COMMISSION ON HUMAN MEDICINES

Title of paper: Conclusions and recommendations of the Expert Working Group on alteplase used in the treatment of acute ischaemic stroke **Type of paper: For advice**

Product: Actilyse 10, 20, 50mg	**Assessors:** Medical assessor: Dr [redacted] Scientific assessors: Dr [redacted] Dr [redacted] Statistical assessor: Dr [redacted] Epidemiological assessors: Dr [redacted] Dr [redacted]
MAHs: Boehringer Ingelheim Limited	**Previous Assessments:** CHM May 2014 EWG: Nov 2014, Jan 2015, June 2015
Active constituents: Alteplase (rt-PA)	**Legal status:** **POM**
Therapeutic classification: Antithrombotic agent, ATC code B01AD02	

1. Introduction

rt-PA is a recombinant human tissue-type plasminogen activator (t-PA). It is produced by expression of the human gene for t-PA in CHO cells. The mechanism of action of rt-PA is understood to be the enzymatic cleavage of plasminogen to plasmin with subsequent increase in fibrinolysis. In the indication of acute ischaemic stroke the recommended dose is 0.9 mg rt-PA/kg body weight (maximum of 90 mg) infused intravenously over 60 minutes with 10% of the total dose administered as an initial intravenous bolus.

The authorised indication in fibrinolytic treatment of acute ischaemic stroke specifies that:

> "treatment must be started as early as possible within 4.5 hours after onset of stroke symptoms and after exclusion of intracranial haemorrhage by appropriate imaging techniques (e.g. cranial computerised tomography or other diagnostic imaging method sensitive for the presence of haemorrhage). The treatment effect is time-dependent; therefore earlier treatment increases the probability of a favourable outcome."

Product information also includes a negative benefit:risk statement for administration beyond the 4.5 hour window, and a further reminder that treatment must be started as early as possible within the 4.5 hours.

In May 2014 the Commission on Human Medicines (CHM) considered an MHRA assessment of new data and specific concerns that had been communicated to the MHRA regarding the use of rt-PA in the treatment of acute ischaemic stroke. CHM concluded that the data presented did not change the favourable balance of benefits and risks for rt-PA, which remains an effective medicine for the treatment of acute ischaemic stroke. However, in order to be assured that all relevant sources of evidence had been taken into consideration, CHM advised that an expert working group should be set up.

The Terms of Reference for the group were agreed by the CHM in June 2014 and endorsed, with a minor amendment (as shown), at the first EWG meeting in November 2014, as follows:

The Expert Working Group on rt-PA will:

- review all sources of evidence on ***efficacy and*** safety of alteplase in clinical use in ischaemic stroke
- advise whether these data have implications for the benefit:risk of alteplase in clinical use for the treatment of ischaemic stroke
- consider whether further measures are necessary to minimise harm in stroke patients
- advise on a communication strategy

This paper provides a summary of the evidence examined by the EWG and its conclusions and recommendations for CHM's consideration. The agreed minutes from the first and second meetings and draft minutes from the third meeting are provided separately.

2. Data considered by the Expert Working Group

The rt-PA expert working group has met three times, in November 2014, January 2015 and in June 2015. The evidence considered by the group and the MHRA's evaluation of it was provided in a series of papers which are summarised below. The individual papers are provided as further annexes to this paper.

Paper 1: Introduction to the papers and background to the current situation

CHM is aware that the MHRA had been contacted by [redacted] who had concerns regarding the balance of benefits and risks of rt-PA when used in the treatment of acute ischaemic stroke. D[redacted] view that an updated independent evaluation was necessary was supported by the [redacted] nd a number of other influential physicians. rt-PA treatment of acute ischaemic stroke has generated debate and strong views since the indication was first approved, and continues to do so within the medical community.

CHM considered an assessment of the data that had become available since the time-window for treatment was extended from 3 hours to 4.5 hours post-symptom onset and the specific concerns that had been raised (in May 2014). CHM concluded that the evidence presented did not affect the favourable balance of benefits and risks for rt-PA. However in order to be assured that all relevant information sources had been taken into consideration, the CHM advised that an expert working group should be set up.

Paper 1A: Regulatory history of rt-PA use in acute ischaemic stroke

rt-PA was first approved in the UK for fibrinolysis in coronary artery occlusion and massive pulmonary embolism in 1988. It was authorised via the European mutual recognition procedure, with Germany acting as the lead member state (Reference Member State, RMS).

In 2002, rt-PA received a conditional licence in the indication of treatment of acute ischaemic stroke from 0-3 hours following symptom onset. The clinical trials that were identified as relevant at this time were NINDS part 1 and 2 (NINDS, 1995), ECASS I (Hacke *et al*, 1995) and II (Hacke *et al*, 1998) and ATLANTIS A (Clark *et al*, 2000) and B (Clark *et al*, 1999). The NINDS part 2 study was the pivotal trial in the application. Several member states expressed divergent opinions during this procedure, and the UK expressed a negative view. The extension of indication was eventually granted following a lengthy arbitration procedure, on condition that the company conducted:

- A further randomised placebo-controlled trial (ECASS III) to assess efficacy and safety within 3-4 hours of symptom onset [later revised to 3-4.5 hours]
- A post-marketing surveillance study (SITS-MOST)

In addition, Periodic Safety Update Reports (PSURs) were to be submitted every 6 months for two years, and then annually for three years.

The ECASS III trial had a positive outcome for rt-PA treatment between 3-4.5 hours of symptom onset. This result was considered by CHMP to indirectly confirm the positive balance of benefits and risks of rt-PA between 0-3 hours of symptom onset, because of the understanding that efficacy diminishes with increasing time to onset of treatment. CHMP also concluded that the SITS-MOST surveillance study results demonstrated that rt-PA can be used safely in an experienced clinical setting.

In 2012, the treatment time-window for rt-PA was extended from 0-3 hours to 0-4.5 hours post-symptom onset. The main data supporting this extension was the ECASS

III trial, with some supporting data from the SITS-ISTR registry and a pooled analysis conducted by the MAH. The UK raised Major Objections during this procedure with a particular concern related to the apparent increase in death rates in the rt-PA group compared with the placebo group in the ECASS III trial. This was resolved when the MAH explained that data on deaths had been collected for unequal lengths of follow-up in the two groups which, when corrected, demonstrated that rt-PA did not have an adverse effect on mortality when given up to 4.5 hours after symptom onset.

Both the initial application for the indication in the treatment of acute ischaemic stroke and the variation to extend the time-window for treatment up to 4.5 hours were extensively discussed both at CHM and within Europe. All data that were available and considered relevant at these times were comprehensively reviewed in the European procedures, and all of the MHRA/CHM's concerns were addressed and resolved via the appropriate means, which included further data submissions/analyses and oral explanations by the company at the meetings of the Committee for Medicinal Products for Human Use (CHMP, previously CPMP).

Paper 2: Stroke care in the UK and a wider perspective since 2000

This paper describes the changes in stroke care in the UK that have taken place during the current and last decade and the impacts of these changes on morbidity and mortality of stroke patients. The paper also considers whether there is evidence for a learning curve within stroke centres and the imaging techniques used in the diagnosis of acute stroke patients and whether there is evidence to support any change to the current product information or clinical guidelines in this respect. The Group was asked to consider whether the introduction of rt-PA for the treatment of stroke had had any noticeable impact on stroke outcome in the UK.

In 1995, before rt-PA was available for the treatment of stroke, a national stroke programme was started by the Royal College of Physicians, which set standards of care for all stroke patients in England, Wales and Northern Ireland. At that time the quality of care of stroke patients in the UK was considered to be poor, with marked regional variations. Specialist care in stroke units was not routinely available despite evidence to show that this reduces death rates and increases proportion of patients able to live independently.

The first National Sentinel Stroke Audit (NSSA) was conducted in 1998, with the first national clinical guideline prepared by the Intercollegiate Stroke Working Party in 2000. The approval of rt-PA for the treatment of acute ischaemic stroke within 3 hours of symptom onset in 2002 resulted in the need for the public to understand the symptoms of stroke and seek medical treatment urgently. A major reconfiguration of stroke services was also necessary so that hospitals could diagnose and treat patients within 3 hours of the onset of symptoms.

Access to thrombolysis has improved since rt-PA was first licensed, with ~11% of all stroke patients now receiving rt-PA, and up to 20% in specialist hyperacute stroke units.

In light of advances in radiological diagnostic techniques for stroke the evidence was evaluated to determine if CT scanning, which is currently recommended in the Summary of Product Characteristics (SmPC), remained the optimal diagnostic method. Evidence suggests that CT imaging is universally accessible, tolerable, quick to perform and excludes haemorrhagic stroke with almost 100% sensitivity. Where uncertainty remains following CT scanning, diffusion weighted MRI imaging techniques would be more sensitive, for example in those with minor strokes.

Paper 3: Usage of rt-PA in acute ischaemic stroke

This paper evaluates the level of use of rt-PA in the UK and more widely, including off-label use.

The British Association for Stroke Physicians (BASP) mandated recording of all thrombolysis patients in the SITS (Safe Implementation of Treatments in Stroke) register in the UK. The number of patients entered into the SITS registry per year increased to a level of ~2000-2500 in 2009 and has remained stable since. Data from the National Sentinel Stroke Clinical Audit indicates that overall ~11% of stroke patients are treated with rt-PA in the UK.

Data extracted from SITS UK between January 2012 and July 2014 indicates that ~70% of patients were treated within the terms of the marketing authorisation, as defined by treatment within 4.5 hours of onset of symptoms in patients aged up to 80 years. Approximately 29% of patients were aged over 80 years (off-label), and ~2% were treated outside of the 4.5 hour time window. This use in patients aged over 80 years, whilst contraindicated, is in line with the current national clinical guidelines

Paper 4: Benefits and risks: new study data

This paper summarises the main clinical trial data that supported the initial approval of the acute ischaemic stroke indication and the extension to the time-window to 4.5 hours post-symptom onset. It also summarises the findings of a re-analysis of the NINDS trial, the data from the IST-3 trial, a meta-analysis by Emberson et al (2014), and further observational data.

Initial approval of the indication:

NINDS part 2 was the pivotal randomised controlled trial supporting the initial licensing application. The primary endpoint was clinical outcome at 3 months, a global measure encompassing scores on the Barthel index (BI), modified Rankin scale (mRS), Glasgow outcome scale (GOS) and NIHSS.

The odds ratio for a favourable outcome (minimal or no disability at 3 months) in the rt-PA group compared with placebo was 1.7 (95% CI [1.2-2.6]). Symptomatic intracranial haemorrhage (ICH) within 36 hours of stroke onset occurred in 6.4% of rt-PA treated patients vs. 0.6% of placebo patients. This did not translate into a significant increase in mortality. Overall mortality at 3 months in the rt-PA group was 17%, vs. 21% in the placebo group (p=0.30).

The NINDS part 1 trial was the first double-blind randomised trial to be conducted with rt-PA in humans. This trial did not reach its primary endpoint of improvement in neurological outcome after 24 hours (improvement in NIHSS by 4 or more points or complete resolution of the deficit), RR 1.2, 95% CI [0.9-1.6]. However a benefit of rt-PA was observed for the global outcome at 3 months following treatment OR 2.1, 95% CI [1.3-3.2].

The ECASS I and II and ATLANTIS A and B studies also failed in their primary endpoints.

A number of differences between the trials were identified as potential reasons for the differences in the results relative to NINDS part 2, including that:

- the primary outcome at 24 hours in NINDS part 1 has since been concluded to be less clinically relevant than the day 90 outcomes (the primary outcome measured in NINDS part 2)

- the dose used in the ECASS I trial was greater than that used in NINDS and which is now the licensed dose (1.1mg/kg body weight compared with 0.9mg/kg body weight) and the primary endpoint measure was different. Whilst there were a significantly greater number of haemorrhagic infarctions (defined as bleeding in the infarcted area but without space occupying effect) in the placebo (n=93, 30.3% 95%CI [25.3-35.8]) than in the rt-PA group (n=72, 23.0% 95%CI [18.5-28.1]), there was a significantly greater number of more severe haemorrhages - parenchymal haemorrhages (bleeding with mild or significant space occupying effect) in the rt-PA group (n=62, 19.8% 95%CI [15.6-24.8]) than in the placebo group (n=20, 6.5% 95%CI [4.1-10.0]) ($p<0.001$).
- the time-window for treatment was 0-6 hours in the ECASS I and II trials, and in the ATLANTIS A and B (amended to 0-5 hours, and then 3-5 hours) compared with 0-3 hours in NINDS.

Extension of the time-window for treatment up to 4.5 hours:

The main study supporting the extension of the time-window to 4.5 hours was the ECASS III trial. The study inclusion and exclusion criteria mirrored the EU SmPC apart from the time-window for treatment, which was set at 3-4.5 hours (compared with the licensed time-window 0-3 hours). The primary outcome was mRS 0-1 at day 90. The odds ratio for a favourable outcome in the rt-PA group compared with placebo was 1.34; 95% CI [1.02-1.76]. A total of 113 patients (27%) in the rt-PA group had intracranial haemorrhages of which 3 were fatal. This compared with 71 patients (17.6%) with ICH in the placebo group of which 0 were fatal.

Re-analysis of the NINDS trials:

This reanalysis was conducted by a review committee that had been established 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". In addition to baseline imbalance the committee reviewed a number of other issues including risk of ICH, effect of onset to treatment time, centre effect and likelihood of a favourable outcome, stroke subtype and effect of rt-PA and efficacy in patients with diabetes mellitus.

Overall, the committee's findings were that despite an increased incidence of sICH in rt-PA treated patients, when rt-PA was administered according to the study protocol there was a statistically significant and clinically important benefit of treatment compared with placebo, measured by an adjusted odds ratio of 2.1, 95% CI [1.5-2.9] for a favourable outcome (using the global outcome measure) at 3 months. The analysis was adjusted for trial centre, time to treatment, study part (1 or 2), age, baseline NIHSS, diabetes and pre-existing disability.

IST-3 trial:

The IST-3 trial was a large randomised open label study with an initial double-blind phase (n=276; total number of patients included in the trial n=3035). Patients were enrolled according to the uncertainty principle, i.e. if a patient had either a clear indication for treatment with rt-PA or for whom benefit-risk of treatment would clearly be negative, the patient was not entered into the trial – patients were only included if treatment was considered promising but unproven, and therefore most patients were treated outside of the EU licensing conditions. Treatment with rt-PA was administered between 0-6 hours, and the primary outcome was the proportion of patients alive and independent at 6 months, OHS 0-2.

The trial failed in its primary outcome. In the rt-PA group at 6 months follow-up, 554 (37%) of patients were alive and independent in activities of daily living (OHS 0-2) compared with 534 (35%) in the control group (adj OR=1.13, 95% CI: 0.95-1.35, p=0.181). A secondary ordinal analysis, treating OHS 0, 1, 2, 3 as distinct outcomes and OHS 4, 5, and 6 as a combined outcome, found a positive outcome (OR=1.27, 95% CI [1.10-1.47]). The increase in mortality observed at 7 days in rt-PA treated patients was not apparent at 6 months post-treatment, with 27% death rate in both rt-PA treated and control patients.

Meta-analysis (Emberson et al, 2014):

The Emberson et al meta-analysis is the most up to date individual patient data meta-analysis available for rt-PA treatment of acute ischaemic stroke. This included data from 9 trials (ATLANTIS A/B, ECASS I/II/III, EPITHET, IST-3, NINDS part 1 and 2), and a total of 6756 patients, 3391 treated with rt-PA.

The primary efficacy outcome was modified Rankin Score (mRS) 0-1 at 3-6 months post-stroke. A significant benefit of rt-PA was observed when given within 3 hours of stroke (adj OR 1.75, 95% CI: 1.35-2.27) and within 3-4.5 hours (adj OR 1.26, 95% CI: 1.05-1.51).

For the overall group, including patients treated >4.5 hours after symptom onset, the risk of sICH was increased with treatment compared with controls, both at 36 hours post-stroke (adj OR=6.67, 95% CI; 4.11-10.84) and at 7 days (adj OR=5.55, 95% CI: 4.01-7.70). The risk of fatal ICH within 7 days was also significantly raised with rt-PA (adj OR=7.14, 95% CI: 3.98-12.8). All-cause mortality within 90 days was numerically increased but not statistically significantly greater in the rt-PA group compared with controls (adj OR=1.11, 95% CI: 0.99-1.25).

Observational data:

The latest data from the SITS-ISTR international registry included patients treated between 2002-2011 and compared patients treated <3 hours with patients treated 3-4.5 hours after the onset of stroke symptoms. The data suggested that the rate of sICH was comparable in patients treated 3-4.5 hours after treatment compared with those treated within 3 hours. However, the 3-month outcomes of no or minimal disability and independence were less favourable for those treated within 3-4.5 hours compared with those treated within 3 hours in the adjusted analyses and significant for 'no or minimal disability': OR 0.90 95%CI [0.82-0.98] and OR 0.92 95%CI [0.83-1.01] respectively.

Interim data from the SITS-UTMOST cohort was presented, including data to November 2013 (see Paper 4A).

The Get With The Guidelines-Stroke registry in the US included data from 58,353 patients treated with rt-PA within 4.5 hours following stroke onset. Data were collected between 2003-2012. Faster onset to treatment time, in 15-minute increments, was associated with reduced in-hospital mortality (OR, 0.96; 95% CI, 0.95-0.98; $P < 0.001$), reduced sICH (OR, 0.96; 95% CI, 0.95-0.98; $P < 0.001$), increased achievement of independent ambulation at discharge (OR, 1.04; 95% CI, 1.03-1.05; $P < 0.001$), and increased discharge to home (OR, 1.03; 95% CI, 1.02-1.04; $P < 0.001$).

The Bade-Wuerttemberg Stroke registry in Germany included 84,439 patients treated between 2008-2012. Unlike the other registries this one included both patients treated with rt-PA (n=10,263, 12.2%) and those not thrombolysed. After adjustment for baseline characteristics, treatment with rt-PA was found to be associated with an increased chance of mRS 0-1 at discharge compared with untreated patients, overall OR 1.70, 95%CI [1.59-1.81], p<0.0001).

The Canadian Alteplase for Stroke Effectiveness Study compared patients treated with rt-PA within 3 hours of stroke onset with those treated between 3-4.5 hours. A total of 1112 patients were included, of which 129 (11.6%) were treated between 3 and 4.5 hours, the rest of the patients were treated at <3 hours. At 90 days, 39.4% of patients treated between 3 and 4.5 hours had achieved mRS 0-1, compared with 36.5% of patients treated <3 hours, a difference that was not statistically significant (adjusted RR 0.98, 95%CI [0.8-1.2]). The size of this study is small, and in particular the number of patients treated between 3 and 4.5 hours is very small (n=129).

Paper 4A: Benefits and risks: new study data - Addendum 1

This paper assesses the final report of the SITS-UTMOST registry. The SITS-UTMOST European registry included patients treated with rt-PA up to 4.5 hours post-symptom onset, and compared patients treated between 0-3 hours with those treated between 3-4.5 hours. Multivariate analysis of the prospective cohorts showed a statistically significant lower adjusted odds ratio for functional independence (mRS 0-2) at 3 months in the 3-4.5 hour cohort compared with the <3 hour cohort (OR 0.81, 95%CI [0.67-0.99], p=0.044). The difference between the time windows for mRS 0-1 was not significant, however (OR 0.87, 95%CI [0.72-1.05], p=0.159).

There was no statistical difference in rate of sICH (using definitions from SITS-MOST, ECASS II or NINDS) or in all-cause mortality at 3 months between the prospective cohorts treated 3-4.5 hours following stroke onset and within 3 hours.

The paper concluded that the final data from the SITS-UTMOST registry demonstrate that treatment with thrombolysis between 3 and 4.5h after acute ischaemic stroke is of similar safety and efficacy compared with treatment within 3h treatment, however efficacy outcomes were more favourable when treatment was initiated within 3 hours of symptoms. These findings are consistent with other data on time-to-onset of treatment. The data suggest that there may be room for improvement of hospital management times and therefore centres should put every effort into improving hospital management of stroke.

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

This paper addresses a wide range of specific concerns that have been raised with MHRA in submissions from Dr [redacted] and Professors Fatovich and Brown. It also discusses the definitions of symptomatic intracerebral haemorrhage used in clinical trials, the choice of primary endpoint and analysis method, and the evidence for the impact of this on the outcome of studies. Endpoint analysis is also discussed following receipt of a submission from Dr Mandava.

CHM considered a paper in May 2014 which included a number of initial concerns raised by Dr [redacted], and concluded that the data presented did not change the favourable balance of benefits and risks for rt-PA. Since this time, Dr [redacted] has submitted additional concerns and we thought it may be helpful to include a brief description of the concerns raised and the MHRA's assessment of them below:

- *the initial trials in animals:* concerns were raised regarding the conduct and design of one animal study with rt-PA.

  This study was published in 1985 and was intended as a proof of concept study. A number of other animal studies have been conducted and their findings were also supportive of a possible role for rt-PA in treatment of stroke. It would not be reasonable to retrospectively apply the present-day

standards for a non-clinical submission to studies conducted prior to the current regulations.

- *data with streptokinase:* concerns were raised regarding the less favourable findings in trials of streptokinase in stroke compared with rt-PA, on the basis that both are thrombolytics.

  Streptokinase has many different properties compared with rt-PA including a lack of fibrin-specificity, a longer half-life, accumulation of fibrinogen-degradation products and high antigenicity. In addition there were differences in the design of the streptokinase clinical trials compared with rt-PA. For these reasons the results with these two thrombolytics are not considered to be interchangeable. Streptokinase is not licensed for the treatment of stroke.

- *specific clinical trials:* concerns were raised relating to all aspects of specific clinical trials, it is not possible to reflect all of these in this summary. However the most important of these included:

  a) NINDS, and the NINDS reanalysis:

  - change of primary outcome between part 1 and 2

    Using a different primary outcome for a second trial based on interim results from a first trial is not a concern, provided the second trial is initially analysed as an independent trial and is positive on its own;

  - modest size of the trial

    Size was based on a sample size calculation and the trial results were positive, and therefore by definition acceptably powered;

  - local not central randomisation

    An FDA analysis of the effect of randomisation errors on the study results found that they would have had no impact. Given that time to treatment needed to be minimised and that the trial was conducted in the early 1990's when communication methods were less efficient, it is perhaps understandable why local randomisation was used;

  - distorted spread of onset to treatment time (50% of patients enrolled in the 0-90 minute time window were treated between 89-90 minutes).

    This may be a reflection on the setting of the trial and investigators focussing on which time category the subject fell in to. Alternatively, investigators may have been rushing to meet the 90 minute cut-off. This distortion does not raise concerns about results or study conduct, but there is little value in analyses using time to randomisation as a continuous variable;

  - primary outcome of part 1 was not significant

    The primary outcome in part 1 had a trend to early improvement, and the secondary endpoint at day 90 was positive – the latter endpoint is considered to be of greater clinical relevance, and part 2 of the trial was positive for this outcome.

  - outcome by centre, smaller centres did not confirm the results from larger centres

    Subgroup analyses by centre did not raise concerns and the largest centres did not always have the largest treatment effects. Confidence intervals between centres had considerable overlap.

- no replication of NINDS

  Whilst replication of the NINDS findings in other trials would enhance confidence in the results, it is not always necessary (as demonstrated by the existence of the CHMP guideline on applications with one pivotal trial) and many products are 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.

- the NINDs re-analysis committee had narrow terms of reference

  The terms of reference covered whether eligible stroke patients may not benefit from rt-PA when given according to the trial protocol, and whether subgroup imbalance in baseline stroke severity invalidates the entire trial. This reflects the main issue debated at the time of the re-analysis, i.e. the impact of the imbalance in baseline stroke severity between the two arms. The reanalysis committee also reviewed several other specific issues including onset to treatment time, centre effect, stroke subtype etc.

b) ECASS III:

- industry funded and managed

  ECASS III was a licensing commitment following approval of the indication and as such was industry funded and managed

- final place of residence not assessed or reported

  This was not defined as an endpoint in the ECASS III trial. Whilst the utility of this parameter as a surrogate for outcome is clear, it has limitations. It is a fairly crude measurement as there will be a full spectrum of care in patients who have returned home, from no help required at all to full time care. Similarly some patients may choose to move to a residential home following stroke but this may be unrelated to the outcome of stroke.

- slow recruitment led to change of inclusion criteria

  The time-window for treatment in the trial was widened from 3-4 hours to 3-4.5 hours, based on the publication of a pooled analysis. Provided the population included is still clinically relevant and the change is not based upon unblinded data from the ongoing trial, this type of change is not considered to be of concern.

