ECASS III 3-4.5 hours), as, for obvious reasons, less severe strokes have generally been shown to present later (Hacke *et al*, 2004). In keeping with this, mortality rates at day 30 were greater in the NINDS study than that observed in the ECASS III study, with 12.8% in the rt-PA group and 15.7% in the placebo group. Likewise, the mortality rates observed in ATLANTIS were 9.5% in the rt-PA arm, and 4.2% in the placebo arm, with a median

baseline NIHSS score of 10, and a treatment window of 3-5 hours (Part B data). Furthermore, in ECASS II, the median baseline NIHSS was 11, day 30 mortality was 8.4% and 8.1% in the rt-PA and placebo groups respectively, and enrolment was up to 6 hours after symptom onset.

Mortality was not the primary endpoint in this study.

- *This study was unusual in not showing increased early mortality.*

This may also be partly related to the population included in the trial. As mentioned above, the baseline stroke severity was lower in the ECASS III trial compared with the NINDS study and a lack of increased early mortality could be suggestive of a more favourable outcome in terms of treatment effect.

3.5.1.4 Legacy of the ECASS III trial:

- *Outcome not replicated by later IST-3 results.*
- *Repeatedly added to pooled and meta-analysed studies.*
- *Licensed drug in Europe to 4.5 h and widely used outside license.*
- *Did not change the 3 hour limit to license with FDA in US.*

See below for assessment of the concerns:

- *Outcome not replicated by later IST-3 results.*

The IST-3 study was specifically designed to include patients who were considered to be outside of the EU licence for rt-PA. In particular, the time-window for treatment used in IST-3 was up to 6 hours post-symptom onset, and the enrolled population included many patients aged >80 years (>50%) (IST-3 collaborative group, 2012). Many subjects had other characteristics that are contraindications in the EU SmPC, for example, severe stroke NIHSS >20 or high blood pressure (systolic ≥165 mm Hg, diastolic ≥90 mm Hg). If a patient had either a clear indication for rt-PA treatment or characteristics that would render the benefit-risk of treatment clearly negative, for example patients with ICH, then the patient was not entered into the trial.

In contrast, ECASS III was conducted following the current terms of the licence, with the exception of the time from symptom onset. Therefore it would not necessarily be expected that the two trials would produce the same results.

As a result of the differences in the baseline populations and time to treatment windows, it is perhaps not surprising that in the IST-3 results differed from the ECASS III trial results.

- *Repeatedly added to pooled and meta-analysed studies.*

This comment is based on the conclusion that the ECASS III trial data are unreliable and therefore should not be considered within pooled or meta-analyses. Alternatively it may be considered that the limitations raised regarding the ECASS III trial are not sufficient to render the results unreliable and that it should be included in pooled analyses.

- *Licensed drug in Europe to 4.5 h and widely used outside licence.*
- *Did not change the 3 hour limit to licence with FDA in US.*

The ECASS III trial was the main study used to support the expansion of the time-window for treatment of acute ischaemic stroke with rt-PA to 4.5 hours post-symptom onset. The time-window for treatment in the US has remained as 0-3 hours, despite a 4.5 hour window being supported by the Scientific Advisory from the American Heart Association Stroke Council. The FDA and the European regulators, as separate agencies, on occasions reach differing conclusions. As an example the

contraindications to treatment with rt-PA in stroke are less extensive in the US than in the EU.

Although Dr [redacted] s submission provides no evidence to support the assertion that rt-PA is widely used outside of the licence, off-label use in the EU is discussed in papers to be considered in January.

3.6 IST-3 trial

Dr [redacted] has a number of concerns regarding the IST-3 trial, many of which relate to the conduct of the trial, the analyses and the presentation of the results. The first phase of the IST-3 trial was a randomised, double-blind design which Dr [redacted] considers was to the highest standard with balanced randomisation and independent telephone assessment. At the end of the initial phase (n=276), the MAH did not continue to support the trial and withdrew supply of the drug and placebo and as a result the study became an open-label randomised study.

3.6.1 Conduct of IST-3

Dr [redacted] has highlighted the following aspects of the IST-3 trial regarding the conduct of the study:

- *Evolved from a double blind to open trial*
- *Refused supply of drug and matching placebo from industry*
- *Recruitment target of 6000 halved during 12 year trial*
- *Steered towards recruitment outside the license*
- *International advisory board eventually lost*
- [redacted]
- [redacted]
- *Subtle change of plan for Cochrane review and meta-analysis*
- *Publication on ordinal analysis by DMC* member*
- *Accurate predictions of older/younger outcome balance*
- *Sponsors in receipt of recurrent funds from drug company*

** DMC Data Monitoring Committee*

The results of the IST-3 trial and a number of specific concerns raised by Dr [redacted] were discussed in the May CHM paper, in section 4.1.

Looking at each of these concerns:

- *Evolved from a double blind to open trial*
- *Refused supply of drug and matching placebo from industry*

Dr [redacted] has elaborated on these points, stating:

"A version of the IST-3 protocol, circulated before commencement, indicated that alteplase would be provided free for the start-up phase, and that industrial collaboration was 'currently being negotiated'. (IST-3 1999) Those marketing alteplase had observer status on the International Advisory Board. The start-up blinded phase had disappointing results in 276 patients with 2% fewer becoming alive and independent. Boehringer Ingelheim discontinued support for IST-3 with no very convincing explanation."

These concerns were discussed in the May CHM paper, section 4.1.5, as follows:

"The fact that Boehringer Ingelheim had observer status on the International Advisory Board is not in itself considered to be an issue, provided that they did not have influence.

[redacted]

It is not possible to ascertain the reasons of the MAH for their decision not to support the IST-3 study from these additional details from the [redacted] The fact that the MAH made no reference to the data collected in the double-blind phase does not confirm that these data were unavailable to them. However, it is perhaps notable that the MAH's decision not to support the trial coincided with the award of the grant from UK Health Foundation.

The important issue that needs to be considered in terms of any assessment of the balance of benefits and risks of rt-PA is what impact this action had on the trial results. In the main, withdrawal of company funding and supply of product led to the change from a double-blind trial to the less rigorous open-label design. However conversely, given the widely acknowledged bias of clinical trials towards positive results when they are conducted/funded by the MAH for the drug in question, the withdrawal of funding by the MAH in this case could be considered to improve the reliability and impartiality of the study's findings.

Whilst the MAH's withdrawal has clearly had an impact on the trial design, it is not clear how the motivations for such an action could impact substantially on the interpretation of the results of IST-3."

- *Recruitment target of 6000 halved during 12 year trial*

The target sample size was reduced in 2007 due to feasibility issues. Further details of this decision were provided by [redacted] states that changes to the sample size were implemented in version 1.93 of the protocol, and the Medical Research Council awarded an extension to funding to permit recruitment to continue to the revised target of 3100 subjects to provide 80% power to detect an absolute difference of 4.7% in the primary outcome.

- *Steered towards recruitment outside the license*

95% of patients enrolled did not meet the terms of the EU licence. The IST-3 trial was designed to enrol patients that were not clearly indicated to receive treatment with rt-PA. Once rt-PA was licensed for the treatment of acute ischaemic stroke, it would have been unfeasible/unethical to enrol patients into a placebo-controlled trial of treatment, as is always the case when performing placebo-controlled RCTs on licensed drugs. Whereas, for most other drugs an alternative treatment can provide a suitable comparator, this is not the case for rt-PA.

- *International advisory board eventually lost*

Further discussion of this point is not provided, and no references have been provided to verify this point. It is not clear what, if any, influence this would have on the results.

[redacted]

- *Subtle change of plan for Cochrane review and meta-analysis*

This point has not been elaborated on.

- *Publication on ordinal analysis by DMC* member*

No details regarding this publication have been provided in Dr [redacted] s submission. It is assumed that this may refer to the paper by Bath *et al* (2007), as this paper is referenced in the article by Sandercock *et al* (2011) which discusses the possible use of an ordinal statistical method for the IST-3 trial, however concludes that the primary

endpoint would remain unchanged (dichotomised OHS endpoint) and an ordinal analysis would be added as a secondary endpoint.

- *Accurate predictions of older/younger outcome balance*

This point has not been elaborated on.

- *Sponsors in receipt of recurrent funds from drug company*

Dr [redacted] elaborates on this point as follows:

"The trialists did declare relevant conflicts of interest, with the University of Edinburgh department running IST-3 receiving funds from Boehringer Ingelheim. (IST-3 2012)"

3.6.2 Analysis of IST-3

Dr [redacted] has highlighted the following aspects of the IST-3 trial regarding the analysis of the study:

- *Considered replacing primary outcome with ordinal analysis*
- *Unexplained pooling of variables to inflate ordinal outcome*
- *Analysis altered from plan - obscuring harm within 3-4.5 h*

Considering each point in turn:

- *Considered replacing primary outcome with ordinal analysis*

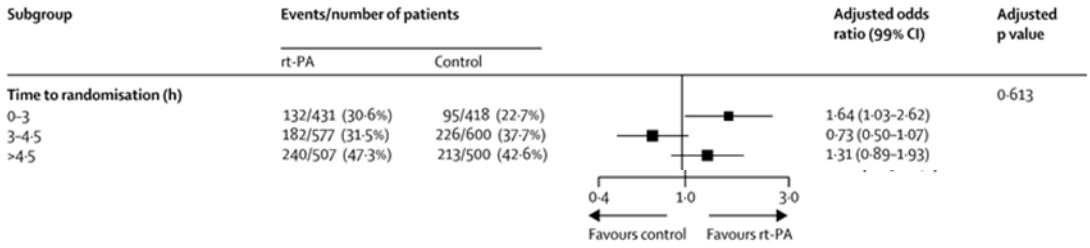
The paper by Sandercock *et al* (2011) which provides the details of the baseline characteristics of the patients in the IST-3 trial as well as an update on the progress of the trial, also discusses the development of the statistical analysis plan. It comments that since the trial started, a number of new approaches to the analysis of ordered functional outcome data have been proposed to provide greater statistical power. Ordinal logistic regression and 'shift analysis' are quoted as a possible primary endpoint, provided that key underlying assumptions are met. However it was determined that the underlying assumptions may not be met and therefore the ordinal analysis was not appropriate for the primary analysis. Given that the primary analysis for the trial was unaffected by these discussions it is not considered that this represents an issue. In any event the primary endpoint was reported as negative, though the abstract in the publication of the trial could be criticised for overemphasising the highly positive ordinal shift analysis (p=0.001).

- *Unexplained pooling of variables to inflate ordinal outcome*

The pre-planned ordinal shift analysis pooled categories 4+5+death of the OHS to create a 5 category scale 0, 1, 2, 3, 4+5+death. This gave an odds ratio of 1.27 p=0.001. An analysis using all 7 categories gave an odds ratio of 1.17, p=0.016. Therefore the pooling did improve the result, but it is difficult to criticise as it was the pre-specified shift analysis, and no one choice is clearly better than any other.

- *Analysis altered from plan - obscuring harm within 3-4.5 h*

The subgroup results for patients treated between 3-4.5 hours, adjusted for age and baseline NIHSS are given in Figure 3 of the IST-3 publication, as follows:



The results as reported clearly show the apparent negative effect in the 3-4.5 hour group.

3.6.3 Presentation of published results for IST-3

Dr [redacted] has highlighted the following aspects of the IST-3 trial regarding the presentation of the study:

- *Contemporaneous posting of Cochrane resembling meta-analysis*
- *Opening sentence undermined main reason for the 12 year trial*
- *Primary outcome not mentioned in 'interpretation'*
- *Comment on trend with outcome and delay to treatment unclear*
- *Kaplan-Meier curve as promised 'Figure 5' omitted in print*
- *Kaplan-Meier curve in webappendix not cited in text*
- *Table of trial final place of residence not published by trialists*
- *Table examining recall bias in 2013 subgroup webappendix*
- *Inseparable long term mortality ignored in 2013 paper*

Looking at each of these concerns:

- *Contemporaneous posting of Cochrane resembling meta-analysis*

This point has not been elaborated on.

- *Opening sentence undermined main reason for the 12 year trial*

Dr [redacted] has commented as follows:

"The first sentence in the final paper strangely declared that the debate was over – 'Thrombolysis is of net benefit in patients with acute ischaemic stroke, who are younger than 80 years of age and are treated within 4.5 h of onset.' (IST-3 2012)"

A similar issue was discussed in the May CHM paper, regarding the expectations for the study, which is discussed by Dr [redacted] in the October submission, and the following assessor's comments are relevant:

"Given the long time period over which the trial took place: 2000 to 2011, it is perhaps not surprising that the suggested importance of the results of IST-3 to the debate over the effectiveness of rt-PA in stroke reduced over time and changed its emphasis. Over the period of the study, rt-PA was approved in the EU for the treatment of acute ischaemic stroke up to 3 hours after onset of symptoms and towards the end of the study the publication of the ECASS III trial resulted in the increase of the time-window for treatment to 4.5 hours following symptom onset. Regulatory approval of use of rt-PA would be a reasonable indication that some level of consensus regarding the balance of benefits and risks had been reached. The individual contribution of one study to a debate on the benefits and risks of a treatment will decrease as increasing amounts of data emerge from other sources. In addition to the lengthy duration of the study, the issues with patient enrolment that ultimately led to a reduction in the final sample size (the final study population was approximately half the intended size) inevitably reduce the contribution that this study can make to the debate."

The first sentence of the IST-3 trial publication reflects the current licence for rt-PA in acute ischaemic stroke, and therefore arguably the benefit-risk balance could be considered to be confirmed in this population of patients (those <80 years of age, treated up to 4.5 hours post symptom onset).

- *Primary outcome not mentioned in 'interpretation'*

This is a criticism of the presentation of the trial results, rather than the conduct of the trial itself. As commented in the May CHM paper, the 'interpretation' section in the abstract of the IST-3 trial publication does arguably give an overly positive slant to the results, however this has no impact on the actual results obtained.

- *Comment on trend with outcome and delay to treatment unclear*

This point has not been elaborated on.

- *Kaplan-Meier curve as promised 'Figure 5' omitted in print*
- *Kaplan-Meier curve in webappendix not cited in text*

These comments are also criticisms of the presentation of the trial results rather than the conduct of the trial itself, or the results obtained.

- *Table of trial final place of residence not published by trialists*

A similar point was raised during the May CHM assessment. Although the supplementary appendix for the IST-3 trial provided data on the place of residence at 6 months and 18 months for patients who were included in the long-term 18 month follow-up, the data on place of residence at 6 months for the entire trial cohort was not included (IST-3 collaborative group, 2013). However, as commented in the May CHM paper, the data as provided by the trialists and included in Dr [redacted] s submission are not considered to affect the conclusions drawn on the published data.

- *Table examining recall bias in 2013 subgroup webappendix*

The potential issue of recall bias was discussed in the May CHM paper, section 4.1.5, and is also discussed below, in section 3.7.5.

- *Inseparable long term mortality ignored in 2013 paper*

Dr [redacted] has elaborated on this, stating:

"Finally, when the inseparable 6 to 18 month survival curves are viewed, the lack of benefit after alteplase, a benefit expected if treated survivors were less disabled, received no comment. (Slot 2008, IST-3 2013)"
The assumption that a reduction in disability in the rt-PA treated arm should lead to a reduction in long-term mortality is reasonable, but may depend on duration of follow-up. This subject is discussed in more detail later (section 3.7.3)

3.6.4 Legacy of IST-3

Dr [redacted] has highlighted the following aspects of the IST-3 trial regarding the legacy of the study:

- *Patients remain uncertain of best treatment option*
- *Clinicians more sceptical but fear openly expressing views*
- *Problems for MHRA, NICE, Cochrane and Royal Colleges*
- *Devalued status of clinical trials and evidence based medicine*

The points highlighted regarding the legacy of the trial relate to a single viewpoint on the evidence and therefore it is not possible to comment on these.

3.7 Specific weaknesses in the evidence of effectiveness

3.7.1 The time window

Dr [redacted] is concerned that there is little good evidence to support the 'time is brain' message – that is, there is little good evidence for loss of efficacy as the time to treatment increases. Dr [redacted] comments that there is good evidence for this effect in coronary thrombolysis, and lack of a similar pattern in stroke would undermine its credibility as an agent in cerebrovascular disease.

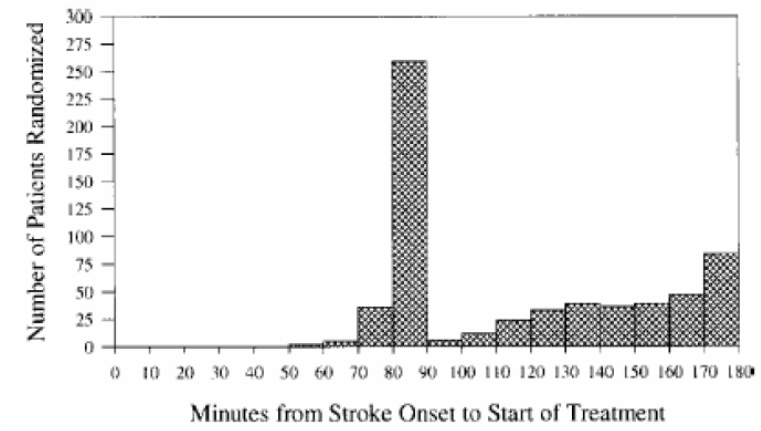
The evidence cited to support this conclusion is as follows:

- *The NINDS re-analysis concluded that the data failed to support a conclusion that the effect of rt-PA diminished with increasing onset to treatment time within the protocol-specified 3 hour time limit.*
- *Time to treatment was not identified as a risk factor for symptomatic intracranial haemorrhage in the review.*
- *A separate re-analysis of NINDS (Hoffman and Schriger, 2009) looked for time trends and concluded that their graphs fail to support the 'time is brain' hypothesis.*
- *IST-3 demonstrated no linear trend, with the authors stating that 'there was insufficient power to examine decay of benefit with time'*
- *A series of other studies show no relation: ATLANTIS, the OR for 3 month mortality was higher in the <3 hour group compared with the 3-6 hour group: OR=3.8, 95% CI [0.64-22.6]; and OR=2.3, 95% CI [1.2-4.4] respectively. ECASS II, the treatment differences (mRS 0-1) were similar whether patients were treated within 3 hours or 3-6 hours, and 'The mortality rate was higher in the rt-PA group than in the placebo group among patients randomised within 3 hours of stroke onset: no such difference was observed in patients treated 3-6 hours after stroke onset.'*
- *In a pooled analysis (Lees, 2010), an excess of 4.2% of large intracerebral haemorrhages with rt-PA was found overall but the absolute rates of haemorrhage were similar across the onset to treatment intervals.*
- *The pooled analysis claimed a treatment effect of onset to treatment time, but failed to highlight the different stroke populations (30 day mortalities ranging from an estimated 5% to 15%) in the individual studies.*
- *The SITS-ISTR mRS curves for similar stroke populations treated within 3 hours and within 3-4.5 hours appear almost identical. An explanation the authors fail to put forward is the minimal benefit of the drug and the absence of any time relation to treatment.*

Generally, a vast body of literature suggests that in theory a shorter onset to treatment time in acute ischaemic stroke patients would be expected to be associated with a more favourable outcome. This is because whilst a core of infarcted tissue develops within a few minutes, surrounding the core or mixed with it are areas of penumbra – hypoperfused but still viable neurones. This penumbral area does not remain viable indefinitely and is affected by factors such as temperature, glucose levels etc. Therefore the sooner the blood supply can be restored to these areas, the better the expected outcome.

- *The NINDS re-analysis concluded that the data failed to support a conclusion that the effect of rt-PA diminished with increasing values of onset to treatment time within the protocol- specified 3 hour time limit.*

The distribution of the onset to treatment times in the NINDS study has been discussed previously in this paper, and in the May CHM paper. The following figure shows the unusual onset to treatment time distribution observed in NINDS, where 50% of patients in the 0-90 minute stratum were treated at 89 or 90 minutes:



The re-analysis committee did conclude that the data fail to support a conclusion that the effect of treatment diminished with increasing onset to treatment time, however importantly, they also concluded *"However, this does not mean that such a relationship does not exist, and further studies are needed to address the question of a differential rt-PA treatment effect related to time from symptom onset to treatment."*

In essence, the data available from the NINDS trial are not sufficient to address this question, the pattern of onset to treatment time means that is it not appropriate to use the time to treatment as a continuous variable.

- *A separate re-analysis of NINDS (Hoffman, 2009) looked for time trends and concluded that their graphs fail to support the 'time is brain' hypothesis.*

Given the above, it is not surprising that the graphical re-analysis (Hoffman and Schriger, 2009) similarly found no support for the 'time is brain' hypothesis, as the re-analysis is based upon the same original data with the limitations as described above. A fuller discussion of the Hoffman graphical re-analysis is provided in section 7, in relation to Professors Fatovich's and Brown's submission.

- *Time to treatment was not identified as a risk factor for symptomatic intracranial haemorrhage in the review*.
- *In a pooled analysis (Lees, 2010), an excess of 4.2% of large intracerebral haemorrhages with rt-PA was found overall but the absolute rates of haemorrhage were similar across the onset to treatment intervals.*

A relationship between the time from onset to treatment time and ICH would not be expected to be observed in the data from NINDS, even if one did exist, because of the unusual pattern of onset to treatment times described above. However, in any case, whilst it might be expected that a shorter onset to treatment time would be related to a more favourable outcome in terms of efficacy, there is less reason to suppose that risk of ICH associated with rt-PA would be related to time to treatment.

The pooled analysis by Lees *et al* (2010) found similar absolute rates of intracerebral haemorrhages across all onset to treatment times, which suggests that any relationship between onset to treatment time and risk of intracerebral haemorrhage is uncertain. However, if anything, the rate of intracerebral haemorrhage appeared to decrease slightly with increasing onset to treatment time in the placebo treated group, whilst it remained fairly constant in the rt-PA treated group (i.e. the risk attributable to rt-PA may increase with time).

- *IST-3 demonstrated no linear trend, with the authors stating that 'there was insufficient power to examine decay of benefit with time'*

The IST-3 trial did not find the expected pattern of benefit over onset to treatment time, with the time-windows of 0-3 hours and >4.5 hours demonstrating the most benefit and 3-4.5 hours the least:

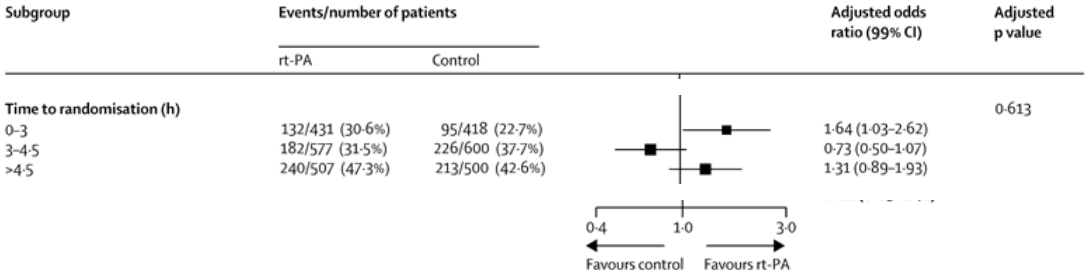


Figure: reproduced from IST-3 collaborative group, Lancet 2012

The IST-3 trial initially intended to enrol 6000 patients, but this was reduced during the trial due to slow recruitment, and in the end 3035 patients were enrolled. Therefore the final size of the subgroups in the subgroup analyses was smaller than originally anticipated, which reduced their power to determine the relationship between time to treatment onset and benefit – as evidence by the implausible pattern of benefit over time.

- *ATLANTIS, the OR for 3 month mortality was higher in the <3 hour group compared with the 3-6 hour group: OR=3.8, 95% CI [0.64-22.6]; and OR=2.3, 95% CI [1.2-4.4] respectively.*

The odds ratios quoted for the ATLANTIS study for mortality do not support higher mortality in the <3 hour group than in the 3-6 hour group. The confidence intervals quoted for the <3 hour group completely overlap the confidence intervals for the 3-6 hours, and are extremely wide, including up to a 36% improvement in mortality for the rt-PA group. This reflects the small size of the ATLANTIS studies, and the number of patients enrolled at <3 hours: n=22 in part A, and n= 39 in part B, 8 of whom were protocol violations after the time-window had been changed to 3-5 hours.

- *ECASS II, the treatment differences (mRS 0-1) were similar whether patients were treated within 3 hours or 3-6 hours, and 'The mortality rate was higher in the rt-PA group than in the placebo group among patients randomised within 3 hours of stroke onset: no such difference was observed in patients treated 3-6 hours after stroke onset.'*

In the ECASS II trial (Hacke *et al*, 1998[1]), only 158 patients were enrolled within the 0-3 hour time-window, and as commented by the authors, the apparent difference in mortality between the two time-windows may be due to the small number of patients in the 0-3 hour subgroup. Likewise the similarity of the treatment differences for the efficacy outcome may be due to the small number of patients in the 0-3 hour subgroup.

- *The pooled analysis claimed a treatment effect of onset to treatment time, but failed to highlight the different stroke populations (30 day mortalities ranging from an estimated 5% to 15%) in the individual studies.*

The pooled analysis by Lees demonstrated a treatment effect in terms of efficacy (mRS 0-1 at day 90), with OR and 95% CI as follows:

0-90 minutes (0 - 1.5 hours): 2.55 [1.44-4.52]

91-180 minutes (1.5 - 3 hours): 1.64 [1.12-2.40]

181-270 minutes (3 - 4.5 hours): 1.34 [1.06-1.68]

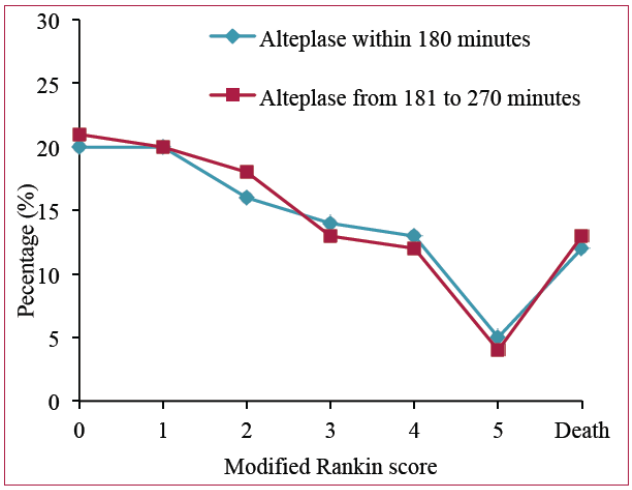
271-360 minutes (4.5 - 6 hours): 1.22 [0.92-1.61]

The 3-4.5 hour treatment window showed a significantly favourable outcome, based on mRS 0-1 at 90 days, for rt-PA relative to placebo. The improvement in outcome in the rt-PA group versus placebo with treatment beyond 4.5 hours was not statistically significant at the 95% level, though the trend was favourable.

No account was taken of the different 30 day mortalities observed between the different individual studies. However, provided that each trial included patients with similar characteristics in each arm the differences between the two arms within the trial are valid. Some of the lowest 30 day mortality rates were observed in the ECASS III study, ~5%, in which patients were treated between 3-4.5 hours. It is likely that patients with more severe strokes will present more quickly and therefore be treated earlier, and this is supported by the data from the ECASS III (with the longer time to treatment window) and NINDS studies (with more severe stroke, 0-3 hours treatment time-window, than in ECASS III). The lesser stroke severity may partly explain the lower mortality rates in ECASS III,

- *The SITS-ISTR mRS curves for similar stroke populations treated within 3 hours and within 3-4.5 hours appear almost identical. An explanation the authors fail to put forward is the minimal benefit of the drug and the absence of any time relation to treatment.*

Regarding the SITS-ISTR mRS curves:



Appendix Figure 2: **Modified Rankin scores 3 months following cerebral infarction after treatment with alteplase within 180 min (n=10,231) and within 181 to 270 min (n=541)**
Data from SITS-ISTR, reference 8.

Figure: [redacted]

In this observational study cohort the authors noted that the patients treated from 3-4.5 hours were slightly younger, and had slightly less severe neurological deficit at baseline than those treated within 3 hours. This would support a possible association between time to treatment and stroke severity. The authors also noted that the patients treated from 3-4.5 hours were less often hypertensive and less frequently had a history of hyperlipidaemia.

3.7.1.1 *rt-PA in 3 to 4.5 hour and 3 to 5 hour window*

Dr [redacted] s main concerns relating to the 3-4.5 hour and 3-5 hour time windows for treatment are summarised as follows:

- *The extension to the time window for treatment to 4.5 hours was approved in Europe shortly before the publication of IST-3. IST-3 found a significantly worse outcome of mRS 0-2 in the 3-4.5 hour time window (unadjusted OR 0.76, 95% CI [0.60-0.97]).*
- *The statistical analysis plan for IST-3 is clear that 95% CI will be presented in the publication, however 99% CI were used which provide a more reassuring visual image (Figure 3 of the IST-3 publication).*
- *The 3-4.5 hour time window has not been studied in the 2009 or 2014 Cochrane reviews or the 2012 Lancet meta-analysis despite many trials having relevant data.*
- *The 3-4.5 hour time window was examined in the pooled analyses (Hacke et al 2004, Lees et al 2010, Emberson et al 2014) but only in relation to symptoms (mRS 0-1) and 90 day mortality, not disability (mRS 0-2) or 7, 10, and 30 day mortality.*
- *rt-PA has only been more widely used since the time window was extended to 4.5 hours and this was critical for drug sales. It is no surprise that the extension was promoted by [redacted] regulators. It would appear there are questions to answer on how the extension occurred in Europe given the expected IST-3 publication and the paucity of sound evidence of effectiveness. Several bodies whose role should be impartial appear to have lacked balance in their role, with potentially serious consequences.*

[redacted]

Dr [redacted] also comments that Professor Joanna Wardlaw was the lead for the Cochrane reviews, and was also a leading member of the IST-3 team.

- *The extension to the time window for treatment to 4.5 hours was approved in Europe shortly before the publication of IST-3. IST-3 found a significantly worse outcome of mRS 0-2 in the 3-4.5 hour time window (unadjusted OR 0.76, 95% CI [0.60-0.97].*

The variation for the extension to the time window considered the relevant information available at the time, and concluded that sufficient evidence was available to demonstrate a positive balance of benefits and risks up to 4.5 hours post onset of stroke. The IST-3 trial was being conducted in a population that was outside of the licence and therefore its relevance to the decision regarding the variation to extend the time window would be questionable.

- *The statistical analysis plan for IST-3 is clear that 95% CI will be presented in the publication, however 99% CI were used which provide a more reassuring visual image (Figure 3 of the IST-3 publication).*

This is another valid criticism of the publication – the wider 99% confidence intervals are more likely to overlap with each-other and 1.00 creating an impression of greater potential similarity between different time-windows.

- *The 3-4.5 hour time window has not been studied in the 2009 or 2014 Cochrane reviews or the 2012 Lancet meta-analysis despite many trials having relevant data.*

- *The 3-4.5 hour time window was examined in the pooled analyses (Hacke et al 2004, Lees et al 2010, Emberson et al 2014) but only in relation to symptoms (mRS 0-1) and 90 day mortality, not disability (mRS 0-2) or 7, 10, and 30 day mortality*

It is difficult to comment on the rationale behind independent analyses of the data. However, a greater number of trials have used mRS 0-1 as their primary endpoint than mRS 0-2 which may have influenced the choice of endpoint in the meta-analysis. The use of day 90 mortality as opposed to day 7/10 or day 30 mortality has been discussed elsewhere.

- *rt-PA has only been more widely used since the time window was extended to 4.5 hours and this was critical for drug sales. It is no surprise that the extension was promoted by [redacted] regulators. It would appear there are questions to answer on how the extension occurred in Europe given the expected IST-3 publication and the paucity of sound evidence of effectiveness. Several bodies whose role should be impartial appear to have lacked balance in their role, with potentially serious consequences.*

This criticism appears to be directed at European regulators. [redacted] The allegation that several bodies lacked impartiality has not been supported with any evidence.

It is not possible to comment on the allegations, or implied criticisms of individuals involved in the trials or meta-analyses.

3.7.1.2 *The 3 hour window*

In this section, Dr [redacted] has provided data on the individual study results for mRS 0-1, and comments that the majority of data in the 0-3 hour time window for the licensed population still mainly comes from NINDS. Dr [redacted] s concerns regarding the NINDS trial have been discussed above. Dr [redacted] considers that the 16% benefit observed at 90 days in NINDS for mRS 0-1 would be reduced by any of the following:

- *using final 12 month data,*
- *adjusting for baseline severity,*
- *using mRS 0-2 as the key endpoint in place of mRS 0-1,*
- *adjusting for recall bias estimated in IST-3 blinded group (7.7%)*

Dr [redacted] considers that the mortality data at the end of study seem reassuring at first, with the two largest datasets (NINDS and IST-3) suggesting improvement, however he considers that a funnel plot displaying these results is consistent with the concern that NINDS was unbalanced and IST-3 was subject to treatment bias because treated patients were known and could be given enhanced care. Mortality after the initial 7 days is better in IST-3 than in the other trials.

Dr [redacted] has also presented a table of harms, in terms of haemorrhage and allergy, for NINDS, 3 industry trials combined (not stated which trials), and IST-3.

- *Impact on efficacy endpoint in 0-3 hour time window, mRS 0-1*

Using 12 month data - Kwiatkowski *et al* (1999) examined outcomes from the NINDS trial at 12 months from treatment, and found that the benefit of rt-PA was sustained:

TABLE 1. OUTCOMES SIX MONTHS AND ONE YEAR AFTER THE ONSET OF STROKE.

TIME POINT AND ASSESSMENT INSTRUMENT	PERCENTAGE OF PATIENTS WITH FAVORABLE OUTCOMES*		ODDS RATIO (95% CI)†	RELATIVE RISK (95% CI)†	P VALUE
	t-PA (N=312)	PLACEBO (N=312)			
6 Months after stroke					
Global test‡	—	—	1.7 (1.3–2.3)	—	<0.001
Barthel index	50	37	1.7 (1.2–2.4)	1.4 (1.1–1.6)	0.001
Modified Rankin scale	41	29	1.8 (1.3–2.5)	1.4 (1.2–1.8)	0.001
Glasgow Outcome Scale	43	31	1.6 (1.2–2.3)	1.4 (1.1–1.7)	0.004
12 Months after stroke					
Global test‡	—	—	1.7 (1.2–2.3)	—	0.001
Barthel index	50	38	1.6 (1.1–2.1)	1.3 (1.1–1.5)	0.005
Modified Rankin scale	41	28	1.8 (1.3–2.5)	1.5 (1.2–1.8)	0.001
Glasgow Outcome Scale	43	32	1.6 (1.1–2.2)	1.3 (1.1–1.6)	0.006

*Scores of 95 or 100 on the Barthel index, 0 or 1 on the modified Rankin scale, and 1 on the Glasgow Outcome Scale were considered to indicate a favorable outcome.

†The Mantel–Haenszel test was used for univariate analyses, with groups stratified according to clinical center and the time to treatment (0 to 90 minutes and 91 to 180 minutes). For the global tests (which used logit-link function), the same stratifying variables were included as covariates. CI denotes confidence interval.

‡There is no published method by which to compute relative risk.

Table: taken from Kwiatkowski *et al,* 1999

Adjusting for baseline severity - The influence of baseline severity on the results of NINDS has been discussed above.

Using mRS 0-2 endpoint - A discussion of endpoints used in clinical trials of acute stroke treatments is provided in section 5 below.

Adjusting for recall bias - As NINDS was a double-blind trial, it would generally not be appropriate to adjust for any recall bias. Concerns have been raised regarding the possibility that treatment was not sufficiently blinded or could have been deduced from bleeding reactions. Even if these are legitimate issues, the former (whether the solution foamed) seems very unlikely to come to the notice of the patient.

- *Mortality in 0-3 hour treatment window*

The numbers of patients in the groups treated between 0-3 hours following onset of stroke symptoms is very small in all rt-PA trials apart from NINDS and IST-3.

- *Harms in patients treated in the 0-3 hour time window*

Dr [redacted] is still trying to obtain data on these parameters, and the table provided in the October submission contains several 'unknowns' with estimates generated from other trials. The estimate for fatal intracerebral haemorrhage at day 7 for the three industry trials combined and for IST-3 has been based on the fraction of fatal cases of all symptomatic intracerebral haemorrhage observed in the NINDS trial (the ratio of symptomatic plus fatal ICHs:fatal ICHs as observed in NINDS). The figure for symptomatic and fatal intracerebral haemorrhage for the three industry trials and IST-3 has been divided by this ratio (2.2) to generate the estimate. However, since the definition of 'symptomatic intracranial haemorrhage' varied in different trials this method of estimation is unlikely to be accurate.

3.7.2 Radiology

Dr [redacted] has a number of concerns relating to radiological findings following thrombolysis with alteplase:

- *Radiological studies of any size, demonstrating less cerebral damage post alteplase have not been published to my knowledge. (Zivin 2011) All radiological studies have demonstrated substantial harm from intracerebral haemorrhage.*
- *In the 755 patient ATLANTIS study, 30 day CT head scans were available. No reduction in infarct size could be demonstrated. In ATLANTIS B (3-5 h, n=613)) infarct size was identical. (ATLANTIS 1999, ATLANTIS 2000)*
- *The EPITHET study of 100 patients found no significant improvement in cerebral infarction judged by MRI scanning in patients given alteplase between 3 and 6 hours post stroke. (EPITHET 2008)*
- *A recent MRI study using diffusion-weighted imaging suggested alteplase only averted an ischaemic stroke in 2 of 231 patients with an initial acute stroke lesion on their scan. (Freeman JW 2013)*

The ATLANTIS A study showed no differences in CT lesion volumes for rt-PA versus placebo at day 30, with both groups showing large variations: placebo 64±74 cm^3 versus rt-PA 45±54 cm^3 (*p*=0.17). The median time to treatment for the rt-PA group was 4 h 36 min with 62% of patients treated after 4 h from onset. ATLANTIS B and EPITHET studies showed no reduction in infarct size for patients who were treated at 3-6 hours after stroke onset.

- CT Imaging

Two post-hoc analyses of the CT scan data from the NINDS study have been published.

Nichols *et al*. (2008) analysed the ischaemic volume data from patients with and without resolution of the hyperdense middle cerebral artery sign to assess the effects of arterial recanalization using the NINDS study results. The hyperdense artery sign was present in the baseline CT scan of 79 (of 604) eligible patients and it resolved in 14 of 37 (38%) patients treated with rt-PA compared with 7 of 43 (17%) treated with placebo (p=0.03). The baseline characteristics and demographics of patients with the hyperdense sign in the rt-PA and placebo groups were similar except those treated with rt-PA were older (rt-PA 69.6 years versus 63.7 years for placebo). For the 23 patients with persistence of the hyperdense artery sign at 24 hours following rt-PA, median (IQR) lesion volumes were: 107.4 (68-229) cm^3 in those treated with rt-PA versus 49 cm^3 (12-139) in the placebo group with persistence of the hyperdense artery sign; for the 14 patients with resolution of the hyperdense artery sign at 24-hours following rt-PA, the median (IQR) lesion volumes were: 16.1 cm^3 (6.6-53) versus 105.6 cm^3 (52-205) in the placebo group with spontaneous resolution of the hyperdense artery sign at 24 hours. Infarct volumes at 24 hours were significantly smaller in patients treated with rt-PA who had resolution of the hyperdense artery sign, compared with those who had persistence of the sign (p=0.004). However, the presence or size of ischaemic lesions on the pre-treatment CT scan was not reported and functional outcomes were not significantly improved based on resolution of the hyperdense artery sign. There were 4 (10.8%) symptomatic intracranial haemorrhages in the rt-PA-treated group compared with 2 (2.4%) in the placebo arm. The mortality rates were high but similar for both groups (24% for rt-PA versus 23% for placebo). It is difficult to reach any firm conclusions on the basis of this post-hoc analysis of a small number of patients.

Another post-hoc analysis of the NINDS study assessed ischaemic lesion volumes on CT at 24 hours, 7 to 10 days, and 3 months after stroke (NINDS rt-PA Stroke Study Group 2000). A reduction in median CT lesion volume was seen at 3 months in the rt-PA group (15 cm^3, interquartile range, IQR 2-87 cm^3) versus placebo (24 cm^3, IQR 4-101 cm^3; p=0.06) with a 11% reduction in cumulative lesion volume. Similar trends towards reduced post-rt-PA volumes were also reported at the other time-points, even after excluding patients that had died or who were lost to follow-up but none were statistically significant and the study was not powered to detect changes in infarct size.

- MR imaging

MRI sequences are more sensitive than CT scanning techniques at detecting acute ischaemia and cerebral perfusion defects. A brief account of the various imaging techniques is given in the individual paper describing stroke care in the UK (Paper 2) for the EWG meeting in November 2014.

In the context of cerebrovascular disease, the diffusion-weighted imaging (DWI) lesion is usually considered to be a surrogate marker of the irreversible ischaemic core (infarct) although a small part of it may be part of the ischaemic penumbra (hypoperfused but viable tissue). The mismatch DWI/PWI hypothesis states that the ischaemic core may be recognised as an area with reduced perfusion and diffusion and that the ischaemic penumbra may be identified as an area with reduced perfusion and normal diffusion. The EPITHET study aimed to test whether rt-PA given 3-6 hours after stroke onset promotes reperfusion and reduces infarct growth in patients with a mismatch in perfusion-weighted MRI and diffusion-weighted MRI. The primary outcomes related to infarct growth. Only 26 patients with the target mismatch profile were treated with rt-PA and reperfusion was achieved in 71% (n=15) of them. The primary outcome of lower infarct growth with rt-PA in mismatch patients was negative but other growth measures supported the hypothesis of reduced infarct growth with rt-PA beyond 3 hours.

However, permanent reversal of the DWI lesion is well established in animal models using temporary arterial occlusion methods. A systematic review of whether DWI represents the ischaemic core identified 18 studies that reported the prevalence of growth and reversal of DWI abnormalities (Kranz and Eastwood 2009). There was substantial variability in the observed rates of DWI lesion reversal (0-83%) with a mean rate of 24% in pooled patients. A combined analysis of the DEFUSE and EPITHET trial data concluded that the amount of DWI reversal was reported as small and unlikely to be clinically significant (Campbell *et al*. 2012).

The extent of DWI reversal in 176 patients thrombolysed with rt-PA within the licensed timeframe (0-4.5 h) from stroke onset was reported by Labeyrie *et al*. (2012). 89 patients had reversible acute DWI lesions (median percentage of reversal was small at 11% and 2.4 ml volume). The percentage reversibility of acute DWI lesions was associated with neurological improvement.

A more recent study quantified DWI and PWI changes at baseline, 2 and 24 hours after intravenous rt-PA was given at 0-3 h after stroke onset in 71 patients (Luby *et al*. 2014). Recanalisation rates were assessed using MR Angiography (MRA). Early reperfusion at 2 hours and a sustained decrease in DWI lesion volume at 24 hours were independent predictors of a favourable clinical outcome after thrombolysis. For those with a favourable outcome (mRS 0 or 1; n=30), the change in acute DWI median volume was -0.1 ml (IQR, -4.5-1.9) at 2 hours (absolute baseline DWI volume was 4.3 ml (IQR, 1.1-17); 0.5 ml (-1.7-9.1) at 24 hours with a reduction in follow-up FLAIR volume at 5-90 days. Those with an unfavourable outcome (mRS ≥ 2; n=41) had median DWI volumes of 18.1 ml (IQR, 3.2-63.1) at baseline; and changes in DWI lesion volume of 2.7 ml (IQR, -0.5-16.6) at 2 hours and 19 ml (1.1-73.1 ml) at 24

hours and an increase in follow-up FLAIR. The complete reperfusion rate was 40% at 24 hours which is consistent with literature values. In all study patients, immediate reperfusion at 2 hours and a subsequent decrease in DWI volume at 24 hours in patients post thrombolysis were predictive of a favourable clinical outcome. These findings suggest that successful treatment with rt-PA reduces expansion of the infarct core into the ischaemic penumbra by timely reperfusion. Similar findings have also been reported using baseline DWI and 90-day FLAIR volume assessments after rt-PA treatment within the 0-3 h time window (Merino *et al*. 2007).

Early reperfusion is associated with a reduction in infarct growth through salvage of ischemic penumbra, and this probably contributes more to the favourable neurological outcomes in these patients than the relatively small volumes of diffusion reversal.

Freeman *et al*. (2012) reported that complete DWI lesion reversal was only observed in 2 of their 231 patients given rt-PA at a median time of 159 minutes from stroke onset. Seventy-eight percent of their patients with persistent DWI lesions had acute perfusion lesions and an additional 4 patients had transient DWI lesion reversal on their 2 hour follow-up scan and acute PWI lesions. No perfusion or clinical outcome data was provided. The 2 cases with averted stroke were young (mean age 44.5 years); one had vascular risk factors; mean acute DWI lesion volumes were small at 1.29 ml; and both had abrupt onset neurological deficits. One patient had a vertebral artery dissection with an acute DWI lesion in the medulla (NIHSS 3) and the other had a small frontal lobe DWI lesion (NIHSS 6) and other area. The duration for symptoms is not given. As one patient had vascular risk factors and the other a vertebral artery dissection and abrupt onset symptoms with no alternative explanation offered, it seems likely that the complete resolution of their acute DWI lesions and clinical symptoms with rt-PA represents successful treatment of a stroke or transient ischaemic attack rather than the initial DWI lesions being false-positives and their symptoms related to a stroke mimic.

So in summary, there is some evidence that the size of DWI abnormalities may reduce slightly after thrombolysis with rt-PA and clinical improvement is associated with reperfusion of the ischaemic penumbra. Complete reversal of DWI lesions is rarely seen or expected in clinical practice when patients may present hours after stroke onset and development of infarction.

Dr [redacted] also states that:

- *IST-3 found symptomatic cerebral oedema on CT scans, as well as haemorrhage, was worse. (IST-3 2012) Data on asymptomatic haemorrhage, highlighted in the trial registry outcomes, are still awaited.*
- *The 2014 Cochrane review looked at symptomatic (including fatal) cerebral oedema on scans post stroke. With 5961 patients the scans indicated oedema in 10.2% given alteplase and 10.4% with controls. Symptomatic intracerebral haemorrhage is around 6% more common in treated patients. (Cochrane 2014, Emberson 2014)*

The results of the IST-3 Study have been discussed in Section 4.1.4 of the May CHM paper.

These concerns will be further addressed at the EWG Meeting in January 2015 when paper 5 will outline the benefits and risks in clinical practice, including off-label use.

3.7.3 Long-term mortality

Dr [redacted] cites a paper by Slot *et al* (2008) which found that increasing stroke severity, as measured by increased mRS at 6 months post-stroke, resulted in increasing risk of stroke-related mortality after long-term follow-up in 3 untreated cohorts of patients with ischaemic stroke.

This leads to a prediction that patients who receive rt-PA and are hopefully less disabled after treatment would have mortality curves that diverge over time from untreated patients. However, in the 18 month follow-up of the IST-3 trial, this divergence was not found. In NINDS, the initially encouraging mortality findings did not translate in 1 year follow-up to a further reduction in death compared with placebo. Between 3 and 12 months, of the 312 patients in each arm of the NINDS study, 22 patients had died in each arm (Kwiatkowski *et al*, 1999).

If rt-PA has a useful impact on disability it will then be associated with improved longer-term mortality, however evidence of this is yet to emerge.

Assessors' comments:

This issue was also discussed in the May CHM paper, section 4.3, in relation to the IST-3 results, and the following conclusions were drawn:

"These prospective cohort studies *[Slot et al, 2008]* consistently illustrate that a lower mRS at 6 months post-stroke is associated with improved survival in the long-term, which is considered to be a logical expectation. As noted in the discussion of the IST-3 trial, the secondary analyses found a (relatively small) improvement of mRS at 6 months post-stroke in the rt-PA treated group but this did not translate into a positive effect on death rate at 18 months of follow-up. This lack of effect on survival may be related to the small impact on mRS that was observed (small shifts in the overall spread of mRS may have been insufficient to result in a measureable impact on mortality) and also the length of follow-up. Whilst 18 months follow-up is much longer than most clinical trials, the cohort studies in this publication provide follow-up of between 7 and 19 years, and at the 18 month time-point the differences in survival are harder to discern particularly for mRS scores <5.

As discussed above, the findings from these three cohort studies of improved long-term survival in patients with a better mRS at 6 months provides some support for the use of mRS at 6 months in clinical trials as a surrogate endpoint for long-term outcomes."

As described above, the prospective cohort studies used mRS at 6 months post-stroke to examine the long-term mortality rates in patients, with data on follow-up for between 7 and 19 years. Although it would be hoped that the beneficial effect of rt-PA on stroke outcomes would lead to a reduction in long-term mortality, the likelihood of observing a difference in mortality rates at 18 months post-stroke, or at 12 months in the case of the NINDS trial is very low. Much longer follow-up of patients would be required. However, it is noted that the paper by Kwiatkowski *et al* found that the beneficial effect on disability observed in the NINDS trial was sustained at 12 months post-stroke.

An additional factor to consider is the other aspects of stroke care that are changing and have changed since the NINDS trial was conducted. For example, general improvements in basic stroke care, and importantly, improvements in secondary prevention of stroke. These factors are also likely to influence long-term mortality following stroke, and the benefit might be expected to be more than additive when combined with successful rt-PA treatment, for example rt-PA itself would have no direct influence on the probability of a patient having a second stroke, even in a patient with an excellent outcome from treatment. Slot *et al* noted a slight

improvement in survival over the time period of the cohorts which would support that improvements in medical care generally are having an impact.

The marketing authorisation for rt-PA does not make, and has never made, claims regarding mortality.

3.7.4 Final place of residence

Dr [redacted] considers final place of residence to be a more robust measure of outcome than mRS. In IST-3 this was a secondary outcome and treatment appeared to have minimal impact. This was omitted from the original Lancet paper of 2012, [redacted]

A table was presented in the supplementary appendix for the subset followed to 18 months, and a consistent benefit is not seen. Dr [redacted] considers that as IST-3 was an open trial with highly selected patients, the highest grade care could have been provided to treated patients, introducing bias, citing Sandercock (2014) which states that rt-PA treated patients, unlike controls, had in some centres a better staffed high dependency clinical pathway.

[redacted] considers there is little data available from other trials on this key cost benefit outcome. In NINDS, rt-PA treated patients were found to be more likely to be discharged to home, than to a nursing home or death. Dr [redacted] considers this could reflect the imbalance in baseline stroke severity. Hospital stay in survivors has been shorter in some trials, possibly due to early mortality.

Assessors' comments:

As discussed above in Section 3.5.1.1, whilst final place of residence may be a useful parameter, it is not without limitations, in particular it is a relatively crude measure in terms of the level of care required by the patient (which will be affected by their personal choice and economic situation) and it has been shown to be affected by other non-stroke related factors such as marital status as well as the baseline severity of the stroke and the effect of any treatment received.

- *Place of residence at 6 months follow-up for the full cohort was not included in the IST-3 publication*

As noted by Dr [redacted], place of residence was provided in the supplementary appendix to the publication for the 18 month follow-up cohort from IST-3. The data provided in the supplementary appendix showed place of residence at 6 months and at 18 months for this sub-group. Dr [redacted] s concern is that the data for the full cohort at 6 months follow-up was not presented.

As previously mentioned, this issue was discussed in the May CHM paper, as follows:

"Similar to the data provided in the supplementary appendix to the paper presenting the 18 month follow-up data, the data on place of residence at 6 months for the whole trial cohort do not suggest a substantial benefit for rt-PA treatment. These data do not affect the conclusions drawn on the published data."

- *rt-PA treated patients, unlike controls, had in some centres a better staffed high dependency clinical pathway*.

This imbalance is mentioned in [redacted] which states that patients treated with rt-PA had a 7% higher use of high-dependency beds than controls because some hospitals required all stroke patients receiving rt-PA to be monitored initially in a high-dependency unit. 89% of patients in both arms received care in a stroke unit, and [redacted] comment that the effect of high-dependency care remains uncertain and therefore it is difficult to judge the significance of this imbalance.

3.7.5 Recall and assessment bias

Dr [redacted] raises concerns that bleeding on treatment and differences between active and placebo as well as 'trials have been populated by treatment enthusiasts (who may perform outcome assessments)' and are usually analysed by sponsors, and that recall or assessment bias is likely. Only IST-3 examined for this bias and some evidence was found.

Details of the contents of active and placebo vials in studies have not been published, and neither has a risk assessment of the success in blinding in relation to the vials, mixing and final solution. Similarly, an assessment of potential unblinding by visible bleeding has not been published. Sandercock *et al* (2014) report that visible bleeding occurs in less than 10% of patients. The NINDS trial publication reports a 20% excess of bleeding in the treatment group, with minor external bleeding occurring in 23% of rt-PA patients and 3% of placebo.

Dr [redacted] comments that if patients were suspected to be in the rt-PA treatment group they may receive enhanced care and give and receive a more positive account of functional status.

Dr [redacted] also comments that the excipient arginine could be hazardous via an effect on nitrous oxide.

…bleeding on treatment and differences between active and placebo as well as 'trials have been populated by treatment enthusiasts (who may perform outcome assessments)' and are usually analysed by sponsors, and that recall or assessment bias is likely. Only IST-3 examined for this bias and some evidence was found.

The data referred to by Dr [redacted] for the IST-3 trial relates to data published in the webappendix of the 18 month follow-up. This was discussed in the May CHM paper, in section 4.1.5:

"The follow-up publication with 18 month data discusses the possibility of recall bias, stating that only 30% of survivors correctly recalled whether or not they received thrombolytic treatment, and accurate recall was associated with better outcome in both treatment groups, and therefore recall bias might have affected the findings. The authors then go on to state that the analysis of recall was based on a variable measured in a subset of survivors after randomisation and so could itself be biased.

The web appendix to the follow-up publication provides the following information:

	rt-PA		control		All	
Recall of thrombolytic therapy	No.	% OHS 0-2	No.	% OHS 0-2	No.	% OHS 0-2
Remembered treatment correctly	273	66.7	156	55.1	429	62.5
Remembered incorrectly or did not know	360	48.6	471	49.9	831	49.3
Question not asked or in double-blind phase[1]	76	44.7	81	38.3	157	41.4

The concern that could be taken from this table is that there is only a difference in success rates in patients that remember their treatment (66.7% vs. 55.1%), whilst in those who could not remember there was no difference (48.6% vs. 49.9%). This

could lead to a conclusion that the treatment difference is entirely driven by bias caused by knowledge of the treatment – people scoring better when they know they have received rt-PA.

However, recall of therapy is a post-randomisation covariate, i.e. it is itself influenced by treatment, because a good response could result in patients being more likely to recall the treatment they received. This is supported by the fact that recall was associated with better outcomes in both groups – control patients who recalled their treatment did better than those that did not. This argues against a conclusion of recall bias, which would be expected to lead to patients who recalled they were on control doing worse as they knew they had not received rt-PA.

In summary, it is not appropriate to draw conclusions from data based upon splitting by a post-randomisation covariate, as the between group comparisons in the sub-groups are misleading when the covariate is itself influenced by treatment. If anything, the table supports a hypothesis that recall is associated with response in both treatment groups and that rt-PA causes both an increased clinical response rate (albeit a small one) and an increased recall rate."

Overall in an open-label trial you would generally assume that there is perfect knowledge of treatment allocation and the primary endpoint should be sufficiently objective that the consequences of running an open-label trial are limited. As, in any case, the trial was negative and has not impacted the licensing decision, this cannot be considered a major concern.

- *Details of the contents of active and placebo vials in studies have not been published, and neither has a risk assessment of the success in blinding in relation to the vials, mixing and final solution has not been published.*

The composition of the placebo used in clinical trials was one of the questions put by the MHRA to the MAH. The MAH has confirmed that for the ECASS II and III studies, the placebo vials contained [redacted] The constituents were: [redacted] The MAH stated that information on the composition of the placebo used in NINDS and the ATLANTIS A and B studies is not currently available from Genentech but that it would be supplied to MHRA when the information was received.

- *Arginine could be hazardous via an effect on nitrous oxide.*

Patients receive up to 90mg of rt-PA (dependent on body weight). This will provide an equivalent dose of [redacted] of L-arginine, of which 10% should be delivered as an initial intravenous bolus, and the rest infused intravenously over 60 minutes.

Arginine could have an impact on outcome in acute ischaemic stroke via more than one mechanism. For example, arginine can increase NO levels because it acts as a substrate for nitric oxide synthase (NOS) including endothelial NOS (eNOS). eNOS generation of NO may increase cerebral blood flow by acting as a vasodilator. This was demonstrated to confer protection from ischaemic stroke in rats (Dalkara *et al*, 1994).

Glyceryl trinitrate (GTN) also raises NO levels and therefore could be expected to have some similar effects compared with arginine. In a small number of stroke patients (n=18), GTN was found to lower blood pressure and did not alter cerebral blood flow or cerebral perfusion pressure (Willmot *et al*, 2006). In a small study of 41 hypertensive stroke or TIA patients (systolic blood pressure >140mmHg) randomised to GTN patch (n=25) or none (n=16) with treatment initiated by paramedics (within 4 hours, median 55 minutes), the GTN patch group was found to have lower blood

pressure, an improved mRS outcome (shift by 1 point) and a lower mortality rate (Ankolekar *et al*, 2013).

A large trial in hypertensive patients with acute ischaemic or haemorrhagic stroke (n=4011), randomised to 7 days of GTN treatment started within 48 hours of the stroke or to no treatment has recently reported its findings (ENOS trial investigators, 2014). This trial was designed to study whether hypertensive stroke patients would have an improved outcome as a result of lowering blood pressure early after stroke (the study was a partial-factorial design, with a subset of patients who were taking antihypertensive medicines prior to the stroke randomised to continue or stop this medication). Whilst GTN was found to significantly reduce blood pressure, this did not translate into an improved functional outcome at day 90, as measured using the mRS with an ordinal analysis. However, patients in this trial began treatment up to 48 hours post symptom onset, and subgroup analyses for different times to randomisation suggest that treatment at ≤6 hours could be beneficial, although the number of patients in this group was relatively low (n=273):

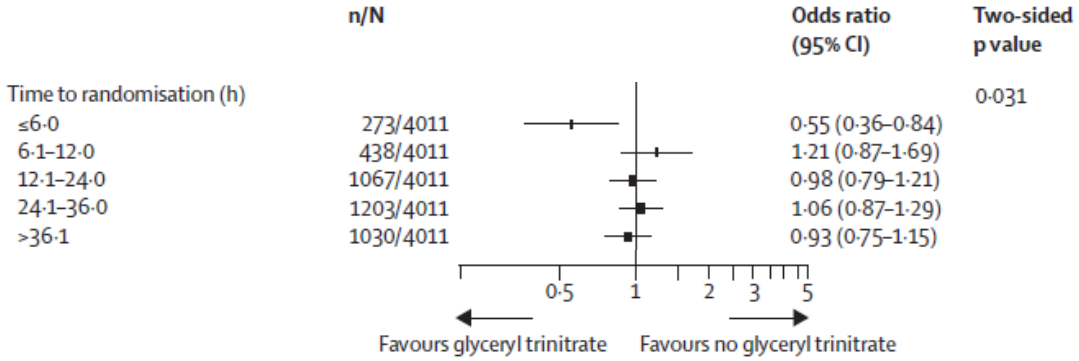


Figure: taken from ENOS trial investigators, 2014

This pattern, if confirmed, would be consistent with the findings of Harston *et al* (2010) who suggested from a systematic review of animal studies that L-arginine may have a neuroprotective role when administered early following ischaemia, but delayed administration may worsen ischaemic damage.

A separate effect of arginine is its potential influence on clot structure, as arginine has been shown *in vitro* to affect fibrin structure and its lysis (Kovacs *et al*, 2014). Arginine is naturally present in the circulation (~100μM), the concentration can vary under different circumstances (e.g. it falls below 50μM in sepsis) and has been shown to be generated during clot thrombolysis (Kovacs *et al*, 2014).

Although arginine is included as an excipient in several other medicines, including some monoclonal antibodies and blood factor concentrates, which demonstrates its overall acceptable safety and tolerability, it is not necessarily the case that arginine is having no effect in all medicines/situations, including ischaemic stroke.

[redacted]

The effects of exogenous arginine [redacted] are likely to be complicated and may depend on several parameters

including the timing of the dose, the underlying condition of the patient, any interaction with rt-PA and the resulting concentration of circulating arginine. On the basis of current knowledge it is therefore difficult to predict whether the inclusion of arginine as an excipient of rt-PA has an effect, what this may be and how it may impact on the balance of benefits and risks of rt-PA.

- *An assessment of potential unblinding by visible bleeding has not been published.*

Although visible bleeding is common with rt-PA treatment, it also occurs in some placebo treated patients and, while possibly increasing the likelihood of a correct guess, it cannot be considered to be a fail-safe method of determining treatment allocations.

For the NINDS trial, the publication states that "the outcome was determined at 24 hours and three months by certified examiners who had not performed the baseline examination and had not been present during the initial treatment". Assuming this protocol was followed, visible bleeding would have to have occurred/be ongoing at the time of the outcome assessments for unblinding to potentially have been an issue.

For the ECASS III trial (Hacke *et al* 2008), it is specified that "patients were assessed by an examiner who was unaware of the treatment assignment. Assessments were made at the time of enrolment, at 1, 2, and 24 hours after administration of the study drug began and on days 7, 30 and 90 after administration of the drug."

It is not possible to know whether sites were always able to ensure that follow-up assessments were carried out by a physician not involved in the acute care of the patient, nor how frequently visible bleeding coincided with follow-up assessments.

We are not aware of any assessment of potential unblinding by visible bleeding.

3.7.6 Generalisability

Dr [redacted] comments that following the procedure in the product licence for correct administration of rt-PA requires detailed history taking and examination and test results including competent reading of scans at any time of day. A paper by Bray et al (2013) is cited to support the view that current NICE and MHRA advice is not being followed in the UK. Protocol violations were even common in the context of trials - ECASS I.

SITS was an observational cohort treated with rt-PA and run by Boehringer Ingelheim, which included a population with a low mortality rate. When compared with a placebo group, a 3% increase in mortality with rt-PA is suggested [redacted].

Four trials of thrombolysis have been halted because of safety concerns (ASK, MAST-I, MAST-E, and ATLANTIS). In NINDS, participant numbers were doubled, in IST-3 they were halved and in ECASS III the time window for treatment was extended mid-trial. Full ATLANTIS results were held back until 2002 by when the FDA and EU regulators had made judgements primarily based on NINDS. Recruitment in all trials was either slow or very slow, never acknowledged in publications as being due to prudent clinicians protecting patients.

Seven and thirty day mortality suggests NINDS and ECASS III were the trials with the best early mortality profile, and regulators mainly based decisions on these trials, others received less attention. Overall 7 day mortality is probably raised ~3-4%, and 30 day mortality by 2-3%, this may persist (if open label IST-3 is not considered).

Trial endpoint mortality is usually increased but was not in NINDS (unbalanced randomisation), ECASS III (low risk population) or IST-3 (open-label design).

- *A paper by Bray et al (2013) is cited to support the view that current NICE and MHRA advice is not being followed in the UK. Protocol violations were even common in the context of trials - ECASS I.*

The paper by Bray *et al* (2013), determined that patients aged over 80 years are being treated with rt-PA, that their treatment is as timely as that of younger patients, that they had similar rates of post-thrombolysis complications as compared with patients <80 years, and mortality was high among older patients whether they were treated with rt-PA or not.

- *SITS was an observational cohort treated with rt-PA and run by Boehringer Ingelheim, which included a population with a low mortality rate. When compared with a placebo group, a 3% increase in mortality with rt-PA is suggested* [redacted]

The SITS registry did not provide placebo data, and comparison of mortality rate from data generated in an observational registry with randomised controlled trial data [redacted] is inappropriate, due to the differences in the populations included in the cohorts.

- *Four trials of thrombolysis have been halted because of safety concerns (ASK, MAST-I, MAST-E, and ATLANTIS).*

The ASK, MAST-I and MAST-E trials were all trials of streptokinase, and as discussed earlier in this paper, data generated with streptokinase are not considered to apply to rt-PA. The ATLANTIS study was not halted because of safety concerns, but rather efficacy, with the DMSB stating after an interim analysis that 'treatment was unlikely to prove beneficial'.

- *Full ATLANTIS results were held back until 2002 by when the FDA and EU regulators had made judgements primarily based on NINDS.*

The ATLANTIS results were published in 1999 and 2000 and were considered as part of the original application for the EU indication in acute ischaemic stroke in 2002.

- *Recruitment in all trials was either slow or very slow, never acknowledged in publications as being due to prudent clinicians protecting patients.*

The main reason for subject exclusion in the NINDS trial was the restriction on time-window for treatment to 0-3 hours following onset of symptoms. Similarly the recruitment during ECASS III was initially slow due to the narrow time-window (initially 3-4 hours following onset of symptoms). The practicalities associated with rt-PA treatment and the lack of an established infra-structure may explain slow recruitment in trials. We are not aware of any evidence that issues with recruitment were instead due to clinician concerns over patient safety and this would seem at odds with another concern highlighted, that the investigators involved in the trials are all 'rt-PA enthusiasts'.

- *Seven and thirty day mortality suggests NINDS and ECASS III were the trials with the best early mortality profile, and regulators mainly based decisions on these trials, others received less attention. Overall 7 day mortality is probably raised ~3-4%, and 30 day mortality by 2-3%, this may persist (if open label IST-3 is not considered). Trial endpoint mortality is usually increased but was not in NINDS (unbalanced randomisation), ECASS III (low risk population) or IST-3 (open-label design).*

Although the NINDS and ECASS III trials were major trials considered during the licensing of the acute ischaemic stroke indication and the extension of the time

window to 4.5 hours respectively, other trial data were equally considered, critically reviewed and discussed during these EU licensing procedures. The overall conclusion regarding mortality associated with rt-PA treatment during these procedures recognised that rt-PA likely had an adverse effect on early (up to day 7) mortality, however this is no longer the case at longer follow-up (day 90 or 6 months). The issues cited with the NINDS, ECASS III and IST-3 studies have been addressed elsewhere.

4. Definitions and frequencies of symptomatic intracerebral haemorrhage (sICH)

The rt-PA trials have used different definitions for sICH, and the choice of definition used affects the magnitude of the results. Therefore when results are compared between trials it is important that consideration is given to the definitions used in those trials. As sICH is the most important adverse effect associated with rt-PA, this section has been included to discuss the different definitions used in clinical trials and whether it is possible to decide which definition is the most clinically relevant.

Haemorrhagic transformation frequently accompanies ischaemic strokes in patients who receive no specific treatment (Khatri *et al*. 2007). The incidence of haemorrhagic transformation is difficult to estimate as it is often asymptomatic and detected incidentally on routine brain imaging performed at varying times after stroke onset. To occur, intracerebral haemorrhages require some degree of reperfusion, either spontaneous or by thrombolysis or via collateral circulation, and vessel weakness (Lyden and Zivin 1993). Anticoagulants and thrombolytic drugs can increase the frequency and severity of haemorrhagic transformation.

Symptomatic intracerebral haemorrhage (sICH) occurs when haemorrhagic transformation is associated with clinical deterioration. However, whether a bleed is symptomatic does not necessarily depend on its size. Even large haemorrhages do not always cause symptoms if a non-eloquent area of the brain is involved whereas small bleeds in critical parts of the dominant hemisphere may have devastating clinical consequences (Dzialowski *et al*. 2007). Most intracranial bleeds occur within 24 to 36 hours of intravenous thrombolysis. However, there is no standard agreed definition for a symptomatic ICH.

4.1 NINDS and ECASS

In the NINDS rt-PA trial, sICH was defined as detection of blood in a computed tomography (CT) scan performed within 36 hours from treatment onset and associated with any neurological decline. Furlan *et al*. (1999) defined neurological deterioration as an increase of ≥4 points on the National Institute of Health Stroke Scale (NIHSS) (or a 1-point deterioration in the level of consciousness item) but did not provide any rationale for the chosen thresholds. Although the NIHSS provides an objective and functional assessment of haemorrhage severity, it is restricted by ‘ceiling effects’ in patients with severe strokes and high baseline NIHSS scores. It can also be difficult to establish a reliable baseline NIHSS score for some patients who progressively deteriorate and close monitoring may be required to follow those patients who are deteriorating before establishing that sICH is present.

The European Cooperative Acute Stroke Study (ECASS) II trial defined symptomatic haemorrhage as blood visible at any site on a brain CT scan (within 7 days) associated with an increase of ≥4 points in the NIHSS (Hacke *et al.*1998[1]). When it was not clear whether cerebral oedema or haemorrhage had resulted in clinical deterioration then an association with haemorrhage was assumed.

For the ECASS III study any intracranial haemorrhage had to be temporally related to neurological deterioration and identified as the predominant cause for it to be categorised as a sICH (Hacke *et al* 2008). The chairs of the safety outcome adjudication committee and steering committee decided if the cause of death or neurological deterioration was related to haemorrhage, brain injury or had an alternative cause.

The NINDS and ECASS II studies also provided radiological definitions of haemorrhagic transformation sub-types (see table below). Haemorrhagic infarction (HI) occurs in the area of infarction but parenchymal haemorrhage (PH) can be observed in areas of infarction or occur in other remote locations. The ECASS II radiological definition classifies HI further on the basis of the degree of haemorrhagic confluence without mass effects and PH by size and degree of mass effects.

The ECASS II definition of PH2 was incorporated into the definition of sICH used in the SITS-MOST study (Wahlgren *et al.* 2007). In the majority of randomised controlled trials (RCTs), those interpreting imaging results were blinded to the clinical outcome but this was rarely true for Registry or cohort studies. The diagnostic sensitivity of the various imaging techniques used in different in studies may also vary.

Table: Radiological classification of haemorrhagic transformation after rt-PA treatment in the NINDS and ECASS II trials

Study	Radiological Definition		Timing of CT scan
	Parenchymal haemorrhage (PH)	**Haemorrhagic Infarction (HI)**	
NINDS	Homogeneous hyperdense lesion with a sharp border with or without oedema or mass effect	Acute infarction with punctuate hypodensity/ hyperdensity with an indistinct border within the vascular territory	36 hours
ECASS II	PH1: blood clots in <30% of infarcted area with some slight space-occupying effect PH2: blood clots in >30% of the infarcted area with substantial space-occupying effect	HI1: small petechiae along the margins of the infarct HI2: confluent petechiae within the infarcted area but no space-occupying effect	7 days

A number of different definitions of sICH based on the extent of haemorrhage, its location (within infarcted area or remote), and the severity of neurological deterioration measured by the National Institutes of Health Stroke Scale (NIHSS) and its association with death have been used in the main rt-PA trials (National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group 1995; Hacke *et al*. 1998[1], 2008; Wahlgren *et al*. 2008) (see table below).

Clinical trials, stroke registries and cohort studies have used different case definitions for sICH with some (e.g. NINDS and SITS-MOST) considering sICH to be attributable to thrombolysis when they occurred within 36 hours of thrombolysis, and others (e.g.

ECASS II) considering haemorrhages to be of clinical importance when they occurred up to 7 days after rt-PA treatment.

4.2 Utility of sICH definitions

The interpretation and comparison of safety data from different rt-PA studies has therefore been complicated by the absence of a standard case definition for sICH. Berger *et al*. (2001) used the ECASS II study data to show that only type 2 parenchymal haemorrhages were associated with an increased risk of deterioration at 24 hours after stroke onset (adjusted odds ratio, 18; 95% confidence intervals (CI), 6 to 56) and death at 3 months (adjusted odds ratio, 11; 95% CI, 3.7 to 36). Studies are increasingly reporting multiple sICH rates using different definitions to simplify comparisons of safety data.

The predictive properties of the common sICH definitions for the clinical outcomes of mortality and disability at 90 days have been reported in 314 patients with anterior circulation ischaemic stroke treated with thrombolytic therapy (Gumbinger *et al*. 2012). Off-label thrombolysis was given to 90 patients aged over 80 years or within an extended time window on the basis of brain imaging findings (9% were treated > 4.5 hours after stroke onset). All patients had a routine CT or MR scan at 24-36 hours after thrombolysis or at any time if clinically indicated. The imaging and clinical data were analysed using the NINDS, ECASS and SITS-MOST definitions for sICH. The odds ratios for patients with and without sICH were calculated for mortality and disability outcomes at 90 days.

Only 34 PHs were detected (PH type 1: 22 patients, 7%; 95% CI, 4.4-10.6%; PH type 2: 12 patients, 3.8%; 95% CI, 1.9-6.7%). The inter-rater agreement rates for haematoma size and variability of PH assessment were low with κ values of 0.61 and 0.74 respectively.

Table: Definitions of symptomatic intracerebral haemorrhage used in the main intravenous rt-PA studies.

Study	Definition
NINDS	Any haemorrhage not observed on a previous CT scan with a suspicion of haemorrhage or any decline in neurologic status
ECASS II	Any haemorrhage with neurological deterioration of 4 points or more on the NIHSS from baseline or from the lowest NIHSS value to 7 days or leading to death
ECASS III	Any haemorrhage with neurological deterioration of 4 points or more on the NIHSS from baseline or the lowest value in the first 7 days or any haemorrhage leading to death. In addition, the haemorrhage must have been identified as the predominant cause of the neurologic deterioration.
SITS-MOST	Local or remote parenchymal haemorrhage (PH) type 2 on the 22-36 hour post-treatment imaging scan combined with neurological deterioration of 4 points or more on the NIHSS value from baseline, or from the lowest NIHSS value between baseline and 24 hours, or leading to death

Key: ECASS, European Cooperative Acute Stroke Study ; NIHSS, National Institutes of Health Stroke Scale; NINDS, National Institute of Neurological Disorders and Stroke; SITS-MOST, Safe Implementation of Thrombolysis in Stroke-Monitoring Study;

The NINDS sICH criteria produced the highest number of cases and the ECASS III study definition gave the lowest. The SITS-MOST sICH definition had the best positive predictive value for death (OR, 14.4; 95% CI, 3.3– 85.9) and the NINDS definition proved the best predictor of an unfavourable outcome (OR, 10.4; 95% CI, 2.49-93.06). However, the large confidence intervals for many of the outcome variables are consistent with an underpowered study which is to be expected given the low number of PHs observed. However, none of the sICH definitions were ideal at predicting mortality and adverse disability outcomes and none showed a high level of inter-rater consistency. The SITS-MOST definition was recommended for the clinical evaluation of mortality and the ECASS II definition for its relatively high inter-rater agreement rate (κ value 0.85).

Another critical review of the different case definitions for sICH following intravenous thrombolysis compared their consistency with mortality rates at 90 days using published data from clinical trials, stroke registries and cohort studies (with > 200 patients) to 2011 (Seet and Rabinstein 2012). The overall mean sICH rate was 5.6% (standard deviation, SD, 2.3%) and the mean mortality rate was 14.7% (SD 4.8%). There was a moderate correlation between the incidence of sICH and mortality (correlation coefficient, r=0.401, p=0.05). Studies that defined sICH as parenchymal haemorrhage with a neurological decline on the NIHSS of ≥ 4 points occurring within 36 hours of thrombolysis reported a higher correlation with mortality rates (r=0.631). Variation in reported sICH rates was highest for studies that used the SITS-MOST criteria than for those using the ECASS II and NINDS criteria.

The MAH states that the most clinically relevant definition of sICH is the SITS-MOST definition as it has a good predictive value for poor outcome and mortality at 90 days after thrombolysis. This conclusion is based on the study by Gumbinger *et al.* (2012) which only identified 34 parenchymal haemorrhages. Others have found that the SITS-MOST criteria produce more variation in reported rates of sICH using mortality data than the NINDS and ECASS II criteria. Until a standard definition of sICH is available which has high inter-rater agreement rates and which correlates well with clinical outcomes, studies should report sICH data using all of the common definitions or provide detailed descriptions of the type of ICHs observed, the extent of NIHSS deterioration and the time intervals after thrombolysis to permit detailed comparison of safety data.

4.3 Estimate of frequencies of sICH

The reported frequencies of sICH from clinical trials, pooled analyses, registry and cohort studies are summarised in the following tables: table A shows the reporting rates of sICH with rt-PA in the main randomised controlled trials, pooled or meta-analyses and observational studies; table B shows the reporting rates of sICH with intravenous alteplase from the large Registries; table C shows the reporting rates of sICH from large cohort studies (n>200).