- imbalance in previous stroke, rt-PA 7.7% vs. control 14.1%

  In subanalyses the treatment effect of rt-PA was greater in patients with prior stroke, and so this imbalance did not appear to bias the result in favour of rt-PA.

- severe disability (mRS 5) was increased by 2.9% with rt-PA

  The percentage of patients who died was slightly lower (1.5%) in the rt-PA group compared with placebo, and the group with favourable outcome, mRS 0-1, was 7.3% higher in the rt-PA group. This issue was also evaluated at the time of the extension to the time-window and 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 a pooled analysis]*, and that the overall net effect of rt-PA was positive.

- endpoint mRS 0-2 shows no benefit and is a more appropriate stratification

  mRS 0-1 and mRS 0-2 are both valid endpoints; both can be justified and mRS 0-1 was pre-specified as the primary endpoint. Although statistical significance is not seen for the mRS 0-2 endpoint, the odds ratio is very similar to mRS 0-1 and the trend clearly favours rt-PA.

c) IST-3:

- 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. Although the pooling improved the result, it is difficult to criticise as it was pre-specified, and neither of the two is clearly better than any other

- overemphasis of the positive secondary endpoint in the presentation of the results

  The publication does not emphasise the negative primary outcome, however, these data are provided. These criticisms do not impact the results obtained.

- lack of improvement in long term mortality, 18 month follow-up of IST-3 trial

  The relatively small improvement in mRS observed in the IST-3 trial may have been insufficient to result in a positive effect on mortality at 18 months. Whilst 18 months follow-up is much longer than most clinical trials, the cohort studies in a publication by Slot et al (2008) that demonstrated worsening survival with increasing stroke severity provide follow-up of between 7 and 19 years; at the 18 month time-point the differences in survival are harder to discern.

- recall bias, due to the open nature of the study

  The data support a hypothesis that recall is associated with response in both treatment groups and that rt-PA causes both an increase in favourable outcome rate (albeit a small one) and an increased recall rate)

Other general issues discussed included:

- *the impact of baseline imbalances on the results of the key studies*

  The concern that baseline imbalance in stroke severity in the NINDS trial may have been the driver for benefit is an issue that has been highlighted many times over the years since the trial was published. This concern was one of the main issues addressed by the reanalysis committee of NINDS data in 2004 (see above). In NINDS there were more patients with NIHSS scores 0-5 in the rt-PA group than the placebo group but sub-group analyses broken down across the baseline NIHSS quintiles demonstrated that the overall benefit of rt-PA was not due solely to the baseline imbalance, and this was confirmed by the committee using covariate adjusted analyses. A matched analysis by Mandava et al also confirmed the positive result of NINDS. Concerns were subsequently raised in a graphical analysis by Hoffman and

Schriger which suggested that the imbalance in stroke severity at baseline was likely responsible for most if not all of the treatment difference in NINDS. However, their analysis focussed on change from baseline (NIHSS) which will be confounded if there is initial confounding by baseline severity. Thus, patients with mild stroke cannot improve post treatment as much as patients with severe stroke and therefore a 'change from baseline' analysis will be biased against the group with more mild strokes at baseline. By contrast, the methods used by the NINDS re-analysis committee take account of the baseline imbalance more effectively. Additional plots of individual results by baseline severity provided by Hoffman and Schriger also achieve this and are consistent with the findings of the NINDS re-analysis committee.
There was also an imbalance of baseline stroke severity in the ECASS III trial favouring the rt-PA group (NIHSS 10.7 ±5.6 vs. 11.6 ±5.9). 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). Analyses excluding patients with baseline NIHSS ≤9 or ≥20 were still in favour of rt-PA.

- *appropriateness of endpoints used in clinical trials*

  Concerns have been raised that day 30 mortality was not examined in the Cochrane reviews or the pooled analyses despite being used widely outside of stroke. It is generally recognised that the longer the duration of follow-up in a clinical trial the better, particularly where the disease under study is chronic, may have long-term consequences, or the patient is expected to change substantially in the weeks/months after the event. Rate of recovery from stroke is usually highest in the first few weeks, and functional improvement may continue for many months/years and an assessment made at 90 days would likely provide a better indication of patient outcome than an assessment at 30 days. The CPMP Points to Consider guidelines (2001) recommend a study duration of 3 months for pivotal trials in acute stroke.

- *loss of blinding (risk of bias)*

  Concerns have been raised regarding the potential loss of blinding due to expectations that the active vials would foam upon reconstitution as a result of the protein content, whilst placebo vials would not. The composition of the placebo used in the NINDS part 1 and 2, ATLANTIS A and B and ECASS II and III trials has been provided by the MAH and potential for foaming has been investigated by NIBSC (paper 5B). No difference in appearance of rt-PA vials and placebo vials were observed, with foaming occurring as a result of the inclusion of [redacted] in both the active and placebo vials.

- *little good evidence to support the 'time is brain' hypothesis*

  Concerns have been raised that a lack of evidence to support the 'time is brain' hypothesis in terms of the efficacy of rt-PA, which would undermine rt-PA as a treatment for acute stroke. The lack of evidence for this effect in the NINDS trial relates to the distorted pattern of onset to treatment time distribution (discussed above) such that the data are insufficient to address this question. The fact that IST-3 did not demonstrate the expected trend in efficacy with time to treatment may have been due to reduced power due to the decrease in enrolment to the trial (6000 patients originally planned, 3035 enrolled). The small numbers of patients enrolled in the 0-3 hour time-window in the ECASS II and ATLANTIS A and B trials precludes any conclusion regarding the effect of time on outcomes. The effect of time to onset of treatment can be clearly seen in the results of the Emberson et al meta-analysis (2014).

- *radiological findings*

  Concerns have been raised that treatment with rt-PA did not demonstrate a significant reduction in infarct size (ATLANTIS and EPITHET studies). However 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.

Paper 5A: Additional submissions received from interested parties

This paper discusses the submission of a (then unpublished) article by Dr Alper which concluded that the balance of benefits and risks of rt-PA used between 3-4.5 hours after onset of stroke is supportive of a recommendation against such use outside of clinical trials.

The authors' main concerns related to:

- the level of recommendation for rt-PA by the 2009 AHA/ASA guidelines, which gave a Class I recommendation with a level of evidence rating of B.

  It was considered that this issue relates to the understanding and interpretation of the Class I recommendation vs. the level of evidence grading, and to the perspective on evidence for harm. Alper et al focus on the comparison between 0-3 hours vs. 3-4.5 hours, whilst arguably a more relevant comparison is treated vs. untreated patients within the 3-4.5 hour time window.

- concerns about balance of benefits and risks in the 3-6 hour time-window in the 2014 Cochrane review.

  This time window is not considered relevant to the licence, because it has previously been concluded that the balance of benefits and risks after 4.5 hours is negative.

- an expectation that the addition of the IST-3 trial data for 3-4.5 hours to the 2010 individual patient data meta-analysis should yield an OR close to 1 and should not be statistically significant, and instead the result was significantly in favour of rt-PA treatment (OR 1.26, 95%CI [1.05-1.51]).

  This apparent anomaly relates to i) the pre-specified reclassification of the time-windows for treatment in the IST-3 trial in the meta-analysis (from 'time to randomisation' to 'time to treatment' in order to make it comparable with the other included trials), and ii) the difference in the primary outcome (OHS 0-2) used in the IST-3 trial compared with the primary outcome used in the meta-analysis (OHS/mRS 0-1).

Paper 5B: Additional information on individuals' concerns on specific aspects of the supporting clinical evidence

This is a report of the investigation by NIBSC of the appearance of the reconstituted rt-PA and placebo (see above). A video demonstrating the reconstitution was also provided to the Group. The investigation of the appearance of the reconstituted rt-PA and placebo is reassuring in terms of the potential for unblinding of the clinical trials, as both formulations led to foaming.

The paper also provides further information on the arginine excipient in rt-PA, which had been suggested to be hazardous via an effect on nitrous oxide. However, the

effects of arginine are complex 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. It was concluded that there are data in favour of both a positive and negative effect of arginine in stroke and it is therefore difficult to predict whether the inclusion of arginine as an excipient has an effect, what this may be and how it may impact on the balance of benefits and risks of rt-PA.

Paper 5C: Further submission by Dr [redacted]

This paper assessed Dr [redacted] concerns that key data relating to the effectiveness of rt-PA on cerebral ischaemia and infarction have not been adequately presented in study publications 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 rt-PA. The other main area of concern relates to cerebral oedema, in particular:

- that in a non-clinical model of stroke, cerebral oedema was exacerbated by reperfusion 2 hours after stroke onset, and that reperfusion should be achieved within 30 minutes of insult

  This was a single historical study of reperfusion (in the absence of rt-PA) in baboons and may have limited relevance to rt-PA. It is also inconsistent with data from other studies.

- that cerebral oedema rates in both arms of the NINDS trial were implausibly high in comparison with other trials

  It is likely that the definition of cerebral oedema used in NINDS differed from the other trials, given the difference in frequencies observed (56-66% in the NINDS trial, compared with 2-7% in the ECASS I, II, III, ATLANTIS B and IST-3 trials)

- that reductions in deterioration due to oedema were expected in IST-3 but not seen

  IST-3 involved higher risk patients treated outside of the marketing authorisation and therefore may not reflect previous findings

- inadequate reporting of oedema in the Emberson et al meta-analysis

  Data on cerebral oedema has not yet been explored in the meta-analysis but will be included in an ongoing programme of secondary publications.

- that early mortality rates may be due to cerebral oedema

  The multi-component outcome measures used in the trials for rt-PA capture adverse outcomes due to cerebral oedema. Overall outcome is what is important for the patient.

Paper 5D: Benefit:risk of rt-PA administered between 3-4.5 hours post-symptom onset

This paper summarises all data relating to the 3-4.5 hours treatment window and discusses the balance of benefits and risks of rt-PA in this time-window. The data presented included:

- results of the ECASS III trial, the main trial supporting the extension to the time-window to 4.5 hours

- the results of an MAH pooled analysis submitted at the time of the extension to the time-window
- data from a number of observational studies (SITS-ISTR, SITS-UTMOST, Get With The Guidelines-Stroke registry, the Baden-Wuerttemberg stroke registry and the Canadian Alteplase for Stroke Effectiveness Study);
- the IST-3 trial;
- the Emberson et al (2014) individual patient data meta-analysis).

Benefits and risks in subgroups of interest – patients with mild or severe stroke, and patients aged over 80 years – were also considered.

The paper concluded that the data available in the 3-4.5 hour time-window are supportive of a positive balance of benefits and risks when used within the conditions of the licence. However, it is clear that benefit decreases with increasing onset to treatment time, and therefore efforts to reduce this time are paramount to improving outcomes. The current SmPC highlights, in several relevant sections including the indication, the need for treatment to be provided as soon as possible following the onset of stroke symptoms and the relationship between decreasing efficacy with increasing onset-to-treatment time.

Paper 6: Clinical use of rt-PA in the UK and feasibility of treating within the conditions of the marketing authorisation

There are many contraindications to treatment with rt-PA for acute ischaemic stroke, and this inevitably raises questions regarding the feasibility of treating within the conditions of the marketing authorisation particularly given the necessity to treat as quickly as possible. Nevertheless, the available data suggest that the main off-label use of rt-PA is the treatment of patients aged >80 years (~30% of thrombolysed patients), whilst use beyond 4.5 hours appears to be very low (~2%). This use in elderly patients is not unexpected because it is in line with the current national clinical guidelines which recommend treatment in all patients regardless of age up to 3 hours post-symptom onset, and that patients should be considered for treatment on an individual basis between 3-6 hours of onset. The data from SSNAP suggested that the most common reason for not thrombolysing a patient relates to time from onset of symptoms (~29% arrived outside of the time window for thrombolysis, ~20% wake-up time unknown). Less common reasons included that the stroke was too mild or too severe, or was haemorrhagic, that the patient's condition was improving, or they had other co-morbidities/conditions/concomitant medications. It is difficult to ascertain the level of off-label use in contraindicated subgroups other than age and time to onset.

The majority of clinical trials had exclusion criteria similar to the contraindications of the current SmPC, and therefore these trial data are most relevant to patients treated within the conditions of the licence. The IST-3 trial is the best source of randomised trial data on patients treated outside of the licence, and the meta-analysis by Emberson *et al* provides the most current summary of all clinical trial data available on rt-PA in acute ischaemic stroke including IST-3. More than 50% of patients enrolled in IST-3 were aged >80 years and thus this trial provides more information in this subgroup than other previously conducted clinical trials. The results for the subgroup of patients aged >80 years were more favourable than the results in patients aged ≤80 years. The information from IST-3 and the Emberson et al meta-analysis, as well as observational sources, do not raise concerns about the off-label use of rt-PA in patients >80 years.

Paper 7: Benefits and risks of rt-PA in clinical practice, including in off-label use, and the occurrence of medication errors

This paper discusses the benefit-risk balance for alteplase when used in clinical practice for the treatment of acute ischaemic stroke. This includes relevant examples of off-label use including in contraindicated patient populations. It also assesses whether recent evidence on the efficacy and harms of alteplase in specific patient sub-groups is appropriately reflected in the relevant sections of the current Summary of Product Characteristics (SmPC) and assesses the potential impact on the SmPC of emerging data on stroke diagnosis and treatment.

Evidence for medication/administration errors and the possible need for clarification of the SmPC are also discussed.

Stroke severity: rt-PA is currently contraindicated in severe stroke (NIHSS>25) and mild neurological deficit. The available evidence demonstrates that baseline stroke severity as measured by NIHSS is an independent predictor of sICH and fatal ICH, following rt-PA given within 4.5 hours of stroke onset. This results in a larger absolute excess risk of sICH for more severe strokes, despite the odds ratio for sICH being similar across subgroups of differing severity. The effectiveness of rt-PA within 4.5 hours does not vary according to stroke severity, however there are few data for patients with NIHSS >22.
There are also limited data for 'mild' stroke, NIHSS ≤5, but the STT Group meta-analysis suggests that rt-PA treatment in mild stroke has a positive effect.

Whilst the benefit-risk balance for rt-PA may be positive at a population level the balance in individual patients is less clear. Ongoing studies may provide further data.

The paper concluded that the SmPC wording is appropriate; there are insufficient data to lift the contraindications in patients with minor neurological deficit or severe stroke as assessed clinically (e.g. NIHSS >25) and/or by appropriate imaging.

Time to treatment: studies have shown that the balance of benefits and risks of rt-PA changes with increasing time to onset of treatment, with risk of ICH staying the same and benefits reducing. There is robust evidence that treatment within 0-3 hours of onset of symptoms is effective with an acceptable risk of ICH and mortality in most patients. More limited evidence supports a positive balance of benefits and risks for treatment between 3-4.5 hours when used according to the licence conditions. Treatment in patients with adverse prognostic factors may be less beneficial in the 3-4.5 hour time period and could result in a negative balance of benefits and risks but these patient subgroups are mostly contraindicated for rt-PA. The balance of benefits and risks beyond 4.5 hours appears to be negative on the basis of the available data.

The paper concluded that the SmPC wording is appropriate; there is extensive wording in the SmPC regarding the need to treat patients as quickly as possible and the decrease in benefit as time from onset of symptoms increases. Treatment is contraindicated beyond 4.5 hours.

Age: recent evidence indicates that the benefits of rt-PA are not age-related particularly when patients are treated early. There is limited evidence to support the use of rt-PA in children.

The paper concluded that the contraindication in children <18 years is acceptable in the absence of positive RCT data. The MAH has indicated that they intend to review the contraindication in patients aged >80 years based on an up to date review of the

benefits and risks of rt-PA treatment in this age group. Until such time the wording in the current SmPC remains appropriate.

Hypertension: the optimal management of high blood pressure soon after acute ischaemic stroke remains controversial, and clinical recommendations are limited due to lack of adequate evidence. The association between blood pressure and mortality appears to be U-shaped, and overall it appears that patients with normal or slightly raised blood pressure may have more favourable outcomes than patients with high blood pressure. However the optimum blood pressure has not been defined and may depend upon several factors including whether the patient has been treated with rt-PA or not, and the reason/mechanism for hypertension in an individual case.

The paper concluded that in the absence of robust evidence the SmPC wording is appropriate, rt-PA is contraindicated in patients with systolic BP >185 mm Hg or diastolic BP >110 mmHg or if intravenous pharmacotherapy is required to reduce blood pressure to these limits.

Prior stroke and concomitant diabetes, blood glucose <2.8 or >22.2 mM: the available data are limited but there are no conclusive data that risks of thrombolysis are substantially higher in these patients. There is currently no clinical evidence that targeting the blood glucose to a particular level during acute ischaemic stroke will improve outcomes after rt-PA.

The paper concluded that the SmPC wording is appropriate, rt-PA is contraindicated in patients with prior stroke and concomitant diabetes and in patients with blood glucose <2.8 or >22.2 mM. The available data are not considered to be sufficient to lift these contraindications.

Concomitant medication:
anticoagulants – rt-PA is not recommended in patients with INR>1.3 as the risk of ICH after rt-PA may depend on baseline INR for warfarin. However, data on the effect of INR on outcome post rt-PA are relatively limited. Likewise there are few data available on safety of therapeutic heparin, direct thrombin and factor Xa inhibitors with rt-PA.

The paper concluded that the current SmPC wording is appropriate, there is no clear evidence for a 'safe' INR and therefore the current recommendations are appropriate (i.e. can consider rt-PA if INR ≤1.3).

antiplatelet therapy – the current SmPC wording describes the risk of ICH with antiplatelet pre-treatment and recommends that antiplatelet drugs should not be started within 24h of rt-PA. However, there is new evidence to suggest that the risk of sICH after rt-PA in patients treated with dual antiplatelet therapy (aspirin and clopidogrel) may be greater than for aspirin monotherapy (SITS-ISTR adjusted OR: 3.2 [95% CI: 1.9-5.2] vs 1.8 [95% CI: 1.5-2.1], n=31,627). The SmPC is currently silent on concomitant dual anti-platelet therapy.

The paper concluded that the data does not suggest a need to alter the established section 4.2 and 4.4 warnings on not starting antiplatelet drugs for 24 h after rt-PA therapy. The increased risk of sICH after rt-PA in those receiving prior antiplatelet drug monotherapy is adequately described. However, the SmPC should be amended to warn of the risk of deleterious synergistic effects of dual antiplatelet therapy as the risks of rt-PA therapy may outweigh any potential benefits if there are any additional adverse prognostic features (e.g. severe stroke, old age).

Seizure at onset of stroke: this contraindication is likely to have been put in place to avoid treatment in patients with stroke mimics e.g. Todd paralysis, a condition that may be difficult to distinguish from ischaemic stroke using CT scan and clinical examination. MR diffusion and perfusion-weighted images or angiography, perfusion

CT or CT angiography can be used to confirm the diagnosis of acute ischaemic stroke.

The paper concluded that the available evidence suggests that treatment of stroke mimics is sufficiently safe to justify rapid treatment of all patients with suspected ischaemic stroke provided that haemorrhagic stroke is excluded.

Medication errors: the relationship between the clinical impact of rt-PA dose, optimal neurological outcome and minimal ICH risk remains unclear. Dosing recommendations are based on limited data (mainly NINDS) which suggest that rt-PA has a relatively narrow therapeutic window. However, as the NINDS trial used weight estimation rather than actual weight, this suggests there is some margin for error. For a number of reasons patient weight tends to be estimated in the acute stroke setting. Medication errors are not systematically recorded, but studies and reports from the National Reporting and Learning System suggest that errors may result from inaccurate estimation of patient weight, incorrect dilution or administration of rt-PA – which may be compounded by time pressures, relatively complex posology and confusion over the different names used for rt-PA (alteplase/Actilyse/t-PA/rt-PA).

The paper concluded that clinical guidance could be updated to provide advice for optimal but realistic weight estimation of stroke patients and data on medication/dosing errors should be routinely recorded and reported in the Sentinel Stroke National Audit Program. Clear instructions for dilution and administration should be provided in stroke centres and in the SmPC, including a weight-based dosing table.

Emerging data:
stroke subtype – there is no consistent evidence that the balance of benefits and risks of rt-PA is altered by stroke subtype.

radiological signs – a small number of observational studies report an increased risk of ICH with rt-PA in patients with severe leukoaraiosis at baseline. The IST-3 trial reported that some combinations of pre-existing radiological signs increased some absolute risks.

The paper concluded that the SmPC should be amended to reflect the increased risk of ICH in patients with severe leukoaraiosis.

Paper 8: Risk minimisation measures

This paper discusses the measures in place within the rt-PA licence, their appropriateness and adequacy.

The MAH has confirmed that they are currently reviewing use in patients aged >80 years, with the intention of potentially submitting a variation within Europe to lift the contraindication. The MAH has noted that this review should better define which patients >80 years are appropriate for rt-PA treatment and in light of these findings it may be necessary to also introduce some restrictions to its use in this subgroup of patients. Given the current level of off-label use in this age group, the total level of use of rt-PA in this group may decrease overall if restrictions are also added.

The paper concluded that no urgent regulatory action is considered necessary, and in the main the conditions for use of rt-PA appear to be appropriate. However the SmPC could be updated with respect to the following, at the next routine opportunity:

- addition of a warning of increased risk of sICH after rt-PA treatment in patients with leukoaraiosis or other established brain lesions at stroke onset
- addition of a warning regarding the risk of deleterious synergistic effects of dual antiplatelet therapy as the risks of rt-PA therapy may outweigh any

potential benefits in those receiving aspirin and clopidogrel if there are any additional adverse prognostic features (e.g. severe stroke, old age)

- inclusion of a weight-based dosing table
- minor clarifications of the dosing and administration section

It would also be appropriate to review the use of rt-PA in patients aged >80 years (as proposed by the MAH).

Paper 9: Communication of risk and benefit to patients

This paper discusses a selection of current national communications and risk estimation/decision tools, and whether there is a need for further materials to aid decision making or understanding of benefits and risks of rt-PA. It provides suggestions for information resources that may be helpful to patients/families and clinicians. These types of information resources are distinct from formal risk minimisation measures which form part of the marketing authorisation.

Patient perspectives on risk of rt-PA treatment were explored in focus groups during the design of the IST-3 trial and patient, family/carer and clinician perspectives on risk communication and decision making on thrombolysis have also been studied via interviews shortly after the stroke event, and by observation of the acute stroke situation and interactions.

Specific points that arose during these studies were:

- Verbal face-to-face discussion is the most important method of conveying information to patients/family
- There are difficulties with providing tailored, individualised information for each patient (dependent upon their baseline characteristics), nuanced information for different subgroups of patients could be developed. A risk estimation tool, such as COMPASS which is based on the Stroke-Thrombolytic Predictive Instrument, could be useful for predicting which patients are likely to benefit and which are likely to be harmed by rt-PA, as well as potentially facilitating the discussion on treatment with the patient/carer.
- Any written information/visual aid specifically designed to aid decision making during the acute stroke event needs to be very concise and simple, and therefore probably pictorial/graphical in format
- There may be a place for written information in the form of a leaflet, for most patients/families, most likely as something they can take away and read later
- There may be a place for communications documents/leaflets that aim to educate members of the general public on stroke: risk factors, signs and symptoms, importance of seeking help as soon as possible, treatment options including thrombolysis and its risks and benefits.

There would be benefits to providing standardised resources, and the currently available examples including NICE decision aids may provide useful suggestions. Generation of any tools should follow a pre-defined structured development and user testing procedure.

The paper concluded that whilst verbal face-to-face discussion is recognised to be the most important method of information provision to patients/families in the acute stroke setting, these interactions might be usefully supported by written information/visual aids. The type of individualised information provided by the

COMPASS tool might be considered to be the most helpful; however this is not yet available. Other more generic resources may be useful in the meantime, and may also be complementary to the output of COMPASS.

A small subgroup could be formed to develop any such materials, and their development, testing and distribution would need to be carefully considered and planned.

Paper 10: Draft conclusions and recommendations to CHM and strategy for communication of the outcome of the group

This paper summarises the conclusions of the group to date and provides draft conclusions and recommendations for the final meeting.

CHM paper, May 2014

This paper was provided to the EWG as background information regarding the issue, including assessment of new data and of initial concerns raised by Dr [redacted] A number of these initial concerns were repeated in the first submission provided by Dr [redacted] to the EWG, together with additional, further concerns. Paper 5 therefore has some overlap with the CHM paper from May 2014.

Oral presentations

The EWG also heard presentations from:

- *Professor Jonathan Emberson on the STT individual patient meta-analysis published in 2014*

  The Group noted that this was a one-stage meta-analysis of individual patient level data, stratified by trial so as to maintain the randomisation in the trials.