Table A: Reported rates of sICH with rt-PA in the main clinical trials, pooled analyses and observational studies (taken and adapted from Lorenzano 2014).

Study	Number of patients	Time from OTT	Time of Assessment	ICH definition	Number or proportion (%) of patients		Odds Ratio (95% CI)	*p* value
					rt-PA (%) (95% CI)	Control (%) (95% CI)		
NINDS 1995	624	≤ 3 hours	≤ 36 hours	NINDS	6.4%	0.6%	-	<0.001
Pooled analysis (NINDS; ECASS I and II; and ATLANTIS) (Hacke *et al*. 2004)	2775	≤ 6 hours	90 days	PH2*	82 (5.9)	15 (1.1)	-	<0.0001
		0-90 mins			5/161 (3.1) (1.6-5.6)	0/150 (0)	-	-
		91-180 mins			17/302 (5.6) (3.9-7.9)	3/315 (1) (0.4-2)		
		181-270 mins			23/390 (5.9) (4.3-8)	7/411 (1.7) (1.0-2.9)		
		271-360 mins			37/538 (6.9) (5.3-8.7)	5/508 (1) (0.5-1.8)		
SITS-MOST Unadjusted analysis (Wahlgren *et al*. 2007)	6483	≤ 3 hours	36 hours	NINDS	468/6438 (7.3) (6.7-7.9)	Pooled RCTs 40/465 (8.6) (6.3-11.6)	-	-
Adjusted analysis (Wahlgren *et al.* 2008a)					8.5% (7.9-9.0)			
ECASS III (Hacke *et al*. 2008)	821	3–4.5 hours	22-36 hours	NINDS	33/418 (7.9)	14/403 (3.5)	2.38 (1.25-4.52)	0.006

Study	Number of patients	Time from OTT	Time of Assessment	ICH definition	Number or proportion (%) of patients		Odds Ratio (95% CI)	p value
					rt-PA (%) (95% CI)	Control (%) (95% CI)		
ECASS III (Hacke *et al.* 2008)	821	3–4.5 hours	22-36 hours	ECASS II	22/418 (5.3)	9/403 (2.2)	2.43 (1.11-5.35)	0.02
				ECASS III	10/418 (2.4)	1/403 (0.2)	9.85 (1.26-77.3)	0.008
				SITS-MOST	8/418 (1.9)	1/403 (0.2)	7.84 (0.98-63)	0.02
Updated pooled analysis (NINDS; ATLANTIS; ECASS I, II and III; EPITHET) (Lees *et al*. 2010)	3670	≤ 6 hours	36 hours	PH2* (likely to affect outcome)‡	96/1850 (5.2)	18/1820 (1)	5.37 (3.22-8.95)	<0.0001
		0-90 mins			5/161 (3.1)	0/151 (0)	-	-
		91-180 mins			17/303 (5.6)	3/315 (1)	8.23 (2.39-28.3)	< 0.0008
		181-270 mins			35/809 (4.3)	10/811 (1.2)	3.61 (1.76-7.38)	< 0.0004
		271-360 mins			39/576 (6.8)	5/542 (0.9)	4.32 (2.84-18.9)	< 0.0001

Study	Number of patients	Time from OTT	Time of Assessment	ICH definition	Number or proportion (%) of patients		Odds Ratio (95% CI)	p value
					rt-PA (%) (95% CI)	Control (%) (95% CI)		
IST-3 (The IST-3 Collaborative Group, 2012)	3035	≤ 6 hours	7 days	IST-3†	104/1515 (7)	16/1519 (1)	6.94 (4.07-11.8)	< 0.0001
SITS-ISTR (Ahmed *et al*. 2013)	25279	≤ 3 hours	22-36 hours	NINDS	1731/24735 (7)	-	-	-
				ECASS II	1140/24845 (4.6)	-	-	-
				SITS-MOST	381/24910 (1.5)	-	-	-
	4056	3–4.5 hours	22-36 hours	NINDS	256/3945 (6.5)	-	1.05 (0.9-1.22)	0.54
				ECASS II	179/3959 (4.5)	-	1.11 (0.93-1.32)	0.26
				SITS-MOST	70/3969 (1.8)	-	1.22 (0.92-1.61)	0.16
	283	4.5-6 hours	22-36 hours	NINDS	14/273 (5.1)	-	0.72 (0.40-1.31)	0.29
				ECASS II	8/273 (2.9)	-	0.54 (0.24-1.23)	0.14
				SITS-MOST	7/273 (2.6)	-	1.57 (0.68-3.64)	0.29

Study	Number of patients	Time from OTT	Time of Assessment	ICH definition	Number or proportion (%) of patients		Odds Ratio (95% CI)	p value
					rt-PA	Control		
Updated Cochrane meta-analysis (Wardlaw *et al*. 2012)	7012	0-3 hours	7 days	sICH (as defined in trials)	72/896 (8)	11/883 (1.2)	4.55 (2.92-7.09)	< 0.0001
		3-6 hours			191/2488 (7.7)	45/2447 (1.8)	3.73 (2.86-4.86)	< 0.0001
		≤ 6 hours			272/3548 (7.7)	63/3463 (1.8)	3.72 (2.98-4.64)	< 0.0001

Key: * PH2 defined as dense blood clot exceeding 30% of infarct volume with substantial space-occupying effect; ‡ Mostly identifies the same patients as ECASS III and SITS definitions for sICH; † defined as significant neurological deterioration accompanied by evidence of significant ICH on the post-randomisation scan (or autopsy if not rescanned and death after 7 days). This included recurrent stroke within 7 days if confirmed to be caused by an ICH). 1= Odds ratio calculated by comparing 3-4.5 h and 4.5-6 h versus within 3 h cohorts. -=not reported.

Abbreviations: ATLANTIS, Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischaemic Stroke; CI, confidence intervals; ECASS, European Cooperative Acute Stroke Study; EPITHET, Echoplanar Imaging Thrombolytic Evaluation Trial; ICH, intracerebral haemorrhage; IST-3, International Stroke Trial; NINDS, National Institute of Neurological Disorders and Stroke; PH, Parenchymal haematoma; RCTs Randomised Controlled Trials; SITS-ISTR, Safe Implementation of Thrombolysis in Stroke-International Stroke Thrombolysis Registry; SITS-MOST, Safe Implementation of Thrombolysis in Stroke-Monitoring Study

Table B: Reporting rates of sICH with intravenous rt-PA from the large Registries to 2011(taken from Seet and Rabinstein 2012).

Study	Number of subjects	Age (years)	Median Baseline NIHSS	Time to treatment (mins)	SICH rates (%)	Mortality rates (%)
STARS (Albers 2000)	389	69	13	164	3.3[a]	13.0[a]
CASES (Hill 2005)	4468	73	14	155	4.6[b]	22.3[e]
SITS-MOST (Wahlgren 2007)	6483	68	12	136	SITS-MOST 1.7; ECASS 4.6; ECASS II 8.8; NINDS 7.3	11.2[e]
SITS-ISTR (Wahlgren 2008b)	12529	68	12	143	SITS-MOST 1.6; ECASS 4.8; NINDS 7.3	12.1[e]
SITS-ISTR (Ahmed 2010)	23942	68	12	146	SITS-MOST 1.75; ECASS 4.85; NINDS 7.13	12.3[e]
GWTG (Fonarow 2011)	25504	70	12	129	5.4[a]	9.9[f]
Canadian Stroke Network (Vergouwen 2011)	1739	75	12	145	5.9[c]	16.3[f]
Total (mean)	**15054**	**70**	**12**	**145**	**3.5[g]**	**13.9**

Key:
a= definition not available; b=defined as any neurological decline and parenchymal haematoma occurring within 24 hours after intravenous thrombolysis; c= defined as any neurological decline and cerebral haemorrhage occurring within 36 hours after intravenous thrombolysis; d= at 30 days after stroke; e= at 90 days after stroke; f=in hospital; g=ECASS II data were used in studies that reported >1 SICH rates.

Abbreviations:
STARS=Standard Treatment with Alteplase to Reverse Stroke study; SITS-ISTR=Safe Implementation of Thrombolysis in Stroke International Study; GWTG=Get With The Guidelines.

Table C: Reporting rates of sICH from large cohort studies (n>200) to 2011 (taken from Seet and Rabinstein 2012).

Study	Number of subjects	Age (years)	Median Baseline NIHSS	Time to treatment (mins)	SICH rates (%)	Mortality rates (%)
Tanne 2002	1205	67	NA	NA	6[a]	13.5[f]
Schenkel 2003	250	63	14	141	8.8[a]	17[g]
Berrouschot 2005	228	68	14	NA	2.6[b]	7.9[g]
Chao 2010	241	66	15	139	SITS-MOST 3.7; ECASS 5.4; NINDS 7.9	10[g]
Grotta 2001	269	68	14	137	5.6[c]	15[h]
Ringleb 2007	468	71	13	148	5.5[a]	16[g]
Sobesky 2007	450	66	11	135	4[d]	11[g]
Uyttenboogaart 2008	252	68	12	174	5.2[c]	17[g]
Seet 2011	212	74	13	141	7.9[b]	20[b]
Strbian 2011	987	71	9	120	SITS-MOST 2.1; ECASS II 7.0; NINDS 9.4	10.2[g]
Total (mean)	**4455**	**68**	**13**	**146**	**5.9[i]**	**14.7**

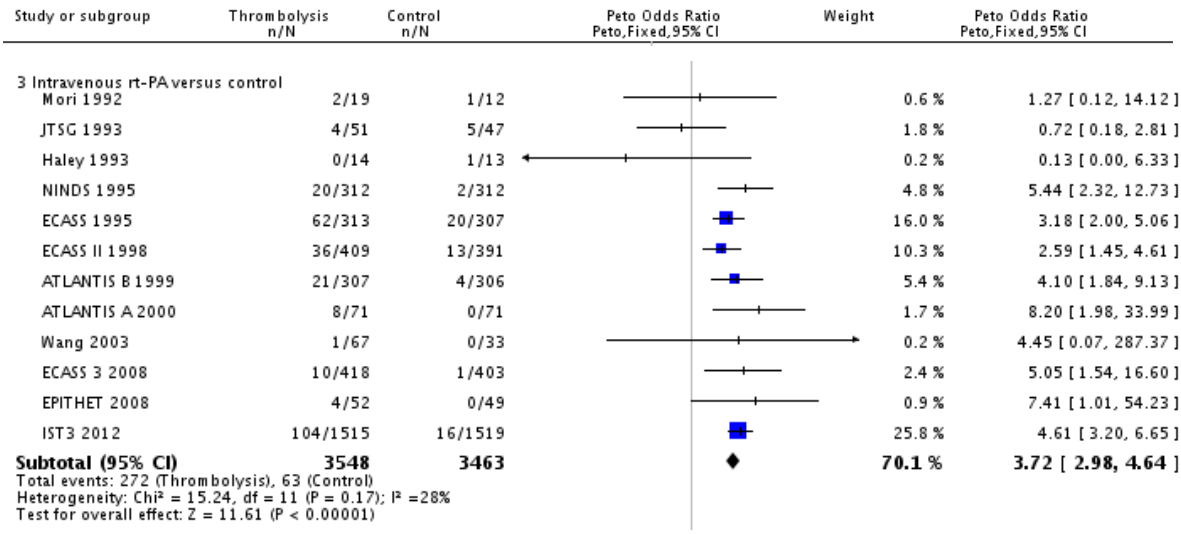
Key:
a= definition not available; b=defined as any neurological decline and parenchymal haemorrhage occurring within 24 hours after intravenous thrombolysis; c= defined as any neurological decline and cerebral haemorrhage occurring within 36 hours after intravenous thrombolysis; d= at 30 days after stroke; e= at 90 days after stroke; f=in hospital; g=ECASS II data were used in studies that reported >1 SICH rates.

In summary, the reported rates of sICH from the main clinical studies ranged from 1.5% using the SITS-MOST definition within 3 hours of stroke onset (Ahmed *et al*. 2013) to 8.5% from the adjusted analysis of the SITS-MOST data using the liberal NINDS definition (Wahlgren *et al*, 2008). Most studies had rates of 5-6%. The corresponding placebo rates of sICH ranged from 0 in certain sub-groups (Hacke *et al*. 2004; Lees *et al*. 2010) to 11.6% from pooled RCT data (Wahlgren *et al*. 2007). Most studies had rates of < 5%. Methodological differences between the studies complicates further detailed analysis.

The reported rates of sICH from larger registry studies reported up to 2011 were 3.3 to 8.8% depending on the definition used. The mean sICH rate was 3.5% which compares favourably with the rates reported from clinical trials. A more complete description of recent registry data was provided earlier.

A recent systematic review of thrombolysis that analysed the results of 12 trials using rt-PA, found there were 60 (95% CI 50 to 70) extra symptomatic ICHs within 7 to 10 days per 1000 participants treated (OR 3.72, 95% CI 2.98 to 4.64, p < 0.00001; 7011 participants) with no heterogeneity between trials (Wardlaw *et al*. 2014) (see figure below). No attempt was made to standardise the definition for sICH between trials, the trial's primary definition for sICH was simply accepted for this meta-analysis.

Figure: Forest plot for rt-PA versus placebo for the outcome of sICH within 7 to 10 days of treatment



The MAH states that symptomatic intracerebral haemorrhage rates are not usually reported after 7 days post treatment. This is accepted as most rt-PA associated ICHs occur within 24-36 hours after treatment due to its short elimination half-life. The incidence of sICHs per the SITS-MOST definition from RCTs (prior to IST-3) is presented according to onset to time of treatment and the frequency of sICH is 2.7%. Data from the SITS-ISTR registry shows that the frequency of sICH was 1.7% if rt-PA was given within a 3 h time window and 2.2% if given within 3-4.5 h. This data is also shown in table 3. This indicates that the rate of sICH appears to be equivalent to that reported from clinical trials if the same sICH definition is used although off-label use is included in the SITS-ISTR. However, it should be noted that use of the SITS-MOST definition for sICH produces the lowest frequencies and the most variable results according to the limited number of studies that have compared the predictive properties of the different sICH definitions for the clinical outcomes of mortality and disability.

5. Discussion of primary endpoints used in rt-PA clinical trials and their appropriateness and implications

The choice of primary endpoint used in acute ischaemic stroke trials is clearly an important one, and this subject is frequently raised as a concern regarding the appropriate interpretation of the studies. This section is intended to provide some background information on the primary endpoints chosen for the key studies of rt-PA, and discussion on the pros and cons of the different possible approaches.

The primary endpoints used in the key trials are described in the following table:

Trial	Primary endpoint
NINDS part I	Evaluation of 'significant early improvement' between treatment groups in 0-90 min; 91-180 min; 0-180 min. *'Significant early improvement' = improvement in baseline NIHSS by 4 points, or complete resolution to score of 0, at 24 hour exam.*
NINDS part II	To assess the hypothesis that: there is a consistent and persuasive difference between the rt-PA treatment group and the placebo group enrolled within 180 minutes of stroke onset in the proportion with a 90 day outcome of: a) Barthel Index ≥ 95 b) Modified Rankin Scale 0-1 c) Glasgow Outcome Scale 1 d) NIHSS 0-1
ATLANTIS part A	1) Significant difference between the rt-PA and placebo treated groups in clinical improvement [defined as decrease of ≥4 points on the NIHSS or complete resolution of symptoms from baseline to 24 hours and from baseline to 30 days] 2) Significant difference between rt-PA and placebo treated groups in volume of cerebral infarction as measured by cerebral CT scanning at 30 days.
ATLANTIS part B	Excellent neurologic recovery at day 90: NIHSS score of ≤1
ECASS I	1) Difference between rt-PA treated and placebo treated groups in activities of daily living, i.e. a difference of 15 points in the Barthel Index at 90 days. 2) Difference between rt-PA and placebo treated patients of one grade in modified Rankin Scale at 90 days
ECASS II	Modified Rankin Scale at 90 days, favourable (0-1) or unfavourable (2-6)
ECASS III	Modified Rankin Scale at 90 days, favourable (0-1) or unfavourable (2-6)
IST-3	Alive and independent at 6 months, defined as Oxford Handicap Score (OHS) of 0-2.

Whilst NINDS part I and II had different primary endpoints, when the studies were published they were analysed as one dataset on account of their otherwise identical protocols. The primary outcome in part I was chosen to test whether rt-PA had clinical activity by improving neurological impairment, as measured by a relatively small change in NIHSS - a sensitive measure to detect a change in neurological deficit. It would also reflect any significant deterioration due to ICH. The primary outcome in part II was intended to measure sustained clinical benefit and this study was considered the pivotal study. The NINDS rt-PA study group considered the

primary endpoint for part II to be more clinically relevant than the primary endpoint used in part I.

5.1 The basic features of the stroke scales used in the key clinical trials

NIHSS: The NIHSS scale was first described in 1989 by Brott *et al*. The scale proposed a 15-item neurologic examination, intended for use in acute stroke therapy trials.

The NIHSS is a non-linear measure that evaluates level of consciousness, language, neglect/inattention, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss. It is scored from 0 (no impairment) to 42, with scores ≥21 usually described as 'severe'. A change of more than 2 points suggests clinically relevant early improvement/deterioration.

Barthel Index: The BI was first described in 1965 by Mahoney and Barthel, and is a 10 item examination that assesses feeding, chair/bed transfer, grooming, toileting, bathing, ambulation, stair climbing, dressing, bowel control and bladder control. The scores range from 0 (dependent) to 100 (independent). A 'good' outcome on the BI has not been fully defined, but often >80 are interpreted as generally independent and usually able to return home, and <40 are very dependent. Other interpretations have suggested that >95 describes an excellent outcome, and <75 describes a poor outcome (Harrison *et al*, 2013).

Modified Rankin Scale: The Rankin Scale was first developed in 1957 to assess the extent of global disability after stroke (Rankin, 1957). The original scale was modified slightly in 1988, and the mRS is a seven point scale ranging from 0 (no symptoms) to 6 (dead). The degrees of disability are described as follows:

0 no symptoms

1 no significant disability; despite symptoms, able to carry out all usual duties and activities

2 slight disability; unable to carry out all previous activities, but able to look after own affairs without assistance

3 moderate disability; requiring some help, but able to walk without assistance

4 moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance

5 severe disability; bedridden, incontinent and requiring constant nursing care and attention

6 dead

Oxford Handicap Score: The OHS is a modification of the mRS:

	Handicap	*Lifestyle*
0	none	no change
1	minor symptoms	no interference
2	minor handicap	some restrictions but able to look after self
3	moderate handicap	significant restriction; unable to lead a totally independent existence (requires some assistance)
4	moderate to severe handicap	Unable to live independently but does not require constant attention
5	severe handicap	Totally dependent, requires constant attention day and night

Glasgow Outcome Score: The GOS allocates patients into broad outcome categories:

1 good recovery – resumption of normal activities even though there may be minor neurological or psychological deficits

2 moderate disability – disabled but independent, patient is independent as far as daily life is concerned. The disabilities found include varying degrees of dysphasia, hemiparesis, ataxia as well as intellectual and memory deficits and personality changes

3 severe disability – conscious but disabled, patient depends on others for daily support due to mental or physical disability or both

4 persistent vegetative state – patient exhibits no obvious cortical function

5 death

5.2 Choice of endpoint scale

Stroke scales have been designed with differing purposes in mind, and measure different aspects of the effect of stroke on the patient. This is described in Harrison *et al* 2013, using the WHO's International Classification of Functioning, Disability and Health as a framework. Stroke scales may measure neurological impairments (direct loss of function), activity limitation (previously called disability) and societal participation (previously called handicap). In addition, there are also measures designed to assess quality of life of the patient. Whilst there are clearly influences and overlaps between these subjects, the focuses of their measurements differ and therefore trial results may vary depending on what scale is deemed to be most appropriate. As is widely discussed, there is no single scale that can provide all relevant information regarding outcome after stroke.

The NIHSS is an example of a scale that measures impairment, or direct loss of function. Advantages of the NIHSS are that it is relatively straightforward and quick to perform (at around 6 minutes), inter-observer reliability has been found to be very good for NIHSS assessment (Goldstein and Samsa, 1997), and training resources are available (online and as DVD) which further improve reliability (Lyden *et al*, 1994).

The NIHSS was designed to be used in acute situations when stroke is initially diagnosed, however baseline severity measured by NIHSS has been demonstrated to be an important predictor of final outcome, at least in terms of 'good' vs. 'bad' outcomes when defined as returning home or remaining in care/dead (Muir *et al*, 1996, Schlegel *et al*, 2003). Kwakkel *et al* (2010) found that NIHSS measured within 72 hours of the onset of stroke was predictive of final outcome at 6 months according to the BI, in a cohort of 159 patients. NIHSS therefore shows predictive validity.

Notably, a cohort study of medical records by Schlegel *et al* (2004) found that the predictive value of the baseline NIHSS was reduced in patients treated with rt-PA especially for patients with moderate stroke (NIHSS 6-15), who were more likely to be discharged to home than to rehabilitation or nursing facility compared with previous studies. This would be expected if rt-PA treatment was having a positive effect. Variability of the final outcome relative to baseline NIHSS was also found to be greater after rt-PA treatment, which may reflect a combination of an increased chance of a good outcome combined with the increased risk of symptomatic ICH (and worse outcome) (Schlegel *et al,* 2004). The NIHSS has also been found to

correlate with objective measures of stroke severity such as infarct size at 7-10 days post stroke (Brott *et al*, 1989).

The limitations of the NIHSS include that there is a tendency for left-hemisphere strokes to be rated higher than right-hemisphere strokes of the same infarction volume (Fink *et al*, 2002; Lyden *et al*, 2004). In addition, the NIHSS is not helpful in the evaluation of infarctions occurring in the brainstem or cerebellum. Patients with such strokes may score low overall NIHSSs whilst their strokes may be disabling or life-threatening. Martin-Schild *et al* (2011) studied all patients presenting at one stroke centre over the course of 5 years with acute cerebral ischaemia and an NIHSS score of 0. Of the 2618 patients with acute cerebral ischaemia, 20 patients had a score of 0 (0.76%). The observed symptoms were truncal ataxia (45% of patients), agitated confusion (10%) and single cases of nystagmus, limb weakness, memory impairment, Horner's syndrome, slow to respond, reduced visual acuity without field cut and tandem gait abnormality. The infarct location was more frequently in the posterior circulation than the anterior, primarily in the cerebellum (32%) and the occipital lobe (16%).

The main limitation of the NIHSS, when used as an outcome measure in clinical trials, is that it is not necessarily representative of the overall functional ability of the patient and is therefore not necessarily reflective of the impact of the stroke on the individual.

The Barthel Index is widely used both within and outside of clinical trials to assess basic activities of daily living in patients with stroke, and is therefore a more functional measure than the NIHSS.

The BI has been extensively studied and its validity (Granger *et al* 1988, Wade and Hewer 1987) and reliability (Duffy *et al* 2013, Shinar *et al* 1987) confirmed. The BI is relatively easy to administer and can be used repeatedly to assess improvements over time.

Whilst the score ranges from 0-100, with 100 classified as 'independent', patients with a score of 100 are not necessarily able to live independently. For example the scale does not take into account any measure of cognition, language, visual function, emotional impairment or pain, all of which may impact on the patient's ability to live independently. Similarly, patients with a score of 0 in an intensive care setting may have significant improvements but still score 0 on the BI. These 'floor and ceiling' effects on the BI mean that this scale is less sensitive to change in condition of patients with very severe or very mild deficits (Kasner 2006; Harrison *et al*, 2013).

The nature of the BI also renders it an inappropriate scale for the assessment of patients in the initial acute phase of stroke. As the majority of patients will be bed-bound at this stage they will initially have very low scores even if the stroke is considered to be minor. Therefore the BI cannot be used to stratify patients by severity in acute stroke trials (Kasner 2006).

The modified Rankin Scale has been described as a measure of global disability with a focus on mobility (Harrison *et al*, 2013). It is widely used both in trials and in clinical practice and has been extensively studied. For studies carried out between 2001 and 2010, mRS has been found to be the most frequently used endpoint measure, and the most frequently used primary outcome (Lees *et al*, ESO outcomes working group, 2012)

The limited range of possible scores (0-6) means that the mRS is likely to be less sensitive to changes in a patient's condition than e.g. the BI or the NIHSS, although it is clear that a single-point shift on the mRS will always be a clinically relevant change. In contrast to the BI, there is a finer grading at the mild/unaffected end of the mRS scale, with options for no symptoms at all (mRS=0), and for no significant

disability - although the patient has some symptoms they are able to carry out all their usual duties/activities (mRS=1). Distinguishing these patients using the BI may not be possible because even patients with the maximum score of 100 are not necessarily able to live independently (Balu, 2009).

Similarly to the BI, the use of mRS during the original hospital admission during the initial acute phase of the stroke is inappropriate, as the patient could not have resumed their usual roles/activities whilst hospitalised.

The validity of mRS has been demonstrated by correlation with measures of stroke pathology (e.g. infarct volume) and agreement with other stroke scales (Harrison *et al*, 2013). The mRS is simple and quick to perform, however whilst strong test-re-test reliability has been reported, inter-observer reliability was found to range from poor to nearly perfect in a meta-analysis by Quinn *et al* (2009) which used a systematic review of all studies that measured mRS reliability. The meta-analysis included 10 studies, all of which were small (median n=47) and of varying methodological quality. Overall, reliability as assessed by this meta-analysis was concluded to be 'moderate', though uncertainty remains, due to the small size of the trials and their methodological flaws. Structured interviews/training have been found to improve inter-observer reliability in some cases (Banks and Marotta, Stroke 2007; Quinn *et al*, Stroke 2009).

The similarity of the OHS with the mRS is such that the advantages and disadvantages are likely to be very similar and the OHS is generally referred to as a variant of the mRS.

The Glasgow Outcome Scale is also a measure of global disability, which includes 5 categories of disability. As such, similar issues regarding the sensitivity of the scale apply to the GOS as to the mRS. The GOS has a key difference with the mRS, which is that there is no distinction between patients with a full recovery and those with mild disability. In addition, some factors that relate to function, social role, cognitive and emotional issues are not measured in the GOS, whereas these are covered by the mRS.

The GOS is widely used and has been demonstrated to be reliable particularly when used with a structured interview, and it has also been demonstrated to be valid when compared with the mRS (Kasner 2006, Brooks *et al* 1986). It is likely that inter-observer reliability will be subject to similar issues as the mRS and would be similarly variable.

5.3 Quality of life

From a patient's perspective, their quality of life is clearly the most important aspect of their stroke recovery. As discussed by Ali *et al* (2013), ideally the assessment tools used in clinical trials will reflect relevant changes to quality of life. This subject is further complicated by the fact that each individual's perception is different and one person's opinion of a 'good' outcome may be very different from another's. Therefore ideally, the assessment scales selected for a trial will reflect changes relating to function/activity that are necessarily measured during the trial, whilst also being as reflective as possible of the patient's quality of life.

Most of the key trials did not include any quality of life outcome scale assessment. ECASS II (Hacke *et al*, 1998[1]) however evaluated the short-form-36 (SF-36) at day 90. This measure uses 36 questions in eight different categories, half of which relate to physical and half to mental health. The total score is on a scale of 0-100, with 100 representing the best health. In ECASS II at day 90, the rt-PA group had a median SF-36 mental score of 49.8, and a median physical score of 38.4. The corresponding results for the placebo group were 48.1 and 36.7. The difference was

very slightly in favour of rt-PA, however these differences were not statistically significant (p=0.18, and p=0.28 respectively).

Data from IST-3 on quality of life have also been published in the form of an abstract (Sandercock *et al* 2013) for rt-PA treated (n=1515) and control (n=1520) patients. Assessments by postal questionnaire or blinded telephone interview were used to determine health related quality of life using the EQ-5D-3L, and its visual analogue scale (VAS). The EQ-5D is a generic quality of life scale measuring five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The VAS records a patient's self-rated health on a vertical analogue scale where the endpoints are 'best imaginable health state' and 'worst imaginable health state'. At 6 months, 27% of each group had died, in the remainder, the mean VAS was significantly higher in the rt-PA group compared with controls (60.7 vs. 57.8, p=0.008). Mean difference in VAS for patients randomised 0-2, 2-3, 3-4, 4-5, 5-6 hours after symptom onset was 8.9, 3.5, 2.0, 0.8, 6.2. For the EQ-5D, rt-PA was associated with reduced adjusted proportional odds of having greater problems with self-care (OR 0.79, 95% CI [0.66-0.94], p=0.008) and usual activities (OR 0.74, 95% CI [0.63-0.88], p=0.001), but not mobility (OR 0.88 95% CI [0.74-1.05]), anxiety/depression (OR 0.94 95% CI 0.79-1.11]) or pain/discomfort (OR1.01 95% CI [0.85-1.20]). These results, although partially positive, should be viewed with caution, given that the overall primary endpoint for this trial was negative, and that the patients that died have not been accounted for in the analysis (see next paragraphs on the quality of life data obtained at 18 months follow-up for the IST-3 trial).

Quality of life measured using the EQ-5D scale was presented in the IST-3 publication detailing the 18 month results (IST-3 collaborative group, 2013). These data were discussed in the May CHM paper (section 4.1.4.2), as follows:

"Another secondary endpoint at 18 months was the EuroQol (EQ-5D) scale which measures health-related quality of life. This was presented as a major focus of the publication detailing the 18 month results with many highly statistically significant differences highlighted.

As always we should be cautious interpreting these results in light of the negative primary endpoint results. However there are additional problems here with the approach to the analysis as described below.

The most extreme result presented was for self-care.

EQ-5D: Self-care at 18 months

	rt-PA (N=1169)	Control (N=1179)
Number analysed	695	689
No problems with self-care	372 (54%)	328 (48%)
Some problems washing or dressing	176 (25%)	191 (28%)
Unable to wash or dress	147 (21%)	170 (25%)
Odds ratio (95% CI)*	1.43 (1.16, 1.78)	
p-value*	p=0.001	
Odds ratio (95% CI)**	1.25 (1.03, 1.53)	
p-value**	p=0.027	

* from logistic regression adjusted for age, NIHS stroke scale score, time from treatment to randomisation and visible infarct on baseline scan
** from unadjusted logistic regression (assessor's calculation)

However the analysis is flawed as it ignores a large part of the cohort, namely those who died before 18 months. For the purposes of this assessment we will have to

focus on the unadjusted analysis as it is not possible to reproduce the adjusted analysis without access to the full data-set.

EQ-5D: Self-care at 18 months

	rt-PA (N=1169)	Control (N=1179)
Number analysed	1103	1103
No problems with self-care	372 (34%)	328 (30%)
Some problems washing or dressing	176 (16%)	191 (17%)
Unable to wash or dress	147 (13%)	170 (15%)
Dead before 18 months	408 (37%)	414 (38%)
Odds ratio (95% CI)**	1.08 (0.93, 1.26)	
p-value**	p=0.318	

** from unadjusted logistic regression (assessor's calculation)

By including the patients who died we get a more appropriate estimate of the proportion of treated patients who might expect to achieve favourable outcomes, and also the statistical significance of the shift across categories is lost, as any shifts seem smaller in the context of the large evenly distributed proportion of deaths.

The improvements in Qol as presented in the paper are not robust to the handling of patients who died in the analysis."

As quality of life measures are infrequently used as primary outcome measures in acute stroke trials, Ali *et al* (2013) evaluated data on the three most commonly used stroke scales (mRS, BI and NIHSS) and their relationship to two quality of life measures, the EQ-5D and two forms of the Stroke Impact Scale (SIS v3.0 and SIS-16). The SIS is a stroke-specific quality of life scale that measures physical problems, memory, emotions, communication, activities of daily living, mobility, participation, hand function and patients' perception of recovery. SIS-16 is based on the physical functioning domains of the SIS. Data from the Virtual International Stroke Trials Archive (VISTA), a repository for anonymised, completed stroke trials, was used to evaluate the strength of association between the quality of life scales and the mRS, BI and NIHSS at 3 months after stroke. Subgroup analyses were conducted to examine possible differences between assessments completed by the subject and by proxy (~22% of responses for the two scales were completed with help from a proxy). A total of 4946 patients were included in the data set.

Ali *et al* found a stronger association between almost all patient-assessed measures of quality of life and the mRS at 3 months. Proxy responses were found to have a stronger association with BI. These data therefore support the use of mRS as the primary outcome measure above the BI and NIHSS in acute stroke trials.

5.4 Guidance

This discussion describes the utility of the different scales used to assess outcomes in patients with stroke, and as previously stated it is clear that no single scale can measure all aspects of stroke impact. The advantages and disadvantages of the different scales have been described, and it is clear that the method of implementation of the selected measure will also impact on the quality of the results.

European regulatory guidance provided in the form of a Points to Consider document on the clinical investigation of medicinal products for the treatment of acute stroke, which was issued in 2001 (subsequent to several key studies having completed/started) advised that "There is currently no ideal single stroke outcome scale available. Indeed, all available outcome scales explore different domains of

recovery and have their limitations. With respect to the heterogeneity of symptoms, severity, and pattern of recovery found in stroke, it is recommended to use a combination of different measurement tools to assess the aforementioned specific domains."

The 'points to consider' document advises that rating scales and instruments to be used in acute stroke trials should be valid, reliable, sensitive to change and as easy and quick to administer as possible. From a regulatory point of view, no specific recommendation is made, and the applicant should justify the choice on the basis of test quality criteria. The guidance notes that if a cut-off is used to define a positive response on the functional or global outcome scales (e.g. BI, mRS, GOS), this should be defined and justified in the study protocol. Such dichotomisation of outcome (positive/negative) is not recommended for neurological assessment scales (e.g. NIHSS), as patients in the same category may be clinically distinct and important information may be lost.

Guidance from the European Committee for Medicinal Products for Human Use (CHMP) on the subject of the clinical investigation of medicinal products for prevention of strokes in patients with non-valvular atrial fibrillation is due to come into effect in December 2014. Whilst this guidance relates to a different area of study in that the occurrence of the stroke event itself is regarded as the outcome to be measured, the guidance also advises that "Final stroke outcome should be assessed at 3-6 months after stroke onset using a validated stroke outcome scale, preferably the widely used modified Rankin scale. A disabling stroke should be defined as a score on the mRS of 3-5, whereas a non-disabling stroke should be defined as a score of 0-2. Other validated stoke outcome scales (e.g. Barthel Index) could be used in sensitivity analyses." The recommendation for the use of the mRS may in part reflect its all-round 'global' analysis of the condition of the patient, as well as its advantages in terms of simplicity, speed of delivery and greater sensitivity in the assessment of patients with mild disability compared with the BI and the GOS.

The European Stroke Organisation Outcomes Working Group (Lees *et al*, 2012) concludes that the mRS is the preferred outcome measure for acute trials and should be assessed at 3 months post-stroke, or later. It is also concluded that although no second measure should be required, correlations with supporting scales may be used to confirm consistency in the direction of effects.

5.5 Clinical considerations when defining positive outcomes in trials

In addition to the selection of the most appropriate stroke scale(s) to be used in the assessment of trial outcome, it is necessary to also prospectively define a 'positive' outcome. In the case of the key clinical trials for rt-PA, the cut-off points used for a positive primary outcome were:

NIHSS – improvement in baseline NIHSS by 4 points (or resolution to 0) [NINDS part I and ATLANTIS part A]; NIHSS 0-1[NINDS part II, ATLANTIS part B]

BI – score of ≥ 95 [NINDS part II]; difference of 15 points [ECASS I]

mRS – score of 0-1 [NINDS part II, ECASS II, ECASS III]; difference of 1 grade [ECASS I]

Glasgow Outcome Scale – score of 1 [NINDS part II]

OHS – score of 0-2 [IST-3]

<u>Analytical approach</u>

Several analytical approaches are possible, a) dichotomised endpoint point for a positive vs. negative outcome b) other groupings – e.g. trichotomisation c) ordinal

analysis looking at overall shifts in outcome scale between the two groups in the study and d) change in score from baseline to the end of study on an individual patient basis (only possible for the NIHSS score for example, as mRS, GOS etc. are not appropriate measurements in the immediate acute phase of stroke).

When a dichotomised endpoint is chosen, it is necessary to justify the threshold for the cut-off of a positive vs. negative outcome. For the mRS, the cut-off has most frequently been defined as mRS 0-1 for a positive outcome, i.e. mRS 2-6 for a negative outcome. From a clinical relevance perspective, arguments for a cut-off of 0-1 would be that this represents an excellent outcome, whereby the patient can return to their pre-stroke life with no changes necessary. However, arguments for a cut-off of 0-2 may also be considered reasonable (as specified in the guidance for a 'non-disabling stroke'), as the patient requires no assistance and would therefore be fully independent.

Stroke severity

When selecting the most appropriate endpoint, consideration may need to be given to the characteristics of the patients enrolled in the trial, for example, patients with very severe strokes are less likely to make a full recovery after treatment than patients with mild-moderate strokes. Therefore setting the threshold for a positive outcome at mRS 0-1 or 0-2 may mean that meaningful clinical improvements in severe stroke are not recognised as such, thereby underestimating the treatment effect. For example, if a treated patient who had experienced a very severe stroke had a final mRS of 3, this would be considered a treatment failure, however if the alternative without treatment would have been mRS of 5, this would still be a significant treatment effect.

Trial design

Another consideration may be the design of the trial, for example the time-window for treatment. As discussed in the STAIR II paper on recommendations for clinical trial evaluation of acute stroke therapies (2001), it may be more appropriate to use for example an endpoint of mRS=0-1 for trials where treatment is administered with a short time to onset (e.g. in NINDS, with a maximum time to treatment onset of 3 hours). For trials with a longer time to treatment onset, it may be unrealistic to have an endpoint of no/minimal deficit, as treatment started later is less likely to be able to completely reverse the neurological deficit (if 'time is brain' is to be believed).

As mentioned above, the CHMP guidance on assessment of strokes occurring in trials of medications for stroke prevention in patients with non-valvular atrial fibrillation recommends dichotomising the mRS at 0-2 for a favourable outcome. The guideline also comments that patients with ischaemic strokes in association with atrial fibrillation who survive are left more disabled by the stroke than patients with other causes of stroke

From a clinical standpoint, the utility of an ordinal analysis, assessing the impact of treatment across the full spectrum of mRS outcomes would appear to be an attractive option, which would solve any issues in terms of underestimating treatment effect for severe stroke patients, or in patients treated at longer time points. Furthermore whilst a dichotomised analysis could show benefit despite an adverse effect at another level of the scale, an ordinal analysis will normally only have a significantly positive result if the overall trend is positive – therefore it protects against inappropriate claims of benefit made in association with harm at other levels (Lees *et al*, ESO outcomes working group, 2012).

Proponents of the ordinal approach also highlight that a shift analysis is the most efficient analytic technique, providing the most information and therefore power

(Saver 2011). However, there are important statistical considerations relating to this approach (see below).

5.6 Statistical considerations when defining positive outcomes in trials

5.6.1 Dichotomous analysis

Ordinal scales such as the mRS are often dichotomised for the purposes of analysis. The dichotomisation could be based on the score observed at endpoint (e.g. percentage of patients with 0 or 1 on the mRS) or by looking at change from baseline (e.g. percentage of patients with at least a 1 point improvement from baseline).

The statistical analysis of a dichotomous endpoint is uncontroversial and does not depend on any particular assumptions.

A dichotomous analysis answers the very specific question of whether the treatment increases the proportion of patients achieving the specified outcome. If the outcome is clinically positive a statistically significant difference between the treatment groups provides evidence of benefit of the treatment.

It can be criticised for discarding information by dichotomisation and therefore lacking power compared to an analysis using the full scale. This is interesting from a design perspective, but from an assessment perspective it does not create any unfair advantage for the treatment and does not complicate the interpretation of results. In the presence of classification errors in the scale the reduction in power from dichotomisation is not as large.

A positive result on a dichotomised endpoint does not rule out that there may be a negative finding in some other aspect of the scale (e.g. if the rate of 0 or 1 is increased but the death rate also increased) and so any assessment must also include a general summary of the whole scale for a full understanding of risks and benefits.

Choice of cut-off in a dichotomous analysis

The choice of dichotomisation point should be based primarily on clinical rather than statistical considerations in order to define a positive outcome for the patient population being studied.

It is important to choose the cut-off to capture the likely benefits of the treatment in the population being studied. For example in a severe population there may not be many patients achieving 0 or 1, but a score of 2 could be a very positive outcome. In this case a criterion of 0-1 would lack power and 0-2 would be a better choice.

Different cut-offs will be subject to different levels of classification error (for the mRS there is high inter-rater variability across the 3-5 range of scores, so classification error is reduced for cut-offs 0, 0-1 or 0-5, and is higher for 0-2, 0-3, 0-4). However the effect of classification error is to increase noise and make it more difficult to detect a difference, not easier, so a positive result is not called into question by this issue.

Provided the choice of cut-off is pre-specified in the protocol, any choice is acceptable statistically, and a positive outcome in the statistical test for a difference between treatment groups provides evidence that the treatment increases the proportion of patients achieving this outcome.

5.6.2 Shift analysis of ordinal scales

A shift analysis has the potential to be a more powerful test for the general question of whether the treatment causes a general shift on the scale. However it is based on

a stronger assumption – the proportional odds assumption – that the odds ratio for the comparison between treatment groups is the same for all possible dichotomisations (e.g. on the mRS these are 0, 0-1, 0-2, 0-3, 0-4 and 0-5), or in other words the chance of the treatment shifting someone who would have scored 5 with no treatment to 4 is the same as shifting from 1 to 0, or from 2 to 1. If this is not the case a shift analysis, which reports a single odds ratio, is not appropriate.

In many trials of rt-PA the proportional odds assumption does not seem to be appropriate, as there is a negligible or even negative effect on mortality or those scoring 5 on the mRS, with a larger positive effect at the 0, 1 end of the score. Figure 1 (below) from De Santis *et al* (2014) illustrates this lack of proportional odds in the NINDS study.

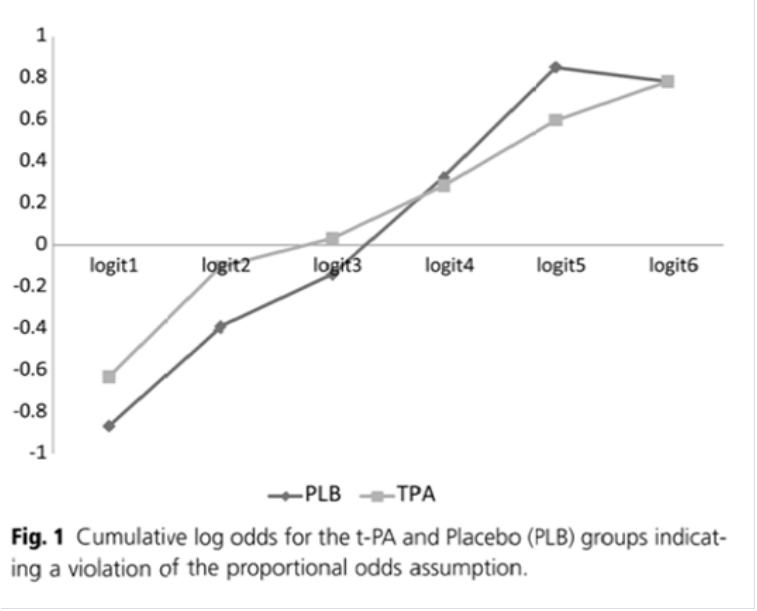


Fig. 1 Cumulative log odds for the t-PA and Placebo (PLB) groups indicating a violation of the proportional odds assumption.

5.6.3 Non-parametric rank-based analyses

Rank based analyses can answer the general question of whether patients on treatment generally have a better outcome than those on control. They are useful tests for generally establishing, with minimal assumptions, that there is an effect of treatment, but the lack of summary statistics means that additional tests and data summaries are required to establish the size of benefit and the nature of the benefit (e.g. which categories are the improvements seen in). They will not be sensitive in a situation where there is improvement at one end of the scale with a worsening at the other and are more sensitive when the effect of treatment is a general shift in one direction.

5.6.4 Continuous analyses

Analysis could be done comparing the average scores in the two treatment groups. Such analyses depend on the assumption that the scores follow a normal distribution and that the effect of treatment is a general shift in a particular direction (as noted above that is not obviously a correct assumption for rt-PA). In addition, to interpret the difference between treatment groups, the meaning of a particular value, e.g. 1 point, would need to be understood. This is not clearly possible, as the clinical meaning of shifting from say 0 to 1 is not the same as going from 3 to 4, or, more starkly, from 5 to 6 (death). Therefore the results from continuous analysis are difficult to interpret.

5.7 Evidence for impact of choice of endpoint on outcome of studies

ECASS I (Hacke et al, 1995): This was a randomised, double-blind placebo controlled trial in patients with moderate to severe neurological deficit and with no or minimal early infarct signs on initial CT scan. A total of 610 patients were enrolled, and a dose of 1.1 mg/kg body weight of rt-PA was employed, with a time-window for treatment of up to 6 hours post-symptom onset.

The primary outcomes as specified in the trial were 1) a difference between rt-PA treated and placebo treated patients in activities of daily living defined as a difference of 15 points in the BI at 90 days post-treatment and 2) a difference between rt-PA treated and placebo treated patients of one grade in the mRS at 90 days post-treatment.

In the intention to treat analysis, there was no significant difference between the groups for either BI or mRS:

Median BI score: placebo 75; rt-PA 85 (p=0.99)

Median mRS score: placebo 3; rt-PA 3 (p=0.41)

In a subsequently published paper (Hacke *et al* 1998[2]), the ECASS Study group presents a post-hoc analysis of the ECASS I data, using the intention to treat data set. This analysis used the NINDS trial analysis methodology, to re-analyse the mRS, BI and NIHSS scores dichotomised according to the NINDS statistical methodology. A favourable outcome was defined as mRS 0-1, BI 95-100, NIHSS 0-1. A global end-point was also evaluated, combining mRS, BI and NIHSS; the GOS was not evaluated in ECASS I. The global end-point used in NINDS was a combination of mRS, BI, NIHSS and GOS.

The following table provides the results obtained for this post-hoc analysis, together with the corresponding results obtained in NINDS:

Global End-Point Analysis, Dichotomized Single End-Point Analysis, and Effect Sizes (ECASS ITT and NINDS Trial)

	ECASS ITT				NINDS Trial, Part II*	
	Placebo n=305	rtPA n=310	OR (95% CI)	*P*	OR (95% CI)	*P*
Global end-point statistics			1.5 (1.1–2.0)	0.008	1.7 (1.2–2.6)	0.008
mRS score of 0 or 1	86 (28%)	111 (36%)	1.4 (1.0–2.0)	0.044	1.7 (1.1–2.6)	0.019
Effect size			+8%		+13%	
BI score of 95 or 100	115 (38%)	137 (44%)	1.3 (0.9–1.8)	0.102	1.6 (1.1–2.5)	0.026
Effect size			+6%		+12%	
NIHSS score of 0 to 1	68 (22%)	111 (36%)	1.9 (1.4–2.9)	0.001	1.7 (1.0–2.8)	0.033
Effect size			+14%		+11%	
GOS score of 0 or 1	...	...	...	...	1.4 (1.0–1.8)	0.025

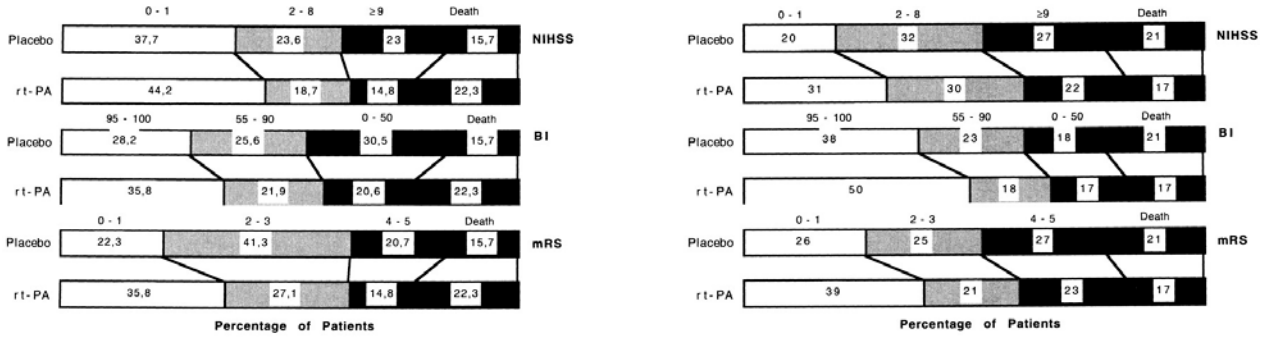
*For further details concerning the NINDS trial, see Reference 2.

The results obtained in ECASS I, when analysed in accordance to the NINDS trial methodology suggest an overall positive result for the global analysis which included mRS, BI and NIHSS. The OR obtained is statistically significantly in favour of rt-PA treatment, and similar to the NINDS trial results. The individual scales, mRS and NIHSS both demonstrated a significant result in favour of rt-PA when analysed according to the NINDS dichotomisation, however the BI only demonstrated a

positive trend. The absolute effects for the three scales were lower than for NINDS for the mRS and BI, but greater for NIHSS.

The results of ECASS I when analysed according to the NINDS methodology present a much more favourable picture of the impact of rt-PA treatment compared with placebo than that obtained according to the original analysis, and the results obtained from these two trials would appear to be consistent with each other.

However, further details of the trial outcomes are obtained from the comparison of the full results across the scales for ECASS I (on the left) and for NINDS (on the right) and are provided by the following diagrams:



The bar diagrams indicate that in the NINDS trial, generally the overall picture for all three scales is a shift towards a more positive result in the rt-PA group compared with placebo. This is not quite as clear cut for the ECASS I trial, where although the best outcomes are increased in the rt-PA group compared with the placebo group, similarly the worst outcome (death) is also increased. It is worth bearing in mind that the time window for treatment in ECASS I was 3 hours longer than for NINDS, with enrolment up to 6 hours post-symptom onset. In addition, the dose of rt-PA employed was higher than that used in NINDS and other clinical trials (1.1 mg/kg body weight, instead of 0.9 mg/kg body weight).

The re-analysis of the ECASS I trial illustrates the impact that different analysis methods can have on the outcome of acute stroke trials. One further consideration is that several of the authors of this re-analysis are employees or connected with Boehringer Ingelheim, as might be expected.

Clearly the use of different analyses or endpoints can lead to different conclusions on whether a trial is “positive”. This should not be seen as a concerning finding – it is more that different endpoints actually provide the answers to different questions. The initial question which ECASS I was set up to answer was whether there was a general improvement in BI and mRS. This was negative and in fact the median score was identical for rt-PA and placebo on the mRS. The subsequent test on mRS, looking at the proportion of patients achieving 0 or 1 asked a different question, whether rt-PA increased the proportion of patients in the better categories, which it seemed to do (although as this was not a pre-specified analysis the finding had to be treated with caution and would need prospective confirmation in subsequent trials).

Neither test was wrong – the pattern of data was that rt-PA increased the number of patients in the good outcomes but at the cost of additional deaths (note that a higher dose of 1.1mg was used and the time-window was up to 6 hours). Therefore it did not induce a general positive shift in outcomes (original question) but did improve the number with good outcomes.

A logical approach in this situation is to use these data to learn about the likely or desired profile of the rt-PA effect and target future studies to be more sensitive to detect that effect (provided this expected efficacy profile is considered to provide a useful benefit). The rt-PA programme did this, changing the dose and the patient

population (time-window) but also changing the endpoint to look for improvement in the good categories rather than a general shift.

Provided future studies using a new primary endpoint are positive it does provide supportive information if past trials also have good information on this new endpoint. The strong classification of trials into “positive” and “negative” is unnecessarily reductive.

It is important in a clinical trial to clearly pre-specify the objectives of a trial (to avoid retrospective focus onto positive endpoints) and to choose an endpoint that will address that specific objective (e.g. does the treatment increase the number of patients in the best categories). The main trial objective should be chosen bearing in mind the likely effect of the treatment in the population in question, however assessment should always take into account the full profile of the data when evaluating the risk-benefit (e.g. are we prepared to accept a small increase in the worse categories to gain an increase in the better ones.)

Bath *et al* (2012) also consider the merits of different endpoints/analysis methods in stroke trials, similarly concluding “Because there is no best approach that will work for all acute stroke trials, it is vital that studies are designed with a full understanding of the type of patients to be enrolled (in particular their case mix, which will be critically dependent on their age and severity), the potential mechanism by which the intervention works (i.e., will it tend to move all patients somewhat, or some patients a lot, and is a common hazard present)”.

Savitz *et al* (2007) compared the shift analysis with the 0-1 and 0-2 dichotomous analyses in the ECASS 2 and NINDS trials. In both trials the 0-1 dichotomy was originally specified as primary). We see that the shift analysis using the full scale did not generally exhibit increased power compared to dichotomisation. NINDS was seen to be positive regardless of the choice of primary analysis. ECASS-2 could have been positive if a 0-2 dichotomisation or the shift test had been used, rather than the 0-1 dichotomy.

Method of Analysis of NINDS Data	Outcome Measure	Choice of Strata for Baseline NIHSS	NINDS Set of All Patients, Range of NIHSS 0–40		
			OR	95% CI	*P* Value
CMH shift test	mRS	No strata	1.60	1.21–2.11	0.001
Logistic regression	Binary mRS 0–2 vs 3–6	No strata	1.62	1.18–2.23	0.003
Logistic regression	Binary mRS 0–1 vs 2–6	No strata	2.05	1.46–2.87	0.001

Method of Analysis of ECASS Data	Outcome Measure	Choice of Strata for Baseline NIHSS	ECASS Set of All Patients, Range of NIHSS 0–40		
			OR	95% CI	*P* Value
CMH shift test	mRS	No strata	1.32	1.02–1.71	0.04
Logistic regression	Binary mRS 0–2 vs 3–6	No strata	1.49	1.11–2.00	0.008
Logistic regression	Binary mRS 0–1 vs 2–6	No strata	1.20	0.89–1.61	0.24

ESASS-2 was also re-analysed by Stingele *et al* (2001) using a rank based bootstrap method, giving p=0.047. It is interesting that this is similar to the result from the shift test, given that both tests look for a general shift.

While it is not appropriate to retrospectively change the primary analysis of this study and these positive results should not be over-interpreted, they do provide supportive evidence that there may be some positive effect of treatment, and it is also reassuring in this context that the original primary endpoint did at least trend in the

positive direction (note most patients in ECASS-2 were treated in the 3-6 hour window rather than 0-3 hours – maybe this is the reason why hoping for a shift to 0 or 1 may have been too optimistic).

6. Submission by Dr Pitchaiah Mandava

Dr Mandava contacted MHRA and Dr [redacted] after reading the [redacted] correspondence, and provided two pieces he has co-authored - an article published in Plos One and a book chapter (previously provided to the group). These relate to the analysis of clinical trials in stroke and the use of ordinal endpoints for analysis. The article and book chapter are reviewed in this section.

6.1 Review of "A critical review of stroke trial analytical methodology: outcome measures, study design, and correction for imbalances" by Mandava *et al* (2012) in the context of the clinical trial data supporting the use of rt-PA

Sections 1, 2 and 3 are of relevance to our deliberations.

Section 2 addresses the issue of baseline imbalance and how this should be addressed. It criticises the use of baseline adjusted analyses for imbalances in baseline characteristics such as NIHSS as such endpoints don't satisfy the assumptions implicit in such analyses, such as the endpoints being normally distributed and there being a linear relationship between the baseline variable and the outcome. They illustrate this concern by showing the double peak distribution of the NIHSS data in NINDS, and the non-linear relationship between probability of a good outcome and baseline NIHSS in the TOAST data.

Figure 40.1 does not seem to show a large departure from the normal distribution for the baseline NIHSS score and figure 40.3 does show a large departure from linearity. Baseline adjusted analyses are fairly robust to assumptions and it is not clear that they would not be useful in this situation, although more than one analysis is preferable to show the robustness of any positive conclusions and a baseline adjusted analysis should not be used alone for this task.

In the introduction (section 1) they look specifically at the baseline imbalance in the NINDS trial. They conclude that rt-PA is an effective treatment, but that the benefit was modestly over-stated because of the baseline imbalances.

Despite their criticisms of the baseline adjusted analysis their own re-analysis using a matching algorithm where each patient on active is matched by the most similar baseline subject in the placebo arm confirms the existence of a treatment effect (Mandava *et al,* 2010). Looking at Table 1 in Mandava *et al* 2010 (see below) we see the original analysis had 42% of patients on rt-PA with 0-1 on the mRS compared to 27% on placebo. After matching, this became 39% vs. 26%. The small difference from the original analysis gives further reassurance that the conclusion of a positive benefit in NINDS is robust to the baseline imbalance. The findings on the other efficacy endpoints were similarly robust, and use of alternative matching algorithms also led to positive results.

Table 1. Matching Based on Elimination of Outliers

	Prematch			Postmatch		
	rtPA (n=312)	Placebo (n=312)	*P*	rtPA (n=283)	Placebo (n=283)	*P*
Median NIHSS	14	15	0.09	14	14	0.61
NIHSS, mean±SD	14.4±7.5	15.2±6.8	0.14	14.4±7.3	14.7±6.7	0.68
Age, years, mean±SD	68.0±11.3	65.9±11.9	0.03	67.9±11.4	66.5±11.8	0.15
Glucose, mean±SD	149±70.7	151±77.9	0.78	144±67.4	145±66.6	0.90
Proportion mRS 0–1	0.42	0.27	<0.001	0.39	0.26	<0.001*
Proportion BI 95–100	0.51	0.38	0.002	0.52	0.40	0.04*
Proportion GOS of 1	0.45	0.31	<0.001	0.45	0.32	<0.001*
Proportion NIHSS ≤1	0.34	0.21	<0.001	0.34	0.21	<0.001*
Mortality	0.17	0.21	0.31	0.17	0.20	0.27*

*McNemar test for matched case–control studies.

Section 3 talks about the modified Rankin Score specifically and notes that the scale is widely used but has some weaknesses. They note that while ratings of 0, 1, 5 and 6 are not usually in question, the other scores are subject to substantial inter-rater variability. This calls into question analyses using the complete mRS also known as shift analyses.

This information would support the use of the 0 or 1 responder criteria as used in NINDS and ECASS-3.

Implications for the study data supplied for rt-PA

This review is generally supportive of the data on which the approval of rt-PA was based. The choice of a dichotomised endpoint is supported and the particular choice of 0-1 on the mRS is endorsed.

In considering the issue of the baseline imbalance in NIHSS scores in the NINDS trial the use of standard baseline adjusted analyses is criticised, however additional analyses based on matching algorithms confirmed the robustness of the positive findings in that trial. This adds to the range of analyses of the NINDS trial data that are already supplied in the O'Fallon *et al* 2004 independent re-analysis.

6.2 Review of "Quantification of Errors in Ordinal Outcome Scales Using Shannon Entropy: Effect on Sample Size Calculations" by Mandava *et al* (2013) in the context of the clinical trial data supporting the use of rt-PA

The paper considers the inter-rater variability in assessment of the modified Rankin Score (mRS) and uses that to look at the likely error rates in terms of patients being attributed with incorrect scores. It then looks at the performance of analyses using the full scale and those using dichotomisations of the scale in the presence of these errors to see whether the loss of information from dichotomisation may be compensated for by a reduction in error rates.

Conclusions from the paper

The paper concludes that misclassification error rates for mRS are lower when looking at dichotomous outcomes (such as percentage of patients scoring 0 or 1 vs. 2-6) rather than looking at the full scale.

This observation is not surprising, and has to be true. When looking at the full scale any misclassification of a patient's score is counted as an error, this includes shifts

from 0 to 1, 1 to 2, 2 to 3 etc. After dichotomisation a shift is only classified as an error if it results in a change of category. So for the percentage of patients scoring 0 or 1 endpoint, only shifts from 0 or 1 to 2, 3, 4 or 5 would be errors, shifts from 0 to 1 or 2 to 3 are no longer counted. The list of errors for any dichotomisation is simply a sub-set of the list of errors on the full scale. Therefore it is not an important or unexpected finding that the error rate is reduced by dichotomisation. The concern is more what effect the errors may have on the analysis.

The impact of the classification errors was found to be that higher sample sizes would be required to detect an effect than if there were no errors.

This is again unsurprising. The introduction of noise into any system makes a signal more difficult to detect. It therefore follows that the existence of classification errors would make it more difficult to identify an effective treatment. From the perspective of licensing of medicines this means that while classification errors may reduce the efficiency of a trial and increase the false negative rate (increasing the chances an effective treatment is declared ineffective) there is no suggestion that the false positive rate (chance an ineffective drug is declared effective) is increased. While the former is sub-optimal, it is the latter that would be the real regulatory concern.

The authors concluded that while there is loss of information from dichotomising (meaning larger sample sizes are required) it is not clear that this advantage from using the wider range will always overcome the noise it appears to generate (from the increased errors).

No conclusion is reached on whether using the full scale or the dichotomisation is the most efficient strategy in the presence of classification errors. It is only concluded that there is not a clear-cut answer to this question. If there were no errors the analysis using the full scale would be more efficient.

Of the various cut-off points investigated the lowest error rate was seen with 0-1 response definition as opposed to other cut-offs. This is consistent with their statement that “the inter-rater reliability of mRS is relatively low, particularly for mid-range (mRS score of 2-4) values”.

This suggests that if dichotomisation is done, the 0-1 cut-off might be the best choice. It seems raters are more consistent in distinguishing between no significant disability and slight disability than they are between slight disability and moderate disability.

Implications for the study data supplied for rt-PA

The main studies upon which the approval was based, NINDS part 2 and ECASS-3 both used the dichotomous outcome of percentage of patients achieving a score of 0 or 1 on the mRS. While the paper does not go so far as to endorse this as the clearly optimal approach, if we consider these models of the likely classification error rates it seems to be the best choice if a dichotomous analysis is to be done, and is not clearly less efficient than an analysis using the full scale. Importantly there is no suggestion that any bias in favour of rt-PA is introduced by the choice of analysis.

7. Submission by Professors Daniel Fatovich and Simon Brown

This section reproduces in its entirety Professors Fatovich’s and Brown’s submission, with assessor’s comments interspersed. Please see the full submission (previously provided) for the details of the references cited by the authors in this section.

Summary of key points

1. The evidence for serious harm (premature death due to intracerebral haemorrhage) from thrombolysis in acute stroke is unequivocal.

2. With regard to any potential benefit from thrombolysis to counterbalance the harms, most studies are *negative*. Three studies are claimed to show an overall benefit from thrombolysis, but there are major issues with each of these:

a. The NINDS study was methodologically and analytically flawed. It was inappropriate to claim proof of benefit on the basis of subsequent re-analysis of data, especially when different approaches/authors supported opposite conclusions.

b. The ECASS-3 analysis was also flawed and the results have been recently altered.

c. Baseline imbalance in stroke severity was most likely responsible for the apparent benefit in both NINDS and ECASS-3.

d. The largest study to date (IST-3) was an open label study with major biases. Despite these biases, the primary outcome was negative and there was very clear evidence of serious harm. Yet benefit continues to be claimed on the basis of one subjective secondary outcome.

3. It is inappropriate to claim that a meta-analysis can overcome these serious flaws in the underlying studies.

4. Registry studies, also used to claim proof of benefit, are full of selection bias and are, at best, only hypothesis-generating.

5. Most analyses of thrombolysis as well as endovascular (neurointerventional) studies indicate that time is NOT brain - i.e. treatment outcomes are not time-dependent. This raises further doubts about there being any plausible mechanism for a benefit from thrombolysis.

6. Further research is urgently required to establish the role, if any, for thrombolysis in acute ischaemic stroke. Until this is done, the use of thrombolysis for stroke must be restricted to properly-conducted (randomised, double-blind, placebo-controlled) clinical trials.

7.1 What is the harm?

The main harm is intracerebral haemorrhage (ICH). This appears to be consistently around 5%. The NINDS trial, the only study with a significant number of patients treated in the 0-3 hours after stroke onset, reported an absolute difference of 5.8% (i.e. 6.4% with tPA vs 0.6% placebo). This presents a major harm of early death with tPA, is statistically highly significant, and has been confirmed in all studies and meta-analyses.

Assessors' comment
The negative effect of rt-PA on early death is not disputed.

7.2 Does tPA in stroke provide benefit - i.e. an improved functional outcome?

Among the 12 commonly cited trials of thrombolysis for stroke, two suggest a statistical benefit on favorable outcomes with tPA[1]. Hence, ten were negative for the primary outcome, and of these, four were stopped early for harm (i.e. increased

mortality). This pattern is typical for a treatment that does not work[2]. The only supposedly positive studies (i.e. they found a statistical benefit) were NINDS-2 and ECASS-3, both of which had significant imbalances in baseline stroke severity (allocation bias) favouring the tPA groups, which could explain the entire statistical benefit[3].

Assessors' comment
This discussion ignores the differences in time-window, dosages, target endpoints and patient populations looked at in the various rt-PA studies and the reasons why the populations were modified after different sets of results were seen in order to identify the population most likely to benefit from treatment.
The two NINDS studies assessed whether a dose of 0.9 mg/kg was efficacious in the first 3 hours after onset of stroke. NINDS part 1 was designed to look at whether rt-PA could provide early improvement for patients (after 24 hours), and no statistically significant differences between rt-PA and placebo were seen at the early time-points. However *post hoc* assessments of the Day 90 endpoints revealed statistically significant improvements for patients on rt-PA and NINDS part 2 was planned to provide prospective confirmation of these findings. This approach of prospective confirmation of promising post hoc results from exploratory endpoints is considered to be a rational approach to a clinical investigation plan.
ECASS I was conducted concurrently with the NINDS trials and demonstrated that a dose of 1.1 mg/kg was too high (resulting in a higher frequency of ICH and increased mortality in the rt-PA group compared with placebo at day 90 (22.4% vs. 15.8%). ECASS II, conducted after the publication of the NINDS trials and ECASS I, used a dose of 0.9 mg/kg and generated the hypothesis that treatment up to 6 hours after onset of symptoms might be too long for treatment to be of benefit.

In addition, the ECASS-3 results change according to the chosen primary endpoint (mRS 0-1 *vs* mRS 2-6). In the original analysis, a mRS of 2 is grouped with a mRS of 6 (dead). This is not an appropriate endpoint stratification[4]. Once the endpoint is reclassified into mRS 0-2 *vs* 3-6, all purported benefits of tPA disappear[5].

Assessors' comment
This is not a fair and accurate summary of the ECASS-3 results. There is no clear ideological preference for the 0-1 or 0-2 definition of mRS response. In the first, responders are those with no significant disability or better. In the second it is those with slight disability or better. In both cases death is considered a negative outcome along with moderate to severe disability – the only difference is the classification of slight disability as a failure (in the 0-1 definition) or success (in the 0-2). Both choices are possible to justify clinically and for this study the 0-1 definition was pre-specified as primary. The results were 52.4% response on rt-PA compared with 45.2% on placebo, OR=1.34, p=0.04. Using the 0-2 definition we see 66.5% vs. 61.6%, OR=1.30, p=0.11. While statistical significance is not seen for the 0-2 definition it is not fair to say that "all purported benefit disappears." In fact the odds ratios are very similar and the trend still clearly favours rt-PA.

It is also noteworthy that the ATLANTIS trial, which had virtually the same design as ECASS-3, caused harm and was stopped early.

Assessors' comment:
ATLANTIS was stopped early because of concerns regarding patients treated in the 5-6 hour window, an unlicensed indication.

More recently, the integrity of the ECASS-3 result has been challenged, with publication of an altered result demonstrating clear allocation bias favouring the tPA group[6,7].

Assessors' comment:
The baseline imbalance in stroke severity was investigated as part of the regulatory assessment when the extended time-window to 4.5 hours was approved. Baseline imbalances for the number of high NIHSS score patients were observed, however, a post hoc sensitivity analysis as well as additionally requested modelling for outcome to address the influence of baseline factors were considered to have corroborated the primary analysis. Regarding the numerical imbalances of the NIHSS at baseline a sensitivity analysis was conducted that analysed the 3 month stroke outcome by baseline stroke severity to exclude the possibility that the results of the ECASS III trial are driven by baseline imbalances of the NIHSS. There were slightly more patients with mild strokes (NIHSS ≤ 9) and fewer patients with very severe strokes (NIHSS ≥ 20) in the rt-PA group compared with placebo at baseline. However, even if these two categories of stroke severity were ignored, efficacy outcomes were still in favour of rt-PA with absolute benefits of 3.3% and 5.8% on favourable outcomes defined as mRS score of 0-1 and NIHSS score of 0-1, respectively. Differences between all matched subgroups were consistently in favour for rt-PA.

The NINDS-1 trial tested neurologic improvement at 24 hours and found no benefit. NINDS-2 then sought "a difference of 20 percentage points" at 90 days between groups, and it is unclear whether this meant a difference between groups in degree of improvement (i.e. the outcome in NINDS-1) or in 'the chance of a good outcome'. The paper reported only the latter. This is unfortunate because subjects in the thrombolytic arm experienced milder strokes than those in the placebo arm. Patients with milder strokes are obviously more likely to achieve a 'good outcome', making degree of improvement a much better comparison. Ultimately NINDS-2 reported that 12% more subjects experienced 'good' outcomes in the thrombolytic group. Despite not being close to the 20% goal, this was reported as a statistically significant difference. The choice not to analyze the data as they had in the first study remains unexplained and non-intuitive."[1,3]

Assessors' comment:
The rationale for analysing a different endpoint in NINDS part 2 than NINDS part 1 is discussed above, and was based upon the objective of confirming the positive *post hoc* 90 day data seen in NINDS part1. The choice not to analyse the data as in the first study is not considered non-intuitive as the first study had suggested a lack of benefit for rt-PA at the early time-points and the second study was set up with a different primary objective. Statistical significance is an assessment of whether an observed difference is likely to be a chance finding – so the result that a 12% difference was statistically significant is not a contradiction to the study objective being a difference of 20%. An observed difference can be of statistical significance (and therefore likely to be a real difference) and then may or may not be considered to be of clinical importance (in this case it was) regardless of whether it is larger or smaller than what was hoped for in the study planning. The imbalance in baseline stroke severity in the trial and the use of the difference between groups in degree of improvement (i.e. change from baseline) as an endpoint intended to adjust for imbalance in baseline stroke severity will be discussed later.

NINDS-1 was originally designed as a 280 patient study with 24-hour endpoints. This trial was sponsored and managed by Genentech, the patent-holder for tPA[8,9]. The primary endpoint of the original study showed no difference in stroke severity at 24

hrs[10]. However, there were more symptomatic intracranial haemorrhages in the tPA arm (8/144 (5.6%) vs 0/147 (0%), p=0.007), four of which were fatal. Instead of publishing this result, and with this information withheld from the study investigators, post-hoc analyses were conducted to find an outcome "that was more informative of clinical benefit" and as a result a second study with a further 300 patients and different endpoints was added on (NINDS-2) [9]. The eventual presentation of the combined results of the two studies to facilitate licensing of tPA for stroke occurred after collaboration between the investigators, Genentech and the FDA. It is known that the baseline imbalance in stroke severity in this study was concealed in the original report and for many years, despite being flagged by the FDA as a potential source of bias[8-10]. This was confirmed in the 2009 re-analysis of NINDS[11].

Assessors' comment:
The circumstances of the FDA licensing cannot be commented on, but the European assessment report specifically considers NINDS- part 1 and NINDS- part 2 as separate studies, as is appropriate when a second study is used to independently confirm promising *post hoc* results from a first study.

Imbalance in baseline stroke severity is a recurring problem in supposedly 'positive' tPA studies, with the placebo group also having more severe strokes in the NINDS study and in ECASS 3[6]. So it is known that systematic bias occurs in RCTs. *A clinical trial has internal validity if and only if inequalities between groups, bias in assessment of outcome, and chance, have been excluded as possible explanations for the observed difference in outcome.*

The IST-3 trial, the largest stroke thrombolysis trial with no imbalance in baseline stroke severity, was negative on its primary outcome[12]. This study was crippled by its biased open-label design that favoured tPA[13,14]. The fact that it was a negative study is telling. However, a recent updated Cochrane review and meta-analysis that incorporates the IST-3 study (thus with nearly half the cases analysed coming from this unblinded low quality/high risk of bias trial and the reminder of cases mainly from two studies with significant allocation bias as outlined above) illogically concludes a proven benefit from tPA and furthermore promotes expansion of the time window[15,16]. As a colleague of ours has said "*simply put, that's where the astute reader ought to stop reading this publication*"[17].

Assessors' comment:
The IST-3 trial is negative and the associated publication does not appear to emphasise this, focussing more on certain positive secondary outcomes.
The patient population recruited in this trial falls outside the licensed indication and it cannot be inferred that the negative result from IST-3 means that all other trials would be negative if there had been no baseline imbalance. The issue of imbalance in NINDS and ECASS-3 is more appropriately dealt with by consideration of the data from those trials and not data from a separate trial in a different population.

Another inconsistency in the available data is that claimed improvements in functional benefit do not seem to be reflected by even a trend towards overall mortality benefit in the longer term, as might be expected given the exuberant claims of benefit by many authors. Indeed, the most recent meta-analyses find the reverse - a 1.4% absolute increase in mortality at 90 days overall, with the 95%CI very close to significance (0.99-1.25) [15], and a 90 day mortality excess in the tPA group that is statistically significant for patients treated 3-6 hrs after stroke onset[16]. Another important point to note is that even if we accept an overall benefit justifying higher mortality from treating patients with tPA within 3 hours of stroke onset, these highly

biased datasets provide compelling evidence *against* expanding the time window for tPA treatment beyond three hours from stroke onset. The OR for a favourable outcome in the 0-3 hr window is 1.53 (95%CI 1.26-1.86), but the OR for a favourable outcome 3-6 hrs after stroke onset is 1.07 (95%CI 0.96-1.20) [16].

Assessors' comment:
The decision to licence a treatment depends upon the risk-benefit balance, and it is a matter of judgement whether other benefits could be considered to outweigh a disadvantage in mortality if one were seen (in the appropriate patient population). The time window for treatment in the EU has not been expanded to 6 hours. The expansion of the time-window to 4.5 hours was based upon ECASS-3 wherein patients treated from 3-4.5 hours (the vast majority treated after 3.5 hours) the odds ratio for a favourable outcome was on mRS (a scale which includes mortality) 1.34 (95%CI 1.02, 1.76). There was also no evidence of worse day 90 mortality (6.7% rt-PA vs. 8.2% placebo).

The re-analyses of NINDS data

In 2004, following major ongoing concerns with the NINDS trial and its analysis, Ingall *et al*[18] published a re-analysis of the trial data. The authors claimed that there was no evidence that the imbalance in the distribution of baseline stroke severity had an effect on trial results. However, given the small number of patients overall, there is essentially no power whatsoever to estimate statistically the impact of such confounding (from the imbalance in baseline stroke severity) in multiple tiny subgroups[19]. Thus, the Ingall *et al* reanalysis merely found no proof of confounding, a foregone conclusion, rather than strong evidence of its absence.

Assessors' comment:
The purpose of the independent 2004 re-analysis was "to address whether there is concern that eligible stroke patients may not benefit from rt-PA given according to the protocol used in the trials and, whether the subgroup imbalance (in baseline stroke severity) invalidates the entire trial as claimed by some of the critics." They concluded "we found that, despite an increased incidence of symptomatic intracerebral hemorrhage in t-PA treated patients and subgroup imbalances in baseline stroke severity, when the drug was administered according to the study protocol, there was a statistically significant, and clinically important, benefit of t-PA treatment measured by an adjusted t-PA to placebo odds ratio of 2.1 (95% CI: 1.5-2.9) for a favorable outcome at three months." This is a much stronger conclusion than "no proof of confounding". They conclude that efficacy was demonstrated even when the baseline imbalance is accounted for. This is the most important conclusion and can be reached even if there were evidence of confounding (which they did not find).

A later graphic reanalysis of the original NINDS data came to a very different conclusion - indicating that imbalance in baseline stroke severity was likely responsible for most, if not all, the difference in outcome between treatment groups[11]. Each group experienced virtually identical change in National Institutes of Health Stroke Scale (delta-NIHSS). Delta-NIHSS improved slightly in almost all NINDS subjects, regardless of treatment, whereas a very few improved a lot, and a few (most of whom were extremely sick at baseline) died.