  The Group noted that all analyses were consistent with better outcomes with rt-PA at shorter time to onset of treatment. The Group was also reassured that all definitions of 'good outcome', in terms of where the endpoint (mRS) had been dichotomised, found a beneficial effect with rt-PA.

  In addition, the Group noted that the effect on mortality by rt-PA was due to the initial risk of fatal ICH, and rt-PA did not impact on other causes of death. The signal of ICH found in the meta-analysis was as expected.

- *Professor Colin Baigent on further analyses of the STT dataset*

  The Group noted that formal tests had demonstrated that the IST-3 'uncertainty principle' trial data were consistent with the other trials, and that no trial had been found statistically to be an outlier.

  The Group noted that the effects of age, treatment delay and baseline stroke severity strongly interacted and required multivariable regression analysis. The analyses found that in general younger patients presented later, that older patients had more severe strokes and that less severe strokes were more likely to be randomised later.

  The Group noted that the OR for ICH does not vary with stroke severity, but that patients with more severe stroke have a larger absolute excess risk due to their higher baseline risk.

The Group heard that when the data from NINDS was removed from the meta-analysis, the results were qualitatively the same, although less robust because NINDS was a positive trial.

A post-hoc analysis of ICH-related and ICH non-related death found an early increase in ICH-related death in rt-PA treated patients and a suggestion of benefit in terms of non-ICH related death. Limited 18 month follow-up data on patients treated <3 hours in IST-3 suggested a possible reduction in later death rates for patients treated with rt-PA, suggesting that preservation of brain tissue may have overall mortality benefits later on. Due to the post-hoc nature of the analysis it was stressed that this is currently a hypothesis.

The Group also noted the results of a 'better than expected outcome' analysis, which showed little difference across the subgroups.

- *Dr [redacted] on his concerns regarding the benefits and risks of rt-PA in the treatment of acute ischaemic stroke*

  The Group noted Dr [redacted] concerns including the asymmetry of funnel plots of mortality, the existence of baseline imbalance in stroke scores in the NINDS trial, the different interpretations of the NINDS results, the results of the IST-3 trial compared with the rate of fatal ICH in the Emberson et al meta-analysis and with the results of NINDS, and the potential for bias in IST-3.

  The Group discussed the points raised at length after this presentation and throughout the discussions of papers 5 and 5C, and whilst the Group questioned the statistical validity of certain aspects of the analyses presented, they agreed that some important issues had been highlighted that merited further evaluation.

- *Professor Gary Ford on his experience of rt-PA and in particular the SITS registry*

  The Group noted that the publication of SITS data and ECASS III appeared to have resulted in increased confidence in rt-PA, resulting in increased use in both the <3 hour time-window and the 3-4.5 hour time-window. The Group also noted that data suggested that door-to-needle times are reducing.

  The Group noted that treatment of patients aged >80 years had also increased since the publication of data from SITS.

- *Professor Peter Sandercock on the IST-3 trial and results and his personal experience with rt-PA*

  The Group noted that the design of the IST-3 trial was such that patients were enrolled only if the clinician was uncertain whether to treat or not; the patient group included would therefore be expected to be at higher risk than patients included in previous clinical trials and treated in the clinic, and essentially off-label use.

  The Group noted that the relative risk of sICH varied little between different patient subgroups, although the absolute risk varied with stroke severity.

  The Group noted that the study found that benefit with rt-PA was greater in patients with more severe stroke and with greater age, whilst the risk for intracranial haemorrhage is likely to be greater in these groups as well.

  The Group noted that the balance of benefits and risks in patients with mild stroke was less clear than in patients with more severe stroke but that IST-3 was underpowered in mild stroke and that ongoing clinical trials in this contraindicated population should provide further information.

- *Professor Keith Muir on the definitions and implications of symptomatic intracranial haemorrhage*

  The Group noted that PH2 bleeds are independently associated with a poor outcome. PH2 bleeds (as used in the SITS-MOST definition of sICH) may therefore be the most clinically relevant.

  The Group heard that trivial amounts of bleeding in an area of ischaemia can indicate successful reperfusion. The Group also noted that the timing of a scan would impact on whether bleeding remained visible or not, and similarly that ischaemic infarcts can be difficult to identify particularly when there is background ischaemia e.g. in older patients.

  The Group noted that the NINDS trial (with the highest rate of sICH) used a very conservative definition of intracranial haemorrhage and that the published scans from the trial showed some of the cases had very small amounts of blood within large areas of ischaemia which would have been classified as ICH but which may have had minimum/no effect on outcome and can indicate successful reperfusion.

- *Dr Gillian Cluckie on the experiences of patients, carers and clinicians with respect to the communication of benefits and risks of rt-PA and the management of uncertainty*

  The Group were interested in the findings of this ethnographic study and were reassured that the clinicians observed consistently informed patients/carers about the risks and benefits of thrombolysis. The Group noted that the participating patients/carers did not recall or value statistical presentation of the benefits and risks but needed to have confidence in the clinician.

  The Group agreed with the finding that the ultimate decision on whether to thrombolyse or not should be the clinician's.

- *Professor Gary Ford and Dr Peter McMeekin on the COMPASS tool and its development and validation*

  The Group noted the work undertaken to develop the COMPASS tool, which provides more individualised estimates of outcomes following stroke both with and without rt-PA treatment.

  The Group provided a suggestion for improving the estimation of excess risk. The Group considered that the COMPASS tool was complex and, based on patient experiences in the ethnographic model, may have limited value for patients. The Group commented that the need to input a series of information into a computer model in an acute stroke setting may be impractical.

3. Conclusions of the EWG

Prognosis of stroke in the UK following the introduction of rt-PA

There are compelling data to suggest that the prognosis of patients with ischaemic stroke has improved in the last decade and the latest mortality data suggests improved outcomes compared with previously. However, the reasons behind these improvements are difficult to ascertain given the many organisational changes that have taken place as a result of the introduction of rt-PA. Furthermore the overall net beneficial effects of the introduction of rt-PA are likely to be small as a result of the small proportion of patients who are eligible for rt-PA treatment (~11% of all stroke patients admitted are thrombolysed, which is considered to be ~80% of those that are eligible (SSNAP)).

Radiological diagnosis

The group concluded that there was insufficient evidence to suggest that another form of radiological detection, including MRI, should be used routinely instead of CT scanning.

New data since extension of the time window to 4.5 hours

The Group concluded that the new data that have become available add substantially to the understanding of the balance of benefits and risks of rt-PA over time and in different patient populations. The balance of benefits and risks of rt-PA in its authorised indication of acute ischaemic stroke remains positive.

Concerns of clinicians on the data

The Group concluded that the concerns expressed and the data presented do not provide sufficient 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. Nevertheless a thorough evaluation of these concerns has been beneficial in furthering understanding of the evidence on which the current marketing authorisation is underpinned and in increasing confidence in rt-PA in thrombolysis.

Feasibility of using rt-PA within the terms of the licence

The Group concluded that it is not possible to obtain every piece of information required to ensure rt-PA is used within the terms of the marketing authorisation and still treat the patient in the shortest possible time. Therefore it is expected that physicians would apply their clinical judgement regarding the need for particular investigations prior to administration of rt-PA. The Group was reassured that rt-PA is being used responsibly.

Medication error

The Group concluded that there is little evidence that rt-PA is mis-used but that where errors occur they mainly relate to dosing and administration. The Group concluded that it may not be practical to make firm recommendations to improve weight estimations of stroke patients, but agreed that inclusion of a weight-based dosing table and clarifying the administration instructions in the SmPC would be helpful.

National tools to facilitate communication of benefits and risks by clinicians to patients

The Group concluded that it was important that physicians are provided with the tools and information they need to better understand the available data, and therefore be confident in their decisions and advice for patients.

The Group concluded that the decision on whether or not a patient should be thrombolysed is for the treating physician rather than the patient/carer, although it is important that all parties have had a discussion and agree with the decision. The Group concluded that any information provided to patients/carers would need

extremely careful consideration, that dot-plots showing outcome probabilities could be helpful but further research is needed, and that it may be more valuable to provide patients/carers with information about stroke and how it might impact their immediate and long-term future.

Overall the Group concluded that some information resources may be helpful to improve the consistency of decision-making and to provide patients/carers with information, but that their development was outwith the remit of the current Group. The Group concluded that the MHRA should determine what information resources relating to stroke generally are currently available and on this basis decide whether further resources are required.

Overall conclusion on balance of benefits and risks, and balance of benefits and risks in different patient populations

The Group concluded that the balance of benefits and risks of rt-PA in the treatment of acute ischaemic stroke is positive when used within the conditions of the marketing authorisation, up to 4.5 hours post-symptom onset. The Group concluded that benefit of rt-PA is highly time-dependent and therefore minimising the time to onset of treatment was critical to ensuring the best possible outcome.

The Group concluded that in general the current contraindications and conditions of the product information remain appropriate, but that there are some areas that should be reviewed by the marketing authorisation holder to determine whether product information accurately reflects the available data.

- Benefits and risks of rt-PA in patients aged >80 years and <18 years, currently contraindications
- Benefits and risks of rt-PA in severe and mild baseline stroke, currently contraindications
- Benefits and risks of rt-PA in patients with INR >1.3, currently not recommended, and contraindicated in 'patients receiving effective oral anticoagulant treatment, e.g. warfarin sodium'
- risk of ICH in patients treated with dual antiplatelet therapy

In addition the posology and method of administration section of the SmPC should be clarified, and a weight-based dosing table included.

The Group concluded that the evidence indicating that the risk of rt-PA-induced ICH is increased in the presence of severe leukoaraiosis was inconsistent and not sufficiently strong to warrant a warning in the SmPC.

Communicating the outcome of the review

The Group considered that a summary of the conclusions providing a clear, consistent message underpinned by the evidence would be required, to be followed by more comprehensive information. The Group considered that transparency was essential and that (after removal of duplication) all data considered by the Group should be made available.

3.1 EWG recommendations to CHM

- On the basis of all the evidence presented the balance of benefits and risks of rt-PA in the treatment of acute ischaemic stroke is positive when used within the conditions of the marketing authorisation, up to 4.5 hours post-symptom onset.
- The Summary of Product Characteristics (SmPC) adequately describes the benefit-risk balance for rt-PA therapy. However, the MAH should conduct a review to determine whether product information accurately reflects the available data with respect to the:
    - benefits and risks of rt-PA in patients aged >80 years and <18 years
    - benefits and risks of rt-PA in severe and mild baseline stroke
    - benefits and risks of rt-PA in patients with INR >1.3
    - risk of ICH in patients treated with dual antiplatelet therapy
- The MAH should clarify the posology and method of administration section of the SmPC and include a weight-based dosing table.
- MHRA should determine what information resources relating to stroke generally are available and on that basis decide whether further resources are required.

4. Next steps: Communication of the outcome of the review

The communication of the final CHM position needs to be carefully considered. The EWG considered that a short, clear, authoritative summary of the Group's conclusions followed by more comprehensive information would be appropriate. The Group emphasised the need for absolute transparency and that all data considered by the Group should be made available (after removal of duplication).

A two-phased approach is therefore proposed, with:

Phase 1

- Summary of the Group's conclusions to be provided as a pro-active press release, to be placed on the MHRA website, provided on request to journalists, other interested stake-holders, and relevant professional bodies (who can communicate the outcome to healthcare professionals that treat acute ischaemic stroke patients).

Phase 2

- Public Assessment Report (PAR) that includes the evidence base which underpins the conclusions of the EWG.
- Short submission to a relevant publication e.g. The Lancet providing the outcome of the review, or a longer more detailed explanation of the review findings.
- Article in the Drug Safety Update bulletin – it may be appropriate to time any DSU article to coincide with any updates to the product information.

5. Key messages

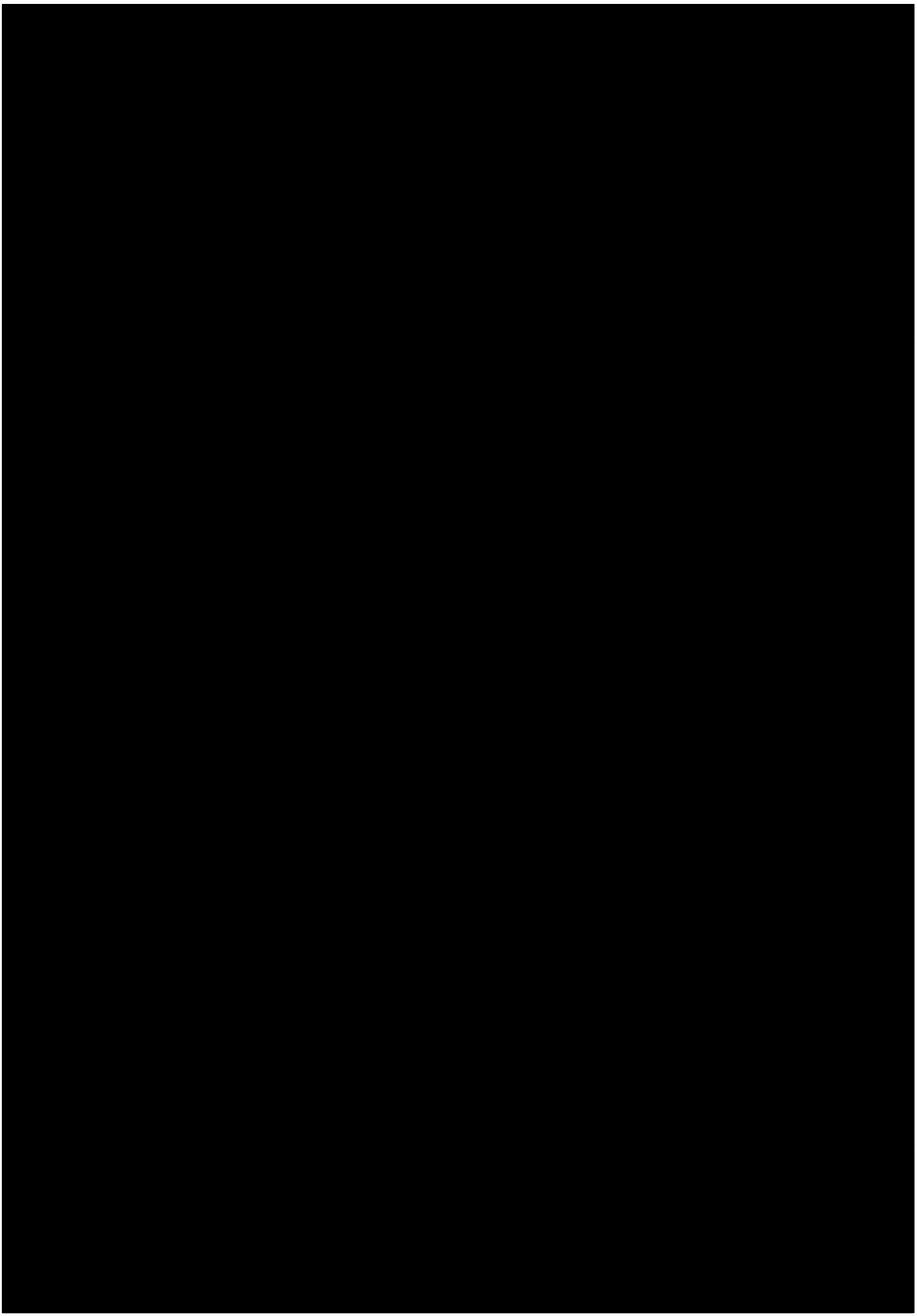
The key messages that need to be communicated following the EWG and CHM review are considered to be as follows:

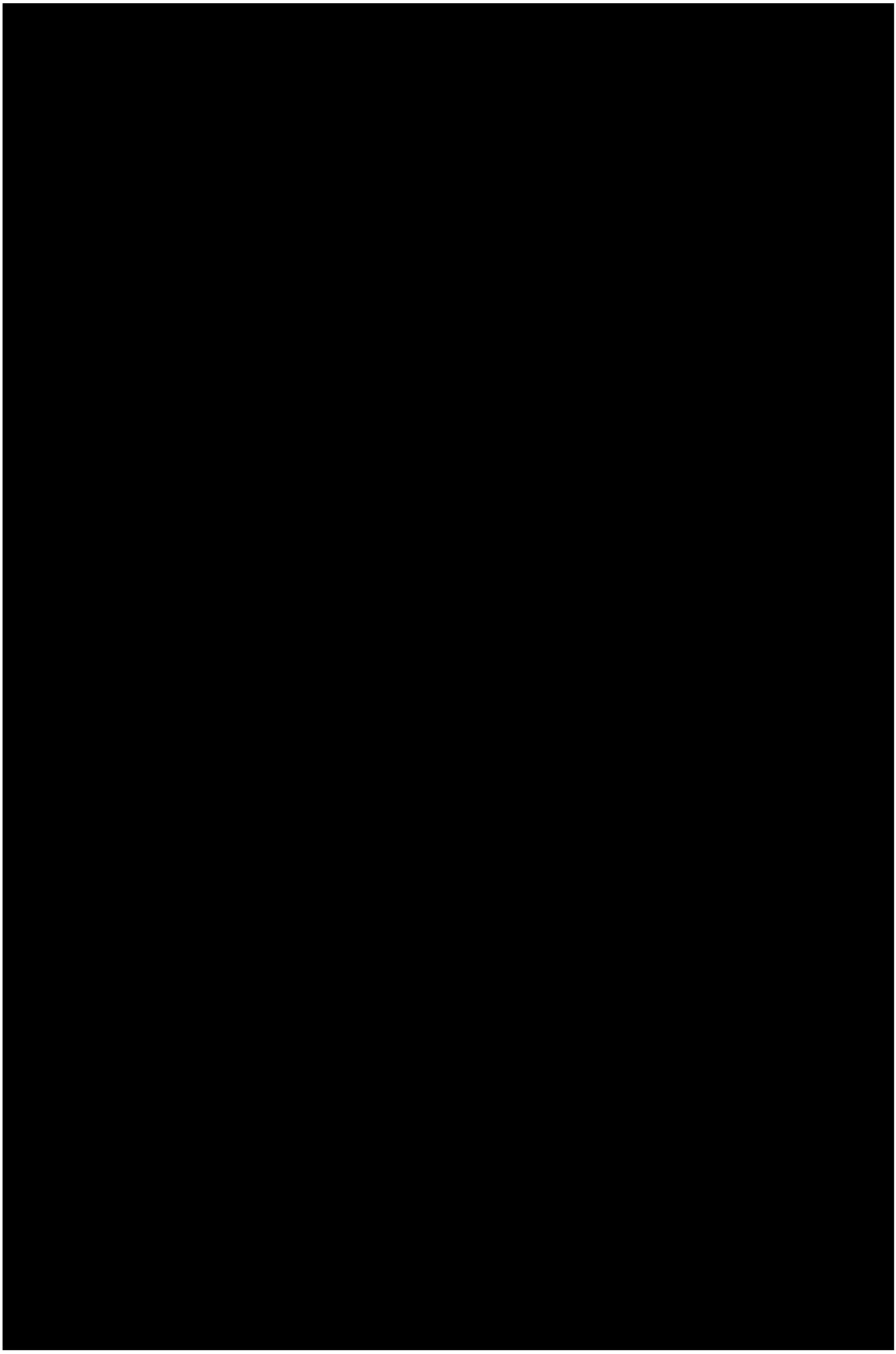
- The balance of benefits and risks of rt-PA are positive up to 4.5 hours post-symptom onset.
- The new data have been rigorously evaluated and add substantially to understanding of the evidence on which the current marketing authorisation is underpinned and the balance of benefits and risks of rt-PA over time and in different patient populations.
- Treatment with rt-PA is highly time-dependent because the benefits decrease over time whilst the risk of ICH remains the same. It is of critical importance that hospitals/stroke units are encouraged to focus on reducing door-to-needle times in order to optimise the balance of benefit and risk.
- Effective communication between clinicians and patients/carers at the time of acute stroke is very important. Patient/carers need to feel confident in the healthcare professional and should be guided in the decision of whether or not a patient should be thrombolysed.

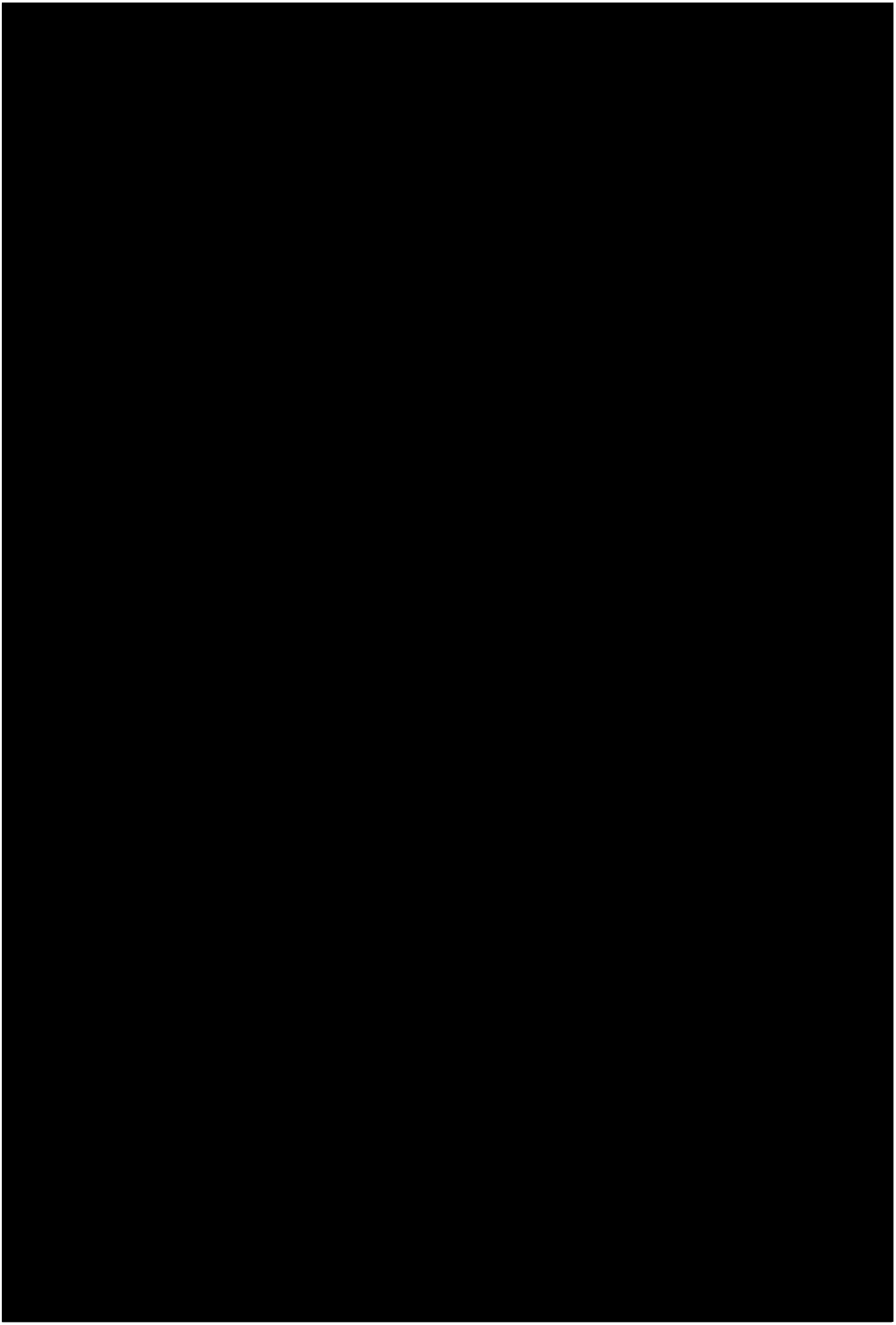
Advice sought

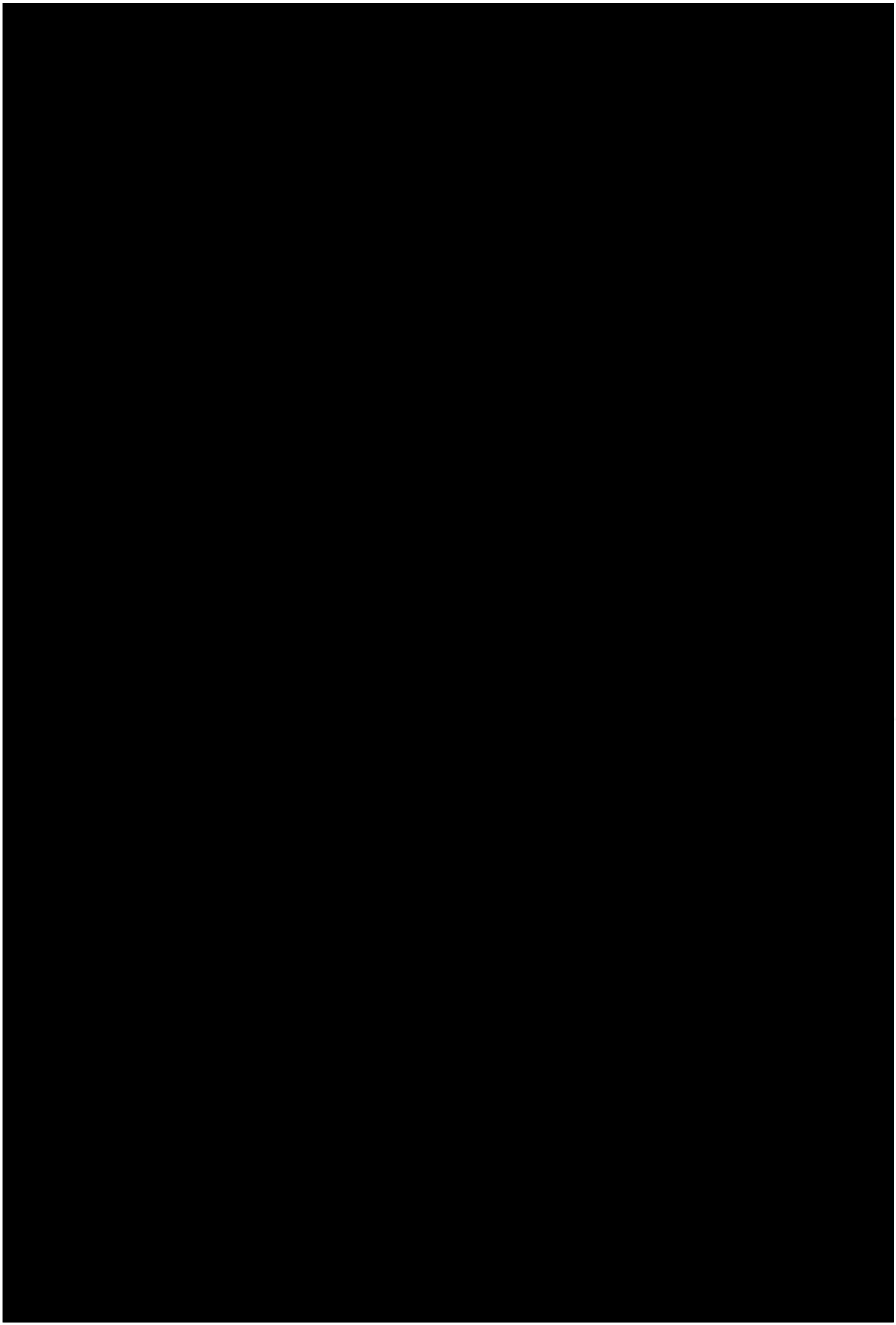
The Commission is asked whether it endorses the conclusions and recommendations of the EWG.

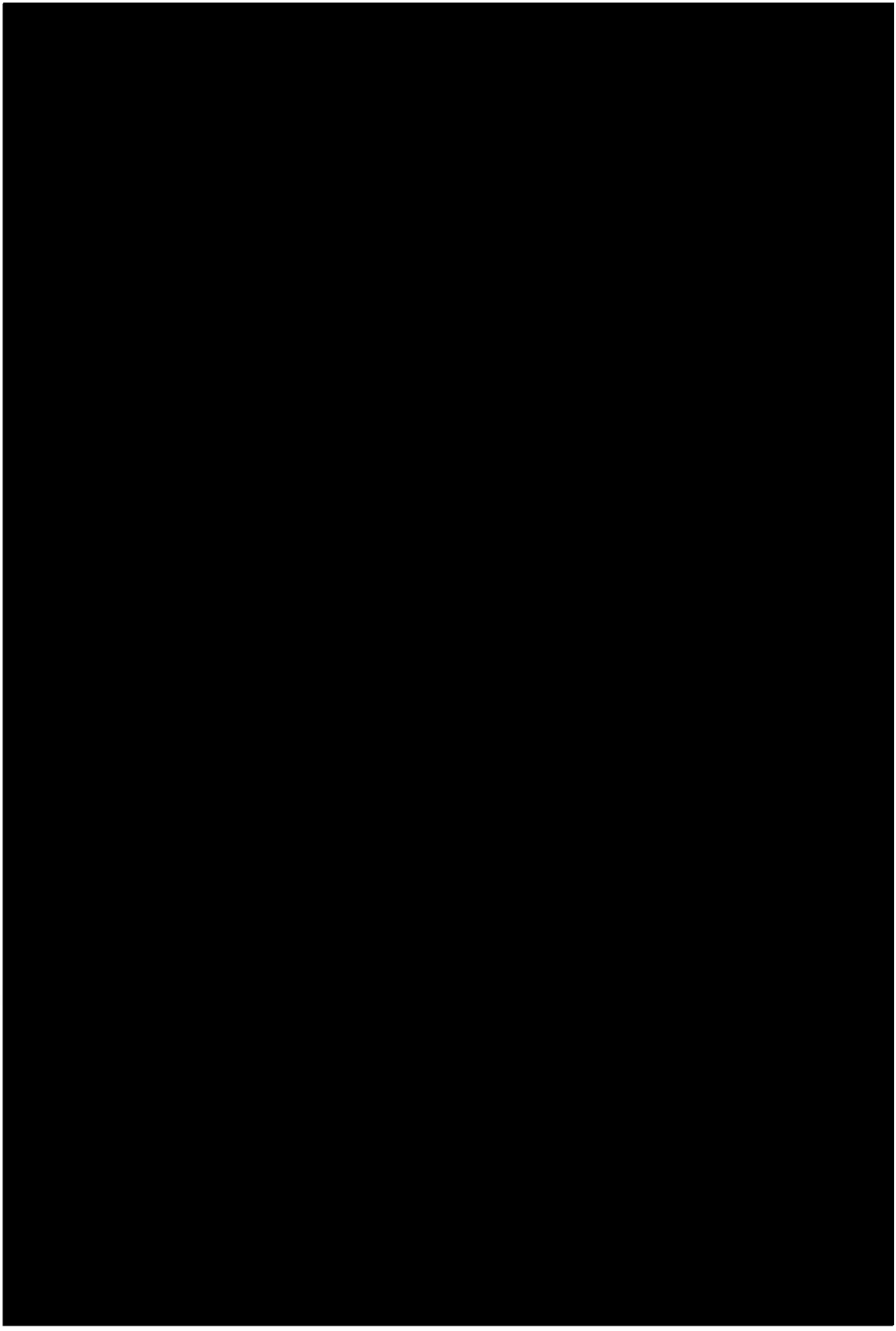
Any comments on the proposals for communicating the outcome of the EWG and CHM discussions would be welcome.