Assessors' comment:
The graphical analysis referred to is in the paper by Hoffman and Schriger (2009). The figure below (figure 1) excerpted from that paper is the basis for the conclusion

that the imbalance in baseline stroke severity is responsible for most of the difference in outcome between treatment groups. The top part of the figure shows the baseline NIHSS data illustrating the imbalance, with a lower median in the rt-PA group, although the range is actually wider in the rt-PA group which includes more extremely severe patients. Next we see the 90-day NIHSS data where there seems to be an increased difference compared to that seen at baseline. However if change from baseline is plotted the distributions seem quite similar, prompting the authors' conclusion that some or all of the difference at day 90 may be due to confounding.

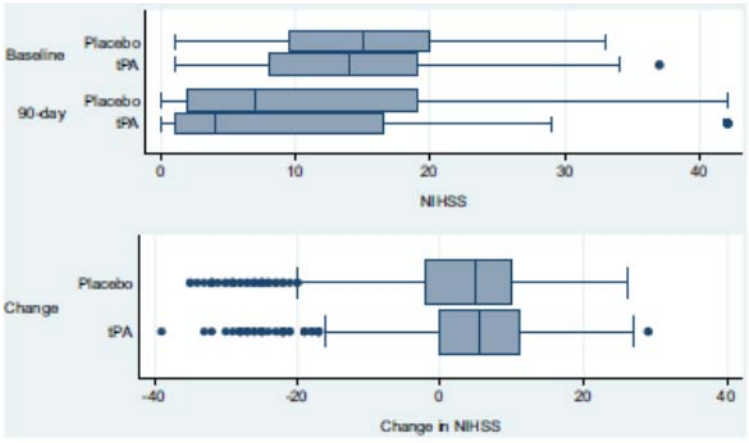


Figure 1

Delta-NIHSS, the only metric recorded both before and after treatment, thus allows for estimation of the effect of treatment independent of confounding by severity. Although more tPA patients did end up with only a small deficit, this paralleled the similarly greater number of tPA patients who started with a very mild stroke.

Assessors' comment:
The authors contend that only a metric recorded before and after treatment allows for estimation of the effect of treatment independent of confounding by severity. However, if there is confounding by baseline severity, a change from baseline measure is also confounded. For example, it might be that patients with mild baseline severity have less room for improvement, in which case a change from baseline analysis would be biased against the group which had milder severity at baseline. Analyses to account for baseline imbalance should attempt to look at the treatment effect while equalising in some way the baseline severity. A simple change from baseline analysis does not achieve this. Possibilities include a covariate adjusted analysis (as performed in the independent 2004 re-analysis) and plots of individual results by baseline severity (as performed in this graphic reanalysis and referenced below).

Graphic depiction of the response to treatment of every single NINDS subject allows readers to judge for themselves whether there was even a hint that treatment modality, or time to treatment, had any independent effect on outcome above that of initial stroke severity[11].

Critics correctly note that delta-NIHSS is not a perfect metric, because it is not truly linear: a change of X points at one end of the scale is not necessarily the same as a similar change at the other end. However, this concern, although theoretically valid, does not appear to be relevant to NINDS, as delta-NIHSS was the same for all the treatment arms at every area of the scale – it changed equally with (early or later)

tPA as with placebo for small strokes, and for moderate ones and for severe strokes[11].

Change in NIHSS also measures only discrete elements of neurological function, rather than the more important overall function of the organism. Still, we might at the very least ask 'just how did tPA lead to better overall outcomes . . . if it had no effect on any element of neurological function?'

A final important point with regard to NINDS is the mere fact that these two re-analyses reaching opposite conclusions emphasizes that post-hoc analyses prove nothing because the method (rather than the actual data) determines the result. The only sensible conclusion to be reached is that NINDS – the only positive RCT for tPA given 0-3 hours after stroke onset - was critically flawed and therefore requires replication, not reanalysis.

Assessors' comment:
The 5 graphs below in figures 2 and 3 (also taken from Hoffman and Schriger) are a far better way to look at the issue of confounding by baseline severity. In these graphs individual day 90 results from the various endpoints are plotted against baseline severity. Any separation between the groups in these graphs is not confounded by baseline severity as by plotting against baseline severity at each point in the curve we are comparing patients with equal severity at baseline.

For all the endpoints considered (including change from baseline in NIHSS) the pattern is the same with the tPA curve showing an advantage over placebo for patients with a baseline NIHSS value between roughly 5 and 22, which represents the vast majority of the recruited population (as can be seen from figure 1 above). Therefore rather than leading to opposite conclusions to the 2004 reanalysis, in fact the conclusion seems consistent and benefit of tPA can still be seen after accounting for confounding.

The apparent reversal of the effect for both very mild and very severe patients in some of the endpoints could be an artefact of the sparsity of data at the extremes or possibly a signal that very mild patients do not require treatment and very severe patients cannot benefit. rt-PA is contraindicated in patients with minor neurological deficit or symptoms rapidly improving before start of infusion and in severe stroke as assessed clinically (e.g. NIHSS>25) and/or by appropriate imaging techniques.

A further observation is that due to the bimodal nature of the data (a clump of data representing those patients still alive and another clump for those who died) fitting curves through the continuous data is not an ideal approach, as the curve sometimes ends up in a space where there are few actual data points, not really representing the average. Therefore a dichotomous analysis, splitting into responders and non-responders as the trial was originally analysed, may be sounder. This is a similar concern to that described above, where it is noted that NIHSS is not truly a linear scale.

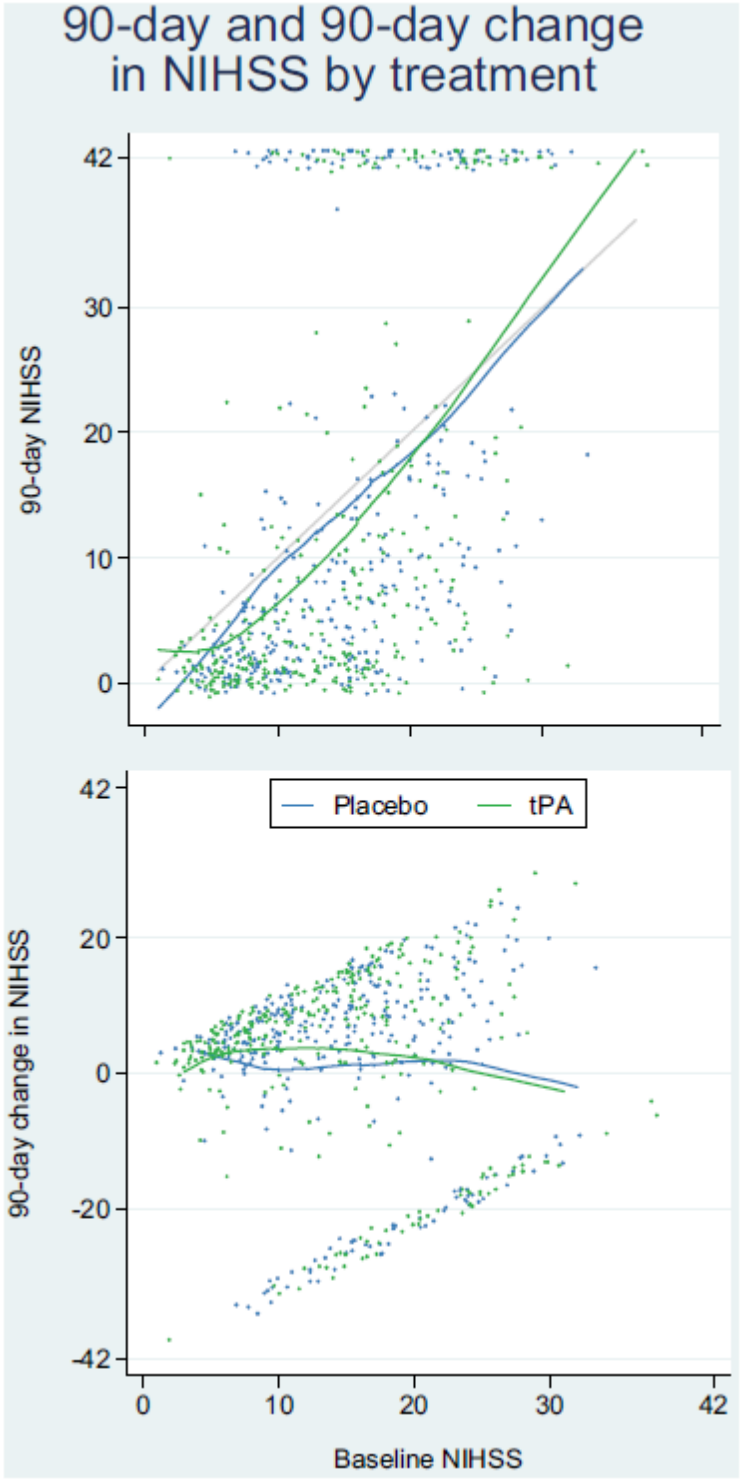
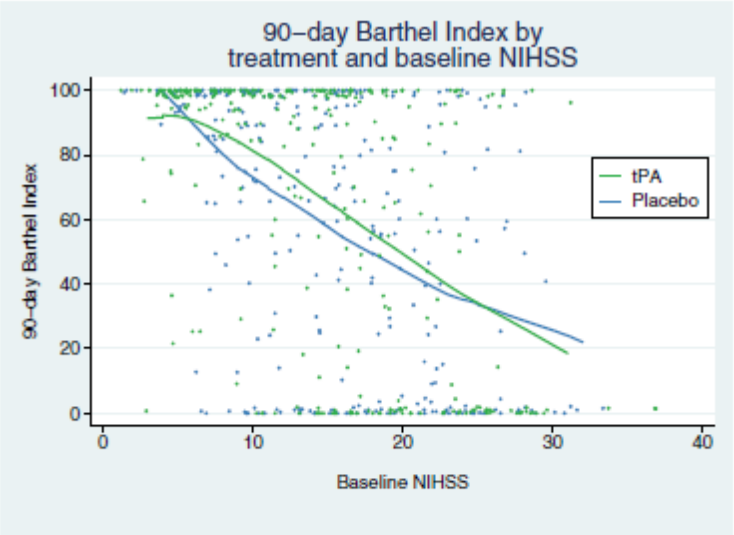
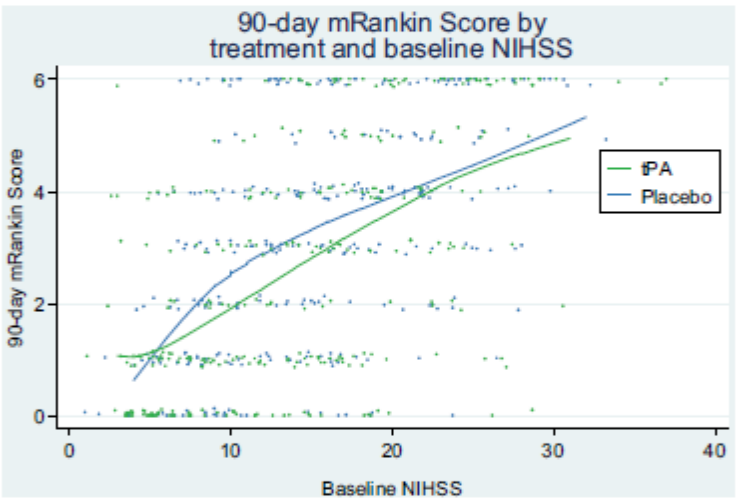

90-day and 90-day change in NIHSS by treatment
90-day NIHSS
42
30
20
10
0
Placebo
tPA
90-day change in NIHSS
42
20
0
-20
-42
0
10
20
30
42
Baseline NIHSS

Figure 2

A


90–day Barthel Index by
treatment and baseline NIHSS
90-day Barthel Index
100
80
60
40
20
0
0
10
20
30
40
Baseline NIHSS
tPA
Placebo

B


90-day mRankin Score by
treatment and baseline NIHSS
90-day mRankin Score
6
4
2
0
0
10
20
30
40
Baseline NIHSS
tPA
Placebo

C

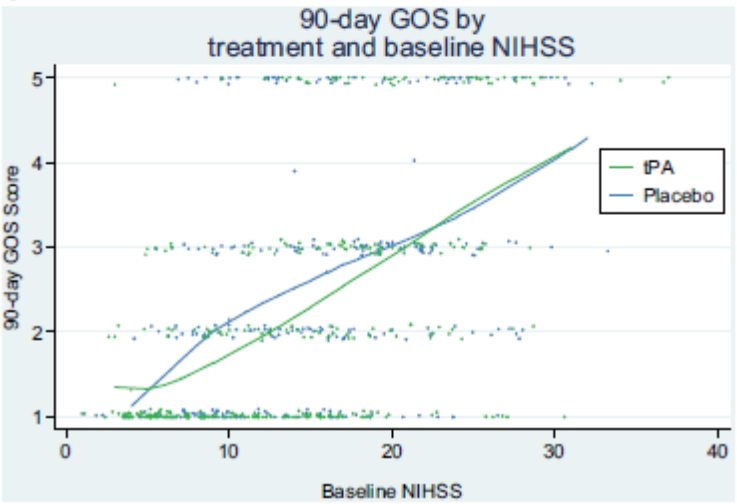

90-day GOS by
treatment and baseline NIHSS
90-day GOS Score
5
4
3
2
1
0
10
20
30
40
Baseline NIHSS
tPA
Placebo

Figure 3

7.3 Is thrombolysis for stroke time-dependent?

The concept that time-to-drug is important for stroke thrombolysis is widely promoted. For acute myocardial infarction it took 60,000 subjects to satisfactorily demonstrate that thrombolysis benefit was present, confined to STEMI (and no other MI subgroup), and time-dependent. In this case all trials and analyses showed statistical benefit and there was an uncontested and consistent relationship between time-to-drug and outcome. There is no theoretical basis, nor any clinical data from the thrombolytics for MI, that would suggest that tPA is less likely to cause ICH or more likely to demonstrate benefit, than any other agent[1]. Stroke data on this matter, which include one sixth as many subjects, are far less uniform.

Assessors' comments:
With regards to the comment that there is no theoretical basis that would suggest that rt-PA is less likely to cause ICH or more likely to demonstrate benefit than any other agent, there are theoretical reasons why it would not be expected that all thrombolytics should have the same properties (see section 3.2), for example the differences between rt-PA and streptokinase in terms of fibrin-specificity and accumulation of fibrinogen-degradation products.

Moreover, there is just one author group that finds time-to-drug to be associated with benefit[15,20], and a number that do not. A Cochrane analysis focusing only on properly blinded RCT data, to the apparent surprise of the authors, found no such association despite a comprehensive examination[21]. This absence of relationship between time-to-drug and effect was confirmed not only in re-analysis of the NINDS trial[11], but also in the largest ever trial of thrombolysis for stroke, IST-3, which reports comparisons of 0-3 hrs, 3-4.5 hrs and >4.5 hrs to be nonsignificant ($p = 0.613$) [12].

So the claim that "time is brain" deserves scepticism, because whether there really is an association between the interval from stroke onset to thrombolysis and magnitude of benefit is important. We would expect to see a clear association if thrombolysis restores blood flow and prevents brain cell death.

The "open artery hypothesis" has been further shattered by three neurointerventional trials, where no improvement in outcome was found despite dramatically higher rates of re-established blood flow to ischaemic brain[22]. As Kidwell states: "the imaging selection hypothesis [ischaemic penumbra] is flawed as conceived."[23] The editors of Annals of Neurology have called this result "humbling indeed" and that we need to be "more humble and less dogmatic."[24]

Assessors' comment:
The important point for the licensing decision is that benefit is shown for the population for which the treatment is indicated. Though scientifically interesting it is not necessary to conclusively establish the relationship between time-to-treatment and response to conclude that we do have enough information for approval in the earlier time period and do not for patients treated later.

7.4 What about the pooled analyses?

We do not trust the pooling of trial data (so-called "individual patient meta-analysis") [15,20]. It is wrong and a child could see that it is silly to blend these patient events together as if these trials represent the same process (they don't), or as if the pooling process somehow magically overcomes the major problems with the trials from which the patients are taken - *it cannot.*

The only thing that makes the pooling of these data psychologically justifiable for some is that they believe that the intervention works. If they did not believe that the intervention worked, they would not accept the pooling of data for the reasons given above. We might suggest to you that calcium channel blockers (which we know increase mortality in MI) are beneficial if you pool the data from the trial patients who received them in the first three hours after arrival. Yet everyone would respond in the same way: interesting perhaps, but you're going to have to prove it with a major, non-funded, valid randomised trial before we would ever believe it or accept it as justification for treatment.

Pooled analyses are not randomized trials. They are selected subgroups, and they are unacceptable as proof of benefit. This is non-negotiable. While we occasionally accept such analyses in situations where there is inadequate data and little harm or expense (eg magnesium is beneficial only in those with severe asthma), we understand that this is a guess and not a proven data-driven reality. We would NEVER accept this for a dangerous and expensive intervention that is shown to increase mortality and has 10 trials with no benefit compared to two that suggest benefit[1].

The early stoppage of trials for harm, means that the number of subjects in studies demonstrating harm might have included over 2400 subjects based on originally intended enrollments[1]. Pooled analyses are therefore missing these phantom data, which would have further eroded any aggregate benefits. In their absence, any pooled analysis is biased toward benefit. Despite this, there remain five times as many trials showing harm or no benefit (n=10) as those concluding benefit (n=2), and 6675 subjects in trials demonstrating no benefit compared to 1445 subjects in trials concluding benefit.

7.5 Problems with meta-analyses

Meta-analysis is promoted as the best method of evaluating data from multiple trials. However, the BMJ has highlighted problems with meta-analysis. Based on methodologic flaws[25], missing patient data from studies stopped early for harm[1], heterogeneity calculations strongly suggesting pooling of data to be inappropriate[1,26], imbalances in baseline stroke severity (NINDS, ECASS-3), the use of a subjective outcome measure with weak inter-rater reliability[27], and the influence of manufacturer involvement in the only two studies suggesting benefit[8,28,29], it is no wonder that the BMJ has called for urgent action to restore the integrity of the medical evidence base[30].

Indeed, it is methodologically flawed to extract subgroups (post hoc) from different trials to combine in meta-analysis, as it removes any possibility of balancing confounders, and magically changes negative studies into a ‘positive’ result[31,32].

7.6 Problems with registry studies

Large industry funded registries have been used to claim a benefit with earlier tPA treatment, but are heavily confounded by selection bias and prove nothing[2]. At best, they can be hypothesis generating, as it is known that most observational studies are incorrect[33].

Outcomes in registry studies are also dependent on multiple statistical adjustments and confuse association with causation[34]. The ORs reported are subtle at best. In

one study these adjustments have such power over the outcomes that the 3-4.5 hr cohort had an unadjusted OR for good outcome of 1.19 (95%CI 1.10-1.29) compared with the <3 hr cohort, but an OR of 0.92 (95%CI 0.89-1.01) after adjustment[35]. With such profound changes resulting from somewhat arbitrary adjustments, and potential for a high degree of bias, it is difficult to accept the validity of such studies.

Several non-randomised registry studies have also claimed a better outcome with earlier treatment. However, a more plausible explanation is that the studies include stroke mimics and TIAs, and that these are more likely to be treated earlier because they resolve quickly and will not be present in the cohort of patients treated later, which will consist almost entirely of established strokes. Stroke mimics have excellent short-term outcomes, thus entirely explaining the supposed "benefit" of early tPA compared to tPA given later in these uncontrolled registry studies.

Prior to the current tPA for stroke paradigm, our emergency and general medicine colleagues would not infrequently observe patients with dense neurologic deficits (including some with large MCA occlusion on CT) that spontaneously resolved before our eyes. A Lazarus effect, without tPA! Such cases were not be seen by neurologists in the Emergency Department. Many in this field are not aware of the aptly titled paper *spectacular shrinking deficit*, which reports on 118 patients with an initial major hemispheric syndrome[36]. Of these, 12% had a spectacular shrinking deficit, attributed to further migration of embolus. The median onset of recovery was 2.5 hrs and the longest was 24 hrs. Now, instead of being recognised as a predictable part of the natural history of the disease, which we see in our clinical practice, such improvements are viewed (and reported in the lay press) as spectacular proof of the benefit of tPA. Such is the power of anecdote.

Another significant issue is that there are problems with the inter-rater reliability of the mRS. Scoring of mRS, even by a neurologist, is only moderately reliable at best when done face to face. We can only imagine how much misclassification occurred in the registry studies, where scoring was based upon telephone interviews or a letter reply form. It is noteworthy that studies with larger numbers of patients and observers reported poorer reliability[27].

Assessors' comment:
The approval was based upon assessment of the individual randomised trials and data from pooled analyses, meta-analyses and registry studies was considered only as supportive data. Such combined analyses are subject to bias (e.g. including the post-hoc day 90 data from NINDS part 1) but provided there are prospective randomised studies with positive results it can be useful to also look back at past data in the relevant population to provide supportive information.

7.7 The importance of replication

For a treatment to be scientifically sound, there needs to be replication of studies, elimination of bias and healthy debate.

Discussions on controversies contribute to our understanding of deficiencies in existing data. Perhaps the most glaring deficiency in data on thrombolysis for stroke is the absence of replication for trials suggesting benefit. It is unethical not to attempt to replicate these data. The medical literature is replete with initially positive studies followed by multiple larger, more reliable, conclusively negative studies, that were considered unethical by some doctors.

There are many examples where doctors refused to enrol patients in studies, because they considered it unethical. When the study is completed, it reaches the exact opposite conclusion to what was expected and practiced, leading to a medical reversal[37-39]. Indeed, most existing evidence suggests thrombolysis to be either unsafe or non-beneficial. In the face of a dangerous and unscientific rush to judgment, and a conspicuous intervention bias (to do something rather than nothing) [40], the MHRA should be commended for conducting this review. Critical evaluation by thoughtful minds is essential to maintaining the integrity of medicine, an attribute that, in the current debate over thrombolysis for stroke, is endangered.

After all, replication is the distinguishing characteristic of scientific knowledge and an essential test of the validity of any scientific statement[41].

Assessors' comments:
It is agreed that replication would enhance confidence in the results, but it is not always considered necessary as demonstrated by the existence of the CHMP guideline on applications with one pivotal trial and many products approved on such a basis. There is some replication in that NINDS part 2 was conducted in order to independently confirm the *post-hoc* day 90 findings from NINDS part 1.

When discussing trial results it is important to distinguish between negative trials which actually prove (to a reasonable degree) that there is no effect, and those that are inconclusive. It is also important to note that a trial that is negative for its primary outcome, or for a particular hypothesis (dose level, treatment strategy), can provide suggestive information about alternative hypotheses that could be tested in future trials, and should not then be weighed in the scales against later positive data as if it is contradictory. Counting the number of positive and negative trials for a treatment may be an overly simplistic approach to assessing benefit.

7.8 So what is the way forward?

A larger trial that essentially reproduces the NINDS study must be performed[4]. This would retest the single hypothesis regarding the possible benefit of tPA that is not clearly inconsistent with the available evidence – that treatment begun within 3 hrs of the onset of symptoms of acute ischaemic stroke might be beneficial[2].

Although some might argue that it would be unethical to retest tPA within 3 hrs, most of the data indicated that it should not be given beyond 3 hrs, and yet this was retested in ECASS-3. They cannot have it both ways!

Most importantly however, is that there is *absolutely no doubt* people with strokes are suffering early deaths because of tPA, but we do not know for sure if the supposed benefit in those that survive is real. Therefore how can we ethically *not* demand another well-designed, independent study? And, given the very high costs of providing safe stroke thrombolysis services, and the proven treatments that will inevitably be denied other patients because of the opportunity costs of providing these services, how can we *not* demand another study simply on economic grounds?

7.9 This treatment is not supported by some organisations

The American College of Emergency Physicians[42] and the Australasian College for Emergency Medicine have not supported this treatment[43].

Assessors' conclusion on submission from Professors Fatovich and Brown:

Aside from providing a general discussion, the main point addressing the data used for approval of rt-PA for use within 0-4.5 hours of stroke onset was the presentation of the Hoffman and Schriger graphical analysis (said to demonstrate that any apparent treatment benefit in the NINDS study is almost entirely the result of the baseline imbalance in stroke severity). For the reasons given above, the graphical analysis may instead be considered to support the independent review of the NINDS study from 2004 which also investigated the effect of the baseline imbalance and concluded that benefit was still seen even after this had been accounted for.

8. Overall conclusion on concerns raised by Dr [redacted], Professors Fatovich and Brown and Dr Mandava

This paper discusses the concerns that have been raised over the data available to support a positive balance of benefits and risks for rt-PA in the treatment of acute ischaemic stroke.

A wide variety of issues have been raised, relating to initial trials in animals, data on streptokinase, concerns regarding specific trials (NINDS, ECASS I, II, III, ATLANTIS and IST-3), and other general concerns, in particular the appropriateness of endpoints used in clinical trials, and the impact of baseline imbalances on the results of the key studies.

This latter point has been highlighted by the medical community in conjunction with the NINDS trials many times since the data were first published. It has been suggested that the positive results from NINDS were driven by the imbalance in baseline stroke severity between the two arms of the trials. Since publication of the study results in 1995, several re-analyses have examined this point and found that this is not the case:

- an independent re-analysis (O'Fallon, 2004) which included a) a baseline adjusted analysis, b) an analysis broken down into baseline categories;
- a graphical re-analysis (Hoffman and Schriger, 2009);
- a matching re-analysis by Mandava *et al* (2010).

In their submission to MHRA Professors Fatovich and Brown have cited the graphical analysis by Hoffman and Schriger as demonstrating that most, if not all, of the difference in the outcome between treatment groups was due to the imbalance of baseline stroke severity. However, when individual day 90 results are plotted against baseline severity the graphical analysis is in fact supportive of the O'Fallon and Mandava re-analyses.

The finding that several re-analysis methods all support the same conclusion would appear to support the robustness of the NINDS trial results.

It is considered that no data has been presented that provides strong evidence to overturn the previous understanding of the positive balance of benefits and risks associated with rt-PA in acute ischaemic stroke, when used in accordance with the Marketing Authorisation.

The following section suggests key points that the EWG may wish to discuss in relation to the topics covered by this paper.

9. Points for discussion

- Relevance of trial data using thrombolytics other than rt-PA in acute ischaemic stroke to assessment of the balance of benefits and risks of rt-PA
- Most clinically appropriate study duration/follow-up measure for endpoints in trials in acute ischaemic stroke (day 7 vs. day 30 vs. day 90 or 6 months etc)
- Most appropriate endpoint/outcome scale when considering results obtained from trials of rt-PA
- Most appropriate analysis method(s) when considering results obtained from trials of rt-PA (eg dichotomisation vs ordinal vs shift, mRS cut-off points etc)
- Is there sufficient evidence to confirm or refute the 'Time is brain' hypothesis and is this relevant for the current indication for rt-PA in acute ischaemic stroke?
- Does an ideal definition of sICH exist? How should ICH rates be reported?
- Does the emerging data from MR diffusion/perfusion studies on ischaemic lesion size support a beneficial effect of rt-PA given for ischaemic stroke within 4.5 hours of stroke onset?
- Is ischaemic lesion size a useful outcome marker that should be measured?
- Is there any evidence that baseline imbalances (particularly in the NINDS and ECASS III trials) have not been accounted for sufficiently.
- Is there a need for the MAH for rt-PA to provide further information regarding the use of arginine as an excipient and its likely effects?

Do any of the issues raised, individually or together, have implications for the authorised indication for rt-PA in acute ischaemic stroke?

10. References

Albers G W, Bates V E, Clark W M, Bell R, Verro P, Hamilton S A. Intravenous tissue-type plasminogen activator for treatment of acute stroke: the Standard Treatment with Alteplase to Reverse Stroke (STARS) study. JAMA 2000; 283:1145-50.

Ali M, Fulton R, Quinn T, Brady M, on behalf of the VISTA Collaboration, How well do standard stroke outcome measures reflect quality of life? A retrospective analysis of clinical trial data. Stroke 2013; 44:3161-3165

Ahmed N, Kellert L, Lees K R, Mikulik R, Tatlisumak T, Toni D, SITS Investigators. Results of Intravenous Thrombolysis within 4.5-6 hours and updated results within 3 to 4.5 hours of onset acute ischemic stroke recorded in the safe implementation of treatment in stroke international stroke thrombolysis register (SITS-ISTR), An observational study. JAMA Neurol. 2013; 70 (7): 837-44.

Ankolekar S, Fuller M, Cross I, Renton C, Cox P, Sprigg N, Siriwardena A N, Bath P M, Feasibility of an ambulance-based stroke trial, and safety of glyceryl trinitrate in ultra-acute stroke: the rapid intervention with glyceryl trinitrate in hypertensive stroke trial. Stroke 2013; 44(11):3120-3128.

Balu S, Differences in psychometric properties, cut-off scores and outcomes between the Barthel Index and modified Rankin Scale in pharmacotherapy-based stroke trials: systematic literature review. Current Medical Research and Opinion 2009; 25(6):1329-1341

Banks J L, Marotta C A, Outcomes validity and reliability of the modified Rankin Scale: Implications for stroke clinical trials: A literature review and synthesis. Stroke 2007; 38:1091-1096

Bath P M, Gray L J, Collier T, Pocock S, Carpenter J, Optimising Analysis of Stroke Trials (OAST) Collaboration, Can we improve the statistical analysis of stroke trials? Statistical reanalysis of functional outcomes in stroke trials. Stroke. 2007; 38(6):1911-5.

Bath P M, Lees KR, Schellinger P D, Altman H, Bland M, Hogg C, Howard G, Saver J L; European Stroke Organisation Outcomes Working Group. Statistical analysis of the primary outcome in acute stroke trials. Stroke 2012 Apr;43(4):1171-8.

Bednar M M, McAuliffe T, Raymond S, Gross C E, Tissue plasminogen activator reduces brain injury in a rabbit model of thromboembolic stroke. Stroke 1990; 21(12):1705-1709

Berger C, Fiorelli M, Steiner T, Schabitz W R, Bozzao L, Bluhmki E, Hacke W, von Kummer R. Hemorrhagic transformation of ischemic brain tissue: asymptomatic or symptomatic? Stroke. 2001;32:1330 –1335.

Berrouschot J, Rother J, Glahn J, Kucinski T, Fiehler J, Thomalla G. Outcome and severe hemorrhagic complications of intravenous thrombolysis with tissue plasminogen activator in very old (≥ 80 years) stroke patients. Stroke 2005; 36:2421-5.

Bond J, Gregson B, Smith M, Lecouturier J, Rousseau N, Rodgers H, Predicting place of discharge from hospital for patients with a stroke or hip fracture on admission. J Health Serv Res Policy 2000; 5(3):133-9.

Bray B D, Campbell J, Hoffman A, Tyrrell P J, Wolfe D A, Rudd A G. Stroke thrombolysis in England: an age stratified analysis of practice and outcome. Age and Ageing 2013; 42:240-45.

Brooks D N, Hosie J, Bond M R, Jennett B, Aughton M, Cognitive sequelae of severe head injury in relation to the Glasgow Outcome Scale, Journal of Neurology, Neurosurgery and Psychiatry 1986; 49:549-553

Brott T B, Marler J R, Olinger C P, Adams H P, Tomsick T, Barsan W G, Biller J, Eberle R, Hertzberg V, Walker M, Measurements of acute cerebral infarction: lesion size by computed tomography; Stroke 1989; 20:871-875.

Campbell B C, Purushotham A, Christensen S, Desmond P M, et al; EPITHET-DEFUSE Investigators. The infarct core is well represented by the acute diffusion lesion: sustained reversal is infrequent. J Cereb Blood Flow Metab 2012; 32:50-6.

Carod-Artal F J, Egido J A, Quality of life after stroke: the importance of a good recovery. Cerebrovascular Disease 2009; 27 Suppl 1:204-214.

Chao A C, Hsu H Y, Chung C P, Liu C H, Chen C H, Teng M M, Peng G S, Sheng W Y, Hu H H;. Outcomes of thrombolytic therapy for acute ischemic stroke in Chinese patients: the Taiwan Thrombolytic Therapy for Acute Ischemic Stroke (TTT-AIS) study. Stroke 2010: 41:885-90.

Chehrazi B B, Seibert J A, Kissel P, Hein L, Brock J M, Evaluation of recombinant tissue plasminogen activator in embolic stroke. Neurosurgery 1989; 24(3):355-360

Cohen D, Macdonald H, UK regulatory body to review alteplase after stroke. BMJ 2014; 349:g5355

Cramer S C, Koroshetz W J, Finklestein S P, The case for modality-specific outcome measures in clinical trials of stroke recovery-promoting agents. Stroke 2007; 38(4):1393-1395

Dalkara T, Morikawa E, Panahian N, Moskowitz M A, Blood flow dependent functional recovery in a rat model of focal cerebral ischemia. American Journal of Physiology – Heart and Circulatory Physiology 1994; 2:H678-H683

De Haan R, Horn J, Limburg M, Van Der Meulen J, Bossuyt P, A comparison of five stroke scales with measures of disability, handicap, and quality of life; Stroke 1993; 24:1178-1181

Desantis S M, Lazaridis C, Palesch Y, Ramakrishnan V. Regression analysis of ordinal stroke clinical trial outcomes: an application to the NINDS t-PA trial. Int J Stroke 2014; 9(2):226-31

Duffy L, Gajree S, Langhorne P, Stott D J, Quinn T J, Reliability (inter-rater agreement) of the Barthel Index for assessment of stroke survivors, systematic review and meta-analysis. Stroke 2013; 44:462-468.

Dzialowski I, Pexman J H, Barber P A, Demchuk A M, Buchan A M, Hill M D; CASES Investigators, Asymptomatic hemorrhage after thrombolysis may not be benign: prognosis by hemorrhage type in the Canadian alteplase for stroke effectiveness study registry. Stroke. 2007 Jan;38(1):75-9. Epub 2006 Nov 22.

Emberson J, Lees K R, Lyden P, Blackwell L, Albers G, Bluhmki E, Brott T, Cohen G, Davis S, Donnan G, Grotta J, Howard G, Kaste M, Koga M, von Kummer R, Lansberg M, Lindley R I, Murray G, Olivot J M, Parsons M, Tilley B, Toni D, Toyoda K, Wahlgren N, Wardlaw J, Whiteley W, Del Zoppo G J, Baigent C, Sandercock P, Hacke W; for the Stroke Thrombolysis Trialists' Collaborative Group. 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. Lancet 2014. pii: S0140-6736(14)60584-5. doi: 10.1016/S0140-6736(14)60584-5. [Epub ahead of print, 5 August]

ENOS trial investigators, Efficacy of nitric oxide, with or without continuing antihypertensive treatment, for management of high blood pressure in acute stroke (ENOS): a partial-factorial randomised controlled trial. Lancet, early online publication doi:10.1016/S0140-6736(14)61121-1

Europe study group – Thrombolytic therapy with streptokinase in acute ischemic stroke. The multicentre acute stroke trial. N Engl J Med 1996; 335(3):145-150

Fink J N, Selim M H, Kumar S, Silver B, Linfante I, Caplan L R, Schlaug G, Is the Association of National Institutes of Health Stroke Scale Scores and acute Magnetic Resonance Imaging stroke volume equal for patients with right and left hemisphere ischemic stroke? Stroke 2002; 33:954-958.

Fonarow G C, Smith E E, Saver J L, Reeves M J, Bhatt D L, Grau-Sepulveda M V, Olson D M, Hernandez A F, Peterson E D, Schwamm L H. Timeliness of tissue-type plasminogen activator therapy in acute ischemic stroke: patient characteristics, hospital factors, and outcomes associated with door-to-needle times within 60 minutes. Circulation 2011; 123:750-8.

Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, Pessin M, Ahuja A, Callahan F, Clark W M, Silver F, Rivera F. Intra-arterial prourokinase for acute ischemic stroke. The PROACT II study: a randomized controlled trial. Prolyse in Acute Cerebral Thromboembolism. JAMA 1999; 282:2003-11.

Goldstein L B, Samsa G P, Reliability of the National Institutes of Health Stroke Scale; Stroke 1997; 28:307-310

Granger C V, Hamilton B B, Gresham G E, The stoke rehabilitation outcome study – Part I: General description. Arch Phys Med Rehabil 1988; 69(7):506-509

Grotta J C, Burgin W S, El-Mitwalli A, Long M, Campbell M, Morgenstern L B, Malkoff M, Alexandrov A V. Intravenous tissue-type plasminogen activator therapy for ischemic stroke: Houston experience 1996 to 2000. Arch Neurol 2001; 58:2009-13.

Guideline on clinical investigation of medicinal products for prevention of stroke and systemic embolic events in patients with non-valvular atrial fibrillation; Committee for Medicinal Products for Human Use (CHMP), 2014; EMA/CHMP/341363/2014 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/07/WC500169918.pdf

Gumbinger C, Gruschka P, Bottinger M, Heerlein K, Barrows R, Hacke W, Ringleb P, Improved prediction of poor outcome after thrombolysis using conservative definitions of symptomatic haemorrhage. Stroke 2012; 43: 240-2.

Hacke W, Kaste M, Fieschi C, Toni D, Lesaffre E, von Kummer R, Boysen G, Bluhmki E, Godehard H, Mahagne M-H, Hennerici M, for the ECASS Study Group.

Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke, The European Cooperative Acute Stroke Study (ECASS). JAMA 1995; 274:1017-1025

[1]Hacke W, Kaste M, Fieschi C, von Kummer R, Davalos A, Meier D, Larrue V, Bluhmki E, Davis S, Donnan G, Schneider D, Diez-Tejedor E, Trouillas P, for the second Euopean-Australasian acute stroke study investigators. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Lancet 1998; 352:1245-51

[2]Hacke W, Bluhmki E, Steiner T, Tatlisumak T, Mahagne M-H, Sacchetti M-L, Meier D, for the ECASS Study Group. Dichotomized efficacy end points and global end points analysis applied to the ECASS intention-to-treat data set. Post-hoc analysis of ECASS I. Stroke 1998; 29:2073-2075

Hacke W, Donnan G, Fieschi C, Kaste M, von Kummer R, Broderick J P, Brott T, Frankel M, Grotta J C, Haley E C Jr, Kwiatkowski T, Levine S R, Lewandowski C, Lu M, Lyden P, Marler J R, Patel S, Tilley B C, Albers G, Bluhmki E, Wilhelm M, Hamilton S; ATLANTIS Trials Investigators; ECASS Trials Investigators; NINDS rt-PA Study Group Investigators. Association of outcome with early stroke treatment: pooled analysis of ATLANTIS, ECASS, and NINDS rt-PA stroke trials. Lancet 2004; 363(9411):768-774.

Hacke W, Kaste M, Bluhmki E, Brozman M, Davalos A, Guidetti D, Larrue V, Lees K R, Medeghri Z, Machnig T, Schneider D, von Kummer R, Wahlgren N, Toni D, for the ECASS Investigators. Thrombolysis with alteplase 3 to 4.5 hours after acute ischaemic stroke. N Engl J Med 2008; 359:1317-1329.

Harrison J K, McArthur K S, Quinn T J, Assessment scales in stroke: clinimetric and clinical considerations, Clinical Interventions in Aging 2013; 8:201-211

Harston G W J, Sutherland B A, Kennedy J, Buchan A M, The contribution of L-arginine to the neurotoxicity of recombinant tissue plasminogen activator following cerebral ischemia: a review of rtPA neurotoxicity. Journal of Cerebral Blood Flow & Metabolism 2010; 30:1804–1816

Hill MD, Buchan AM. Thrombolysis for acute ischemic stroke: results of the Canadian Alteplase for Stroke Effectiveness Study. CMAJ 2005; 172:1307-12.

Hoffman J R, Schriger D L, A graphic reanalysis of the NINDS trial. Ann Emerg Med 2009; 54(3):329-336

Hudson I, Alteplase for ischaemic stroke – responses. Lancet 2014; 384(9944):662-663

IST-3 collaborative group. The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial. Lancet 2012; 379(9834):2352-63.

IST-3 collaborative group. Effect of thrombolysis with alteplase within 6 h of acute ischaemic stroke on long-term outcomes (the third International Stroke Trial [IST-3]): 18-month follow-up of a randomised controlled trial. Lancet Neurol. 2013; 12(8):768-76.

Kasner S, Clinical interpretation and use of stroke scales, Lancet Neurol 2006; 5:603-612

Khatri P, Wechsler L R, Broderick J P. Intracranial haemorrhage associated with revascularization therapies. Stroke 2007; 38:431-40.

Kissel P, Chehrazi B, Seibert J A, Wagner F C Jr, Digital angiographic quantification of blood flow dynamics in embolic stroke treated with tissue-type plasminogen activator. J Neurosurg 1987; 67(3):399-405

Kovacs A, Szabo L, Longstaff C, Tenekedjiev K, Machovich R, Kolev K, Ambivalent roles of carboxypeptidase B in the lytic susceptibility of fibrin. Thromb Res Suppl 2014; 133(1):80-87.

Kranz P G, Eastwood J D. Does diffusion-weighted imaging represent the ischemic core? An evidence-based systematic review. AJNR Am J Neuroradiol 2009;30:1206-12.

Kwakkel G, Veerbeek J M, van Wegen E E H, Nijland R, Harmeling-van der Wel B C, Dippel D W J, the EPOS investigators Journal of the Neurological Sciences 2010; 294:57-61

Kwiatkowski T G, Libman R B, Frankel M, Tilley B C, Morgenstern L B, Lu M, Broderick J P, Lewandowski C A, Marler J R, Levine S R, Brott T. 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(23):1781-1787.

Labeyrie M A, Turc G, Hess A, Hervo P, Mas J L, Meder J F, Baron J C, Touze E, Oppenheim C, Diffusion lesion reversal after thrombolysis: a MR correlate of early neurological improvement. Stroke 2012; 43:2986-91.

Lees K R, Bluhmki E, von Kummer R, Brott T G, Toni D, Grotta J C, Albers G W, Kaste M, Marler J R, Hamilton S A, Tilley B C, Davis S M, Donnan G A, Hacke W; for the ECASS, ATLANTIS, NINDS and EPITHET rt-PA Study Group Investigators, Time to treatment with intravenous alteplase and outcome in stroke: an updated pooled analysis of ECASS, ATLANTIS, NINDS, and EPITHET trials. Lancet 2010; 375(9727):1695-1703.

Lees K R, Bath P M W, Schellinger P D, Kerr D M, Fulton R, Hacke W, Matchar D, Sehra R, Toni D, Contemporary outcome measures in acute stroke research: choice of primary outcome measure. Stroke 2012; 43:1163-1170

Lenzer J, Controversial stroke trial is under review following BMJ report. BMJ 2002; 325(7373):1131.

Longstreth W T Jr, Katz R, Tirschwell D L, Cushman M, Psaty B M, Intravenous tissue plasminogen activator and stroke in the elderly. Am J Emerg Med 2010; 28(3): 359-363.

Lorenzano S and Toni D, Benefits and harms of pharmacological thrombolysis in *Rational Basis for Clinical Translation in Stroke Therapy*, editors: Guiseppe Micieli and Diana Amantea, CRC Press: Florida 2014.

Luby M, Warach SJ, Nadareishvilli Z, Merino JG. Immediate changes in stroke lesion volumes post thrombolysis predict clinical outcome. Stroke 2014; 45:3275-9.

Lyden P D, Brott T G, Tilley B, Welch K M A, Mascha E J, Levine S, Haley E C, Grotta J C, Marler J R. Improved reliability of the NIH Stroke Scale using video training. Stroke 1994; 25:2220-2226

Lyden P, Claesson L, Havstad S, Ashwood T, Lu M. Factor analysis of the National Institutes of Health Stroke Scale in patients with large strokes. JAMA Neurology 2004; 61(11):1677-1680

Lyden P D, Zivin J A. Hemorrhagic transformation after cerebral ischemia: mechanisms and incidence. Cerebrovasc Brain Metab Rev 1993; 5:1-16.

Mahoney F I, Barthel D W. Functional evaluation: the Barthel Index. Md State Med J. 1965; 14:61-65.

Mandava P, Kalkonde Y V, Rochat R H, Kent T A. A matching algorithm to address imbalances in study populations: application to the National Institute of Neurological Diseases and Stroke Recombinant Tissue Plasminogen Activator Acute Stroke Trial. Stroke 2010; 41:765-770

Mandava P, Krumpelman C S, Murthy S B, Kent T A, Chapter 40: A critical review of stroke trial analytical methodology: outcome measures, study design, and correction for imbalances. From: Translational Stroke Research from target selection to clinical trials; eds. Lapchak P A, Zhang J H, 2012. Springer Series in Translational Stroke Research, DOI 10.1007/978-1-4419-9530-8_40

Mandava P, Krumpelman C S, Shah J N, White D L, Kent T A. Quantification of Errors in Ordinal Outcome Scales Using Shannon Entropy: Effect on Sample Size Calculations. PLoS ONE 2013; 8(7):e67754. doi:10.1371/journal.pone.0067754

Martin-Schild S, Albright K C, Tanksley J, Pandav V, Jones E B, Grotta J C, Savitz S I. Zero on the NIHSS does not equal the absence of stroke. Ann Emerg Med 2011; 57:42-45.

MAST-I group – Randomised controlled trial of streptokinase, aspirin, and combination of both in treatment of acute ischaemic stroke. Multicentre Acute Stroke Trial – Italy (MAST-I) Group. Lancet 1995; 346(8989):1509-1514

Merino J G, Latour L L, Todd J W, Luby M, Schellinger P D, Kang D W, Warach S, Lesion volume change after treatment with tissue plasminogen activator can discriminate clinical responders from nonresponders. Stroke 2007; 38:2919-23.

Muir K W, Weir C J, Murray G D, Povey C, Lees K R, Comparison of neurological scales and scoring systems for acute stroke prognosis; Stroke 1996; 27:1817-1820

National Institute for Health and Care Excellence (NICE), Alteplase for treating acute ischaemic stroke (review of technology appraisal guidance 122). TA264; September 2012 https://www.nice.org.uk/guidance/ta264

National Institute of Neurological Disorders and Stroke (NINDS) rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischaemic stroke. N Engl J Med 1995; 333:1581-1587.

NINDS rt-PA Stroke Study Group. Effect of intravenous recombinant tissue plasminogen activator on ischaemic stroke lesion size measured by computed tomography. Stroke 2000; 31:2912-2919.

Nichols C, Khoury J, Brott T, Broderick J. Intravenous recombinant tissue plasminogen activator improves arterial recanalization rates and reduces infarct volumes in patients with hyperdense artery sign on baseline computed tomography. J Stroke Cerebrovasc Dis 2008; 17:64-8.

O'Fallon WM, Asplund K, Goldfrank LR, Hertzberg VS, Ingall TJ, Louis TA. Report of the t-PA Review Committee 2004. http://stroke.nih.gov/resources/t-pa-review-committee.htm

Papadopoulos S M, Chandler W F, Salamat M S, Topol E J, Sackellares J C, Recombinant human tissue-type plasminogen activator therapy in acute thromboembolic stroke. J Neurosurg 1987; 67(3):394-398

Parsons M, Spratt N, Bivard A, Campbell B, Chung K, Miteff F, O'Brien B, Bladin C, McElduff P, Allen C, Bateman G, Donnan G, Davis S, Levi C, A randomised trial of tenecteplase versus alteplase for acute ischaemic stroke. N Engl J Med 2012; 366(12):1099-1107

Phillips D A, Davis M A, Fisher M, Selective embolization and clot dissolution with tPA in the internal carotid artery circulation of the rabbit. Am J Neuroradiol 1988; 9(5):899-902

Points to Consider on clinical investigation of medicinal products for the treatment of acute stroke; Committee for Proprietary Medicinal Products (CPMP), 2001 CPMP/EWP/560/98.
http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003342.pdf

Quinn T J, Dawson J, Walters M R, Lees K R, Reliability of the modified Rankin Scale: A systematic review; Stroke 2009; 40: 3393-3395.

Qureshi A I, Harris-Lane P, Kirmani J F, Janjua N, Divani A A, Mohammad Y M, Suarez J I, Montgomery M O. Intra-arterial reteplase and intravenous abciximab in patients with acute ischemic stroke: an open-label, dose-ranging, phase I study. Neurosurgery 2006; 59(4):789-796

Rankin J, Cerebral vascular accidents in patients over the age of 60: prognosis. Scott Med J 1957; 2:200-215

Ringleb PA, Schwark C, Kohrmann M, Kulkens S, Jüttler E, Hacke W, Schellinger P D. Thrombolytic therapy for acute ischaemic stroke in octagenarians: selection by magnetic resonance imaging improves safety but does not improve outcome. J Neurol Neurosurg Psychiatry 2007; 78:690-3.

Sandercock P, Lindley R, Wardlaw J, Dennis M, Innes K, Cohen G, Whiteley W, Perry D, Soosay V, Buchanan D, Venables G, Czlonkowska A, Kobayashi A, Berge E, Slot K B, Murray V, Peeters A, Hankey G J, Matz K, Brainin M, Ricci S, Cantisani T A, Gubitz G, Phillips S J, Antonio A, Correia M, Lyrer P, Kane I, Lundstrom E and the IST-3 collaborative group, Update on the third international stroke trial (IST-3) of thrombolysis for acute ischaemic stroke and baseline features of the 3035 patients recruited. Trials 2011; 12:252

Sandercock P A, Wardlaw J M, Lindley R I, Dennis M S, IST-3 Collaborative Group. Abstract WMP22: Third international stroke trial (IST-3): Effect of iv rt-pa <6 hours in acute ischaemic stroke on living circumstances and health related quality of life at six months. Stroke 2013; 44:AWMP22

Sandercock P, Lindley R, Wardlaw J M, Murray G, Whiteley W, Cohen G. Alteplase for ischaemic stroke – responses. Lancet 2014; 384:660-661

Saver J L, Optimal endpoints for acute stroke therapy trials: best ways to measure treatment effects of drugs and devices; Stroke 2011; 42(8):2356-2362.

Schenkel J, Weimar C, Knoll T, Habert R L, Busse O, Hamann G F, Koennecke H C, Diener H C; German Stroke Data Bank Collaborators. R1-systemic thrombolysis in German stroke units-the experience from the German Stroke data bank. J Neurol 2003; 250:320-4.

Savitz S I, Lew R, Bluhmki E, Hacke W, Fisher M. Shift analysis versus dichotomization of the modified Rankin scale outcome scores in the NINDS and ECASS-II trials. Stroke 2007; 38(12):3205-12.

Schlegel D J, Tanne D, Demchuk A M, Levine S R, Kasner S E, Prediction of hospital disposition after thrombolysis for acute ischemic stroke using the National Institutes of Health Stroke Scale, Arch Neurol 2004; 61:1061-1064.

Seet R C, Rabinstein A A. Symptomatic intracranial haemorrhage following intravenous thrombolysis for acute ischemic stroke: a critical review of case definitions. Cerebrovasc Dis 2012; 34:106-14.

Seet R C, Zhang Y, Moore S A, Wijdicks E F, Rabinstein A A. Subtherapeutic international normalized ratio in warfarin-treated patients increases the risk for symptomatic intracerebral hemorrhage after intravenous thrombolysis. Stroke 2011; 42:2335-5.

Shinar D, Gross C R, Bronstein K S, Licata-Gehr E E, Eden D T, Cabrera A R, Fishman I G, Roth A A, Barwick J A, Kunitz S C, Reliability of the activities of daily living scale and its use in telephone interview. Arch Phys Med Rehabil 1987; 68(10):723-728

Shinton R, Questions about authorisation of alteplase for ischaemic stroke. Lancet 2014; 384:659-660.

Slot K B, Berge E, Dorman P, Lewis S, Dennis M, and Sandercock P, and the Oxfordshire Community Stroke Project, the International Stroke Trial (UK) and the Lothian Stroke Register. Impact of functional status at six months on long term survival in patients with ischaemic stroke: prospective cohort studies. BMJ 2008; 336(7640):376-79.

Sobesky J, Frackowiak M, Zaro Weber O, Hahn M, Möller-Hartmann W, Rudolf J, Neveling M, Grond M, Schmulling S, Jacobs A, Heiss W D. The Cologne stroke experience: safety and outcome in 450 patients treated with intravenous thrombolysis. Cerebrovasc Dis 2007; 24:56-65.

Stingele R, Bluhmki E, Hacke W, Bootstrap statistics of ECASS II data: just another post hoc analysis of a negative stroke trial? Cerebrovasc Dis. 2001;11(1):30-3.

Strbian D, Sairanen T, Meretoja A, Pitkaniemi J, Putaala J, Salonen O, Silvennoinen H, Kaste M, Tatlisumak T. Patient outcomes from symptomatic intracerebral haemorrhage after stroke thrombolysis. Neurology 2011; 77:341-8.

Stroke Therapy Academic Industry Roundtable II (STAIR II). Recommendations for clinical trial evaluation of acute stroke therapies. Stroke 2001; 32:1598-1606

Tanne D, Kasner S E, Demchuk A M, Koren-Moraq N, Hanson S, Grond M, Levine S R. Markers of increased risk of intracerebral haemorrhage after intravenous recombinant tissue plasminogen activator therapy for acute ischemic stroke in clinical practice: the Multicenter rt-PA Stroke Survey. Circulation 2002; 105(14):1679-85.

Uyttenboogaart M, Koch MW, Koopman K, Vroomen PC, Luijckx GJ, De Keyser J. Lipid profile, statin use, and outcome after intravenous thrombolysis for acute ischaemic stroke. J Neurol 2008; 255:875-80.

Vergouwen M D, Casaubon L K, Swartz R H, Fang J, Stamplecoski M, Kapral M K, Silver F L. Subtherapeutic warfarin is not associated with increased haemorrhage rates in ischemic strokes treated with tissue plasminogen activator. Stroke 2011; 42:1041-5.

Wade D T, Hewer R L, Functional abilities after stroke: measurement, natural history and prognosis. Journal of Neurology, Neurosurgery and Psychiatry 1987; 50:177-182

Wahlgren N, Ahmed N, Davalos A, Ford G A, Grond M, Hacke W, Hennerici M G, Kaste M, Kuelkens S, Larrue V, Lees K R, Roine R O, Soinne L, Toni D, Vanhooren G; SITS-MOST investigators. Thrombolysis with alteplase for acute ischaemic stroke in the Safe Implementation of Thrombolysis in Stroke-Monitoring Study (SITS-MOST): an observational study. Lancet 2007; 369:275-82.

Wahlgren N, Ahmed N, Davalos A, Hacke W, Millán M, Muir K, Roine R O, Toni D, Lees K R; SITS investigators. Thrombolysis with alteplase 3-4.5 h after acute ischaemic stroke (SITS-ISTR): an observational study. Lancet 2008; 372:1303-9

[a]Wahlgren N, Ahmed N, Eriksson N, Aichner F, Bluhmki E, Dávalos A, Erilä T, Ford G A, Grond M, Hacke W, Hennerici M G, Kaste M, Köhrmann M, Larrue V, Lees K R, Machnig T, Roine R O, Toni D, Vanhooren G; Safe Implementation of Thrombolysis in Stroke-MOnitoring STudy Investigators. Multivariable analysis of outcome predictors and adjustment of main outcome results to baseline data profile in randomized controlled trials: Safe Implementation of Thrombolysis in Stroke-MOnitoring STudy (SITS-MOST). Stroke 2008; 39:3316-22.

Wardlaw J M, Murray V, Berge E, del Zoppo G, Sandercock P, Lindley R L, Cohen G. Recombinant tissue plasminogen activator for acute ischaemic stroke: an updated systematic review and meta-analysis. Lancet 2012; 379(9834):2364-72.

Wardlaw J M, Murray V, Berge E, del Zoppo G J, Thrombolysis for acute ischaemic stroke. Cochrane Database of Systematic Reviews 2014; Issue 7. Art. No.: CD000213. DOI: 10.1002/14651858.CD000213.pub3.

Willmot M, Ghadami A, Whysall B, Clarke W, Wardlaw J, Bath P M, Transdermal glyceryl trinitrate lowers blood pressure and maintains cerebral blood flow in recent stroke. Hypertension 2006. 47(6):1209-1215.

World Health Organisation, Towards a common language for functioning, disability and health. Geneva: WHO; 2002
http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf

Zivin J A, Fisher M, DeGirolami U, Hemenway C C, Stashak J A, Tissue plasminogen activator reduces neurological damage after cerebral embolism. Science 1985; 230(4731):1289-1292

Zivin J A, Simmons J G, t-PA for stroke: The story of a controversial drug. 1st edition 2010; OUP USA, New York.

EXPERT WORKING GROUP

ACTILYSE (A LTEPLASE) B ALANCE OF BENEFI TS AND RISKS WHEN USED IN THE TREATMENT OF ACUTE ISCHAEMIC STROKE

Title of paper: Paper 5C: Further submission by Dr [redacted].

Product: Actilyse 10, 20, 50mg	**Assessors:** Medical assessor: Dr [redacted] Statistical assessor: Dr [redacted]
MAHs: Boehringer Ingelheim Limited	**Previous Assessments:** CHM May 2014 EWG: Nov 2014, Jan 2015
Active constituents: Alteplase (rt-PA)	**Legal status:** **POM**
Therapeutic classification: Antithrombotic agent, ATC code B01AD02	

Contents

1. Further submission by Dr [redacted]

Dr [redacted] s submission of February 2015 has been circulated to the group and the additional concerns [redacted] has raised are discussed in this section (see Annex 1 and 2).

Extracts and summaries of Dr [redacted] s submission are included below in blue italic font, followed by our assessment of each concern.

[redacted] main concerns are that key data relating to the effectiveness of alteplase on cerebral ischaemia and infarction have not been adequately presented and that a number of systematic reviews and meta-analyses have recycled incomplete data over the last 5 years to provide reassuring evidence on the benefit-risk balance of alteplase. The other main area of concern is early mortality rates due to cerebral oedema. Additional concerns regarding symptomatic haemorrhage rates by baseline stroke severity and baseline imbalances in IST-3 and final place of residence in IST-3 by subgroup according to onset-to-treatment time were considered in paper 5.

1.1 Concerns about MAST-I

MAST-I was a controlled, randomised, open trial of both streptokinase and aspirin in ischaemic stroke. It was terminated on safety grounds in January 1995.

Within 10 days of the stroke, 83 of 313 (26.5%) thrombolysed patients had died compared with 36/309 (11.7%) patients not thrombolysed. This 14.8% increase in early death comprised 7% which were cerebral, but not obviously caused by intracranial haemorrhage.

Full radiological data on those who died within 10 days have not been provided in the main trial publication. (MAST-I 1995) It seems plausible that thrombolysis led to an increase in the number of patients who died of ischaemic cerebral oedema.

Paper 5 has already considered the relevance of streptokinase clinical trial data and concluded that the balance of benefits and risks of streptokinase in the unlicensed indication of acute ischaemic stroke will not be considered further as the effects of streptokinase cannot be directly extrapolated to alteplase purely on the basis that they are both thrombolytics. Streptokinase possesses different biological and pharmacological properties to alteplase and there were important differences in the design of the clinical trials with streptokinase compared with alteplase. The scope of this review includes only alteplase.

1.2 Nonclinical studies

Experiments on 41 baboons provided good evidence that reperfusion following 100 mins of impaired cerebral blood flow increased oedema formation (Bell et al. 1985).

They concluded: 'If these results are applicable to man, restoration of flow should not be attempted after an ischaemic insult that reduces flow to less than 40% of normal unless it can be accomplished within 30 mins of the insult.'

'In baboons, cerebral oedema can be exacerbated by reperfusion 2 hours after stroke onset but it is unknown whether this occurs in humans (Bell et al., 1985).'

Normal cerebral blood flow (CBF, approximately 60 ml/100 g brain/min) is maintained by autoregulation at mean systemic arterial blood pressures of 60 to 150 mmHg (Numan et al. 2014). Cerebral autoregulation is impaired by acute ischaemia. After arterial occlusion, CBF is variably reduced in the territory of the occluded artery depending on the extent of occlusion and degree of collateral blood flow. Focal ischaemia can irreversibly damage a portion of cerebral tissue producing a core of infarction when its regional blood flow is reduced to less than 10-15 ml/100 g brain/min. An area of ischaemic penumbra contains inadequately perfused brain tissue at risk of infarction but which is potentially salvageable with reperfusion. The penumbra has regional blood flows in the range 10 to 25 ml/100 g brain/min. Altered sodium (Na^+) - potassium (K^+) ion pump function and ATP depletion result in neuronal dysfunction at the higher CBF threshold and cell death due to impaired active membrane transport occurs at the lower CBF threshold (Symon et al. 1977). However, the exact CBF threshold values that define the ischaemic core, penumbra and benign oligaemia regions are variably reported (Bandera et al. 2006). Neurones in the penumbral region do not survive indefinitely but the exact duration of survival in human ischaemic tissue is unknown but it is likely related to the degree of hypoperfusion and other factors such as glucose levels and temperature.

Focal cerebral ischaemia and post-ischaemic reperfusion result in cerebral capillary dysfunction that may increase blood-brain barrier permeability and cause cerebral oedema (Simard et al. 2007, Bai and Lyden 2015). Infarcted tissue will die and produce cytotoxic oedema following intracellular/extracellular shifts in ionic and water content but active blood flow is necessary for cerebral oedema to form. The early stages of capillary endothelial dysfunction are associated with upregulation of non-selective cation channels (eg transient receptor potential and the sulfonylurea receptor 1 (SUR1)-regulated $NC_{Ca\text{-}ATP}$ channels) that allow sodium and water accumulation in the extravascular space.

The second phase of endothelial dysfunction (vasogenic oedema) is associated with increased permeability of the blood-brain barrier allowing leakage of plasma proteins into the extracellular space. Various mechanisms have been proposed for this second phase including: reverse pinocytosis, formation of inter-endothelial gaps, partial degradation of the basement membrane, and disrupted calcium (Ca^{++}) signalling. Haemorrhagic transformation is not usually considered as a severe form of cerebral oedema although progressive ischaemic damage to endothelial capillaries causes death and loss of their structural integrity which allows red cells to enter the brain parenchyma. Animal and human studies have shown a close relationship between blood-brain barrier compromise and haemorrhagic conversion including intracerebral haemorrhage (ICH) (Simard et al. 2007).

Bell et al. (1985) examined the effects of ischaemia and reperfusion using a middle cerebral artery (MCA) baboon stroke model. They took 41 baboons and occluded the right MCA using a Scoville clip inserted via a transorbital approach for 30 (n=21) or 100 minutes (mins; n=20). In 10 animals from each group, the period of ischaemia was followed by 60 mins of reperfusion. Regional cerebral blood flow (CBF) was determined by hydrogen clearance and oedema was measured using microgravimetry. The animals received induction anaesthesia with phencyclidine and thiopentone, were paralysed, intubated, maintained on a mixture of chloral hydrate

and glucose (alpha-chloralose) and ventilated with 100% oxygen. Arterial carbon dioxide, rectal temperature and blood pressure were maintained within normal values during the study. All animals were euthanised at the end of the study with intravenous potassium chloride and the average specific gravity of the brain samples was determined to calculate the water content. Cytotoxic oedema was measured as the difference between the water content of the sample and of brain remote from the ischaemic area. Regional CBF was calculated from hydrogen clearance using intracranial electrodes over 2 mins at the start and end of the periods of ischaemia and perfusion.

The mean CBF (±SD) in the normal baboon cortex was 46.7 ±12.7 ml/100 g brain/min. During the period of MCA occlusion, the mean regional ischaemic CBFs blood were reduced to: 23% of control CBF in the right Sylvian area cortex; and to 28% of control CBF in the right intermediate cortical area (vs mean 93% for left hemisphere cortex). Removal of the arterial occlusion immediately restored blood flow to the ischaemic areas to approximately 70% of normal and to 89-100% by the end of reperfusion period. There was no significant oedema in any cortical area when ischaemia was limited to 30 mins. When the duration of ischaemia was extended to 100 mins, a significant increase in cortical water content of 12.9 mg/g cortex was noted in the right Sylvian area (the control water content was 804.5 mg/g cortex, $p<0.01$) and in the subcortical water content (increase of 6.8 mg/g white matter, $p<0.01$). Cortical oedema formation was maximal at a CBF of 5 ml/100 g brain/min after 100 mins of ischaemia.

Cortical water accumulation after 100 mins of ischaemia was only apparent when CBF fell below a threshold value which was unaltered when reperfusion followed (figure 1). The apparent inflection point defining the threshold value was 19 ml/100 g brain/min (40.5% control CBF) with a threshold value of 7.4 ml/100 g brain/min for the subcortical areas.

Figure 1: Amount of oedema (mg water/g brain) in samples of cortex obtained from: baboons exposed to 100 mins ischaemia (group 1) or 100 mins ischaemic followed by 60 mins reperfusion (group 2). The difference between the water content of the brain sample and brain remote from the ischaemic and flow electrodes is shown. Values are plotted against the flow (ml/100 g brain/min) during the ischaemic period (Bell et al. 1985).

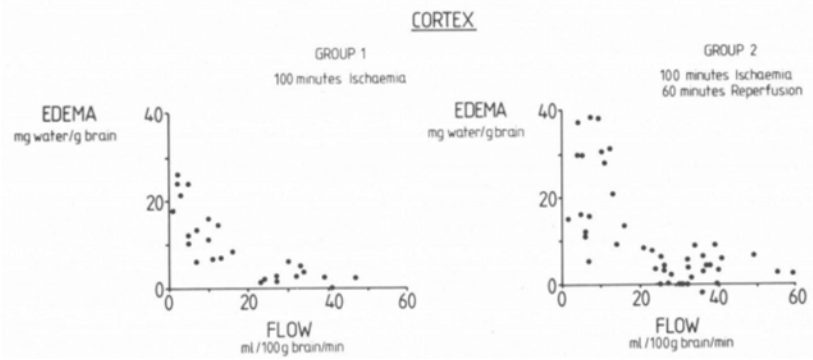
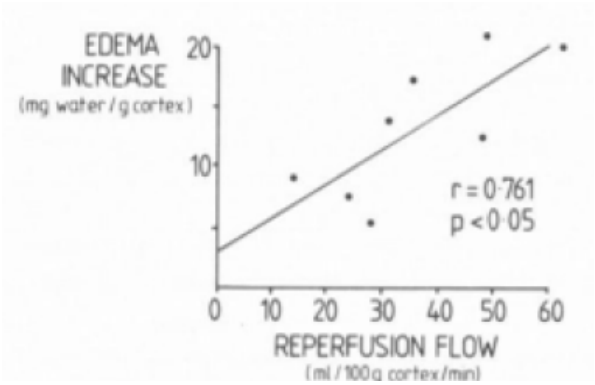


Figure 2: Effect of reperfusion on oedema in cortex that has been below the oedema threshold for 100 mins. Increase in oedema representing the change in water in cortical samples occurring after reperfusion is compared with cortical samples that had been subjected to an identical degree of ischaemia without reperfusion. Oedema values are plotted against the flow during reperfusion (Bell et al. 1985).



Reperfusion of the cortex for 1 h after 100 min of ischaemia increased oedema by a mean of 7.2 mg/g in the Sylvian area which was not statistically significant from the increase in water content observed after ischaemia alone. However, there was marked heterogeneity in the degree of ischaemia within an area and the increase in water in cortical samples subjected to the same degree of ischaemia depended on the flow rate during reperfusion. A positive correlation between oedema increase and reperfusion flow was noted for cortical regions only (p<0.05, figure 2). The largest increases in oedema were around 20 mg water/g cortex at maximal reperfusion rates of 60 ml/100 g cortex/min (figure 2).

The authors concluded that blood flow must be reduced below approximately 40% of normal before ischaemic oedema develops in baboons and the cortical blood flow threshold for oedema formation of 19 ml/100 g cortex/min is just above the reported threshold for the loss of sensory evoked potentials (16 to 18 ml/100 g brain/min) and higher than the flow threshold for loss of homeostasis of extracellular Ca^{++} and K^{+} levels (10 ml/100 g brain/min).

The baboon model used has a number of potential disadvantages: subtotal or complete MCA infarctions only occur in up to 10% of anterior circulation stroke patients limiting the generalisability of any findings to human patients (Huttner et al. 2009); baboons have a more extensive collateral circulation than humans (Cook and Tymianski 2012); the anaesthetic drugs used may be vasoactive or have neuroprotective properties; there is marked inter-individual variability in the degree of ischaemia produced; animal studies using carbon tracer techniques and fluorescent-labeled intravascular markers have reported no reperfusion despite recanalization of the MCA (Bai and Lyden 2015); ischaemic stroke may induce a compensatory increase in mean arterial blood pressure after arterial occlusion but the blood pressure of anaesthetised animals was maintained in the normal range; the continuous supply of glucose in the intravenous anaesthetic may have altered individual susceptibility to ischaemia; and the surgical procedures involved are

technically demanding with vascular complications which can produce findings that are difficult to replicate.

The findings of this article are consistent with the ischaemic penumbra concept as it reports CBF and time-thresholds that may define penumbral viability in a particular primate stroke model. Others have reported similar CBF threshold effects in awake baboons with moderate to large infarcts after 2 to 3 h of reversible MCA territory ischaemia (Jones et al. 1981). Once an ischaemic core has developed then its extent is expected to be related to the risk of developing cerebral oedema during reperfusion. There was marked inter-individual variability in CBF reductions observed consistent with variable arterial collateralisation.

Bell's statement "*In baboons, cerebral oedema can be exacerbated by reperfusion 2 hours after stroke onset but it is unknown whether this occurs in humans"* is supported by the study results but the absolute increases in cortical water content appear relatively modest (maximum 20 mg water/g cortex increase over the baseline value of 804.5 mg water /g cortex represents a less than 3% increase over the amount of oedema observed with no reperfusion). The authors did not comment on the clinical significance of the observed increases in cerebral oedema or if they resulted in any major mass effect so the applicability of these results to human patients is unclear. This animal study was conducted before the introduction of good laboratory practice so would not meet current regulatory standards, no primary or secondary outcome measures would have been specified in advance so it is not clear whether the reported data is derived from pre-planned or exploratory analyses. The methods section does not describe any statistical techniques such as correction for multiple comparisons.

Others have also reported their experience of primate stroke models:

- Crowell et al. (1970) used a retro-orbital approach to clip the right MCA at its origin for 1 to 24 h in 43 macaque monkeys (1-2 h, n=11; 4 h, n=10; 6-8 h, n=12; 24 h, n=10). One to 2 h clippings caused little or mild neurological deficits, 4 h clippings caused mild to moderate deficits and 6 to 24 h clippings resulted in extensive areas of infarction with a high incidence of haemorrhagic transformation. Occlusion times of ≤4 h often resulted in areas of cerebral infarction limited to subcortical structures. Cerebral swelling, defined as a shift of midline structures, was observed only when the clipping had produced medium-sized or large infarcts which were reported in a third of monkeys after 1-2 h total MCA occlusion and in 5 monkeys (63%) after 4 h of occlusion. The authors concluded that an increased incidence of haemorrhagic cerebral infarction might be avoided if reperfusion is achieved within about 4 h of occlusion.

- Jones et al. (1981) temporarily occluded the MCA for 15 or 30 mins, 2 to 3 h, or permanently in an awake primate model and serially monitored neurological function and local CBF. Clinical improvement occurred after up to 3 h of MCA territory ischaemia associated with the development of moderate to large-sized infarcts.

- Crowell et al. (1981) transiently occluded the right MCA transorbitally using a snare ligature in 38 macaque monkeys for 30 mins (n=4), 4 h (n=6), 8 h (n=3), 16 h (n=3), 24 h (n=3) or permanently (n=5). Three monkeys died from thrombosis within 24 h. Eleven monkeys were excluded from the study after developing post-operative complications. The monkeys were sedated with phencyclidine and pentobarbitone for the surgical procedure and then restrained and examined every 15 mins during the day of ischaemia. The monkeys were then euthanized at 2 weeks after MCA occlusion and underwent angiography and neuropathological examination to determine the time needed for reversible ischaemia to develop into irreversible infarction. Focal cerebral ischaemia required 4 to 8 h to evolve into maximal infarction. Thirty mins to 4 h of occlusion was tolerated with little or no infarction and only a single infarct showed evidence of haemorrhage. Poorly circumscribed infarcts were noted in less than a third of monkeys but were often associated with pronounced cerebral oedema but the duration of occlusion was not specified in these animals.

- The development of endovascular techniques to reversibly occlude the MCA allows longitudinal monitoring of ischaemic lesions in conscious baboons (Wey and Duong 2012). The reproducibility of endovascular primate models has not been established. The spatial-temporal characteristics of perfusion-diffusion mismatch in baboons have been reported following up to 90 mins of temporary MCA occlusion. The size of the perfusion-diffusion mismatch tissue (equivalent to the ischaemic penumbra) gradually decreased for up to 6 h post-occlusion (Wey et al. 2011) implying time-dependent reduction in the viability of the ischaemic penumbra. The average final infarct volume was ~17% of total brain volume and no altered consciousness, consistent with clinically significant cerebral oedema, was reported.

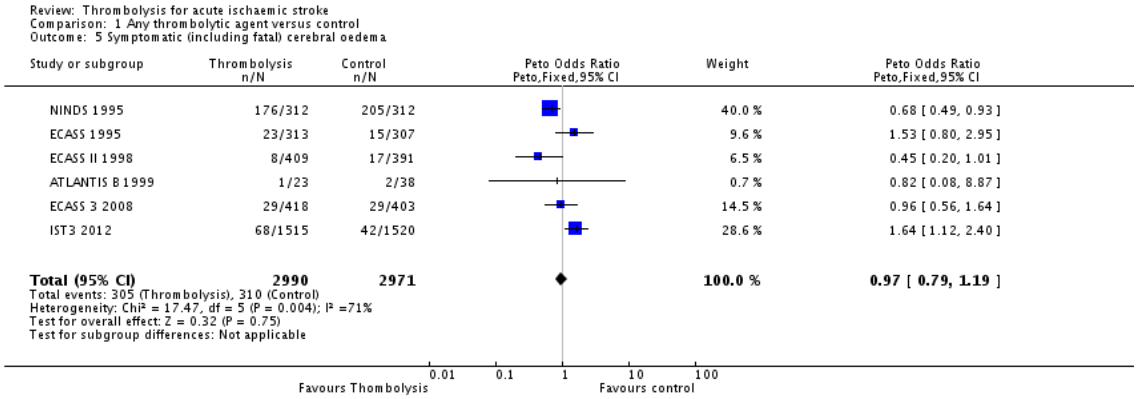
The pilot study of alteplase administered within 90 mins of ischaemic stroke states that evidence from animal studies suggests that ischaemic brain injury will occur when arterial occlusion persists for more than 2 to 3 h (Brott et al. 1992; Jones et al. 1981).

The incidence and severity of post-thrombolytic haemorrhagic infarction was not increased if alteplase was administered within 3.5 h of MCA occlusion and followed by 30 mins reperfusion in awake baboons (del Zoppo et al. 1990). The incidence of cerebral oedema was not reported but no neurological deterioration occurred after alteplase infusion suggesting that clinically significant oedema did not develop.

1.3 Thrombolysis for acute ischaemic stroke, Cochrane reviews 2009 and 2014

The data on symptomatic and fatal oedema from the NINDS study in Figure 5 of Cochrane Reviews 2009 and 2014 are of concern (Analysis 1.5 of the Cochrane Review 2014 is shown as figure 3)*. The data are highly inconsistent statistically and not plausible from a clinical perspective (see figure). I have been unable to find evidence that the NINDS data on post randomisation symptomatic (including fatal) oedema have ever been published in a peer reviewed journal.* [redacted]

Figure 3: Analysis 1.5 showing the outcome of symptomatic (including fatal) cerebral oedema for the comparison of any thrombolytic agent versus control (Wardlaw et al. 2014).



Analysis 1.5 for the previous Cochrane review (Wardlaw et al. 2009) was identical to the Cochrane review of 2014 except the weightings were higher for: NINDS 1995 (55.8%); ECASS 1995 (13%); ECASS II 1998 (8%); ATLANTIS B 1999 1%; and ECASS III 2008 (20.5%) as the IST-3 study, a large study with 28.6% weight in the 2014 analysis, was not included. The earlier Cochrane review also included data for MELT 2007 which was an uncompleted trial of intra-arterial urokinase within 6 h of stroke onset that randomised 114 patients to treatment (n=57 for urokinase and control) (Ogawa et al. 2007). The reason for the omission of MELT 2007 from the 2014 review is not clear.