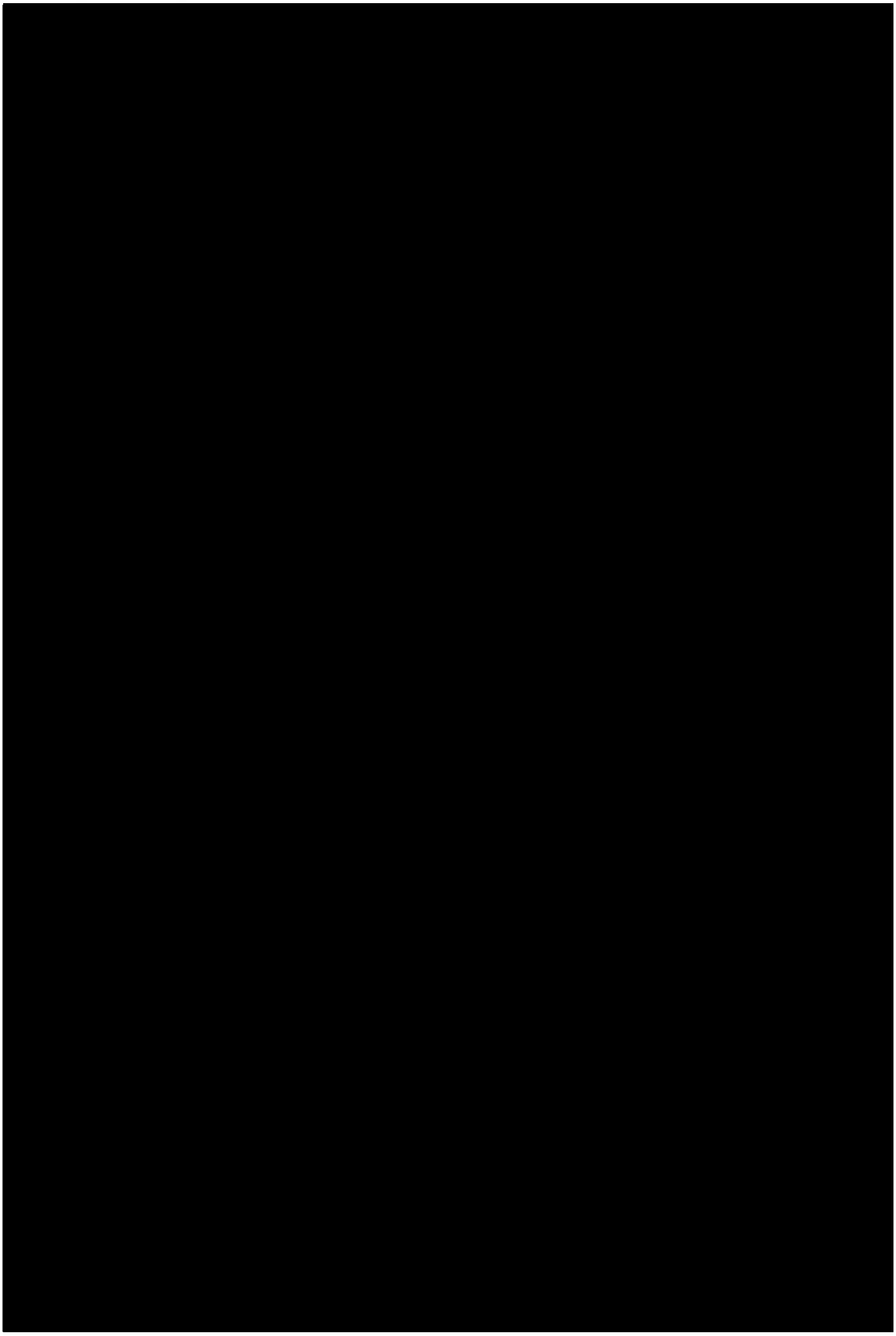


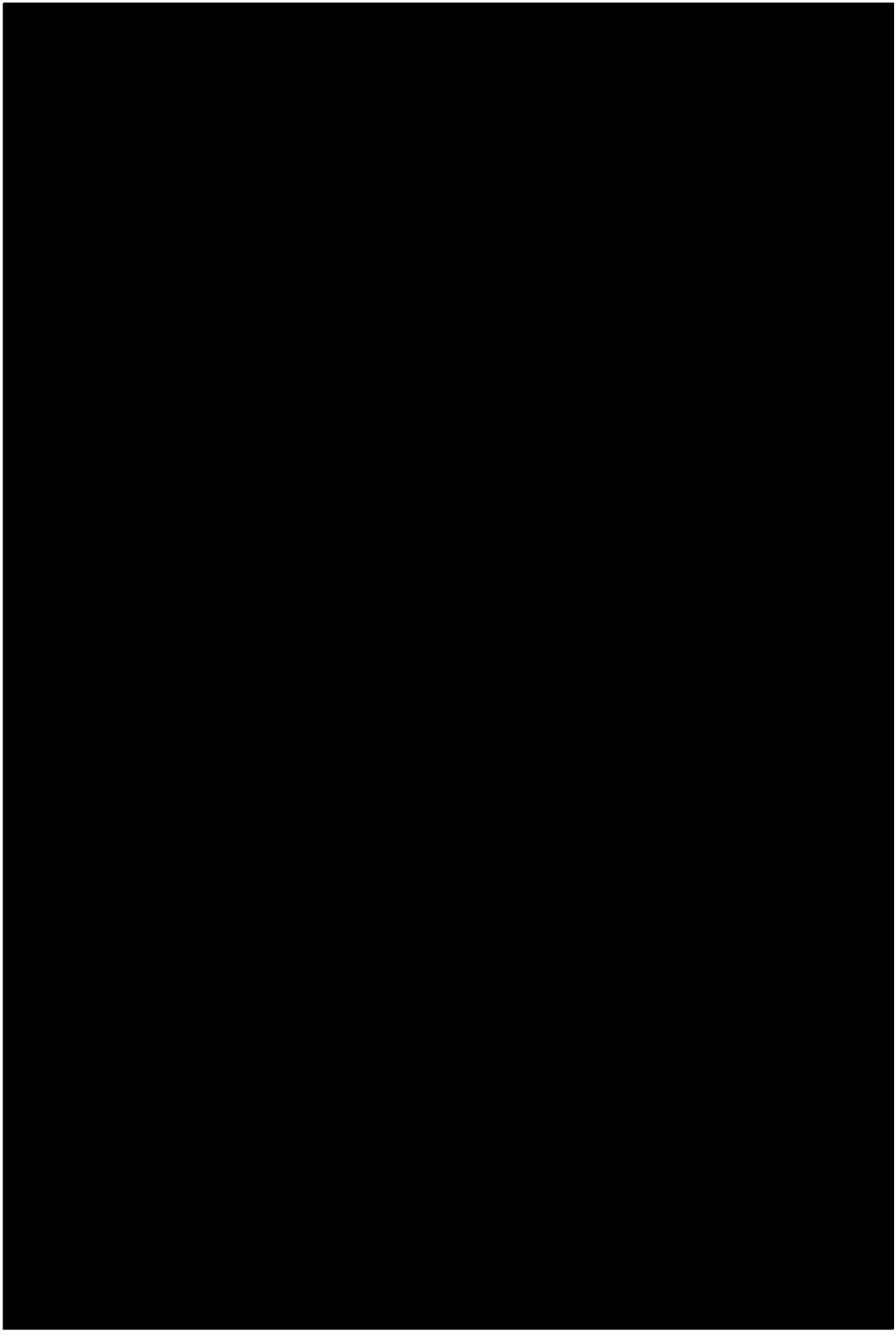


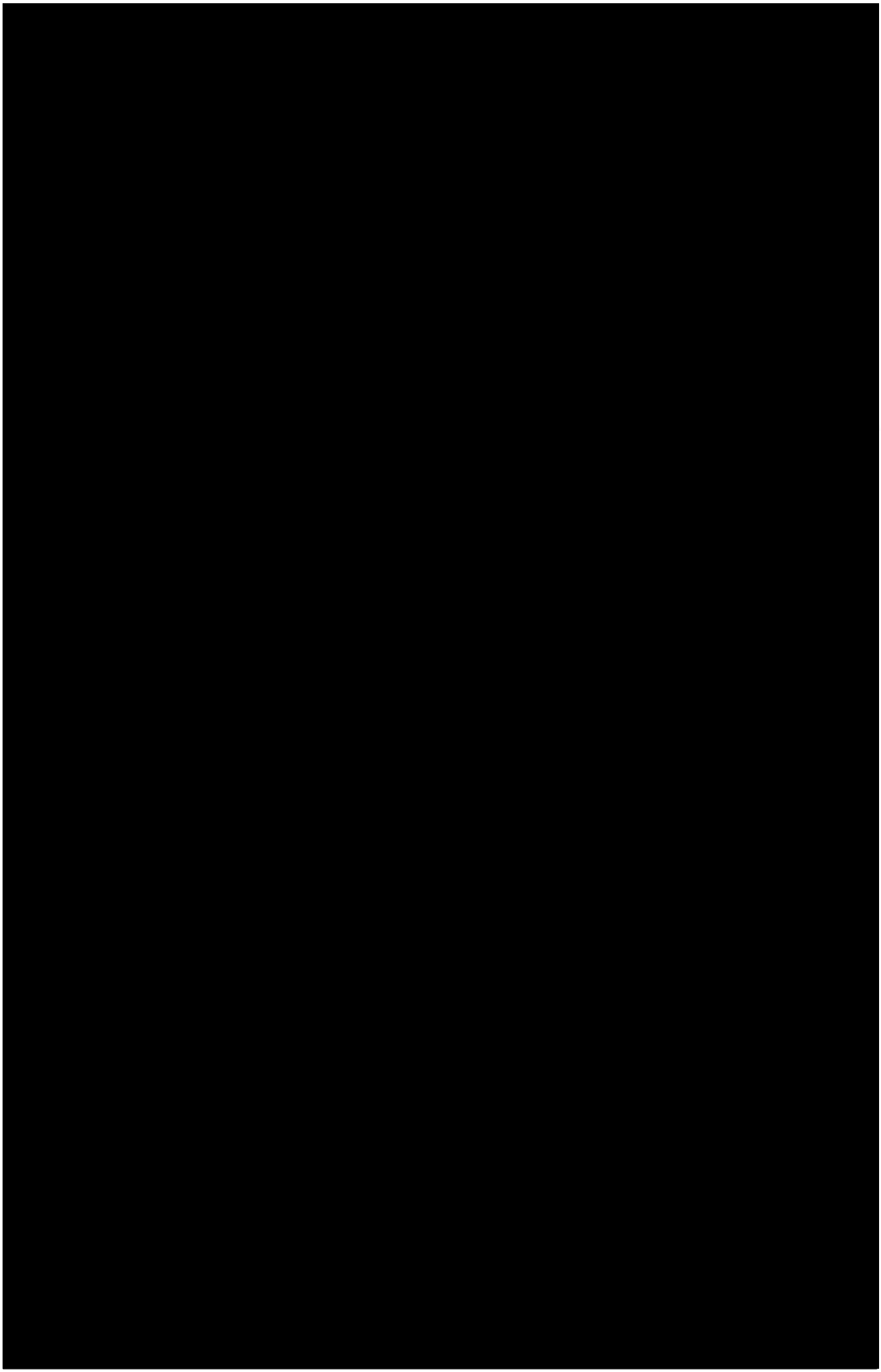


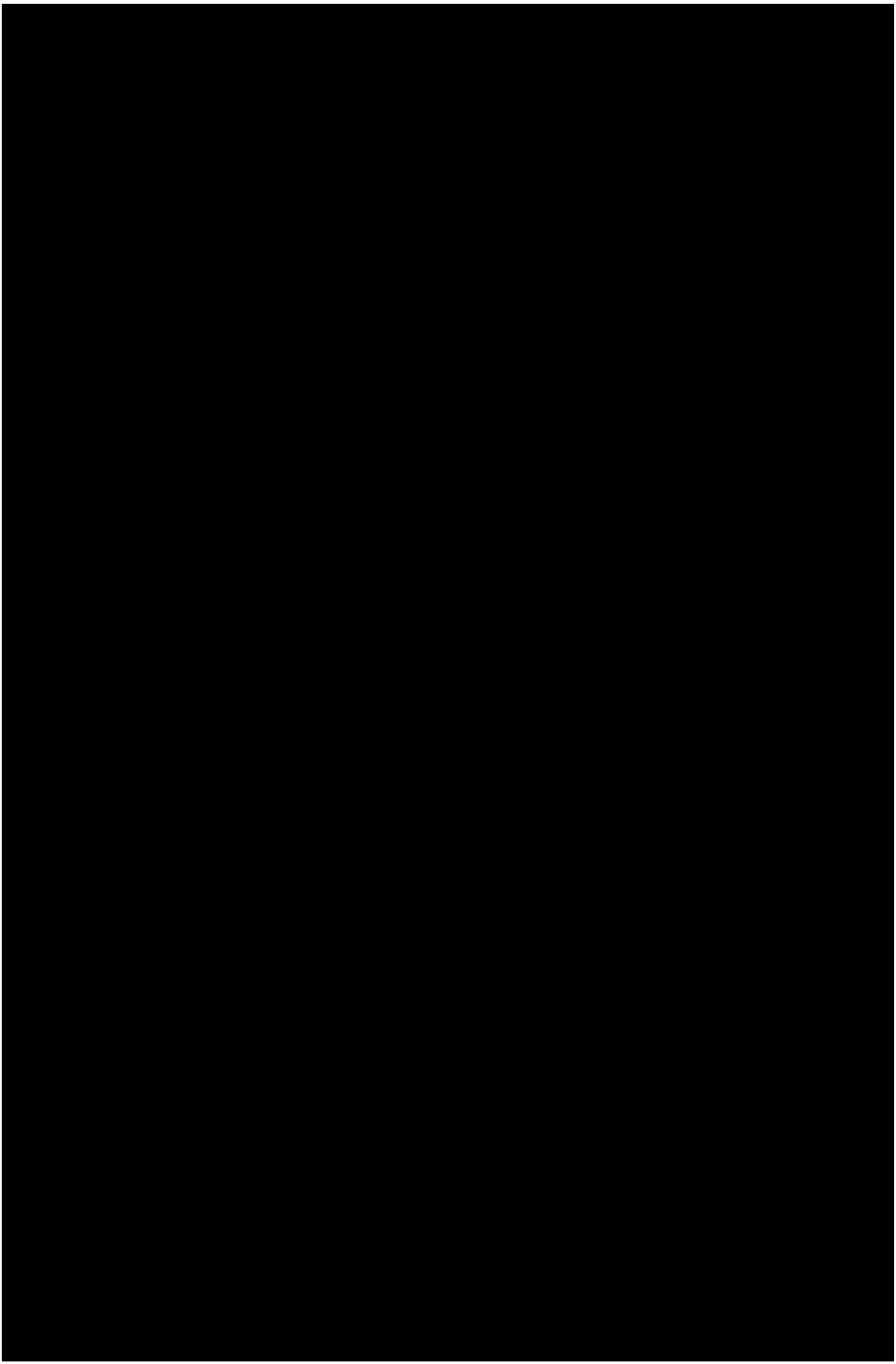


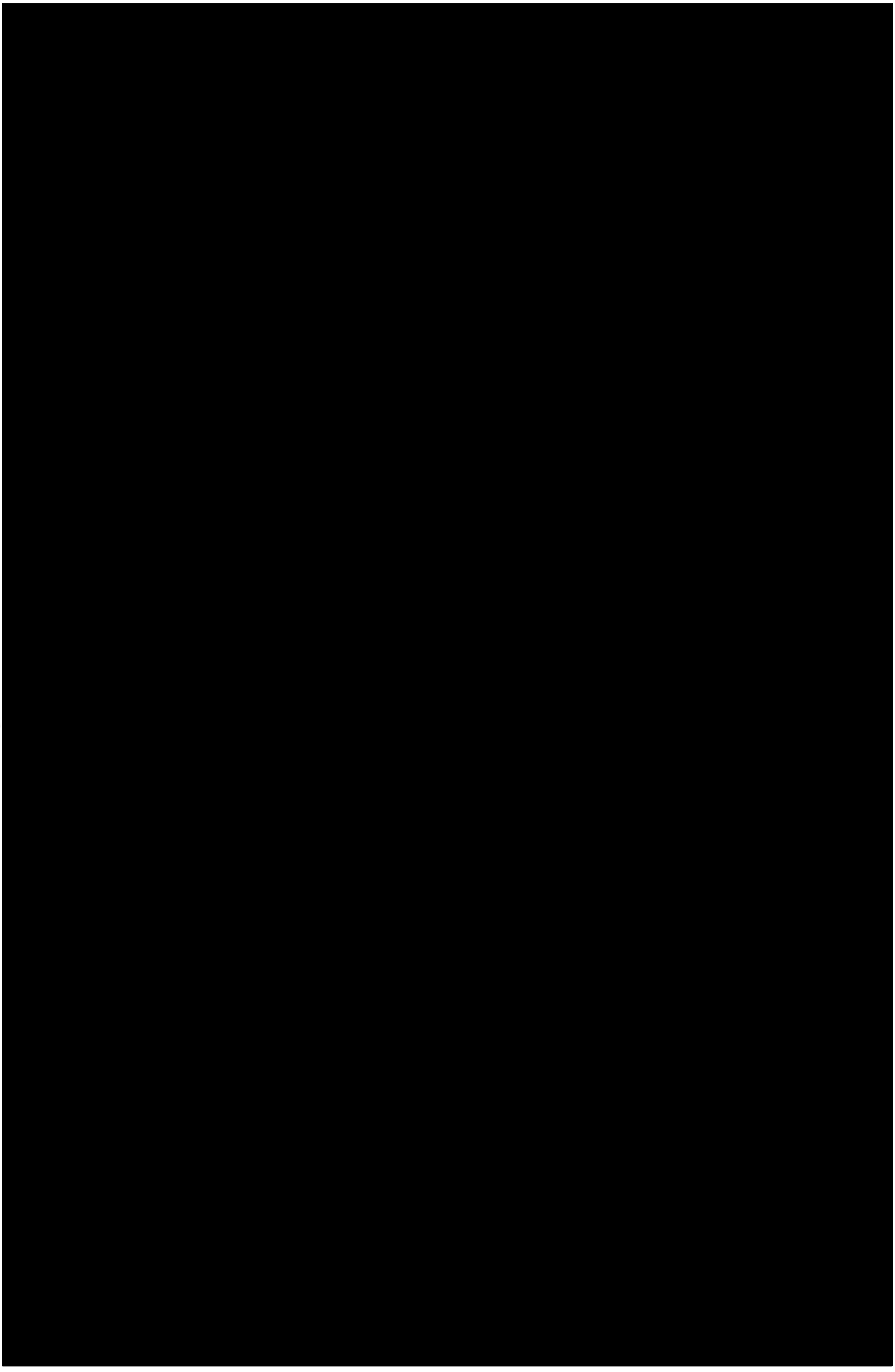


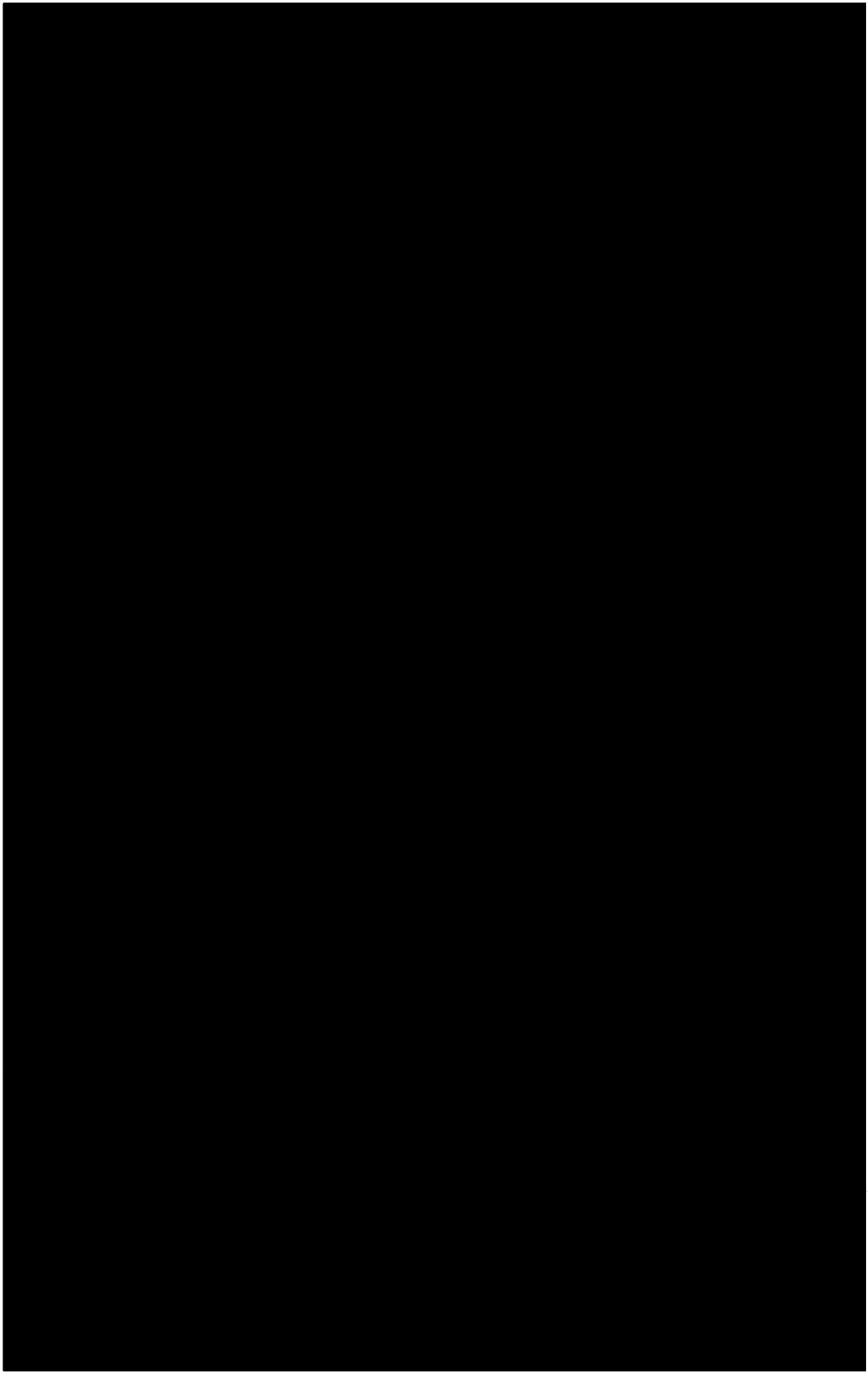


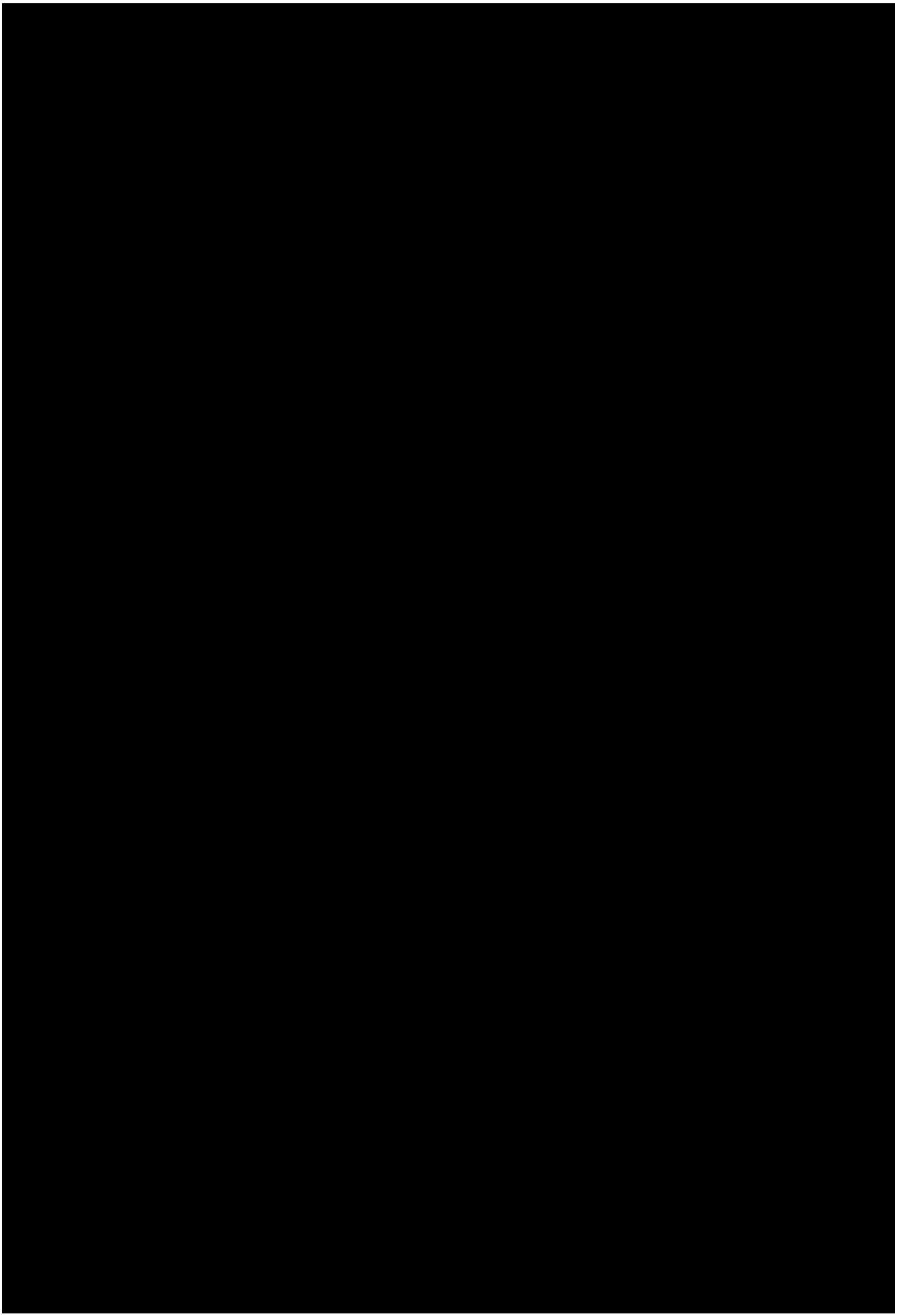


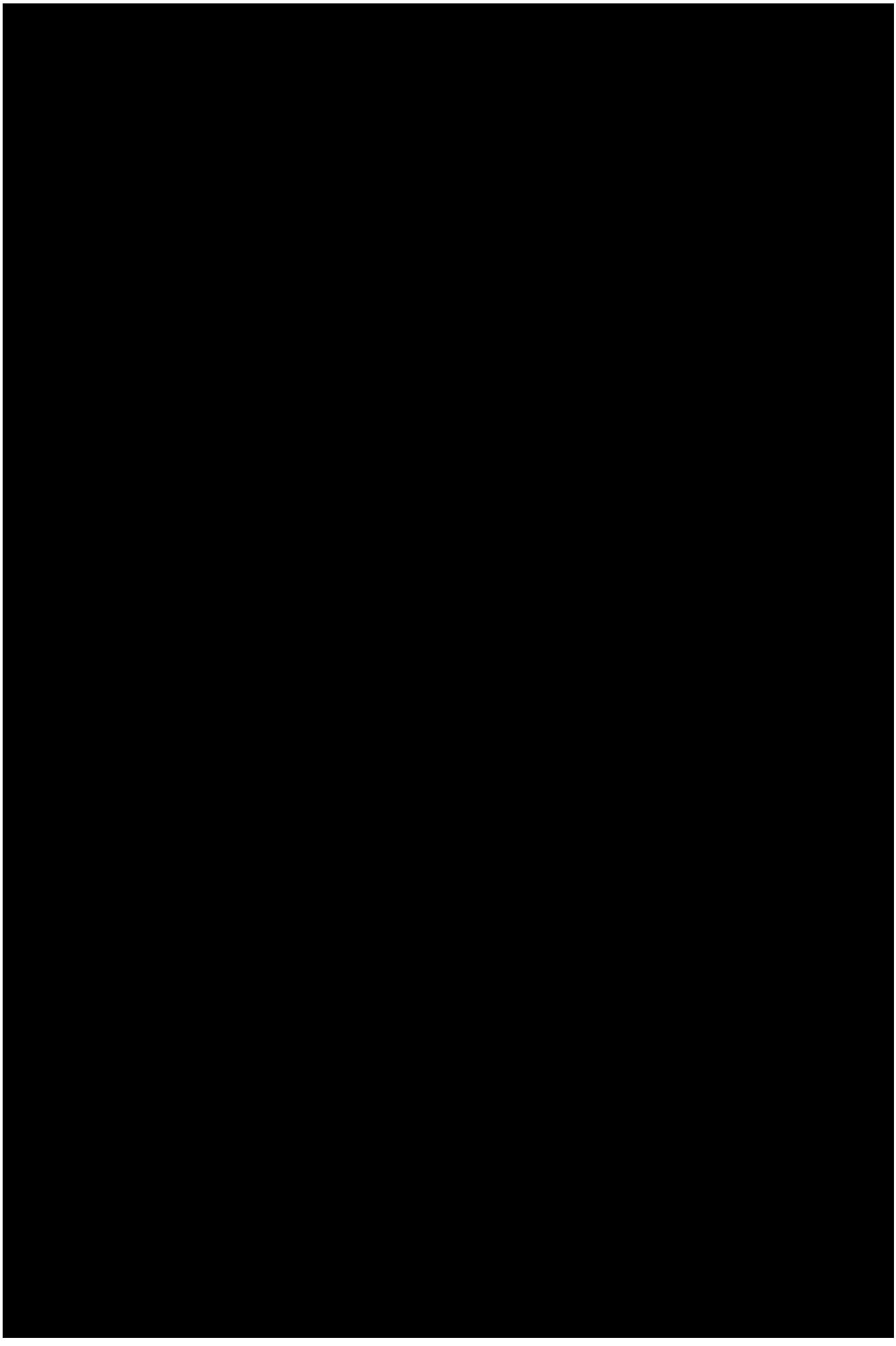


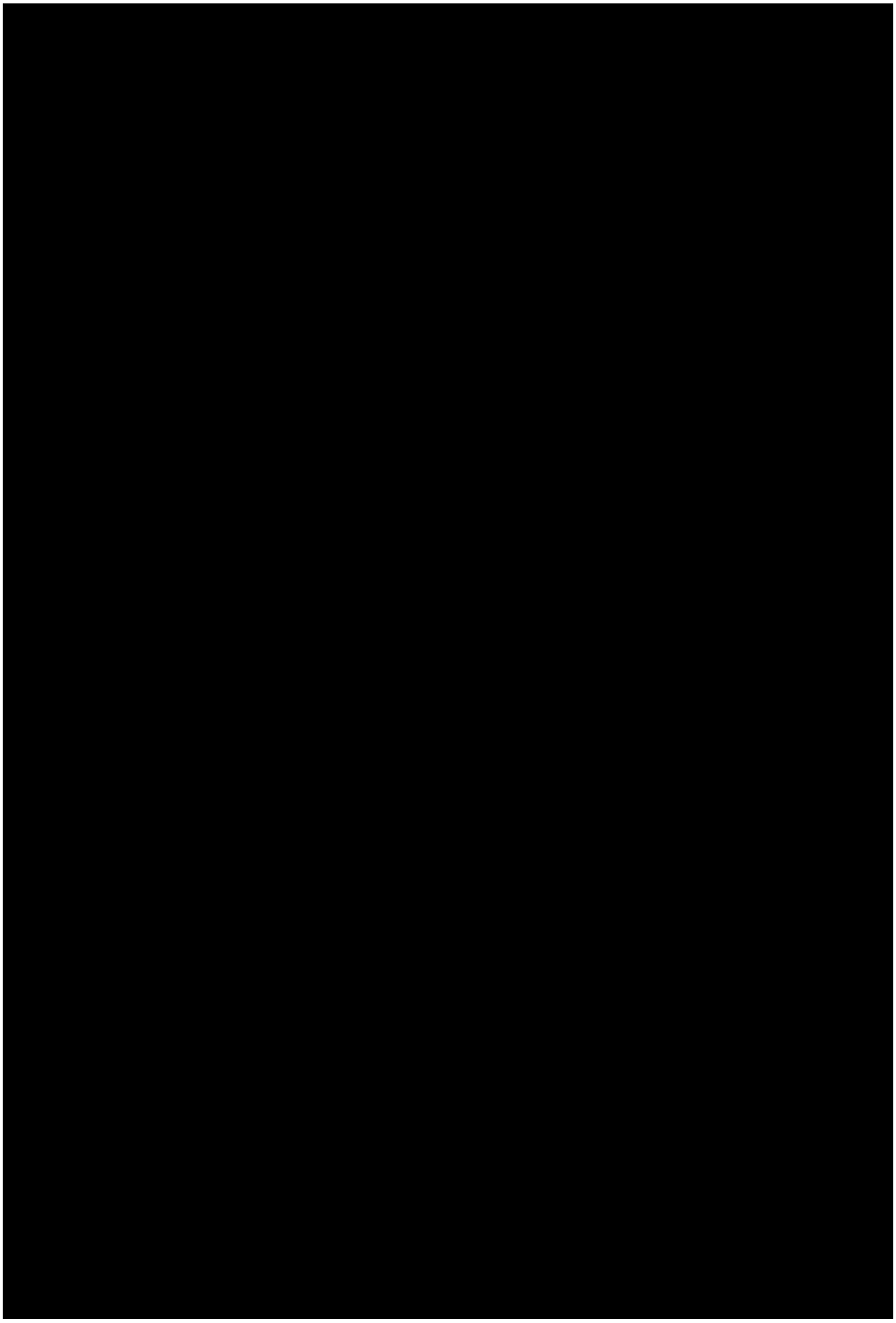


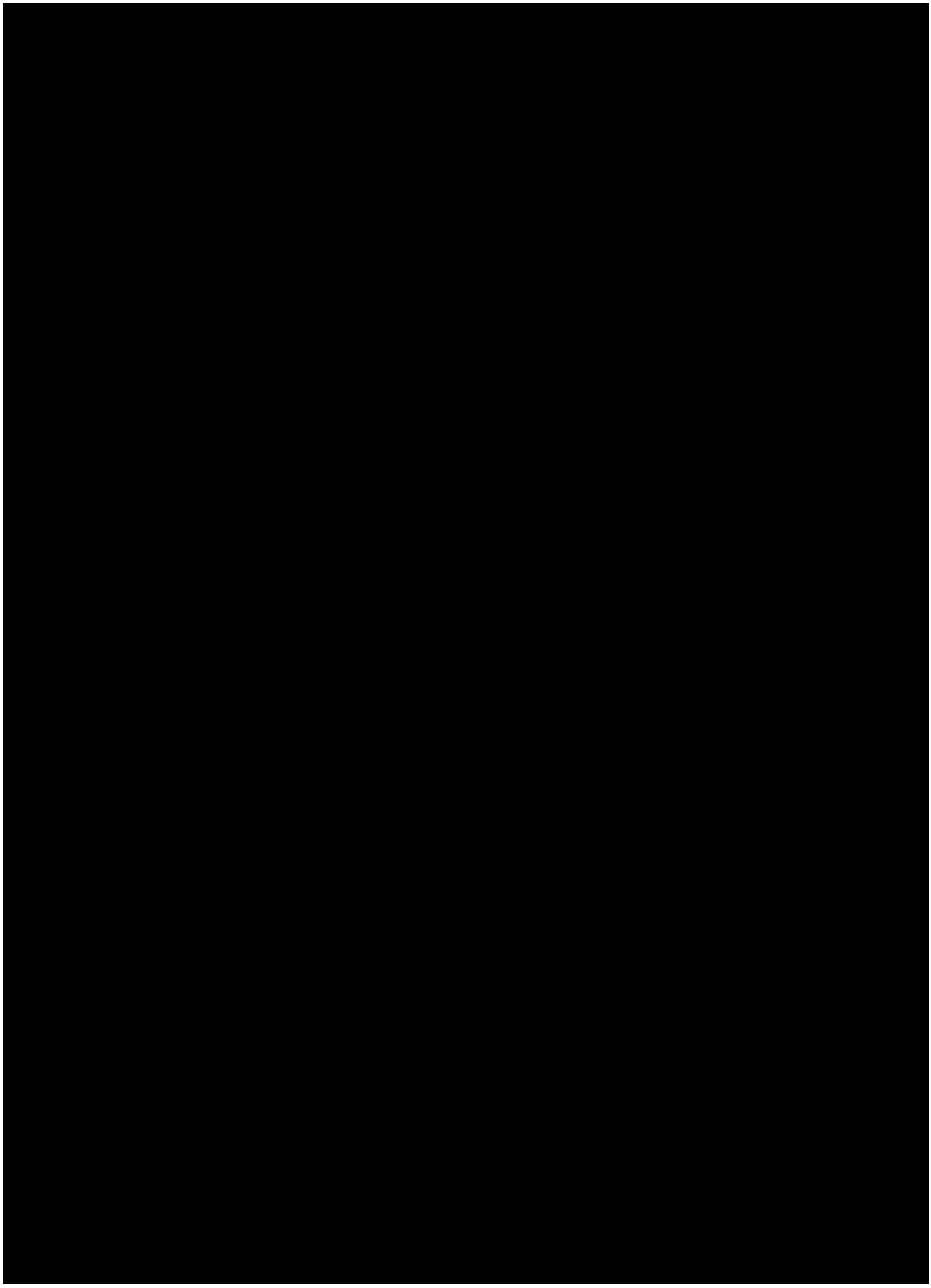


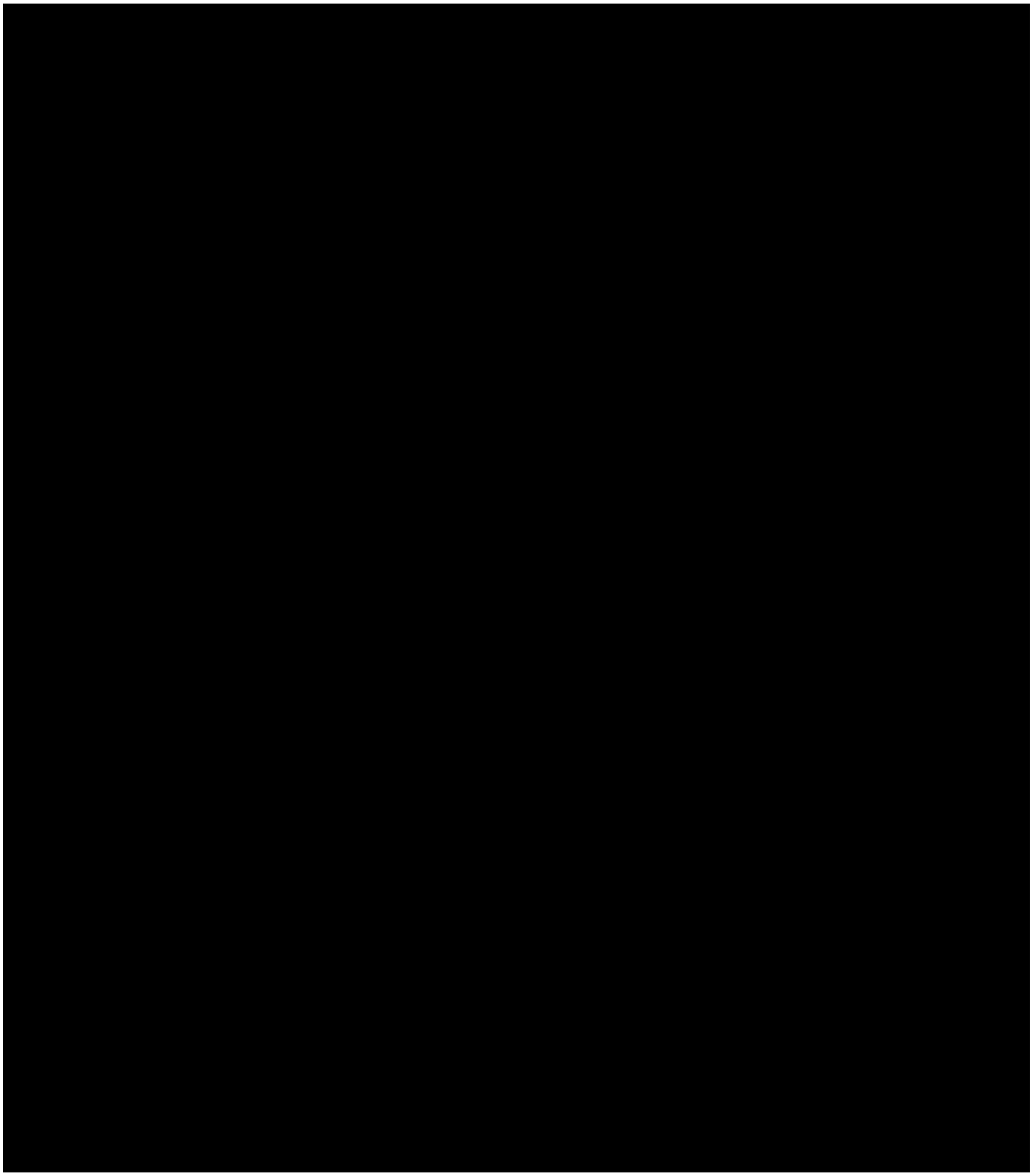












CHM 2014/5th MEETING

COMMISSION ON HUMAN MEDICINES

CARDIOVASCULAR, DI ABETES AND RENAL, RESPIRATORY AND ALLERGY EAG

Title of p aper: Actilyse (alteplase): Update on t he b alance of be nefits a nd r isks w hen used in the treatment of acute ischaemic stroke **Type of paper: For Advice**

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

Contents

Acknowledgements

We are very grateful for the contributions of [redacted] (usage data), and [redacted] (GCP inspectorate).

Executive Summary

Alteplase (rt-PA) is authorised via the mutual recognition procedure with Germany acting as RMS. It was initially licensed in the UK in October 1988 for fibrinolysis in coronary artery occlusion and massive pulmonary embolism.

In 2002 the indication was extended to include treatment of acute ischaemic stroke from 0-3 hours. Divergent opinions between member states were expressed during the procedure, with the UK expressing a negative view. The indication was eventually approved following a lengthy arbitration procedure.

UK raised further Major Objections when the time window for treatment was extended in March 2012 to allow treatment from 0-4.5 hours.

Since the approval of the extension to the time window for treatment, a large trial, the third international stroke trial (IST-3) has been published[1,2], and the current paper considers the additional evidence provided by this study as well as other data sources not previously considered.

Overall, it is considered that the majority of data do not impact on the previous conclusions made regarding rt-PA. The IST-3 trial[1,2] however failed in its primary endpoint, the proportion of patients alive and independent at 6 months post-stroke, following treatment with rt-PA within 6 hours of symptom onset, compared with control patients. In particular the results of IST-3 cast some doubt on the benefit of treating between 3-4.5 hours post symptom onset. However, in light of the substantial differences between the trial population included in the IST-3 trial and the time window for treatment, compared with the licensed conditions for use, the regulatory implications of these data for the current marketing authorisation are not completely clear.

Advice sought

Does the Commission consider that the data discussed have implications for the balance of benefits and risks of rt-PA as currently authorised in the EU?

The Commission is also asked whether they consider that the issue needs to be further explored by an ad-hoc expert working group.

1. Issue

The MHRA has been contacted by a [redacted] who has concerns regarding the balance of benefits and risks for rt-PA when used in ischaemic stroke. Dr [redacted] s request for an updated evaluation of the evidence has the support of [redacted] and a number of other influential physicians.

Despite the fact that rt-PA is licensed for the treatment of thrombolysis in acute ischaemic stroke, and has been for a considerable length of time, its use in this indication remains a source of controversy within the scientific literature and the medical community at large. Opinions on the use of rt-PA in stroke tend to be polarised and firmly held and both sides of this debate have many supporters. A poll following a recent 'head-to-head' in the BMJ on the question "Do risks outweigh benefits in thrombolysis for stroke?" was won narrowly (54% of 612 votes cast) by the 'Yes' response. The 'Yes' argument was that thrombolysis for stroke outside of the clinical trial setting presents too great a risk at present because the benefits, and the best patients to treat, have not yet been defined with a high level of scientific certainty. In addition to the opposing views regarding the scientific interpretation of the available data, the overall situation is further complicated by suggestions that clinical guideline committees recommending use of rt-PA may have too many links to industry sponsors. This latter issue is not limited to rt-PA[1].

Nevertheless, UK usage data obtained for rt-PA suggests that in 2013 nearly 15,000 single treatment doses of rt-PA were dispensed, based on a defined daily dose of 100mg. This has progressively increased from ~7,800 single treatment doses in 2009. These data represent the use of rt-PA in all indications (acute myocardial infarction (MI) and massive pulmonary embolism (PE) as well as acute ischaemic stroke) as it was not possible to obtain indication-specific information.

It is easy to understand why rt-PA treatment of stroke generates strong, polarised views – with the risk of leaving a patient with a potentially devastating ischaemic stroke untreated needing to be weighed against the risk of causing a potentially devastating intracranial haemorrhage in a patient who might otherwise have had a reasonable outcome if left untreated.

Whilst the balance of benefits and risks of use of rt-PA in ischaemic stroke have been concluded to be positive during previous European licensing procedures, the discussions during these procedures have been complex (a reflection that randomised clinical trial data has not demonstrated a substantial beneficial effect).

It is therefore particularly important that any new data or new major concerns about the data that were critical to the current marketing authorisation are thoroughly investigated, whilst giving due regard to the regulatory considerations described in the next section.

The current paper provides a summary of the previous assessments of rt-PA in the indication of acute ischaemic stroke and an assessment of the data that has become available since the treatment window was extended from 3 to 4.5 hours. It also considers data that were not discussed in detail in previous procedures but that Dr [redacted] considers pertinent to this issue.

The Commission is reminded that the following data have already been reviewed within the context of the original licence application or the extension to the treatment window:

- NINDS trials, part 1 and 2[2]
- ECASS I[3] and II[4]

- ATLANTIS part A[5] and B[6]
- ECASS III[7]
- SITS-ISTR observational registry[8]
- Some additional data from other observational sources[9-11]

The additional data considered in this paper are:

- IST-3 trial[12,13]
- Re-analysis of the NINDS trial data[14]
- Observational cohort data from the Oxfordshire community stroke project, the Lothian stroke register and the first international stroke trial (in the UK)[15]
- Unpublished data from a meta-analysis presented recently at the American Stroke Association meeting

This paper does not evaluate how rt-PA is being used in everyday clinical practice in the UK and whether or not the conditions for use as described in the SmPC are being adhered to.

The Commission is asked to consider whether these new data potentially impact on the overall balance of benefits and risks of rt-PA in the treatment of acute ischaemic stroke and whether they consider that the issue needs to be further explored by an ad-hoc expert working group.

2. Regulatory considerations

In 1988 rt-PA was authorised in the UK via the mutual recognition procedure with Germany acting as RMS. It was initially licensed for fibrinolysis in coronary artery occlusion and massive pulmonary embolism.

In 2002 the indication was extended to include treatment of acute ischaemic stroke from 0-3 hours. Divergent opinions between member states were expressed during the procedure, with the UK expressing a negative view. The indication was eventually approved following a lengthy arbitration procedure.

UK raised further Major Objections when the time window for treatment was extended in March 2012 to allow treatment from 0-4.5 hours.

The indication for use of rt-PA in ischaemic stroke and the subsequent variation to extend the time window for treatment to 4.5 hours have therefore been extensively debated both at CHM and within Europe, with member states expressing divergent viewpoints during the procedures. All the data available, and considered relevant, at the time of licensing and during the extension to the indication have been comprehensively reviewed within Europe and all the MHRA's/Commission's concerns that were raised at the time were addressed through the appropriate means, including submission of further data and oral explanations by the company at CHMP.

As any re-examination of rt-PA in the indication of ischaemic stroke within Europe would have to be based on new evidence, we have restricted this assessment to data that has since become available or that have a bearing on assessments conducted previously. If there is sufficient concern about any aspect of the existing marketing authorisation we can trigger a review within Europe using the Article 31 referral procedure.

3. Background: previous assessments

Alteplase (rt-PA) is a recombinant human tissue-type plasminogen activator (t-PA). It is produced by expression of the human gene for t-PA in CHO cells. The mechanism of action of rt-PA is understood to be the enzymatic cleavage of plasminogen to plasmin with subsequent increase in fibrinolysis. In the indication of acute ischaemic stroke the recommended dose is 0.9 mg rt-PA/kg body weight (maximum of 90 mg) infused intravenously over 60 minutes with 10% of the total dose administered as an initial intravenous bolus. Product information currently includes a negative benefit:risk statement for administration beyond the 4.5 hour window, and a reminder that treatment must be started as early as possible within the 4.5 hours.

As the Commission has already considered in detail the data in support of the original licence application and the extension to the indication a high level summary only is provided below.

3.1 Initial approval of the indication in acute ischaemic stroke (0-3 hours)

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. Approval of the indication had been granted in Germany in 1999.

The studies that formed the basis of the assessment were NINDS part 1 and 2[5], ECASS I[6] and II[7], ATLANTIS[8].

NINDS part 1 and 2 (0-3 hours of symptom onset)[2]

These two studies were conducted in the US by the US National Institute of Neurological Disorders and Stroke (NINDS). Both studies were placebo-controlled, used a dose of 0.9mg/kg and treatment was within 0-3 hours of symptom onset. NINDS part 1 was a phase II study (n=291), and NINDS part 2 was phase III (n=333).

The primary endpoint for NINDS part 1 was neurological outcome after 24 hours measured as an improvement from baseline in the NIH stroke scale (NIHSS; see Glossary) of 4 or more points or complete resolution of neurologic deficit. The primary endpoint was not reached as there was no significant difference between the rt-PA and the placebo groups at 24 hours. However there was a benefit observed at 3 months after treatment in the rt-PA group.

NINDS part 2 was the pivotal randomised, placebo controlled trial supporting the application. The primary endpoint was clinical outcome at 3 months, according to scores on the Barthel index (BI; see Glossary), modified Rankin scale (mRS; see Glossary), Glasgow outcome scale (GOS; see Glossary) and NIHSS. The odds ratio for a favourable outcome (minimal or no disability at 3 months) in the rt-PA group compared with placebo was 1.7 (95% CI [1.2-2.6]). The absolute increase in number of patients with minimal disability was 12% at 90 days. Symptomatic intracranial haemorrhage (ICH) within 36 hours of stroke onset occurred in 6.4% of rt-PA treated patients vs. 0.6% of placebo patients. This did not translate into an increase in mortality. Overall mortality in the rt-PA group was 17%, vs. 21% in the placebo group. Stratifying patients by the time window for treatment suggested that the benefit of rt-PA over placebo is greater in the first 1.5 hours compared with the second 1.5 hours.

ECASS I (0-6 hours of symptom onset)[3]

The European Co-operative Acute Stroke Study (ECASS I) was a randomised double-blind placebo controlled trial in 14 European countries, designed to evaluate efficacy and safety of rt-PA in patients with acute ischaemic stroke with moderate to

severe neurological deficit and with none or minimal early infarct signs on the initial CT scan. This was a phase III study, n=610.

The primary endpoints included BI and mRS at 90 days. Secondary endpoints included combined BI, mRS and Scandinavian stroke scale (SSS) at 90 days and 30 days mortality. The dose of rt-PA used was 1.1 mg/kg body weight within 6 hours of symptom onset.

There was no difference between the groups in the intention to treat (ITT) analysis of the primary endpoints. The secondary endpoint of combined BI and mRS demonstrated a difference in favour of rt-PA treatment. Mortality at 90 days was higher in the rt-PA population (22.4% vs. 15.8% for placebo) and parenchymal haemorrhages were significantly more frequent in the rt-PA group (ITT: n=62 for rt-PA vs. n=20 for the placebo group).

ECASS II (0-6 hours of symptom onset)[4]

The European-Australasian Acute Stroke Study (ECASS II) used a lower dose of rt-PA (0.9mg/kg body weight) to match that used in NINDS. This was a phase III study, n= 800. Treatment was given within 6 hours of symptom onset, stratified into 0-3 hours and 3-6 hours. Due to issues with patients receiving early treatment such as the time taken between onset of symptoms and arrival at hospital, only 81 out of 409 rt-PA treated patients were included in the early stratum.

The primary endpoint was the proportion of patients with a favourable outcome (score 0 or 1) on the mRS at day 90 after treatment. No significant difference was found between rt-PA and placebo for the primary endpoint. The study found no evidence that efficacy depends upon administration within 3 hours of symptom onset, however there were only a small number of patients in the 0-3 hour time window.

Symptomatic ICH occurred in 8.8% of rt-PA patients and in 3.4% of placebo patients, but no increase in morbidity or mortality at day 90 was observed in the rt-PA group compared with placebo.

ATLANTIS part B (3-5 hours of symptom onset)[6]

The Alteplase ThromboLysis for Acute Non-interventional Therapy in Ischaemic Stroke (ATLANTIS) study was a placebo-controlled, double-blind, randomised study conducted in North America. It was initially designed to assess rt-PA administered from 0-6 hours following onset of symptoms. Two years into the study the DSMB halted enrolment and the time-window for treatment was changed to 0-5 hours due to safety concerns in the 5-6 hour group. At this point the trial was re-started as part B[6], with the previously enrolled patients to be considered separately as part A[5]. Part B was further modified 2 years later to a time window of 3-5 hours in light of the NINDS trial results for the 0-3 hour window. 31 patients in part B had been enrolled from 0-3 hours at the time of this change.

Part B was a phase III study, n=613. The dose of rt-PA used was 0.9mg/kg, as used in NINDS. This study was not included in the initial submission by the applicant for this indication. The trial endpoints were changed during the study[1] (for reasons unknown to this assessor), at the time of publication of part B the primary endpoint was the number of patients with an excellent neurological recovery at day 90 (score 0-1 on the NIHSS). Secondary endpoints were excellent recovery on BI, mRS and GOS scales at days 30 and 90.

[1] The ATLANTIS part A publication[5] states primary hypotheses as: 1. Significant difference between rt-PA and placebo groups in clinical improvement, (decrease of ≥4 points on the NIHSS or complete resolution of symptoms from baseline to 24 hours/30 days); 2. Significant difference between rt-PA and placebo groups in volume of cerebral infarction as measured by CT scanning at 30 days.

For the primary outcome, 32% of placebo and 34% of rt-PA patients had an excellent recovery at 90 days. There were no differences in secondary outcome measures.

In the first 10 days following treatment, the rate of ICH was higher in the rt-PA treated group than in the placebo group (symptomatic ICH: 11.4% rt-PA vs. 4.7% placebo). Mortality at 90 days was 11% in the rt-PA group, vs. 6.9% in the placebo group.

This trial was stopped prematurely after a pre-planned interim analysis, as the DSMB considered a beneficial effect seemed unlikely.

Part A included 142 patients and also found no significant benefit for any of the planned efficacy endpoints, and an increased risk of ICH.

3.1.1. Initial comments and conclusions of medical and statistical assessors

When these data were initially assessed, the UK assessors commented:

- Of the total of 2647 patients in the 5 studies, 1341 received rt-PA treatment. However, only 405 patients received the dose that is proposed for marketing (0.9mg/kg) within the proposed time window (0-3 hours).
- NINDS part 2 demonstrated a positive balance of benefits and risks however ECASS I, II and ATLANTIS did not.
- A direct comparison across these trials is not possible due to the differences in time window for treatment, dosage used, target parameters and patient population.
- In NINDS part 2, the benefit of rt-PA treatment over placebo appears to be greater for the patients treated in the first 1.5 hours compared with patients treated from 1.5-3 hours. The benefit in NINDS part 2 was considered to be rather modest.
- Statistical assessment concluded that all of the studies were well-designed and conducted.
- No formal dose finding study was done, the dose was pragmatically derived from treatment of acute MI and PE, as well as previous studies in stroke.
- The frequency of ICH appeared greater at the higher dose used (1.1mg/kg) and is still cause for some concern at the proposed dose (0.9mg/kg).
- The applicant submitted an integrated analysis of the NINDS and ECASS studies, however this was not a pre-planned analysis, and the primary endpoints, protocols, dosages and stroke severity at baseline varied across these trials. ATLANTIS was not included.

CSM agreed that the application could not be granted and they agreed the points raised as serious public health concerns and points for clarification.

3.1.2 Arbitration procedure for the initial application for treatment of acute ischaemic stroke (treatment within 0-3 hours of symptom onset)

The UK raised serious public health concerns during the application for treatment of acute ischaemic stroke, and an arbitration procedure was required. The Netherlands, France, Spain and Greece also raised concerns.

CPMP was requested to consider “whether there was sufficient clinical data with reference to efficacy as well as safety e.g. risk of intra-cranial bleeding, to grant a Marketing Authorisation for the proposed new indication for fibrinolytic treatment of acute ischaemic stroke without putting the public health at risk. In particular, the major concern for most of the CMS was the lack of favourable results of the EU studies (ECASS I and II), nor in another US trial (Atlantis)”.

The Rapporteur for the procedure was Netherlands, with France acting as Co-Rapporteur. The main issues addressed during the arbitration procedure were:

- Differences between the NINDS, ECASS and ATLANTIS studies and the characteristics of the patients: for patients treated within 0-3 hours, European patients had lower stroke severity compared with US patients, and also had other favourable prognostic factors in terms of blood glucose level, prior MI and prior hypertension. Mean and median time to treatment was higher in the European patients.
- Discrepancy in the results of ECASS II [4% more patients with favourable outcome under rt-PA, but 5.8% more deaths] compared with NINDS part 2 [13% more patients with a favourable outcome under rt-PA, and 3% less deaths].
- A number of additional information sources were cited by the Co-Rapporteur to support the argument in favour of granting the indication, including:
    - Improvement from 25% of patients to 40% of patients with complete recovery when rt-PA is used [German centre experience, Koennecke *et al* Stroke 2001][9]
    - Review of stroke patients in Indianapolis showed that when the restrictions of the NINDS trial are observed, the ICH and mortality rates are similar to NINDS [Lopez-Yunez *et al*, Stroke 2000][10]
    - Other post-marketing studies e.g. STARS [JAMA, 2000][11], ICH was lower than in NINDS
    - European Stroke Initiative Recommendations [Cerebrovascular Disease 2000][16] of use of rt-PA.
- The Co-Rapporteur also considered that granting of the indication in acute ischaemic stroke would result in development of dedicated stroke centres.

An EMA Scientific Advice Group (SAG) meeting considered that rt-PA was already a widely accepted therapy amongst neurologists and that the results from US trials could be extrapolated to the EU setting. They believed that it is possible to use rt-PA in accordance with the SmPC restrictions in order to maximise benefit. They also considered that the meta-analysis (conducted by the MAH) findings were trending in the direction of the US trials but patient numbers were too small. Finally they considered that a placebo-controlled trial in patients treated up to 3 hours after symptom onset would likely face recruitment difficulties and may not be feasible; however either a study with the time window of 3-4 hours may be possible or a study in high-risk subgroups within 3 hours of symptom onset.

After several rounds of assessment, oral explanations by the company and the SAG meeting, over a period of 18 months, the CPMP voted by majority in June 2002 that the indication should be conditionally granted. Post-licensing commitments were required:

- A further randomised placebo-controlled trial (ECASS III) to assess efficacy and safety within 3-4 hours of symptom onset
- A post-marketing surveillance study (SITS-MOST)

The UK remained negative at the time of this vote but the outcome of the arbitration procedure was legally binding in all member states.

3.2 Extension of time-window for treatment (0-4.5 hours)

In February 2010, the MAH applied to extend the treatment administration window from 0-3 hours to 0-4.5 hours after onset of symptoms. This application was based

mainly on the ECASS III trial data, with some supportive data from SITS-ISTR (observational registry) and a pooled analysis.

All member states involved in the procedure other than the UK were positive. The UK raised a major concern for public health supported by the Neurology EAG and CHM:

"The benefit:risk balance is not considered positive for patients treated between 4-4.5 hours.

From the pooled analysis the point where the estimated mortality is against active treatment is 4 hours. Because of the way the data is presented from the pooled trials only those in the treatment group with OTT [onset to treatment] of 271-360min have a marked increase in OR estimate of 1.69 for death in the active treatment arm.

In line with the pooled data, the ECASS III study subgroup analysis for mortality (ITT) showed a higher incidence of mortality with increasing time to treatment.

Because OTT is a continuous variable, and completely accurate times for onset of stroke are rarely available, a cut-off of 270 mins needs further justification.

The gap in sICH (symptomatic ICH, NINDS definition) between active and placebo group widens over the three time periods (3-3.5 hrs, 3.5-4 hrs and 4-4.5 hours) such that the rate of ICH remains the same in the active treatment group but progressively reduces in the placebo group. These factors combined with the modest benefit in the group overall from 3-4.5 hours, does not support a positive benefit: risk ratio for those treated at 4-4.5 hours. The MAH are requested to provide a full analysis of the cases treated between 4-4.5 hours in terms of baseline characteristics, efficacy and safety."

ECASS III (3 - 4.5 hours of symptom onset)[7]

The European Cooperative Acute Stroke Study III (ECASS III) was a randomised, multi-national, double-blind, placebo controlled trial in patients treated with rt-PA between 3 and 4.5 hours after stroke onset. The study took place between 2003 and 2008 and enrolled 821 patients. Inclusion and exclusion criteria were identical to the EU SmPC for rt-PA in ischaemic stroke, except for the time-window for treatment.

rt-PA was administered at a dose of 0.9 mg/kg body weight (the licensed dose). The primary outcome was mRS 0-1 at Day 90. The secondary outcome was a global measure combining Day 90 results for the mRS score 0-1, the BI score ≥95, the NIHSS score of 0-1 and the GOS score of 1.

The safety endpoints included overall mortality at day 90, stroke-related and neurological deaths, all ICH, symptomatic ICH, and symptomatic brain oedema.

ECASS III results

Treatment with rt-PA was significantly associated with a favourable primary outcome (mRS = 0-1 at day 90) compared with placebo in the intention to treat population:

OR 1.34; 95% CI [1.02-1.76]
RR 1.16; 95% CI [1.01-1.34]

The more global secondary endpoint for the intention to treat population was also found to have a statistically significant difference in favour of rt-PA treatment:

OR 1.28; 95% CI [1.00-1.65].

The per-protocol results found similar, slightly greater ratios. The effect on the distribution of mRS scores is shown in the following figure.

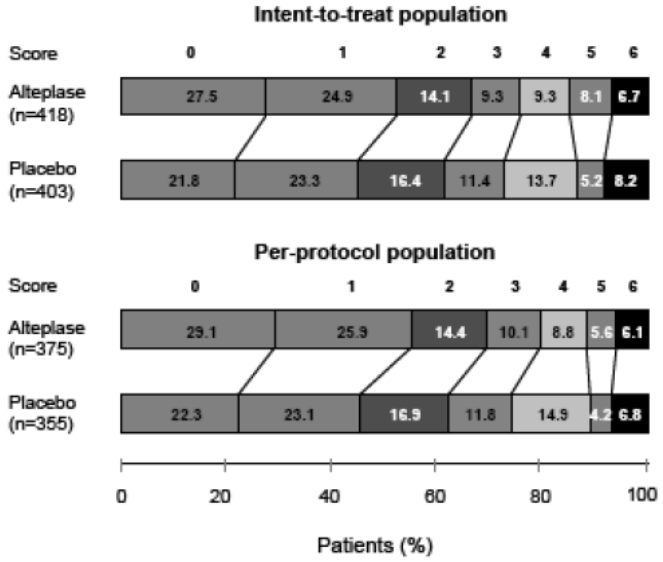


Figure 1: Overall distribution of scores on the mRS at the day 90 visit

A total of 113 patients (27%) in the rt-PA group had intracranial haemorrhages of which 3 were fatal. This compares with 71 patients (17.6%) with ICH in the placebo group of which 0 were fatal. Most ICH occurred within 24 hours of receiving treatment. The OR for any ICH was 1.73 95% CI [1.24-2.42]. Symptomatic ICH was defined as any blood in the brain or intracranial associated with a clinical deterioration of ≥4 points of the NIHSS for which haemorrhage has been identified as the dominating cause. All symptomatic ICH occurred within the first 22 to 36 hours after initiation of treatment. Symptomatic ICH frequencies were also estimated using definitions from previous studies[2].

	Alteplase n (%)	Placebo n (%)	Odds Ratio (95% CI)	p value[1]
ITT population	418 (100.00)	403 (100.00)		
As per ECASS III definition	10 (2.39)	1 (0.25)	9.85 (1.26-77.32)	0.0076
As per ECASS II definition	22 (5.26)	9 (2.23)	2.43 (1.11-5.35)	0.0228
As per SITS-MOST definition	8 (1.91)	1 (0.25)	7.84 (0.98-63.00)	0.0219
As per NINDS definition	33 (7.89)	14 (3.47)	2.38 (1.25-4.52)	0.0064
PP population	375 (100.00)	355 (100.00)		
As per ECASS III definition	7 (1.87)	1 (0.28)	6.73 (0.82-55.01)	0.0398
As per ECASS II definition	19 (5.07)	9 (2.54)	2.05 (0.92-4.60)	0.0751
As per SITS-MOST definition	5 (1.33)	1 (0.28)	4.78 (0.56-41.15)	0.1157
As per NINDS definition	28 (7.47)	14 (3.94)	1.97 (1.02-3.80)	0.0410

Table 1: Overall incidence of symptomatic ICH in ECASS III (according to different definitions).

The percentage of patients with symptomatic ICH was found to remain the same for the rt-PA treated group across the three time periods of treatment (3-3.5, 3.5-4 and

[2] Definition of sICH according to:
ECASS II: Any intracranial bleed and at least 4 points worsening on the NIHSS score (the same as the **ECASS III** protocol definition except that the causal relationship between haemorrhage and clinical deterioration was not required).
SITS-MOST: Local or remote parenchymal haematoma type 2 on the 22- to 36-hour posttreatment 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 hours, or leading to death.
NINDS: A haemorrhage was considered symptomatic if it was not seen on a previous CT scan and there had subsequently been either a suspicion of haemorrhage or any decline in neurological status. To detect intracranial haemorrhage, CT scans were required at 24 hours and 7 to 10 days after the onset of stroke and when clinical finding suggested haemorrhage.

4-4.5 hours) using NINDS criteria. For the placebo group, the rate of symptomatic ICH reduced with longer time to treatment.

Overall mortality rates were similar between the two groups, in the ITT population a total of 32 (7.7%) of patients in the rt-PA arm died whilst 34 (8.4%) in the placebo group died. A trend for increasing mortality with increasing time to treatment was found.