An individual study gets its weight from the number of events rather than the number of patients in an odds ratio analysis for a rare event such as symptomatic cerebral oedema. In analysis 1.5, The NINDS study has weight much above its sample size and is very influential. If the NINDS study data is removed from the analysis then the total odds ratio increases from 0.97 to 1.23. However, the percentage of treated patients developing symptomatic (including fatal) cerebral oedema is much higher in the NINDS study (alteplase 56.4% vs 65.7% placebo) compared with the other studies listed (alteplase 2-7% vs placebo 3-7%) which is consistent with major differences between the studies so a combined analysis is highly questionable. It is difficult to believe that all studies are looking at the same clinical phenomenon when there is an event rate over 50-60% in one study and around 2-7% in all of the others. The NINDS study OR 0.68 (95% CI: 0.49-0.93) shows a favourable outcome as does the ECASS II study (OR 0.45 [95% CI: 0.20-1.01]) which has a more extreme point estimate and almost reaches statistical significance. So the results of the NINDS study are not really that extreme in terms of the odds ratio and are in line with ECASS II in that respect but the heterogeneity between the study findings in terms of event rates (as evidenced by the highly significant Chi^2 test, p=0.004) suggests that combining the findings of the studies in this respect may not be optimal. No time limit for the outcome measurement is specified.

NINDS differed from later studies in that the only radiological feature at baseline that excluded patients was intracranial haemorrhage on a CT scan and symptomatic ICH (sICH) was liberally defined as any haemorrhage associated with neurological

decline (NIHSS score ≥1) or that led to death within 7 days. Any sICH occurring within 36 h from treatment onset was considered attributable to study medication (The NINDS t-PA Stroke Study Group 1997). Symptomatic cerebral oedema was not defined in the primary publication. The definition of sICH is broad and would allow symptomatic cerebral oedema to be misclassified as sICH if any trace of haemorrhage was present, even if remote.

The risk of an adverse outcome associated with baseline early ischaemic changes (EICs) has been reported in a post hoc analysis of the NINDS study data (Patel et al. 2001). SICH was defined as any neurological deterioration thought to be due to haemorrhage on a CT scan although neurological deterioration at 24 h was defined as an increase in the baseline NIHSS score ≥4. All baseline CT scans were examined by a single neuroradiologist who was blinded to treatment allocation but who had access to all other clinical details (including individual component scores of the NIHSS). The EICs were classified into 3 categories: loss of grey/white matter distinction (focal or diffuse area in cerebral or cerebellar hemispheres); hypodensity or hypoattenuation; and compression of cerebrospinal fluid spaces (focal and/or diffuse brain swelling). A visual inspection was conducted to determine the extent of any ischaemic changes in the MCA territory. EIC was classified as more than one third of the MCA territory if ischaemic changes were seen in 2 or more different lobes of the cerebral hemisphere and basal ganglia plus insular cortex. The accuracy of lesion identification was assessed by reviewing the location and presence of hypodensity on the 24 h CT scans for those patients allocated placebo. Subjects with old lesions on baseline imaging were excluded from the analysis. Follow-up CT scans were done at 24 h, 7 to 10 days and 3 months after stroke. The distribution of EICs at baseline is shown in table 1. There was no statistically significant difference

Table 1: Distribution of early ischaemic changes at baseline (n=616*) (Patel et al. 2001).

	Type of EIC, No. (%)
Any EIC	194 (31)
Loss of GWMD	164 (27)
Presence of hypodensity	54 (9)
Compression of CSF spaces	89 (14)
Loss of GWMD >1/3 MCA	77 (13)
Hypodensity >1/3 MCA	14 (2)
Compression of CSF >1/3 MCA	54 (9)
Extent of EIC	
>1/3 MCA†	84 (14)
≤1/3 MCA	110 (18)
None	422 (69)

Key: *EIC (early ischaemic change) indicates loss of grey/white matter distinction (GWMD), presence of hypodensity, or compresion of cerebrospinal fluid (CSF) spaces; MCA, middle cerebral artery territory. †For a patient to be counted as having >⅓ MCA involvement, 1 or any combination of 3 components of EIC had to involve >⅓ of the MCA distribution.

in the distribution of EICs between alteplase and placebo groups (28% vs 35%, p=0.09). Table 2 shows the associations of EICs with baseline variables. EIC was significantly associated with baseline NIHSS scores (ρ=0.23, p<0.001) and time from stroke onset to imaging (ρ=0.11, p=0.007). An adjustment was made for any baseline variables associated with treatment outcome. The relationship between EICs on the

baseline CT and clinical outcomes, unadjusted and adjusted for other baseline variables using 4 different models are shown in table 3. The reference group was the placebo group in the 'no EIC' category. Thirty four percent of patients receiving alteplase with EICs affecting more than one third of the MCA territory at baseline were dead at 90 days compared with a mortality rate of 26% for placebo (model 4: adjusted OR 1.2 [95% CI: 0.5-2.9] for alteplase vs 1.1 [95% CI: 0.5-2.6] for placebo, *p*=0.82 for EIC x treatment interaction). Twenty one percent of patients receiving alteplase with EIC affecting more than one third of the MCA territory at baseline had deteriorated (defined as ≥4 points increase from the baseline NIHSS score) at 24 h compared with 20% receiving placebo (adjusted OR 1.3 [95% CI: 0.6-3.1] for alteplase vs 1.1 [95% CI: 0.5-2.4] for placebo, *p*=0.25 for EIC x treatment interaction). There was no increased risk of adverse outcome in the alteplase group after adjustment for baseline variables. However, the analysis was insufficiently powered to detect interactions with ORs less than 1.4 due to small subgroup sizes. This paper does not report the number of patients with symptomatic cerebral oedema but it provides reassurance that alteplase treatment did not markedly increase the risk of neurological deterioration at 24 h or of death at 90 days due to the presence of cerebral oedema at baseline compared with placebo.

Table 2: Association of selected baseline variables with extent of early ischaemic change* (Patel et al. 2001).

	Baseline Computed Tomography (CT) Scan Status						Comparison of Baseline Variables and EIC	
	EIC >1/3 MCA		EIC ≤1/3 MCA		No EIC			
	rt-PA	Placebo	rt-PA	Placebo	rt-PA	Placebo	ρ Value	*P* Value
Selected baseline variables, No.†	38	46	49	61	220	202		
Edema and mass effect, No. (%)	6 (16)	13 (28)	5 (10)	3 (5)	5 (2)	3 (1)	NA	.001
Edema, mass effect, and/or hyperdense artery sign (baseline CT finding), No. (%)	14 (37)	21 (46)	13 (27)	13 (21)	22 (10)	21 (10)	NA	.001
Diabetes, No. (%)	10 (27)	12 (26)	12 (24)	10 (16)	43 (20)	40 (20)	NA	.38
Presence of old lesion, No. (%)	9 (24)	4 (9)	13 (27)	21 (34)	60 (27)	62 (31)	NA	.03
Aspirin prior to treatment, No. (%)	14 (37)	7 (15)	18 (37)	24 (39)	95 (43)	57 (28)	NA	.11
Presumptive stroke type, No. (%)								
Small vessel	0 (0)	1 (2)	4 (8)	6 (10)	46 (21)	22 (11)		
Cardioembolic	24 (63)	27 (59)	21 (43)	26 (43)	88 (40)	84 (42)	NA	.001
Large vessel	14 (37)	17 (37)	23 (47)	28 (46)	79 (36)	89 (44)		
Other	0 (0)	1 (2)	1 (2)	1 (1)	7 (3)	7 (3)		
NIHSS, median (interquartile range)‡	19 (15-23)	17 (13-22)	15 (12-19)	16 (10-20)	12 (6-18.5)	14 (9-19)	0.23	<.001
Age, mean (SD), y	67 (14)	61 (12)	69 (10)	67 (13)	68 (11)	67 (11)	−0.06	.16
Admission mean arterial blood pressure, mean (SD), mm Hg	115 (16)	109 (16)	114 (18)	109 (19)	113 (18)	113 (17)	−0.04	.36
Time from stroke onset to baseline CT scan, mean (SD), min	88 (35)	97 (57)	89 (55)	81 (34)	80 (34)	78 (31)	0.11	.007

Key: *EIC (early ischaemic change) indicates loss of grey/white matter distinction, presence of hypodensity, or compresion of cerebrospinal fluid spaces; MCA, middle cerebral artery territory; NIHSS National Institutes of Health Stroke Scale; and NA, not applicable. *P* values computed from χ^2 tests for trend (binary variables) or Spearman rank correlation (ρ) (continuous variables) combining alteplase and placebo groups. †Selected baseline variables shown to be associated with at least 1 of the outcomes specified above. ‡Higher values indicate greater stroke severity.

The accuracy of the NINDS data presented in the Analysis 1.5 of the Cochrane reviews could not be corroborated as the primary data source(s) could not be located. The introduction of the review states that both published and unpublished data were used and that the data was verified with the principal investigators of all

major trials so it possible that unpublished data on symptomatic cerebral oedema was provided to the Cochrane reviewers.

Table 3: Baseline computed tomography scan status by treatment associated with clinical outcomes* (Patel et al. 2001).

							P Value	
	EIC >1/3 MCA		EIC ≤1/3 MCA		No EIC		EIC × Treatment Interaction‡	EIC Effect§
EIC Type†	rt-PA	Placebo	rt-PA	Placebo	rt-PA	Placebo		
No. of EIC findings	38	46	49	61	220	202		
Model 1: 3-month favorable outcome								
Unadjusted OR (95% CI)	0.5 (0.2-1.1)	0.5 (0.2-0.9)	0.9 (0.5-1.7)	0.8 (0.4-1.4)	2.1 (1.5-2.9)	1.0	.40	<.001
Adjusted OR (95% CI)	1.1 (0.4-2.5)	0.5 (0.2-0.9)	1.2 (0.6-2.5)	0.8 (0.5-1.5)	2.1 (1.5-3.1)	1.0	.52	.07
Rankin score = 0 or 1‖ Outcome, %	21	15	33	28	48	29		
Unadjusted OR (95% CI)	0.7 (0.3-1.5)	0.4 (0.2-1.0)	1.2 (0.6-2.3)	0.9 (0.5-1.8)	2.3 (1.5-3.4)	1.0	.41	<.001
Adjusted OR (95% CI)	1.7 (0.6-4.6)	0.4 (0.2-1.1)	1.7 (0.8-3.7)	1.0 (0.5-2.1)	2.4 (1.5-3.9)	1.0	.53	.18
NIHSS score = 0 or 1‖ Outcome, %	13	13	22	21	40	22		
Unadjusted OR (95% CI)	0.5 (0.2-1.5)	0.5 (0.2-1.3)	1.0 (0.5-2.1)	0.9 (0.5-1.9)	2.3 (1.5-3.5)	1.0	.21	.001
Adjusted OR (95% CI)	1.1 (0.4-3.6)	0.5 (0.2-1.4)	1.3 (0.6-3.0)	1.0 (0.5-2.2)	2.3 (1.4-3.8)	1.0	.59	.13
Barthel Index >95‖ Outcome, %	26	26	39	33	59	43		
Unadjusted OR (95% CI)	0.5 (0.2-1.1)	0.5 (0.2-1.0)	0.8 (0.4-1.6)	0.6 (0.4-1.2)	1.9 (1.3-2.8)	1.0	.42	<.001
Adjusted OR (95% CI)	0.8 (0.3-2.1)	0.4 (0.2-0.9)	1.1 (0.5-2.3)	0.6 (0.3-1.3)	2.1 (1.3-3.3)	1.0	.91	.003
Glasgow = 1‖ Outcome, %	24	17	35	34	51	34		
Unadjusted OR (95% CI)	0.6 (0.3-1.4)	0.4 (0.2-0.9)	1.0 (0.5-2.0)	1.0 (0.6-1.9)	2.0 (1.4-3.0)	1.0	.29	<.001
Adjusted OR (95% CI)	1.6 (0.6-4.1)	0.4 (0.2-1.0)	1.4 (0.7-3.1)	1.2 (0.6-2.4)	2.1 (1.3-3.4)	1.0	.32	.17
Model 2: deterioration at 24 hours Outcome, %	21	20	20	15	11	18		
Unadjusted OR (95% CI)	1.3 (0.5-3.0)	1.1 (0.5-2.5)	1.2 (0.5-2.6)	0.8 (0.4-1.8)	0.6 (0.3-1.0)	1.0	.22	.40
Adjusted OR (95% CI)	1.3 (0.6-3.1)	1.1 (0.5-2.4)	1.2 (0.6-2.7)	0.8 (0.4-1.9)	0.6 (0.4-1.1)	1.0	.25	.47
Model 3: 3-month lesion volume Median (interquartile range)	74 (41-159)	81 (29-181)	24 (8-89)	18 (6-121)	8 (1-64)	18 (2-75)		
Unadjusted	1.6 (1.3-1.9)	1.6 (1.3-1.9)	1.1 (1.0-1.4)	1.2 (1.0-1.4)	0.9 (0.8-1.0)	1.0	.68	<.001
Adjusted	1.2 (1.1-1.4)	1.4 (1.2-1.6)	1.1 (0.9-1.3)	1.1 (1.0-1.3)	0.9 (0.8-1.1)	1.0	.58	<.001
Model 4: death at 90 days Outcome, %	34	26	18	18	14	20		
Unadjusted OR (95% CI)	2.2 (1.0-4.7)	1.4 (0.7-3.0)	0.9 (0.4-2.0)	0.9 (0.4-1.9)	0.7 (0.4-1.1)	1.0	.28	.03
Adjusted OR (95% CI)	1.2 (0.5-2.9)	1.1 (0.5-2.6)	0.7 (0.3-1.6)	0.8 (0.4-1.8)	0.7 (0.4-1.3)	1.0	.82	.48

Key: *EIC (early ischaemic change) indicates loss of grey/white matter distinction, presence of hypodensity, or compresion of cerebrospinal fluid spaces; MCA, middle cerebral artery territory; NIHSS National Institutes of Health Stroke Scale; and CI, confidence interval. Adjusted ORs are in reference to the placebo-no EIC group. For a favourable 3month outcome, OR ≤1 indicates that patients in the EIC subgroup had lesser odds of a favourable outcome than those given placebo with no EIC; for deterioration and death, an OR >1 indicates that patients in the EIC subgroup had greater odds of having the event than patients given placebo with no EIC. †Model 1 includes baseline age, diabetes, NIHSS, admission mean arterial pressure (MAP), age x NIHSS, age x admission MAP. Model 2 includes aspirin use prior to randomisation. Model 3 includes old lesion volume, baseline NIHSS, baseline age, NIHSS x old lesion volume, presumptive stroke subtype (small and large vessel, cardioembolic), age x treatment, and time strata. Model 4 includes baseline age, diabetes and NIHSS. ‡After adjustment for the other variables in models 1-4.

§Adjusted based on models 1-4, including the EIC and treatment main effect, but not the EIC x treatment interaction term.

‖ The definitions of 3month favourable outcomes: values >1 indicate unfavourable outcomes for the modified Rankin Scale, NIHSS, and Glasgow outcome scale, and values <95 indicate unfavourable outcomes for the Barthel Index.

The NINDS study group did provide the following data for the Cochrane review: deaths from all causes during follow-up; and information on deaths or dependency at the end of follow-up. The primary publication of the NINDS study contains a survival analysis that appears to show early mortality before day 30 and the number of fatalities resulting from ICH within the first 36 h is stated. The Cochrane review states that the they *examined the number of deaths occurring between the first seven to 10 days and the end of follow-up in the 13 trials that provided data for both early and late deaths* so it would appear that the early mortality data was available but not provided. The only way to establish the true facts would if the Cochrane reviewers answered a request for further information.

It should be noted that the Alteplase Expert Working Group was informed by Professor Baigent that the results of the STT meta-analysis were qualitatively the same, although less robust, when the data from the NINDS trials was removed from the STT meta-analysis. The effect of alteplase on 90 day mortality, overall and by period of follow-up and effect on a good stroke outcome were reviewed.

The adverse IST-3 results (seen in the later 2014 review) on frequency of symptomatic (including fatal) oedema seem plausible; alteplase 4.5%, control 2.8%, odds ratio 1.64 (95% CI 1.12 to 2.40), they contrast with the NINDS results (Cochrane 2009 and 2014) suggesting 9.3% benefit; 56.4% alteplase, 65.7% placebo. Data from several of the other studies are also suspect. The time frame is missing in the title of Figure 5 in both reviews.

The incidence and outcome figures for symptomatic swelling of the original infarct in IST-3 have been published (The IST-3 collaborative group 2012). For the alteplase group, the number of patients with non-fatal symptomatic swelling was 21(1%) vs 17 (1%) on placebo (adjusted OR 1.23 [95% CI: 0.64-2.35, *p*=0.539]). The number of deaths due to symptomatic swelling were 47 (3%) for alteplase and 25 (2%) for placebo (adjusted OR 1.89 [95% CI: 1.14-3.14, *p*=0.013]). The total number of non-fatal and fatal cases due to symptomatic swelling of the original infarct were 68 (4%) for alteplase vs 42 (3%) for placebo (adjusted OR 1.66 [95% CI: 1.11-2.49, *p*=0.0.14]). A recent publication reported an adjusted OR 1.55 (95% CI: 1.17-2.06, *p*=0.002) for the association of early swelling on a baseline CT scan and death within 7 days using logistic linear regression analysis (The IST-3 collaborative group 2015).

Atlantis data are wrongly labelled as from Atlantis B. The small data set (n=61) matches a combined analysis of Atlantis A and B examining the within 3 hour window. (ATLANTIS 2002) This comprises 8.1% of the whole study (n=755). No time frame was given with original publication on the 3 hour data. (ATLANTIS 2002) The data cover only fatal events and thus are likely to be within 10 days.

The information provided by Dr [redacted] is factually correct as the paper reporting the combined analysis of ATLANTIS A and B for the 3 h window shows that 2 patients (5.3%) in the placebo group had fatal brain herniation without haemorrhage vs 1 on alteplase (4.3%, *p*=0.40) with no time frame specified. (Albers et al. 2002). Mortality at days 30 and 90 was a secondary outcome measure for ATLANTIS A (Clark et al. 2000) but was not a primary, secondary or additional outcome measure for ATLANTIS B (the safety analyses included overall mortality at 30 and 90 days; Clark

et al. 1999). Overall mortality data at 30 and 90 days is presented in both of the primary ATLANTIS publications.

ECASS I presented fatal oedema/herniation within 7 and '30 and 90 days'. Fatal cerebral oedema was worse with alteplase at all time points. No data on symptomatic oedema were presented. The mortality data seem plausible but the 7 day data are not those presented in the meta-analyses. (ECASS 1995)

ECASS I included 620 patients with moderate to severe hemispheric stroke syndromes (defined as a Scandinavian Stroke Scale (SSS) score of <50 total points, moderate to high-grade hemiparesis, sensory disturbance, dysarthria or nonfluent aphasia, and occasionally hemianopia) with none or minor early infarct signs on the baseline CT scan. Major EICs were defined as diffuse swelling of the affected hemisphere, parenchymal hypodensity, and/or effacement of the cerebral sulci in more than a third of the MCA territory. The dose of alteplase was 1.1 mg/kg (up to a maximum of 100 mg). The secondary endpoints included the mortality rate at 30 days and SSS scores on days 1 (2 h, 8 h, 24 h), 7 and 30 and NIHSS scores on days 1 and 90. Prespecified safety parameters included overall mortality and frequency of death related to space-occupying infarction. Follow-up CT scans were done at 24 h and between 6 and 8 days after stroke onset.

The most frequent protocol violation was the inclusion of patients with major EIC (n=31 alteplase vs n=21 placebo). Table 4 shows the mortality figures for cerebral oedema or herniation over time (patients randomised: n=313 alteplase; n=307 placebo). The death rates for cerebral oedema in the intention-to-treat population are: 3.8% (12/313) for alteplase until day 3 (vs 2.6% [8/307] placebo); 5.4% (17/313) for alteplase until day 7 (vs 4.2% [13/307] placebo); 7.3% (23/313) for alteplase until days 30 and 90 (vs 4.9% [15/307] placebo). The overall cerebral oedema death rates until days 30 to 90 are presented in the Cochrane analysis. This seems appropriate as no time limit is specified in the figure for Analysis 1.5 but this figure does not include cases of non-fatal symptomatic cerebral oedema.

Table 4: Further aspects of mortality: neurological death at different time points and causes of death (Hacke et al. 1995).

	Intention-to-Treat Population		Target Population		Protocol Violations	
Variable	Placebo	rt-PA	Placebo	rt-PA	Placebo	rt-PA
Until day 3						
Hemorrhage	5	12	5	8	0	6
Edema, herniation	8	12	4	7	4	5
Until day 7						
Hemorrhage	8	16	6	9	0	7
Edema, herniation	13	17	6	9	7	8
Until days 30 and 90						
Hemorrhage	7	19	7	10	0	9
Edema, herniation	15	23	8	13	7	10
Neurologic death						
Hemorrhage	7	19	7	10	0	9
Edema, herniation	15	23	8	13	7	10
Other neurologic	4	3	4	3	0	0
Nonneurologic death						
Vascular	6	7	6	6	0	1
Nonvascular	5	4	5	4	0	0
Undetermined	2	0	1	0	1	0
Total No. of deaths	39	56	31	36	8	20

Some ECASS I data on symptomatic cerebral oedema was reported in a subsequent publication using logistic regression analyses (Davalos et al. 1999). Early progressing stroke (defined as a decrease of ≥2 points in consciousness or motor power or a decrease of ≥3 points in speech scores in the SSS at the 24 h evaluation) and late progressing stroke (when 1 of these decreases occurred between the day 1 and 7 evaluations. Early progressing stroke occurred in 231 patients (38% of alteplase group vs 37% placebo, *p*=0.68). Focal hypodensity (OR 1.9 [95% CI: 1.3-2.9]) on the baseline CT scan was an independent prognostic factor for early progression. A multivariate analysis showed that extent of hypodensity >33% of the MCA territory (OR 2.5 [95% CI: 1.6-4.0]) and brain swelling (OR 1.8 [95% CI: 1.1-3.2]) on CT at 24 h were associated with early progression. Late progression of stroke occurred in 20.3% of patients (17% of alteplase group vs 23% placebo, *p*=0.111) and was independently predicted by older age, a low SSS score and brain swelling at admission. The authors concluded that early and late progression of acute stroke was related to baseline radiological evidence of cerebral oedema but that alteplase treatment did not influence the early clinical course.

The published ECASS 2 results are incomplete and inconsistent. A clear time frame and definitions are missing from the original paper (ECASS 2 1998). We are not told who died of cerebral oedema due to haemorrhage or who from oedema due to infarction. Time frames are also ambiguous or absent. Symptomatic cerebral oedema is not presented.

The inclusion/exclusion criteria for ECASS II were similar to those for ECASS I except patients with diffuse swelling of an entire hemisphere were excluded from ECASS I and those with brain swelling exceeding 33% of the MCA territory at baseline were excluded from ECASS II (Hacke et al. 1998). The secondary endpoints included change in stroke severity (NIHSS scale) from baseline to day 30 and further endpoints included the SSS score on day 90. Safety variables included mortality on days 30 and 90, haemorrhagic infarction, parenchymal haemorrhage as defined in ECASS I, and symptomatic haemorrhage. CT scans of the brain were scheduled for 22–36 h after study medication and on day 7. Symptomatic cerebral oedema is not defined and no data is presented.

The primary ECASS II publication states:

During the first 7 days, there were more deaths in the alteplase group than in the placebo group from intracranial haemorrhage alone (11 vs two) or from the combination of cerebral oedema and intracranial haemorrhage (seven vs two). Cerebral oedema was the commonest cause of death in the placebo group (n=17); this complication was found in eight of the alteplase-treated patients who died within the first 7 days. After day 7, the causes of death in the two groups were similar, and most were non-cerebral (cardiac arrest, pulmonary embolism, pneumonia).

So it appears that fatal cerebral oedema occurred in 4.3% of patients receiving placebo (n=391) during the study and in 2.0% of patients receiving alteplase (n=409) up to day 7. SICH was defined as blood at any site in the brain on the CT scan so it is assumed that the deaths due to combined ICH/oedema were counted as sICH

deaths. It is accepted that the mortality data is not presented as clearly as for ECASS I but the information is reported.

ECASS 3 results are incomplete as no time frame is given and there are no data on either fatal cases or on CT frequency of oedema. (ECASS 3 2008) 'Symptomatic edema' was added to the trial registry safety outcomes 3 months before printed publication. At the outset, 'cerebral oedema' on scan had been the registry outcome.

Patients in ECASS III were assessed with the NIHSS on days 1, 7, 30, and 90. In addition, the patients' clinical condition was closely monitored for the first 24 h (formally at 1 and 2 h) (Hacke et al. 1998). Initial assessments included a physical examination, CT or MRI, and an assessment of stroke severity using the NIHSS. The radiological exclusion criterion was a stroke involving more than one third of the MCA territory (using CT or MR imaging). CT or MRI was performed before treatment and at 22 to 36 h after treatment. Safety end points included overall mortality at day 90, any ICH, sICH, symptomatic oedema (defined as brain oedema with mass effect as the predominant cause of clinical deterioration), and other serious adverse events. In the ECASS III protocol, sICH was defined as any apparently extravascular blood in the brain or within the cranium that was associated with clinical deterioration, as defined by an increase of 4 points or more in the score on the NIHSS, or that led to death and that was identified as the predominant cause of the neurologic deterioration. The primary publication states:

The rate of symptomatic oedema did not differ significantly between the study groups: 6.9% in the alteplase group and 7.2% in the placebo group (29 patients in each group; odds ratio, 0.96; 95% CI, 0.56 to 1.64; p=0.88).

The total and sICH mortality data is presented although it is not stated how many fatalities were due to cerebral oedema. The timeframe is assumed to be 90 days as specified in the safety end points.

Symptomatic cerebral oedema is accepted as a more clinically relevant end point than cerebral oedema but the reason for changing the outcome cerebral oedema to symptomatic cerebral oedema is not known and only an answered request for this information from the ECASS III group would help.

This is the only brain scan data presented in the reviews looking at the expected radiological benefits from intravenous stroke thrombolysis, and so is pivotal. The true picture is likely to be provided by the adverse IST-3 result (n=3035) backed up by the adverse ECASS I (n=620) mortality data. The likely hazard has, in effect, been smothered in Cochrane Figure 5, 2014 – implying a non significant benefit with alteplase (odds ratio 0.97). This has been possible because of the high weighting provided by the outlying NINDS data, with large numerators, despite the IST-3 trial having 5 times more participants.

Dr [redacted] does not state what the expected radiological benefits from alteplase are. Infarct volumes are not reduced by alteplase as thrombolysis aims to restore flow to the ischaemic penumbra rather than to the infarct core (The IST-3 collaborative group 2015). ECASS I used a higher dose of alteplase than is currently licensed and

there was a high number of protocol violations related to the presence of major EICs at baseline. IST-3 largely reflects off-label use of alteplase.

I have re-calculated fatal and symptomatic (including fatal) cerebral oedema frequencies within 7 days using ECASS I and IST-3 data.

Fatal cerebral oedema due primarily to infarction within 7 days affects 3.5% of alteplase patients and 2.1% of controls. It is 1.4% worse with alteplase. Chi-squared 6.803, df=1, 2 tailed p=0.009.

Symptomatic (including fatal) cerebral oedema within 7 days is only presented in IST-3. (IST-3 2012) Frequencies are 4.5% with alteplase and 2.8% in controls. The difference is an adverse 1.7%. Chi squared 6.466, df=1, 2 tailed p=0.011.

These calculations appear correct.

Cerebral oedema requires active blood flow to develop so the frequency of symptomatic cerebral oedema after thrombolysis would be expected to be increased if alteplase is effective. The higher early mortality rate due to cerebral oedema in treated patients is not disputed. Section 4.4 of the SmPC for Actilyse states that reperfusion of the ischaemic area may induce cerebral oedema in the infarcted zone. It should be noted that ECASS I and IST-3 permitted thrombolysis within 6 h of stroke onset and the inclusion/ exclusion criteria used to select patients are not consistent with the information contained in the current SmPC.

The above criticisms are re-iterated and elaborated on in a submission from Dr [redacted] It would appear that a request for a detailed discussion of the review has been sent to the Cochrane Group, as is appropriate. The findings of the Cochrane reviews have not been used to support any regulatory decision-making.

1.4 International Stroke Trial (IST)-3

Cerebral oedema was classified as 'Neurological deterioration attributed to swelling of the initial ischaemic stroke' and defined as follows: 'In patients with relevant clinical deterioration, the presence of significant cerebral oedema (i.e. complete ventricular effacement, midline shift or obliteration of the basal cisterns) on a post-randomisation CT scan (or MR) performed within 7 days of randomisation.'. Both fatal and non-fatal outcomes were defined secondary end-points. (Sandercock 2008)

Fatal ischaemic cerebral oedema within 7 days occurred in 47 of 1515 (3.1%) thrombolysed patients and 25 of 1520 (1.6%) untreated patients, a 1.5% increase (95% CI 0.4 to 2.5%). Symptomatic (including fatal) ischaemic oedema affected 68 of 1515 (4.5%) treated and 42 of 1520 (2.8%) untreated stroke patients. This increase was 1.7% (95% CI 0.4 to 3.1%).

This information is correct. The primary publication contains the following information on symptomatic swelling of original infarcts at 7 days (table 5).

Table 5: Extract of table showing fatal and non-fatal cerebral and non-cerebral events within 7 days of randomisation (taken from table 3; The IST-3 collaborative group 2012).

	rt-PA (n=1515)	Control (n=1520*)	Adjusted analysis†		Absolute difference per 1000 (95% CI)‡
			Odds ratio (95% CI)	p value	
Cerebral events					
Symptomatic swelling of original infarct§					
Non-fatal	21 (1%)	17 (1%)	1·23 (0·64 to 2·35)	0·539	3 (−5 to 11)
Fatal	47 (3%)	25 (2%)	1·89 (1·14 to 3·14)	0·013	15 (4 to 25)
Total	68 (4%)	42 (3%)	1·66 (1·11 to 2·49)	0·014	17 (4 to 31)

In the 'Discussion' there is comment on the adverse effect on ischaemic cerebral oedema: 'We also anticipated a reduction in fatal and non-fatal neurological deterioration due to swelling of the initial infarct, so the clear 17 per 1000 excess was unexpected, and inconsistent with data from previous trials. (Wardlaw 2009)'

The Cochrane review (Wardlaw et al. 2009) also states:

There was a marginal reduction in symptomatic infarct swelling with thrombolysis which did not quite reach statistical significance: 15.7% of those allocated thrombolysis had symptomatic infarct swelling compared with 17.9% of those allocated control (OR 0.79, 95% CI 0.62 to 1.01, P=0.06) with no heterogeneity (I^2 = 34%, P = 0.18).

The I^2 statistic measures heterogeneity by assessing the differences in odds ratios between studies and we can see why that might result in a conclusion of no heterogeneity here (I^2=34%, *p*=0.18). There is much overlap in the confidence intervals for the odds ratio and no striking outliers in terms of OR. However if we consider the absolute percentages of events in each trial NINDS is a clear outlier – to the extent that it seems that something different must be being measured. In terms of odds ratio however, NINDS is not an outlier. When comparing studies in this way it is important to be confident that the studies are actually measuring the same thing and this large difference seems to suggest otherwise.

Wardlaw et al. (2014) go on to state:

There was no overall reduction in symptomatic infarct swelling with thrombolysis: 10.2% of those allocated thrombolysis had symptomatic infarct swelling compared with 10.4% of those allocated control (OR 0.97, 95% CI 0.79 to 1.19, P = 0.75) with significant heterogeneity (I^2 = 71%, P = 0.004). Due to the heterogeneity we undertook an analysis according to a random-effects model. This gave very similar results (OR 0.79, 95% CI 0.62 to 1.51, P = 0.88), and identical heterogeneity compared with the fixed-effect model.

Inclusion of IST-3 resulted in significant study heterogeneity consistent with methodological differences related to the inclusion/exclusion criteria, baseline patient characteristics, thrombolysis windows and definitions of sICH and cerebral oedema. Given these significant differences, there is no logical reason why the IST-3 study would be expected to replicate the results of earlier clinical trials. The authors of the

primary IST-3 publication do not discuss why their data is not consistent with that previously reported but one of the primary objectives of their trial was *to establish the balance of benefits and harms of thrombolytic therapy in patients who did not exactly meet the licence criteria (particularly elderly patients)* so previous knowledge derived from the licensed use of alteplase may not be applicable.

1.5 Meta-analysis of individual patient data from randomised trials (Emberson 2014).

Missing data are presented for the baseline measures. It looks reasonably robust. It is not presented for outcome data. It would not look robust. For example in Atlantis A, 57% of modified Rankin Score (mRS) data at 90 days are missing (alteplase investigators 2004). Many follow up brain scans may be missing in studies such as Atlantis – if only because of death. These scan numbers are often not presented in original publications. (ATLANTIS 1999, ATLANTIS 2000)

Hacke et al. (2004) state in the statistical analysis section of their pooled analysis of ATLANTIS, ECASS, and NINDS alteplase stroke trials:

We focused primarily on the 3month favourable outcome defined by three neurological function scores of modified Rankin Scale (0 or 1), Barthel Index (95 or 100), and NIHSS (0 or 1)… All trials were missing one or more of these outcome measures at 90 days for some patients. A conservative algorithm assigning outcomes based on measurement made earlier than 90 days for these patients was developed and applied to all investigations. If no measurements were available after baseline, the worst score for the modified Rankin Scale, NIHSS, or Barthel Index was assigned. The algorithm allowed all patients with known OTT to be included in the final intention-to-treat (ITT) analysis. 12, 11, 33, 37, 83, and 47 patients from NINDS part 1, NINDS part 2, ECASS I, ECASS II, ATLANTIS A, and ATLANTIS B, respectively, were missing one or more outcomes at 3 months based on the ITT algorithm and were given the worst outcomes. In the original report on the NINDS trials only one patient in part 1 and four in part 2 were reported as having missing outcomes. For the other 18 patients, outcomes after 90 days were available.

ATLANTIS A randomised 142 patients to treatment so 58% of subjects were missing one or more functional outcome data at 90 days but this was the smallest study contributing data to the pooled analysis. The amount of missing functional outcome data varied but was <10% for the other larger studies: NINDS part 1 (12/291 enrolled=4%), NINDS part 2 (11/333 enrolled=3%), ECASS I (33/620 included=5%), ECASS II (37/800 enrolled=5%), and ATLANTIS B (47/613 enrolled=8%). These trials seem better than many with missing data rates only in the 5% range. It would be better if we could see the missing data rates by treatment group, and the reasons for missing data. If many of them are deaths then it is actually quite an easy situation to handle (and questionable whether this even counts as missing data) – as it is clear that imputing the worst value for such patients is appropriate. We are not told the treatment allocation for these subjects with missing functional outcome data but those patients with missing data were assigned the worst outcomes in the statistical analysis.

The primary papers reporting the findings of the ATLANTIS A and B studies do not describe the extent of missing data for the primary or secondary outcome measures but the 'last observation carried forward' method with death given the worst outcome score on all of the measures was used for missing data (Clark et al. 1999 and 2000). The volume of cerebral infarction measured by CT scan at day 30 was a primary outcome measure for ATLANTIS A (no difference was seen on CT lesion volume at day 30, with both groups showing large variations: placebo 64 ±74 cm^3 versus alteplase 45 ±54 cm^3 [*p*=0.17]) and a secondary outcome for ATLANTIS B (no difference was seen on CT lesion volume at day 30, with both groups showing large variations: placebo 47 ±71 cm^3 versus alteplase 47 ±66 cm^3 [*p*=0.98]). The prognostic significance of the extent of MCA territory involvement on the baseline CT scan was reported in a subsequent publication that analysed a random sample of 50 scans of patients allocated to alteplase and placebo (Marks et al. 1999). According to the latest Cochrane review (Wardlaw et al. 2014), data on all 619 participants randomised in ATLANTIS B has not yet been presented, only the data from 547 patients randomised between 3 and 5 h has been reported.

In the absence of any conclusive information such as the final clinical study reports, it is not possible to conclude that many follow-up brain scans may have been lost in clinical trials.

Emberson et al. (2014) imputed missing data with rules that were prespecified in the statistical analysis plan (The Stroke Thrombolysis Trialists' Collaborative Group 2013). None of their results changed substantially according to the choice of imputation for missing data, including exclusion of those with missing data. This data was not shown in the primary publication.

The general approach employed regarding missing data seems reasonable, and it is reassuring that the results did not change substantially when different approaches were used.

The radiological methodology is not clear to either the Lancet reader or the reader of either of 2 protocols. (The Stroke Thrombolysis Trialists' Collaborative Group 2013, Emberson 2014) It is critically important.

A range of planned analyses are not presented in the final publication. We are not told why.

'Symptomatic ischemic brain edema (brain swelling associated with neurological deterioration by ≥ 4 NIHSS points)' is listed for assessment as a secondary outcome in the initial protocol but missing from the results. (The Stroke Thrombolysis Trialists' Collaborative Group 2013, Emberson 2014) Missing too is 'Early edema, effacement and/or midline shift'. (The Stroke Thrombolysis Trialists' Collaborative Group 2013, Emberson 2014)

The Stroke Thrombolysis Trialists' Collaborative Group's (2013) statistical analysis plan (SAP) describes the analyses that were agreed prior to becoming unblinded to the results from the IST-3. It includes the following 'other' secondary outcomes:

Further analyses will be done to assess the effect of allocation to alteplase on

- *Symptomatic ICH, defined using PH2 or PH2 with the SITS-MOST criterion of deterioration of ≥4 NIHSS points).*
- *Fatal ICH (PH2 and death within 7 days).*
- *Symptomatic ischaemic brain oedema (brain tissue swelling associated with neurological deterioration by ≥4 NIHSS points).*
- *Early oedema, effacement and/or midline shift.*

Time to event outcomes will be analysed using Cox proportional hazards regression, stratified by trial, with failure time set to time from randomisation to outcome. Where there are a sufficient number of events (at least 10 per predictor variable), the potential for effect modification will be assessed by the addition of interaction terms to the model.