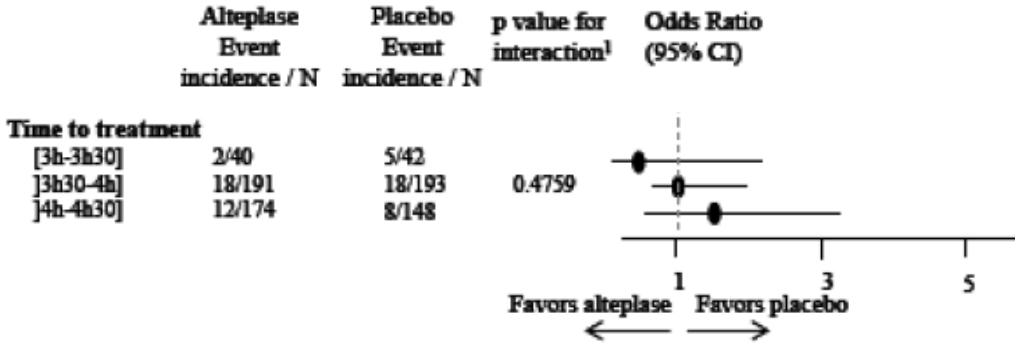


Figure 2: Subgroup analysis (time to treatment) for mortality (ITT population)

Elderly patients (≥65 years) were found to have a trend for increased mortality, an increased risk for symptomatic ICH and a trend to lower efficacy.

SITS-ISTR (observational registry)[8]

SITS-ISTR was a prospective multi-national registry study for patients given rt-PA following stroke. Data were collected from 2002 to 2007, and a later update to 2008. Two cohorts were compared, patients treated between 0-3 hours of symptom onset (n=11,865) and those treated between 3-4.5 hours (n=664). The main efficacy endpoint was functional independence (mRS ≤2) at day 90. Safety endpoints were symptomatic ICH within 24 hours and mortality at day 90.

SITS-ISTR results

Patients in the later time period had a median age that was 3 years younger than the earlier treatment time-point and a stroke severity 1 point lower on the NIHSS. In the 3-4.5 hour cohort, ~60% of patients were treated in the first 20 minutes, and only ~8% in the last 30 minutes. For the 3-4.5 hour cohort compared with the 3 hour cohort:

Rate of sICH (SITS-MOST definition) 2.2% vs. 1.6%; OR 1.18 [95% CI 0.89–1.55], p=0.24; adjusted OR 1.32 [95% CI 1.00–1.75], p=0.052;

Rate of sICH (ECASS II definition): 5.3% vs. 4.8%; OR 1.06 [95% CI 0.89–1.26], p=0.54;

Rate of sICH (NINDS definition): 8.0% vs. 7.3%; OR 1.06 [95% CI 0.91–1.22], p=0.46.

Mortality rate: 12.7% vs. 12.2%; OR 1.02 [95% CI 0.90–1.17], p=0.72; adjusted OR 1.15 [95% CI 1.00–1.33]; p=0.053.

Independence (mRS ≤2 at day 90): 58.0% vs. 56.3%, OR 1.04 [95% CI 0.95–1.13], p=0.42; adjusted OR 0.93 [95% CI 0.84–1.03], p=0.18.

Pooled analysis

This pooled analysis combined data from ECASS III with ECASS II, ATLANTIS A and B, NINDS 1 and 2. ECASS I was not included as it used a higher dose. Overall a total of 2958 patients (1490 rt-PA and 1468 placebo) treated within the 0-6 hour time

window were included. Of these, 1355 patients were treated in the 3-4.5 hour time window.

The primary efficacy endpoint was mRS 0-1 at day 90. A benefit in favour of rt-PA was observed for the 3-4.5 hour time window, OR 1.31; 95% CI [1.06-1.63] p=0.014. For the subset of patients who fulfilled the inclusion/exclusion criteria other than time to treatment, according to the SmPC, a significant benefit was also found, n=1251; OR 1.42; 95% CI [1.13-1.78], p=0.002.

Rates of symptomatic ICH in the 3-4.5 hour cohort were comparable with those from ECASS III, according to the SITS-MOST definition. The pooled analysis found the excess risk with rt-PA compared with placebo was slightly smaller for 0-3 hours compared with 3-4.5 hours, and was markedly higher for patients treated after 4.5 hours. Results from ECASS III also found a numeric trend of increasing risk of sICH with increasing time to treatment onset.

No difference was found in the rate of all-cause mortality in the rt-PA compared with the placebo arms, and mortality rate was similar to that in the ECASS III trial. Risk of mortality in the rt-PA arm increased with increasing time to treatment.

3.2.1 UK major objection and final approval of extended time window

At a meeting between the MAH and MHRA Licensing Division, the MAH presented their arguments for a favourable balance of benefits and risks in the 3-4.5 hour time window.

A particular concern had been the increase in death rates observed in the rt-PA group compared with placebo in the sub-group 4-4.5 hour time window, rt-PA: 12/147 (6.8%) vs. placebo 8/148 (5.4%). It transpired that the MAH had included additional deaths reported after day 90, such that deaths were not collected over an equal length of follow-up in the two arms. When these deaths (5 in the rt-PA group and 2 in the placebo group) were excluded, rt-PA was not found to have an unfavourable effect on mortality at 4-4.5 hours.

A further argument put forward during the meeting was that even though there is no difference in severe disability and death, many patients will be willing to accept the risk of treatment if they have even a small chance of complete recovery.

Following this meeting and discussion of the data and MAH responses, including the pooled analysis, CHM agreed that there is some benefit in the group overall and there was no negative effect on mortality. The variation to extend the indication to 4.5 hours was considered acceptable and finally granted in March 2012.

4. Update: Assessment of newly available data

4.1 IST-3 trial[12,13] (see Annex 3 for publications)

4.1.1. Study description

The third international stroke trial (IST-3) was intended to determine whether a wider range of patients than the licensed population would benefit from thrombolysis, in particular patients >80 years of age and patients treated up to 6 hours after symptom onset.

IST-3 was an international, randomised, open-label trial. The initial phase was double-blind and placebo controlled (n=276), and in total for both phases of the trial 3035 patients were enrolled (rt-PA: n=1515; control: n=1520). The study period was from 2000 to 2011.

The primary trial hypothesis was that 0.9 mg/kg rt-PA within 6 hours of symptom onset increased the proportion of people who were alive and independent at 6 months, as measured by the Oxford Handicap Score (OHS; see Glossary). Patients with OHS of 0, 1 or 2 were classed as independent.

Patient outcome was recorded at 7 days and 6 months, and a proportion of patients were additionally followed up at 18 months.

4.1.2. Study population

The eligibility criteria included: symptoms/signs of clinically definite acute stroke, time of stroke onset was known, treatment could be started within 6 hours of onset, and CT or MRI had reliably excluded both intracranial haemorrhage and structural brain lesions.

It should be noted that if a patient had either a clear indication for treatment with rt-PA or for whom benefit-risk of treatment would clearly be negative (e.g. those with ICH), the patient was not entered into the trial: patients were only included if treatment was considered promising but unproven.

The contraindications for rt-PA (section 4.3 of the SmPC) include the following:

- Symptoms beginning longer than 4.5 hours prior to treatment
- Adults over 80 years of age (stated as 'not indicated for treatment in…')
- Severe stroke assessed clinically e.g. NIHSS >25 or by imaging
- Systolic blood pressure >185 mmHg or diastolic >110 mmHg
- Blood glucose <50 or >400 mg/dL [<2.8 mmol/L or >22 mmol/L]

In line with the intention of enrolling patients that were not clearly indicated to receive treatment with rt-PA, the baseline characteristics of patients included in IST-3 show that a sizable proportion had characteristics close to or included within these contraindications, as shown in the table below:

Baseline variable	**rt-PA (n=1515)**	**Control (n=1520)**
Delay in randomisation 4.5-6.0 h >6 h	 507 (33%) 0	 500 (33%) 2 (<1%)
Age (years) 81-90 >90	 706 (47%) 111 (7%)	 701 (46%) 99 (7%)
NIHSS >20	213 (14%)	214 (14%)
Systolic blood pressure ≥165 mm Hg	530 (35%)	510 (34%)
Diastolic blood pressure ≥90 mm Hg	500 (33%)	480 (32%)
Blood glucose ≤5 mmol/L ≥8 mmol/L	 254 (18%) 455 (33%)	 285 (21%) 456 (33%)

Table 2: Selected baseline characteristics of patients included in IST-3

The authors report that 95% of enrolled patients did not meet the terms of the EU approval for treatment.

4.1.3. Results

Outcomes at 7 days

There was a statistically significant increase in the number of deaths in patients who received rt-PA compared with control patients by day 7 post-stroke. 163 (11%) of patients in the rt-PA group had died compared with 107 (7%) of control patients, OR =1.6 [95% CI 1.22-2.08].

Fatal intracranial haemorrhage was significantly increased in the rt-PA group compared with control, as was fatal swelling of the original infarct. There was also a significant increase in non-fatal events of neurological deterioration not due to swelling or haemorrhage in the rt-PA group compared with control. No significant difference between the groups was seen in rates of recurrent ischaemic stroke, recurrent stroke of unknown type, or myocardial infarctions. Table 2 provides the rates of ICH and swelling of the infarct.

	rt-PA (n=1515)	**Control (n=1520)**	**Odds ratio* [95% CI]**	**p value**
Symptomatic swelling of original infarct				
- Non-fatal	21 (1%)	17 (1%)	1.23 [0.64-2.35]	0.539
- Fatal	47 (3%)	25 (2%)	1.89 [1.14-3.14]	0.013
- Total	68 (4%)	42 (3%)	1.66 [1.11-2.49]	0.014
Symptomatic ICH				
- Non-fatal	49 (3%)	9 (1%)	5.56 [2.72-11.4]	<0.0001
- Fatal	55 (4%)	7 (<1%)	8.12 [3.68-17.9]	<0.0001
- Total	104 (7%)	16 (1%)	6.94 [4.07-11.8]	<0.0001
Total deaths (cerebral causes)	145 (10%)	87 (6%)	1.76 [1.32-2.34]	0.0001
Total deaths (non-cerebral causes)	18 (1%)	20 (1%)	0.89 [0.47-1.69]	0.717
Total deaths overall	163 (11%)	107 (7%)	1.60 [1.22-2.08]	0.001

*Table 3: Day 7 outcomes for rt-PA vs. control. ICH = intracranial haemorrhage. *Odds ratio adjusted for age, NIHSS, time, and presence/absence of visible acute ischaemic change on baseline scan.*

Outcomes at 6 months (including primary outcome)

Vital status at 6 months was known for 99% (3011 of 3035) of patients. Patients recruited within 1-2 hours of symptom onset were significantly more likely to have more severe baseline deficit than those recruited later, and were significantly more likely to be older than those recruited later.

Primary outcome: In the rt-PA group at 6 months follow-up, 554 (37%) of patients were alive and independent in activities of daily living (OHS 0-2) compared with 534 (35%) in the control group. Adjusted OR 1.13, 95% CI [0.95-1.35] (p=0.181); OR adjusted for age, NIHSS, time, and presence/absence of visible acute ischaemic change on baseline scan.

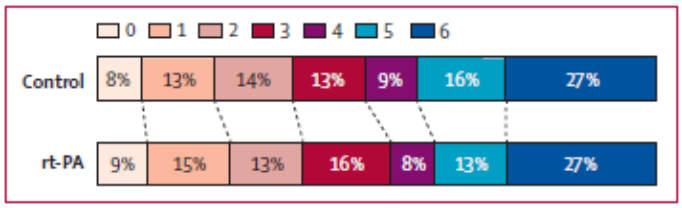


Figure 3: Outcome at 6 months, OHS by treatment group.

A secondary ordinal analysis found a favourable shift in distribution of OHS scores at 6 months. This analysis was adjusted for age, NIHSS, delay and presence/absence

of visible acute ischaemic changes on baseline scan. With OHS levels 4, 5 and 6 grouped and 0, 1, 2, 3 all discrete, the odds ratio was 1.27 [95% CI 1.10-1.47]. The corresponding OR with all levels discrete was 1.17 [95% CI 1.03-1.33].

Pre-defined sub-group analyses of the primary outcome (alive and independent at 6 months) were conducted, adjusted for age, NIHSS and time, to take into account that for a specific prognostic factor the distribution of other factors might differ between sub-categories.

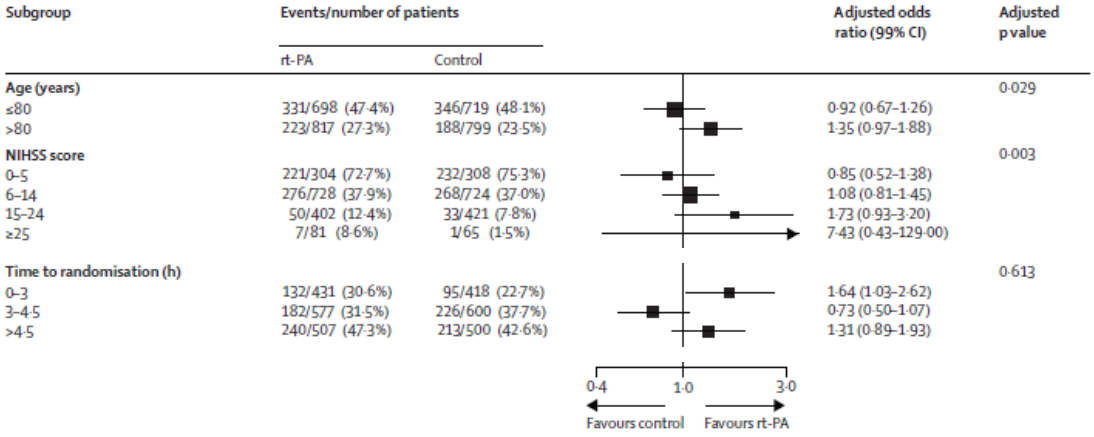


Figure 4: Selected subgroup analyses of the primary outcome: alive and independent (OHS 0-2) at 6 months; OR adjusted for age, NIHSS and time [see annex 3 for publication with all subgroups]

The imbalance in number of deaths observed at 7 days was not found at the 6 month time point, with 408 (27%) rt-PA treated patients compared with 407 (27%) control patients having died by 6 months.

Outcomes at 18 months

A total of 2348 patients (77% of the study population) were eligible for follow-up at 18 months of which 1169 were in the rt-PA group and 1179 in the control group.

The primary outcome measure at 6 months was subjects alive and independent, OHS score 0-2. These criteria were met at 18 months for 391 (35.0%) of patients in the rt-PA group, compared with 352 (31.4%) in the control group; adjusted OR 1.28, 95% CI [1.03-1.57], p=0.024. OR adjusted for age, NIHSS, time, and presence/absence of visible acute ischaemic change on baseline scan.

Survival was similar between the two groups, with 408/1169 (34.9%) patients in the rt-PA group vs. 414/1179 (35.1%) patients in the control group having died (see statistical assessment for Kaplan-Meier survival curves).

At the 18 month follow up, the EuroQoL (EQ) instrument was also used to measure quality of life. The EQ utility index was calculated for 1341 (91.3%) of the 1468 patients who were alive at 18 months. Whilst statistically significant improvements were reported for most of the EQ measures in the rt-PA group compared with control, these data have not been represented correctly – see Statistical assessment.

4.1.4. Statistical Assessor's assessment of Study IST-3

4.1.4.1. Statistical assessment of efficacy

This was a pragmatic, multi-centre, randomised, controlled, open-label trial where patients with acute ischaemic stroke were allocated to 0.9 mg/kg intravenous recombinant tissue plasminogen activator (rt-PA) plus standard care or to standard

care alone. The initial pilot phase (276 patients) was randomised and placebo controlled.

Patients were eligible according to the following criteria: they had symptoms and signs of clinically definite acute stroke; the time of stroke onset was known; treatment could be started within 6 h of onset; and CT or MRI had reliably excluded both intracranial haemorrhage and structural brain lesions, which could mimic stroke (e.g. cerebral tumour).

Treatments

Patients allocated to the control group were to avoid treatment with rt-PA but otherwise received stroke care in the same clinical environment as those allocated to the rt-PA group.

Assessor's comment: The initial phase was double-blind and placebo-controlled, so it is certain that the background standard of care was not influenced by knowledge of the treatment allocation. Once the trial became open-label it is possible that background care could have been influenced by knowledge of treatment. One specific difference is that in the double-blind phase both groups were to avoid antiplatelet or anticoagulant therapy for 24 h. In the open phase, patients allocated to the control group were to start aspirin immediately. Therefore in the open-phase patients in the control group received more appropriate treatment that in the double-blind phase. Hence the title being labelled "pragmatic" as it may in some way more closely reflect the real difference that rt-PA may make when added in as an additional treatment option. Provided there is good conduct from all those responsible for administering treatment this would not bias the results in favour of rt-PA.

Timing of treatment

Part of the inclusion criteria was that treatment should be administered within 6 hours of onset of stroke symptoms.

Assessor's comment: Note than when the indication in stroke was first approved in 2002 it was noted in the SPC that treatment must be started within 3 hours of the onset of the stroke symptoms. The license was amended in 2012, expanding the time window to 4.5 hours. This trial was initiated before any of these licensing decisions and employed a 6 hour time window. When considering whether the results of this trial support the licensed indication it will be useful, when possible, to look at the sub-group results for those treated <4.5 hours after treatment, as poor results for those treated later than 4.5 hours would not be a regulatory concern given the current licensed indication. However treatment comparisons in subgroups will be underpowered.

Sample size and powering

When the study was initially powered in 2000 it was estimated that to detect a 10% difference between the treatment arms for the primary endpoint and to have sufficient power for reliable analyses in the planned sub-group analyses, a sample of 6000 patients would be needed. Given the powering for sub-groups there would be 80% power to detect a 3% difference for the primary endpoint in the overall population at the 5% level.

By 2007 it was decided that a sample size of 6000 was no longer feasible, and the number to be recruited was reduced to 3100. Based upon the pooled event rates currently observed this gave 80% power to detect an absolute difference of 4.7% at the 5% level.

Treatment allocation

Patients were entered into the trial using a central randomisation system. Allocation to rt-PA or control was decided by a minimisation algorithm.

The study centres were classified into eight world regions (North-west Europe, Scandinavia, Southern Europe, Eastern Europe, Australasia, Americas, Asia and Rest of world (in fact only the first five of these regions had significant recruitment). Within each region, the algorithm balanced the number of patients in each arm of the trial according to the following variables: age (>= or < 70); sex; NIH stroke score (0-5, 6-10, 11-15, 16-20 or >20); time from onset to randomisation (<= or > 3 hours); use of antiplatelet agents within 48 hours pre-randomisation (Yes, No or Unknown); and stroke subtype (LACI or other). There was an 80% chance of the allocation being to the treatment group that achieved the greatest balance.

Primary endpoint

The primary endpoint as specified in the protocol and the statistical analysis plan (SAP) was the proportion of patients alive and independent after 6 months. This was measured by the Oxford Handicap Score (OHS). This is a 6 point scale (running from 0-5) with classifications 0 (no symptoms at all), 1 (symptoms, but these do not interfere with everyday life), 2 (symptoms that have caused some changes in lifestyle but patients are still able to look after themselves), 3 (symptoms that have significantly changed lifestyle and patients need some help looking after themselves), 4 (severe symptoms requiring help from other people but not so bad as to need attention day and night) and 5 (severe handicap needing constant attention day and night).

Patients with a score of 0, 1 or 2 were classified as independent and therefore as a success for the primary endpoint.

Assessor's comment: The primary endpoint is a subjective scale, which could be an issue given the open-label design of the trial. The 6 month data was collected via a postal questionnaire which was assessed by staff blinded to treatment allocation (in Italy and Austria the questionnaire was conducted by phone and in Portugal in-clinic) however patients were not blind to the allocation, so there is potential for bias based upon their knowledge of treatment allocation, though the demarcation between categories 2 and 3 seems reasonably robust.

Primary analysis

The primary endpoint was to be analysed by logistic regression adjusting for age, initial stroke severity as measured by NIH stroke score, time from stroke onset to randomisation, and presence or absence of visible acute ischaemic change on baseline scan as judged by the expert reader.

In the statistical analysis plan (SAP) it was noted that these covariates were chosen because "The analysis of the baseline characteristics of the patients in the trial showed clear trends in key prognostic factors (age, stroke severity, degree of ischaemic change on baseline CT/MR) among patients randomised at different times after stroke onset that might complicate the estimation of the effect of treatment overall and in subgroups."

This refers to the fact that the patients recruited with the shortest time from onset of symptoms were (i) more likely to have a more severe neurological deficit than those recruited at later time-points; (ii) older; (iii) less likely to have a definitely visible ischaemic lesion. All these associations were statistically significant. Therefore an unadjusted analysis comparing across subgroups, would be misleading because

comparing the effect across, for example, time of recruitment would be confounded by differences in all the other factors.

An unadjusted analysis was also performed. This simply compared the percentages using the normal approximation – a standard analysis for comparing two binomial proportions.

A secondary analysis of OHS was planned using ordinal logistic regression, with the OHS being compared between groups as 5 categories; 0, 1, 2 and 3 being kept separate and 4, 5 and death combined into one group.

The same covariates used in the primary analysis were specified as the key sub-group analyses.

For patients with missing data at month 6 the OHS from day 7 was carried forward and used in the analysis. A sensitivity analysis was conducted including only patients with known OHS score.

Assessor's comment: The adjusted analysis is considered a reasonable supportive analysis, for the reasons quoted in the SAP. However it is a questionable choice as the primary analysis because the covariates seem to have been selected based upon the study data (albeit baseline data only and blinded to treatment allocation) rather than before the data were seen.

This lack of separation between analysis plan and data is a symptom of the fact that the plan was written in January 2012, practically simultaneously with the analysis being conducted – the results were published in May 2012.

The CHMP guideline on adjustment for baseline covariates states that stratification factors should usually be used as covariates in the statistical analysis. It would therefore have been easy to justify the choice if all the randomisation factors had been included (factors that were anticipated as being of importance) – the randomisation factors not included were region, use of anti-platelet agents and stroke subtype. An additional factor used in the analysis that was not used in the randomisation was degree of ischaemic change on baseline CT/MR.

Given the data-driven choice of the covariates it will be important that the results are robust to the choice of covariates. In particular the unadjusted analysis for the primary endpoint should be looked at, as this is still valid despite the relationships between covariates noted by the company. The adjusted model will be necessary if comparing across sub-groups.

Patient accountability

	rt-PA	Control
Randomised	1515	1520
Received randomised treatment	1488 (98%)	1508 (99%)
Treated in opposite arm*	26 (2%)	7 (<1%)
Unknown treatment status	1 (<1%)	5 (<1%)
Assessed at day 7	1515 (100%)	1520 (100%)
Provided OHS data or dead at 6 months	1473 (97%)	1466 (96%)
Follow-up planned at 18 months**	1169	1179
Provided OHS data or dead at 18 months	1117 (96%)	1122 (95%)

* rt-PA patients who did not receive any rt-PA or control patients who received some rt-PA.

**In three countries (Australia, Norway, and Sweden), all recruited patients were to be followed up to 18 months. In seven countries (Austria, Belgium, Canada, Italy, Mexico, Poland, and UK) follow-up had to cease on Jan 30, 2012; therefore, any patients from these countries who were recruited after June 30, 2010 were not included in the 18 month analysis. Two countries (Portugal and Switzerland) followed up patients to 6 months only therefore patients from those countries were not included in the 18 month analysis.

Assessor's comment: The follow-up was good with a very high proportion of patients being assessed for the primary endpoint at month 6 and month 18, with no imbalance between the groups in terms of missing data. Therefore results should be fairly robust to the handling of missing data.

Analysis populations

All analyses were done using the ITT population. This included all patients in the group they were randomised to, no matter what treatment they received and regardless of any protocol violations.