The definition for symptomatic ischaemic oedema is similar to that used to define sICH in the ECASS studies. It is not explicitly stated that cerebral oedema must have been identified as the predominant cause of the neurological deterioration or if it includes patients with ICHs that were not type 2 parenchymal haemorrhages (PH2; Hacke et al. 1998). SICH was defined using PH2 or PH2 with the SITS-MOST criterion of a deterioration ≥4 NIHSS points (on the 22–36 h post-treatment imaging scan, combined with a neurological deterioration of 4 points or more on the NIHSS from baseline, or from the lowest NIHSS value between baseline and 24 h, or leading to death) (Wahlgren et al. 2007).

Emberson et al. (2014) state that a full description of the analyses is provided in the pre-specified SAP and that the key secondary outcomes were fatal ICH within 7 days, any sICH, and 90 day mortality (separated by cause where possible). The key secondary analyses given in the SAP were: effect of treatment allocation on death within 90 days; and the effect of treatment allocation on modified Rankin Scale so it would appear that the key secondary outcomes reported did not correspond to those pre-specified in the published SAP and that there was selective reporting of the 'other' secondary outcomes. No data relating to cerebral oedema is presented.

In a license application it would be seen as suspicious if a key endpoint was in the protocol and the data was collected yet not presented. The solution in those cases would be to ask for data on that endpoint. We may only speculate on the reason for excluding the results from the publication. The only way to really find out would be to ask for the results or an explanation from a member of the Stroke Thrombolysis Trialists' Collaborative Group (see below).

In a second protocol published after IST-3, two issues are prominent but contradictory. (Emberson 2014) Firstly, consensus is requested from 'ALL' the trialists to assist in 'silencing critics who have doubted the previous data.' Second and more concerning is the plan: 'BUT if agreement cannot be reached, then trials have the right to remove their data from an analysis.' This selective presentation of data is a key concern of those critical of the evidence. Again 'symptomatic ischemic

brain oedema and early oedema, effacement and/or midline shift' are both proposed as 'key' outcomes prior to analysis commencing.

[redacted]

A second protocol is available on-line (The Stroke Thrombolysis Trialists' Collaborative Group et al. 2014) at the University of Oxford's Clinical Trial Service Unit and Epidemiological Studies Unit website. It refers to the published version of the SAP and states in the background section:

...results from the IST-3 trial (and, if possible, the TESPI (Thrombolysis in Elderly Stroke Patients in Italy) trial will be included to help address several key questions, including:

1. *After what treatment delay is benefit (defined by modified Rankin Score [m RS] 0-1 at final follow-up at 3-6 months) lost or does harm begin, and do age or stroke severity modify the proportional effect of alteplase on stroke outcome?*

2. *What are the effects of alteplase on a range of other secondary outcomes, including: death within 90 days; symptomatic ICH, fatal s ICH, symptomatic ischaemic brain oedema and early oedema, effacement and/or midline shift*

The rationale section states:

Robust data from an updated individual patient meta-analysis would not only provide the highest level of evidence, but consensus from ALL the trialists would be enormously powerful in promoting a substantial increase in the appropriate use of alteplase and in silencing critics who have doubted the previous data.

And the publication policy section reads:

All publications will be in the name of the Stroke Thrombolysis Trialists' (STT) Collaboration, with the names of collaborators listed at the end of the paper. All collaborators will be expected to participate fully in manuscript preparation and editing, and will be expected to consult with, and collate comments from, colleagues from the trials they represent. Publications will be circulated for comments and approval before submission to peer review. The principles for agreeing the text of papers are that any such papers should:

- *Focus on conveying clear findings on which all trial groups are agreed, with controversial findings labelled as such;*

- *Where there is disagreement, the aim should be to moderate language to try to reach agreement; BUT*

- *If agreement cannot be reached, then trials have the right to remove their data from an analysis.*

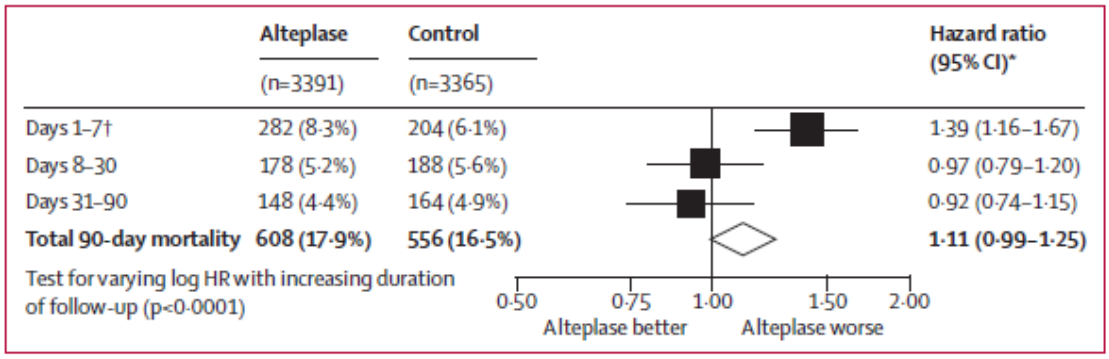
The published meta-analysis of individual patient data from randomised trials included all eligible completed randomised phase 3 trials of intravenous alteplase for the treatment of acute ischaemic stroke for which data were available (Emberson et al. 2014). It would appear that no trial data was removed from the analyses due to disagreement. Individual data were not made available to the Collaborative Group at the start of the project for five trials involving 270 participants (Haley et al. 1993; Mori et al. 1992; Wang et al 2003; Yamaguchi et al 1993; Hemmen et al. 2010).

The STT Collaborative Group has recently confirmed that they have an ongoing programme of secondary publications in progress although the data concerning cerebral oedema has not yet been explored (personal correspondence).

The IPD meta-analysis suggests 'Alteplase did not increase the risk of other early causes of death (ie, those other than intracranial haemorrhage)'. This is potentially misleading given the clearly adverse IST-3 and ECASS I cerebral oedema outcomes.

The quoted statement refers to figure 4 for justification. It states that there were 191(n=3391, 5.6%) deaths from other causes in the alteplase group versus 191 (n=3365, 5.7%) deaths in the control group. No *p* value is quoted.

Figure 4: Effect of alteplase on 90-day mortality by follow-up period (Emberson et al. 2014).



Patients can only contribute to a particular risk period if they have already survived any preceding periods. *Estimated by Cox proportional hazards regression stratified by trial (and adjusted only for treatment allocation). †Includes 91 versus 13 deaths caused by intracranial haemorrhage (with evidence of parenchymal haemorrhage type 2) and 191 versus 191 deaths from other causes.

1.6 Stroke registry data

No data on cerebral oedema was reported in the following large registry or cohort studies: Canadian Alteplase for Stroke Effectiveness Study (CASES; Hill et al. 2005); SITS-MOST and SITS-ISTR safety registries (Wahlgren et al. 2007 and 2008); Get With the Guidelines-Stroke Program (Fonarow et al. 2011); and the Canadian Stroke Network (Vergouwen et al. 2011).

The Multicenter alteplase Stroke Survey reported that early ischaemic changes affecting more than a third of the MCA territory were seen in 39 patients (4%) with no ICH (Tanne et al. 2002). The Standard Treatment with Alteplase to Reverse Stroke Study safety registry reported a 6% rate of cerebral oedema on the baseline CT scan

of 389 patients with a 2% rate of mass effect (defined as the presence of an acute hypodensity greater than one third of the MCA territory) related to alteplase treatment (n=389, Albers et al. 2000).

Strbian et al. (2013) reported the impact of cerebral oedema on the outcome of 943 patients receiving alteplase for ischaemic stroke at the Helsinki University Central Hospital from 1995 to 2008. Cerebral oedema was graded by neuroradiologists into 3 categories according to the SITS-MOST protocol (2002) on CT scans done at 24-72 h after thrombolysis:

CED-1 = focal brain swelling up to one third of the hemisphere
CED-2 = focal brain swelling greater than one third of the hemisphere
CED-3 = brain swelling with midline shift

Early infarct signs were defined as any of the following: hypoattenuation of less than a third of the MCA territory; obscuration of the lentiform nucleus; loss of basal ganglion outline; loss of insular ribbon; obscuration of the Sylvian fissure; or cortical sulcal effacement.

Table 6: Univariate and multivariable analyses testing associations between baseline characteristics and the development of cerebral oedema (CED) in thrombolysis-treated ischaemic stroke patients (Strbian et al. 2013).

	No edema	CED-1	CED-2	CED-3		Multivariable	
Parameter	*n* = 683	*n* = 167	*n* = 40	*n* = 53	*P*	OR (95% CI)	*P*
Age, years, median (IQR)	70 (60–77)	69 (59–77)	65 (57–76)	73 (60–78)	0·30	–	–
Males (%)	54·9%	50·3%	53·5%	55·0%	0·55	–	–
OTT, min, median (IQR)	115 (86–154)	132 (102–170)*	122 (87–167)	130 (98–153)	<0·01	1·01 (1·00–1·01)	<0·001
NIHSS, score, median (IQR)	8 (5–12)	13 (9–18)**	15 (11–18)**	18 (13–21)**	<0·001	1·15 (1·12–1·19)	<0·001
Early infarct signs (%)	25·1%	50·3%***	55·8%***	70·9%***	<0·001	2·20 (1·53–3·03)	<0·001
HCAS (%)	13·0%	28·0%***	37·2%***	47·3%***	<0·001	2·15 (1·44–3·20)	<0·001
BLOOD PRESSURE (BP), mmHg, median (IQR), or mean (±SD)							
Systolic BP prior to rtPA	155 ± 21	156 ± 22	149 ± 22	157 ± 24	0·34	–	–
Diastolic BP prior to rtPA	82 ± 14	83 ± 14	80 ± 16	84 ± 15	0·48	–	–
Systolic BP after rtPA	147 ± 21	148 ± 22	147 ± 22	155 ± 25	0·11	1·01 (0·99–1·01)	0·18†
Diastolic BP after rtPA	77 ± 14	79 ± 15	76 ± 14	81 ± 16	0·12	1·01 (1·00–1·02)	0·07†
LABORATORY PARAMETERS, median (IQR)							
Glucose, mmol/l	6·6 (5·7–7·8)	6·7 (5·9–8·0)	6·7 (5·7–7·7)	7·1 (6·4–8·5)	0·053	1·05 (0·98–1·12)	0·16
Leukocytes, E6/l	7·2 (6·0–8·7)	7·4 (6·0–8·8)	7·4 (6·1–9·5)	7·6 (6·3–8·9)	0·36	–	–
Platelets, E9/l	214 (178–256)	218 (178–254)	212 (178–278)	217 (168–258)	0·97	–	–
MEDICAL HISTORY (% of patients)							
Hypertension	58·5%	57·7%	48·8%	65·5%	0·43	–	–
Diabetes mellitus	13·5%	16·0%	7·0%	18·2%	0·37	–	–
Atrial fibrillation at base	16·8%	21·1%	32·6%*	27·3%*	0·015	1·09 (0·72–1·66)	0·68
Hyperlipidemia	39·3%	31·4%	32·6%	43·6%	0·18	0·91 (0·65–1·29)	0·61
Congestive heart failure	10·7%	13·7%	20·9%*	18·2%	<0·001	0·89 (0·54–1·46)	0·64

Key: *<0.05, **<0.001 in post-hoc analysis of ANOVA; †because of multicollinearity, these 2 variables were added to the model separately. CED, cerebral oedema; IQR, interquartile range; OTT, onset-to-treatment time; HCAS, hyperdense cerebral artery sign; NIHSS, National Institutes of Health Stroke Scale.

CED-1 was observed in 167 (17.7%) of patients, CED-2 in 40 (4.2%) and CED-3 in 53 (5.6%) and cerebral oedema was already present on the 24 h CT scans of 95% of patients with it. The presence of early infarct signs on the baseline CT scan were independently associated with the development of cerebral oedema but the report does not explicitly detail the severity or extent of ischaemic changes at baseline. Forty nine patients with cerebral oedema were treated with glycerol, mannitol, hypertonic saline and 3 required neurosurgical intervention (2 decompressive craniectomy and 1 ICH evacuation). A multivariate model adjusted for known stroke

prognostic variables showed the associations of cerebral oedema with poor 3 month outcomes (mRS 3-6) and mortality versus those for patients without cerebral oedema (table 7). Progressive brain swelling was associated with poorer functional outcomes and increased mortality: the ORs for mortality at 3 months varied from 2.78 (95% CI: 1.48-5.23, p=0.001) for focal brain swelling up to one third of the hemisphere (CED-1) to 14.81 (95% CI: 6.40-34.27, p<0.001) for brain swelling with midline shift (CED-3) ; and the ORs for a 3month mRS of 3-6 varied from 1.60 (95% CI: 1.04-2.47, p=0.031) for focal brain swelling up to one third of the hemisphere (CED-1) to 18.96 (95% CI: 5.03-71.53, p<0.001) for brain swelling with midline shift (CED-3). When compared to patients without cerebral oedema (n=683), several baseline parameters were independently associated with the development of cerebral oedema: increasing baseline NIHSS score; increasing onset-to-treatment times; presence of a hyperdense cerebral artery sign; or early infarct signs on CT (table 6). The distribution of 3 month mRS scores among the subtypes of cerebral oedema are shown and compared to those with sICH (defined using ECASS-II criteria) and cerebral

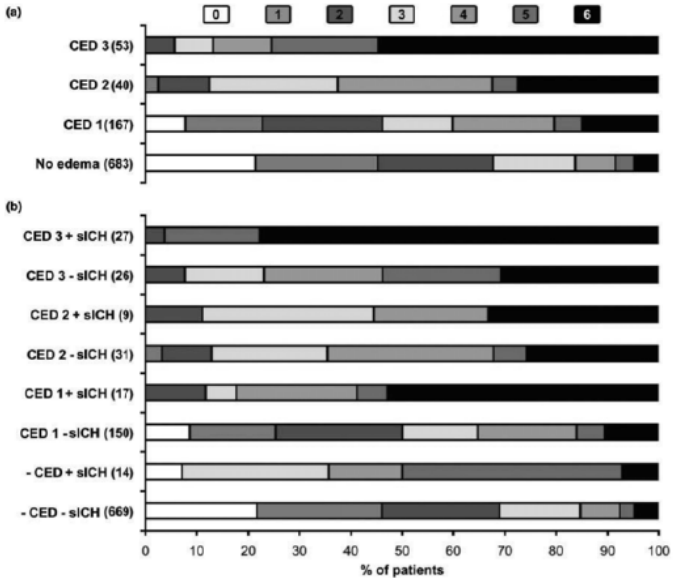
Table 7: Multivariable model testing associations of types of cerebral oedema (CED) with poor 3 month outcome (m RS 3-6) and with mortality (Strbian et al. 2013).

	Three-month mRS 3–6		Mortality	
	OR (95% CI)	P	OR (95% CI)	P
HCAS	1·64 (1·04–2·58)	0·032	0·71 (0·37–1·38)	0·32
Baseline glucose level	1·13 (1·06–1·21)	<0·001	1·11 (1·01–1·22)	0·028
Onset-to-treatment time	1·00 (1·00–1·01)	0·075	1·00 (0·99–1·00)	0·36
Baseline NIHSS score	1·18 (1·14–1·22)	<0·001	1·12 (1·06–1·17)	<0·001
mRS above 1 on admission	2·69 (1·41–5·15)	0·003	2·82 (1·32–6·02)	0·007
Age	1·07 (1·05–1·09)	<0·001	1·08 (1·05–1·11)	<0·001
No edema	Reference category		Reference category	
CED-1	1·60 (1·04–2·47)	0·031	2·78 (1·48–5·23)	0·001
CED-2	12·08 (4·18–34·90)	<0·001	9·09 (3·74–22·05)	<0·001
CED-3	18·96 (5·03–71·53)	<0·001	14·81 (6·40–34·27)	<0·001

Key: m RS modified Rankin Scale; HCAS, hyperdense middle cerebral artery sign; NIHSS, National Institutes of Health Stroke Scale.

oedema in figure 5. Cerebral oedema was present in 28% of patients after thrombolysis and severe forms were seen in 9.8%. These findings are similar to those reported in untreated cohorts of ischaemic stroke patients (Hacke et al. 1996) but the incidence of cerebral oedema appears higher than in the reported randomised clinical trials or safety registries. It should be noted that none of the median NIHSS scores for the cerebral oedema groups were >25 at baseline, there were small numbers of patients in the sub-groups with the more severe forms of oedema and we are not told how many patients had ischaemic changes exceeding a third of the MCA territory at baseline. The median onset-to-treatment times were also significantly longer than those for patients without cerebral oedema (table 6). The authors noted that higher baseline NIHSS scores may reflect larger areas of infarct, but longer onset-to-treatment times and the presence of hyperdense cerebral artery sign are consistent with reperfusion injury after thrombolysis that could worsen cerebral oedema.

Figure 5: (a) Distribution of three month modified Rankin Scale among patients with cerebral oedema. (b) Cerebral oedema plus symptomatic intracerebral haemorrhage (Strbian et al. 2013).



However, this registry report did not assess the following patient characteristics: size of cerebrospinal fluid spaces, infarct size, or the extent of perfusion-diffusion mismatch on multimodal MR imaging.

1.7 Conclusions

A number of concerns have been raised by Dr [redacted] has identified a number of inconsistencies in the analysis and reporting of cerebral oedema data following alteplase therapy.

The results from a single historical nonclinical study of reperfusion in baboons (Bell et al. 1985) may have limited relevance to alteplase and are inconsistent with data from other studies and subsequent human safety data.

There are a number of factual discrepancies and methodological issues in the Cochrane reviews of thrombolysis for acute ischaemic stroke (Wardlaw et al. 2009 and 2014). However, regulatory agencies would have had access to the complete datasets for any pivotal studies that were the basis for initial licensing of alteplase and subsequent amendments.

The Stroke Thrombolysis Trialists' Collaborative Group have not reported all of the key secondary outcomes described in the specified statistical analysis plan for their meta-analysis of individual data from RCTs involving alteplase. Further information has been received.

The most specific sign of significant cerebral oedema after an ischaemic stroke is a reduced conscious state resulting from disruption of the ascending reticular projections in the brainstem and thalamus by swelling (Wijdicks et al. 2014). The clinical features of a large MCA infarction include: hemiplegia, global or expressive

dysphasia, severe dysarthria, neglect, gaze preference, visual field defect and cerebral ptosis. The initial NIHSS score reflects stroke severity and infarct volume and is often ≥20 for dominant hemisphere infarcts and ≥15 for nondominant hemisphere anterior circulation strokes in patients that develop massive ischaemic oedema with hemispheric stroke (Kreiger et al. 1999). The NIHSS score increases with worsening cerebral oedema. Krieger et al. (1999) reported mean [±SD] baseline NIHSS scores of 20.5 [±3.4] in 23 cases of cerebral oedema due to anterior circulation stroke that progressed to coma then death (vs 20.5 ±3.7 in control patients with anterior circulation strokes of similar severity); rising to 31.8 [±7.4] at 48 h in 19 cases [vs 18.3 ±6.8 in controls, $p<0.0001$]).

There is no validated clinical feature that reliably measures the level of consciousness after stroke but randomised controlled trials assessing the efficacy of decompressive surgery for the treatment of malignant MCA infarction have used a score of ≥1 on the level of consciousness item (1a) of the NIHSS scale to detect neurological deterioration due to cerebral oedema (Juttler et al. 2007; Vahedi et al. 2007) or a reduction in the Glasgow coma scale (Hofmeijer et al. 2009). Both scales are in routine clinical use in the UK. The Scandinavian Stroke Scale (SSS) also contains an item to assess consciousness. Neurological deterioration due to cerebral oedema usually occurs within 24 to 96 h after stroke onset although some patients may worsen at 4 to 10 days.

There are a number of predictors for the development of cerebral oedema after ischaemic stroke: severe stroke severity measured on the NIHSS, delayed treatment; radiological factors include hypoattenuation on the baseline CT scan within the first 6 hours, involvement of more than a third of the MCA territory, the presence of a hyperdense artery sign or midline shift ≥5 mm within the first 2 days and a DWI volume of ≥80 ml on MR imaging (Wijdicks et al. 2014).

The randomised controlled trials of alteplase have mainly defined neurological deterioration as an increase in the NIHSS scale of ≥4 and permitted unscheduled repeat CT imaging if clinically needed to detect symptomatic haemorrhagic transformation/ICH. All acute deaths are recorded. It would appear that the design of these studies would detect the development of acute symptomatic (including fatal) cerebral oedema and the functional outcome measures would capture subsequent recovery. It is possible that thrombolysis may be associated with an increased risk of developing cerebral oedema due to reperfusion injury; however, it is not established that haemorrhagic transformation is an extreme form of cerebral oedema. The available imaging, mortality and functional data do not indicate that the benefit-risk balance of alteplase is consistently or substantially altered after analysis of the specific data related to cerebral oedema.

The National Stroke Guideline recommends a neuroscience service delivering neurosurgical interventions should be commissioned to manage major intracerebral haemorrhage, malignant cerebral oedema, and hydrocephalus (The Intercollegiate Working Party for Stroke 2012). The European Stroke Organization ischaemic stroke Guidelines recommend surgical decompressive therapy within 48 h after symptom onset in patients up to 60 years with evolving malignant MCA infarction and that osmotherapy can be considered prior to surgery if necessary.

The product licence (Summary of Product Characteristics, SmPC) for Actilyse does not specifically discuss the risk of cerebral oedema but states the following:

- it should be started as early as possible within 4.5 h of stroke onset (section 4.2)
- severe stroke as assessed clinically (eg NIHSS>25) and/or by appropriate imaging techniques and prior stroke within the last 3 months are contraindicated (section 4.3)
- thrombolytic treatment requires adequate monitoring and alteplase should only be used by trained and experienced physicians with the facilities to monitor that use (section 4.4).
- with later time-to-treatment from onset of stroke symptoms the net clinical benefit is reduced and may be associated with a higher risk of ICH and death compared to patients treated earlier (section 4.4)
- patients with very severe stroke are at higher risk for intracerebral haemorrhage and death and should not be treated. Patients with extensive infarctions are at greater risk of poor outcome including severe haemorrhage and death. In such patients, the benefit/risk ratio should be thoroughly considered (section 4.4).
- reperfusion of the ischaemic area may induce cerebral oedema in the infarcted zone (special warning in section 4.4).

Cerebral oedema is not listed in section 4.8 as an adverse drug reaction.

The current SmPC wording adequately describes the risk factors that are associated with the development of clinically significant cerebral oedema and contraindicates delayed use in patients with severe strokes who are at the highest risk of progressing to cerebral oedema.

1.8. Points for discussion

Regarding cerebral oedema, the Group is asked whether:

- any of the issues discussed in this paper have implications for public safety
- any of the issues discussed in this paper have implications for the current authorised indication for alteplase in acute ischaemic stroke
- is the risk of cerebral oedema after alteplase adequately covered in the SmPC? Do any of the issues raised, individually or together, have implications for the authorised indication for alteplase in acute ischaemic stroke?

1.9. References

Albers GW, Bates VE, Clark WM, Bell R, Verro P, Hamilton SA. Intravenous tissue-type plasminogen activator for treatment of acute stroke: the Standard Treatment with Alteplase to Reverse Stroke (STARS) study. JAMA 2000; 283:1145-50.

Albers GW, Clark WM, Madden KP, Hamilton SA. ATLANTIS trial: results for patients treated within 3 hours of stroke onset. Stroke 2002; 33:493-5.

Bai J, Lyden PD. Revisiting cerebral postischemic reperfusion injury: new insights in understanding reperfusion failure, hemorrhage, and edema. Int J Stroke 2015; 10:143-52.

Bandera E, Botteri M, Minelli C, Sutton A, Abrams KR, Latronico N. Cerebral blood flow threshold of ischemic penumbra and infarct core in acute ischemic stroke: a systematic review. Stroke 2006; 37:1334-9.

Bell BA, Symon L, Branston NM. CBF and time thresholds for the formation of ischemic cerebral edema, and effect of reperfusion in baboons. J Neurosurg 1985; 62:31-41.

Brott TG, Haley EC Jr, Levy DE, et al. Urgent therapy for stroke. Part I. Pilot study of tissue plasminogen activator administered within 90 minutes. Stroke 1992; 23:632-40.

Clark WM, Wissman S, Albers GW, Jhamandas JH, Madden KP, Hamilton S. Recombinant tissue-type plasminogen activator (Alteplase) for ischemic stroke 3 to 5 hours after symptom onset. The ATLANTIS Study: a randomized controlled trial. JAMA 1999; 282: 2019-26.

Clark WM, Albers GW, Madden KP, Hamilton S. The rtPA (alteplase) 0- to 6-hour acute stroke trial, part A (A0276g): results of a double-blind, placebo-controlled, multicenter study. Thrombolytic therapy in acute ischemic stroke study. Stroke 2000; 31:811-6.

Cook DJ, Tymianski M. Nonhuman primate models of stroke for translational neuroprotection research. Neurotherapeutics 2012; 9:371-9.

Crowell RM, Olsson Y, Klatzo I, Ommaya A. Temporary occlusion of the middle cerebral artery in the monkey: clinical and pathological observations. Stroke 1970; 1:439-48.

Crowell RM, Marcoux FW, DeGirolami U. Variability and reversibility of focal cerebral ischemia in unanesthetized monkeys. Neurology 1981; 31:1295-302.

Davalos A, Toni D, Iweins F, Lesaffre E, Bastianello S, Castillo J. Neurological deterioration in acute ischemic stroke: potential predictors and associated factors in the European Cooperative Acute Stroke Study (ECASS) I. Stroke 1999; 30:2631–6.

Del Zoppo GJ, Copeland BR, Anderchek K, Hacke W, Koziol JA. Hemorrhagic transformation following tissue plasminogen activator in experimental cerebral infarction. Stroke 1990; 21:596-601.

Emberson J, Lees KR, Lyden P, et al.; Stroke Thrombolysis Trialists' Collaborative Group. 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. Lancet 2014; 384:1929-35.

Fonarow GC, Smith EE, Saver JL, et al. Timeliness of tissue-type plasminogen activator therapy in acute ischemic stroke: patient characteristics, hospital factors, and outcomes associated with door-to-needle times within 60 minutes. Circulation 2011; 123:750-8.

Hacke W, Kaste M, Fieschi C, et al. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS). JAMA 1995; 274:1017-25.

Hacke W, Schwab S, Horn M, Spranger M, De Georgia M, von Kummer R. 'Malignant' middle cerebral artery territory infarction: clinical course and prognostic signs. Arch Neurol 1996; 53:309-15.

Hacke W, Kaste M, Fieschi C, et al. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators.Lancet 1998; 352: 1245-51.

Hacke W, Donnan G, Fieschi C, et al.; ATLANTIS Trials Investigators; ECASS Trials Investigators; NINDS rt-PA Study Group Investigators. Association of outcome with early stroke treatment: pooled analysis of ATLANTIS, ECASS, and NINDS rt-PA stroke trials. Lancet 2004; 363:768-74.

Hacke W, Kaste M, Bluhmki E, et al.; ECASS Investigators. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med 2008; 359: 1317-29.

Haley EC Jr, Brott TG, Sheppard GL, et al. Pilot randomized trial of tissue plasminogen activator in acute ischemic stroke. The TPA Bridging Study Group. Stroke 1993; 24:1000-4.

Hemmen TM, Raman R, Guluma KZ, et al.; ICTuS-L Investigators. Intravenous thrombolysis plus hypothermia for acute treatment of ischemic stroke (ICTuS-L): final results. Stroke 2010; 41:2265-70.

Hill MD, Buchan AM; Canadian Alteplase for Stroke Effectiveness Study (CASES) Investigators. Thrombolysis for acute ischemic stroke: results of the Canadian Alteplase for Stroke Effectiveness Study. CMAJ 2005; 172:1307-12.

Huttner HB, Schwab S. Malignant middle cerebral artery infarction: clinical characteristics, treatment strategies, and future perspectives. Lancet Neurol 2009; 8:949–58.

Jones TH, Morawetz RB, Crowell RM, et al. Thresholds of focal critical ischemia in awake monkeys. J Neurosurg 1981; 54:773-82.

Juttler E, Schwab S, Schmiedek P, et al.; DESTINY Study Group. Decompressive Surgery for the Treatment of Malignant Infarction of the Middle Cerebral Artery (DESTINY): a randomized, controlled trial. Stroke. 2007; 38:2518–25.

Krieger DW, Demchuk AM, Kasner SE, Jauss M, Hantson L. Early clinical and radiological predictors of fatal brain swelling in ischemic stroke. Stroke 1999; 30:287-92.

Marks MP, Holmgren EB, Fox AJ, Patel S, von Kummer R, Froehlich J. Evaluation of early computed tomographic findings in acute ischemic stroke. Stroke 1999; 30:389-92.

Numan T, Bain AR, Holland RL, Smirl JD, Lewis NC, Ainslie PN. Static autoregulation in humans: a review and reanalysis. Med Eng Phys 2014; 36:1487-95.

Ogawa A, Mori E, Minematsu K, et al.; MELT Japan Study Group. Randomized trial of intraarterial infusion of urokinase within 6 hours of middle cerebral artery stroke: the middle cerebral artery embolism local fibrinolytic intervention trial (MELT) Japan. Stroke 2007; 38:2633-9.

Patel SC, Levine SR, Tilley BC, et al; National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Lack of clinical significance of early ischemic changes on computed tomography in acute stroke JAMA 2001; 286:2830-8.

Sandercock P, Lindley R, Wardlaw J, et al.; IST-3 Collaborative Group. The third international stroke trial (IST-3) of thrombolysis for acute ischaemic stroke. Trials 2008; 9: 37. doi: 10.1186/1745-6215-9-37

Simard JM, Kent TA, Chen M, Tarasov KV, Gerzanich V. Brain oedema in focal ischaemia: molecular pathophysiology and theoretical implications. Lancet Neurol 2007; 6:258-68.

Safe Implementation of Treatments in Stroke Monitoring Study (SITS-MOST). Final study protocol dated 18.12.2002). www.acutestroke.org/SM_Protocol/SITS-MOST_final_protocol .pdf, accessed 25 May 2015.

Strbian D, Meretoja A, Putaala J, Kaste M, Tatlisumak T; Helsinki Stroke Thrombolysis Registry Group. Cerebral edema in acute ischemic stroke patients treated with intravenous thrombolysis. Int J Stroke 2013; 8:529-34.

Symon L, Branston NM, Strong AJ, Hope TD. The concept of thresholds of ischemia in relation to brain structure and function. J Clin Pathol Suppl (R Coll Pathol) 1977; 11:149-54.

Tanne D, Kasner SE, Demchuk AM, et al. Markers of increased risk of intracerebral hemorrhage after intravenous recombinant tissue plasminogen activator therapy for acute ischemic stroke in clinical practice: the Multicenter rt-PA Stroke Survey. Circulation 2002; 105:1679-85.

The Intercollegiate Working Party for Stroke. National clinical guidelines for stroke – fourth edition. London: Royal College of Physicians, 2012.

The IST-3 collaborative group. The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]: a randomised controlled trial. Lancet 2012; 379:2352-63.

The IST-3 collaborative group. Association between brain imaging signs, early and late outcomes, and response to intravenous alteplase after acute ischaemic stroke in the third International Stroke Trial (IST-3): secondary analysis of a randomised controlled trial. Lancet Neurol 2015; 14:485-96.

The NINDS t-PA Stroke Study Group. Intracerebral hemorrhage after intravenous t-PA therapy for ischemic stroke. Stroke 1997; 28:2109-18.

The Stroke Thrombolysis Trialists' Collaborative Group. Details of a prospective protocol for a collaborative meta-analysis of individual participant data from all randomized trials of intravenous rt-PA vs. control: statistical analysis plan for the Stroke Thrombolysis Trialists' Collaborative meta-analysis. Int J Stroke 2013; 8: 278-83.

The Stroke Thrombolysis Trialists' Collaborative Group. A collaborative meta-analysis of individual participant data from all randomised trials of intravenous rt-PA versus control. The Stroke Thrombolysis Trialists' (STT Collaboration) Protocol. 2014. www.ctsu.ox.ac.uk/ research/meta-trials/stt-1/2014-07-14STTmainprotocol.pdf, accessed 25 May 2015.

Vahedi K, Vicaut E, Mateo J, et al.; DECIMAL Investigators. Sequential-design, multicenter, randomized, controlled trial of early Decompressive Craniectomy in Malignant Middle Cerebral Artery Infarction (DECIMAL Trial). Stroke 2007; 38:2506–17.

Vergouwen MD, Casaubon LK, Swartz RH, et al.; Investigators of the Registry of the Canadian Stroke Network. Subtherapeutic warfarin is not associated with increased hemorrhage rates in ischemic strokes treated with tissue plasminogen activator. Stroke 2011; 42:1041-5.

Wahlgren N, Ahmed N, Davalos A, et al.; SITS-MOST investigators. Thrombolysis with alteplase for acute ischaemic stroke in the Safe Implementation of Thrombolysis in Stroke-Monitoring Study (SITS-MOST): an observational study. Lancet 2007; 369:275-82.

Wahlgren N, Ahmed N, Davalos A, et al.; SITS investigators. Thrombolysis with alteplase 3-4.5 h after acute ischaemic stroke (SITS-ISTR): an observational study. Lancet 2008; 372:1303-9.

Wang SY, Wang XL, Zeng H, et al. [Early intravenous thrombolysis with recombinant tissue plasminogen activator for acute cerebral infarction]. Zhonqquo Wei Zhong Bing Ji Jiu Yi Xue 2003; 15:542-5.

Wardlaw JM, Murray V, Berge E, del Zoppo GJ. Thrombolysis for acute ischaemic stroke. Cochrane Database Syst Rev 2009; 4:CD000213.

Wardlaw JM, Murray V, Berge E, del Zoppo GJ. Thrombolysis for acute ischaemic stroke Cochrane Database Syst Rev 2014; 7:CD000213.

Wey HY, Kroma GM, Li J, Leland MM, Jones L, Duong TQ. MRI of perfusion-diffusion mismatch in non-human primate (baboon) stroke: a preliminary report. Open Neuroimag J 2011; 5:147-52.

Wey HY, Duong TQ. Multimodal MRI of nonhuman primate stroke. Transl Stroke Res 2012; 3:84-9.

Wijdicks EF, Sheth KN, Carter BS, et al.; American Heart Association Stroke Council. Recommendations for the management of cerebral and cerebellar infarction with swelling; a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2014; 45:1222-38.

Yamaguchi T, Hayakawa T, Kiuchi H. Intravenous tissue plasminogen activator ameliorates the outcome of hyperacute embolic stroke. Cerebrovasc Dis 1993; 3:269-72.

Annex 1: [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

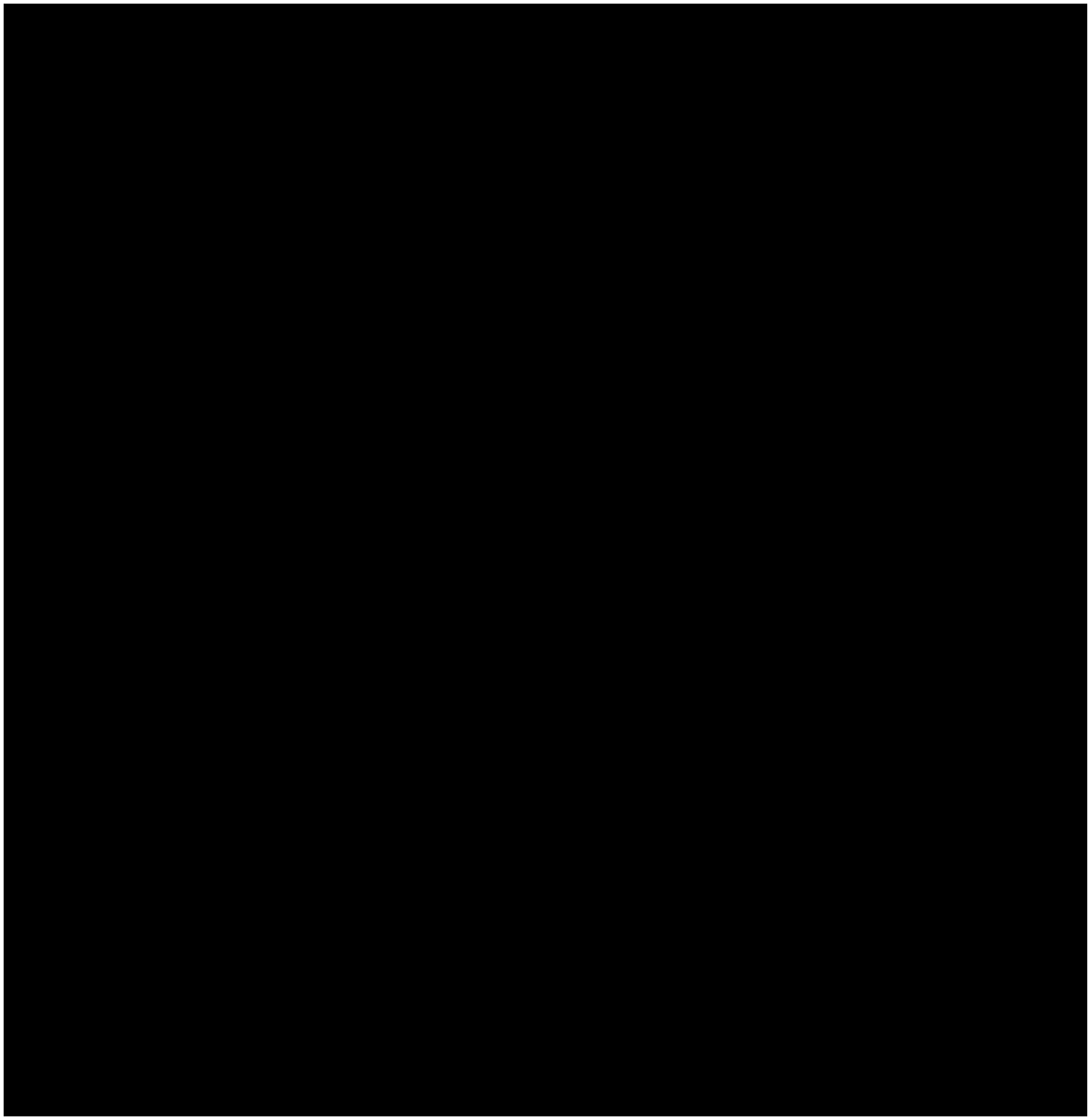
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]