Baseline characteristics

	rt-PA	Control
	(N=1515)	(N=1520)
Region		
Northwest Europe (UK, Austria, Belgium, Sweden)	792 (52%)	797 (52%)
Scandinavia (Norway, Sweden)	251 (17%)	250 (16%)
Australasia	89 (6%)	90 (6%)
Southern Europe (Italy, Portugal)	204 (13%)	204 (13%)
Eastern Europe (Poland)	174 (11%)	173 (11%)
Americas (Canada, Mexico)	5 (<1%)	6 (<1%)
Age		
18-50	59 (4%)	68 (4%)
51-60	98 (6%)	104 (7%)
61-70	188 (12%)	177 (12%)
71-80	353 (23%)	371 (24%)
81-90	706 (47%)	701 (46%)
>90	111 (7%)	99 (7%)
Sex		
Female	782 (52%)	788 (52%)
NIHSS		
0-5	304 (20%)	308 (20%)
6-10	422 (28%)	430 (28%)
11-15	306 (20%)	295 (19%)
16-20	270 (18%)	273 (18%)
>20	213 (14%)	214 (14%)
Time to randomisation		
0-3 hours	431 (28%)	418 (28%)
3-4.5 hours	577 (38%)	600 (39%)
4.5-6 hours	507 (33%)	500 (33%)
>6 hours	0	2 (<1%)
Treatment with antiplatelet drugs in previous 48 hours	775 (51%)	787 (52%)

Stroke clinical syndrome		
TACI	639 (42%)	666 (44%)
PACI	596 (39%)	551 (36%)
LACI	168 (11%)	164 (11%)
POCI	110 (7%)	136 (9%)
Other	2 (<1%)	3 (<1%)
Expert reader's assessment of acute ischaemic change		
Scan completely normal	140 (9%)	129 (8%)
Scan not normal but no sign of ischaemic change	743 (49%)	781 (51%)
Signs of acute ischaemic change	624 (41%)	600 (40%)
Missing	8 (<1%)	10 (<1%)

Assessor's comment: The groups were balanced for all the factors included in the treatment allocation algorithm, so the procedure performed well.

4.1.4.2. Statistical assessment of Results

Primary endpoint

OHS at 6 months*

	rt-PA	Control
	(N=1515)	(N=1520)
0	138 (9%)	116 (8%)
1	225 (15%)	204 (13%)
2	191 (13%)	214 (14%)
3	235 (16%)	193 (13%)
4	115 (8%)	140 (9%)
5	203 (13%)	246 (16%)
Died before 6 months	408 (27%)	407 (27%)
Alive and independent (0+1+2)	554 (37%)	534 (35%)

* Values imputed for 42 (2.7%), 55 (3.6%) patients on rt-PA and control respectively by carrying forward from day 7

Analysis of alive and independent at 6 months

Unadjusted analysis		Adjusted analysis	
Difference (95% CI)	p-value	Odds ratio (95% CI)	p-value
1.4% (-2.0, 4.8%)	p=0.409	1.13 (0.95, 1.35)	p=0.181

Unadjusted odds ratio: 1.06 (0.92, 1.24)

Ordinal analysis of 5 category OHS, 0, 1, 2, 3, 4+5+death

Adjusted analysis	
Odds ratio (95% CI)	p-value
1.27 (1.10, 1.47)	p=0.001

Ordinal analysis of 7 category OHS (not a pre-planned analysis)

Adjusted analysis	
Odds ratio (95% CI)	p-value
1.17 (1.03, 1.33)	p=0.016

Assessor's comment: The trial failed on its primary endpoint, the proportion of patients alive and independent at 6 months. The difference between treatment groups was only 1.4% and this was not statistically significant in either the adjusted

or unadjusted analyses, and is clearly much smaller than the 10% difference hoped for when the trial was planned.

After a failed primary analysis, the results of secondary endpoints should be treated with caution, but there does appear to be some evidence that treatment with rt-PA was generally associated with an improvement in OHS scores, but the clinical relevance of any improvement would have to be questioned, as evidenced by the small difference in the primary endpoint (assessing the highly clinically important factor of being independent) and the similarly small differences seen in all the individual OHS categories. As noted above, there are no differences even approaching the 10% hoped for when the trial was being planned.

OHS at 18 months

	rt-PA	Control
	(N=1169)	(N=1179)
Number analysed*	1117 (96%)	1122 (95%)
0	119 (11%)	83 (7%)
1	135 (12%)	141 (13%)
2	137 (12%)	128 (11%)
3	132 (12%)	138 (12%)
4	81 (7%)	107 (10%)
5	105 (9%)	111 (10%)
Died before 18 months	408 (37%)	414 (37%)
Alive and independent (0+1+2)	391 (35%)	352 (31%)

* Missing data was ignored and only patients with OHS status at month 18 were included.

OHS at 6 months in 18 month cohort

	rt-PA	Control
	(N=1169)	(N=1179)
Number analysed*	1140 (98%)	1138 (97%)
0	115 (10%)	89 (8%)
1	170 (15%)	155 (14%)
2	157 (14%)	173 (15%)
3	184 (16%)	154 (14%)
4	82 (7%)	105 (9%)
5	152 (13%)	193 (17%)
Died before 6 months	309 (27%)	310 (27%)
Alive and independent (0+1+2)	437 (38%)	409 (36%)

* Missing data was ignored and only patients with OHS status at month 6 were included.

Analysis of alive and independent at 18 months

Unadjusted analysis		Adjusted analysis	
Difference (95% CI)	p-value	Odds ratio (95% CI)	p-value
3.6% (-0.3, 7.5%)	p=0.068	1.28 (1.03, 1.57)	p=0.024

Unadjusted odds ratio: 1.18 (0.99, 1.40)

Analysis of alive and independent at 6 months in 18 months cohort

Unadjusted analysis		Adjusted analysis	
Difference (95% CI)	p-value	Odds ratio (95% CI)	p-value
2.4% (-1.6, 6.4%)	p=0.237	1.18 (0.97, 1.45)	p=0.101

Unadjusted odds ratio: 1.11 (0.93, 1.31)

Assessor's comment: We should note that missing data has not been accounted for in these analyses and all patients with missing data were just excluded. However the amount of missing data was small and was balanced across groups. The assessor performed an analysis using missing=failure and this did not alter the results. Any conclusions seem robust to the handling of missing data.

Ordinal analysis of 5 category OHS, 0, 1, 2, 3, 4+5+death at 18 months

Adjusted analysis	
Odds ratio (95% CI)	p-value
1.30 (1.10, 1.55)	p=0.002

Ordinal analysis of 5 category OHS, 0, 1, 2, 3, 4+5+death at 6 months in 18 month cohort

Adjusted analysis	
Odds ratio (95% CI)	p-value
1.35 (1.15, 1.59)	p=0.0004

Assessor's comment: As noted previously, we should be cautious when interpreting secondary analyses after the primary analysis of a trial has failed. The results at 18 months are similar to those seen after 6 months (particularly if we look at the 6 month results from the same cohort rather than referring back to the full 6 month analysis where the treatment effects were slightly smaller).

Subgroup analysis by time from stroke symptoms to randomisation

The indication for rt-PA is currently only granted for use if treatment can be given within 4.5 hours of symptoms emerging. The trial population included patients receiving treatment for up to 6 hours, so sub-group analyses are necessary if we wish to evaluate the support (or otherwise) given by this trial for the licensed indication.

Analysis of alive and independent at 6 months by time to randomisation

Time to randomisation	rt-PA	Control	Adjusted odds ratio (99% CI)
0-3 hours	132/431 (30.6%)	95/418 (22.7%)	1.64 (1.03,2.62)
3-4.5 hours	182/577 (31.5%)	226/600 (37.7%)	0.73 (0.50,1.07)
>4.5 hours	240/507 (47.3%)	213/500 (42.6%)	1.31 (0.89,1.93)
0-4.5 hours	314/1008 (31.2%)	321/1018 (31.5%)	

Analysis of alive and independent at 18 months by time to randomisation

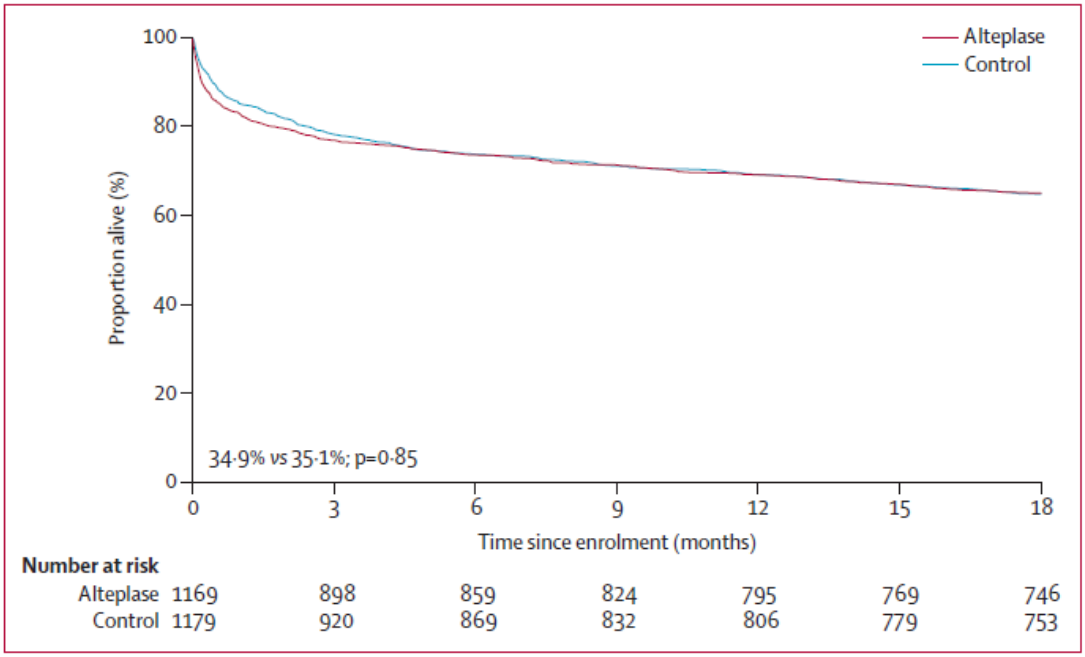
Time to randomisation	rt-PA	Control	Adjusted odds ratio (99% CI)
0-3 hours	92/308 (29.9%)	63/300 (21.0%)	1.54 (0.96,2.47)
3-4.5 hours	144/450 (32.0%)	145/449 (32.3%)	1.16 (0.81,1.66)
>4.5 hours	155/359 (43.2%)	144/371 (38.8%)	1.32 (0.91,1.92)
0-4.5 hours	236/758 (31.1%)	208/749 (27.8%)	

Assessor's comment: It appears that the overall prognosis in both groups is better when treatment is administered later, given the higher response rates in both groups compared with earlier treatment. However this is likely to be due to the confounding across covariates previously mentioned – namely that the patients recruited earliest were (i) more likely to have a more severe neurological deficit than those recruited at later time-points; (ii) older; (iii) less likely to have a definitely visible ischaemic lesion.

Looking at the difference between treatment groups there does seem to be evidence of a benefit in terms of being alive and independent when treatment is given within 3 hours. There is no evidence of a benefit for treating between 3 and 4.5 hours. However it should be noted there was no clear pattern for decreased efficacy across time.

Mortality

Kaplan-Meier curve from 18 month cohort



Proportion of patients who died before:

	rt-PA	Control	Difference	p-value
7 days	163/1515 (10.8%)	107/1520 (7.0%)	3.7%	p=0.0004*
6 months	408/1515 (26.9%)	407/1520 (26.8%)	0.2%	p=0.924
18 months	408/1169 (34.9%)	414/1179 (35.1%)		p=0.85**

* from unadjusted test of difference in percentages
** from log-rank test

Assessor's comment: A difference of about 4% in early deaths was anticipated by the authors based on the product information for rt-PA, and that is what occurred. By around the 4 month point the proportion of patients dead has evened out between the treatment groups and it remains even for the rest of the study. When the 6 month data were published, the Trialists hoped that the better disability status at 6 months (by their conclusion) would mean that mortality by 18 months would show an advantage for rt-PA, but this did not occur and there is no suggestion of a long-term mortality advantage.

EuroQol at 18 months (EQ-5D)

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.

Assessor's comment: 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.

Assessor's comment: The improvements in QoI as presented in the paper are not robust to the handling of patients who died in the analysis.

4.1.4.3. Statistical assessor's conclusions on IST-3

Based upon the pre-specified primary analysis, benefit of rt-PA when given within 6 hours of the occurrence of stroke symptoms has not been demonstrated. Therefore the trial is negative from a statistical perspective.

The primary analysis at 6 months did not show a difference between rt-PA and control in the proportion of patients alive and independent. The clearest finding was the disadvantage in early mortality, with a clear difference in mortality rates at 7 days. The difference disappeared by 6 months, but some other advantage of treatment

would be expected to compensate for this early mortality disadvantage, and this was not seen from the primary endpoint.

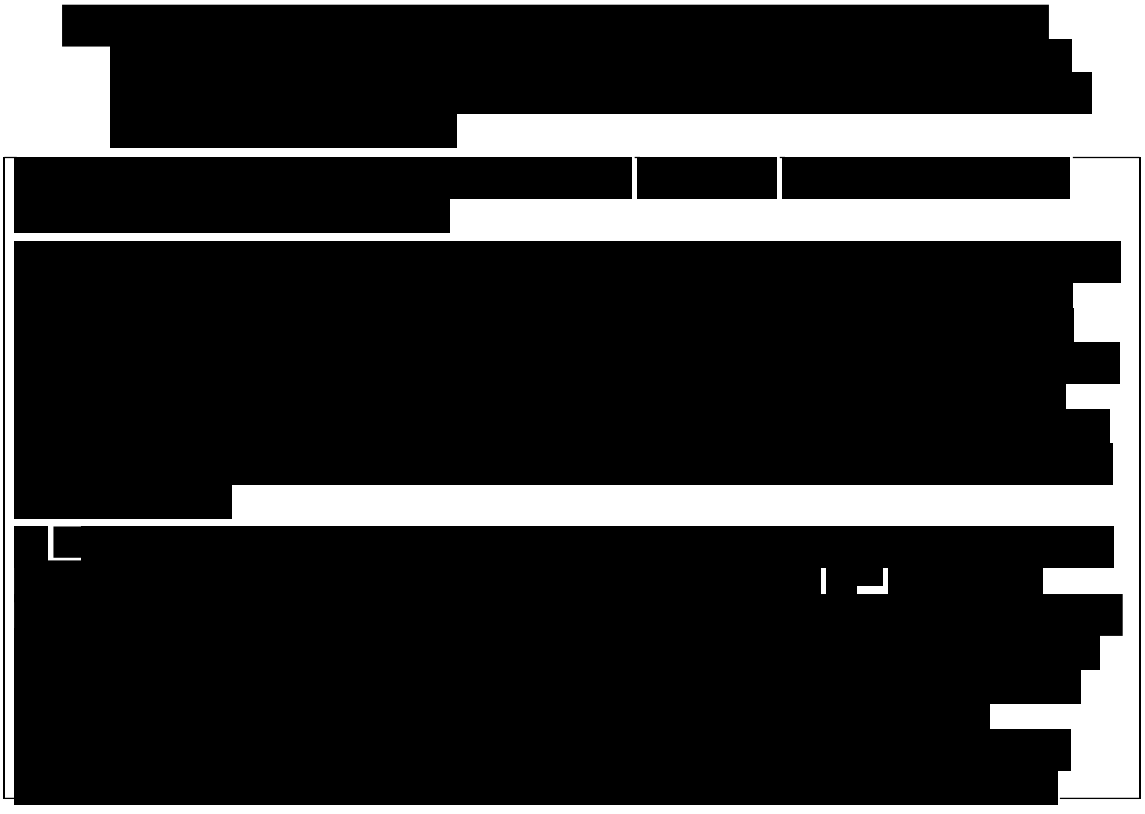
It is difficult to interpret secondary endpoints in the light of a failed primary analysis, but there was some suggestion of a shift in the OHS in favour of rt-PA but even if the finding were true the clinical relevance of the size of the shift would be questionable. The advantages claimed in QoI are not supported as the analysis excluded patients who died.

If evaluated in the context of the approved licence it might be possible to use the data from this trial to support the initial restriction to treat patients only within 3 hours of stroke symptoms emerging. In this sub-group there was an 8% difference in the proportion of patients alive and independent at 6 months which was maintained at 18 months. This could be weighed up in a risk-benefit discussion against the early mortality disadvantage. It would be difficult to make a case for the extension of the time-window to 4.5 hours based on these data, as no benefit was seen once the 3-4.5 hour group were included.

4.1.5 Specific concerns raised by Dr [redacted] regarding IST-3

Dr [redacted] has outlined his specific concerns relating to rt-PA in a submission to [redacted]

The concerns that relate to the IST-3 trial are as follows:



[redacted]

2. “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’. Those marketing alteplase had observer status on the International Advisory Board. The start-up phase had disappointing results and Boehringer Ingelheim discontinued support for IST-3.”

Assessor’s comments: 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 [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 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.

4. "Once an open design was in place, managing expectations seemed the priority. Earlier on, leading trialists had highlighted that 'large trials are *essential'* to examine the benefits and harms of thrombolysis for acute ischaemic stroke and assess whether or not it is cost effective."…… "The aim of resolving a major debate over the effectiveness of alteplase in stroke was progressively undermined in a 'thread' of publications. The approach increasingly downplayed the importance of the results to the overall debate and changed the focus towards those older than 80 years and those presenting after 4.5 h."

Assessor's comments: 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.

[redacted]

6. "The revised IST-3 analyses highlighted complex statistical secondary outputs and marginalised key straightforward outcomes. For example the far from reassuring Kaplan-Meier survival curve was promised as a primary outcome to be depicted in 'Figure 5' of the main publication. It emerged in unconventional form in the much less accessible online appendix as a 'web figure'"

Assessor's comments: This is a criticism of the presentation of the trial results, rather than the conduct of the trial itself. It is agreed that certain aspects of the publication, in particular the 'Interpretation' in the abstract, arguably give an overly positive slant to the results, however this has no impact on the actual results obtained.

7. "Full trial data for place of residence at six months, surely of interest in an open trial, is still missing (but showed no benefit in the subsequent publication examining the subset followed for 18 months)."

Assessor's comments: [redacted]

	rt-PA	**Control**
Own home	805 (53.1)	799 (52.6)

Relative's home	74 (4.9)	79 (5.2)
Residential home	64 (4.2)	53 (3.5)
Nursing home	115 (7.6)	119 (7.8)
Still in hospital	6 (0.4)	9 (0.6)
Died	408 (26.9)	407 (26.8)
Question not answered	1 (0.1)	nil (0.0)
Form not returned	42 (2.8)	54 (3.6)
All	1515 (100)	1520 (100)

Place of residence at 6 month completion of the IST-3 trial (%).

Similar to the data provided in the supplementary appendix to the paper presenting the 18 month follow-up data, these 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.

8. "Additionally, the likely impact of recall bias, expected from open trials, was not fully explored in the main publication. When the webappendix of the trial subset followed to 18 months is eventually located, it presents data which suggest that the expected recall bias was both present and substantial."

Assessor's comments: 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.

4.1.5.1 Assessor's overall comments on Dr [redacted] s concerns regarding the IST-3 trial

The [redacted] raises a number of issues relating to the IST-3 trial. [redacted]

It would appear that some aspects of the trial conduct were non-ideal, for example the near simultaneous publication of the statistical analysis plan and final study results. Whilst it is important to bear this in mind, it does not unduly influence the interpretation of the trial results, particularly as the trial had a negative result for the primary endpoint – see also regulatory implications below.

The withdrawal of financial support of the MAH resulted in an open-label design which is undeniably less robust than a double-blind trial. However, it could also be viewed as reducing one potential source of bias. The [redacted] explanation for why it was necessary to revert to an open-label trial design seems reasonable.

Another issue raised is the presentation of the study data in the final publication. It is agreed that the presentation of results was given a more positive slant than perhaps justified by the data, however the publication does provide the relevant data and the assessors' conclusions drawn from these data are unaffected by this issue.

4.1.6 Discussion and conclusions on IST-3 and regulatory implications

IST-3 failed in its primary outcome of an improvement in the proportion of patients alive and independent (OHS 0-2) at 6 months. There appears to be some supportive evidence of benefit in terms of overall improvement in OHS scores, however as discussed in the statistical assessment, this is difficult to interpret in light of the failed primary outcome and the limitations of the analysis discussed above.

Mortality data showed an increased death rate in the rt-PA group compared with the control group in the first 7 days after treatment. This was mainly due to an increase in intracranial haemorrhage and swelling of the initial infarct. The authors have commented that the proportion of fatal intracranial haemorrhages in the rt-PA group was as expected based on the product information and they consider that a higher rate could have been expected given the trial population included. This adverse effect on mortality is not observed at the 6 month time point, at which the proportion of patients who had died was equal in the two groups.

As well as the primary endpoint failing to provide a compensatory benefit for this early adverse effect on mortality, the small apparent improvement in OHS scores in the rt-PA group (as measured by secondary analyses at 6 months) was not found to translate into a mortality benefit at 18 months. The expectation that an improved functional outcome at 6 months should translate into a longer-term mortality benefit is reasonable (see Slot *et al*, section 4.3), however this was not observed in the IST-3

trial. This in itself casts doubt on the suggested benefit found in the secondary analyses.

With respect to time to treatment the sub-group analyses suggest some benefit in the patients treated with rt-PA within 3 hours of the onset of symptoms, whilst no benefit was found in those treated from 3-4.5 hours, and a non-significant trend for benefit was observed in those treated >4.5 hours after symptom onset. In addition, the overall proportion of patients (including those in the control group) with a good outcome was found to increase with increasing time to treatment. This latter finding is likely due to the more severe (and older) patients presenting earlier. The ORs observed (see figure 2) are adjusted for age and NIHSS at baseline. The unexpected pattern of ORs for these two groups may have been influenced by the lack of power (sample size was revised to recruit half the initially intended number of subjects).

Finally, the results of the quality of life assessments at 18 months appear more encouraging on face-value, however as discussed in the statistical assessment these analyses did not include those patients who had died prior to 18 months. Inclusion of this group of patients results in a loss of the beneficial effect.

Regulatory implications

The results of the IST-3 trial were overall negative for the primary endpoint. However from a regulatory perspective, the following should be considered:

- A mortality benefit for rt-PA has never been claimed, and therefore the lack of improvement in mortality cannot be considered to impact negatively upon the balance of benefits and risks for the licenced use of rt-PA. However an improvement in mortality at the longer time points would have been re-assuring.
- The IST-3 trial evaluated treatment with rt-PA up to 6 hours after the onset of symptoms, whilst the licence specifies treatment up to 4.5 hours. Therefore the failure of the primary outcome per se cannot be considered to be truly reflective of a negative balance of benefits and risks for use under the terms of the licence. The sub-group analyses provided some support for treatment from 0-3 hours, but were not supportive of treatment between 3-4.5 hours. However, the failure of the study to meet its original recruitment target meant that it was underpowered for the primary outcome – and therefore also for the sub-group analyses.
- The IST-3 trial purposefully enrolled patients for whom rt-PA was not specifically indicated, and who did not meet the prevailing licence criteria (95% of enrolled patients). Over half of patients were over the age of 80 years, and many patients had high blood pressure, high blood glucose or severe symptoms on presentation. Patients aged over 80 years are contraindicated in the current EU product licence, as are patients with very severe strokes, very high blood pressure/high blood glucose. A positive result in such a trial population might be considered to provide supporting evidence that the balance of benefits and risks in the licenced population is positive.
- Interestingly, and perhaps unexpectedly, the sub-group analyses suggested that benefit from rt-PA treatment was greater in patients aged over 80 compared with younger patients, and greater with increasing severity of stroke (although confidence intervals were overlapping).

Whilst the data generated by the IST-3 study (in particular the 3-4.5 hour time window) is clearly of interest, the regulatory implications of a negative trial for the

licensed population are unclear, when the trial population itself is outside that of the licence and might reasonably be considered to include higher risk patients.

4.2 Re-Analysis of NINDS study[14] (see Annex 4 for committee report)

4.2.1. Background and re-analysis description

An rt-PA review committee was established in May 2002 at the request of NINDS, to address concerns raised by doubts expressed in the medical literature regarding the result of the NINDS studies. The committee was requested 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'.

The committee declined to explore a secondary issue of whether pharmaceutical company participation had biased the results of the trial, because it considered 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.

NINDS appointed a chair[3], and invited him to appoint the rest of the committee. The final committee consisted of three statisticians and three clinicians. The committee had full access to the NINDS dataset via an independent contractor.

Four clinical outcome measures were used, three are measures of functional status (BI, mRS, Glasgow outcome scale) and one is a measure of neurologic deficit (NIHSS). These represent all the measures used in the original trials (NINDS part 1 and 2). The committee evaluated placebo vs. rt-PA for each measure individually as well as a global analysis of favourable response which included all four measures.

4.2.2 Re-analysis findings

Overall, the committee's findings were that despite an increased incidence of symptomatic intracerebral haemorrhage in rt-PA treated patients, when rt-PA was administered according to the study protocol there was a statistically significant and clinically important benefit of treatment compared with placebo, measured by an adjusted odds ratio of 2.1, 95% CI [1.5-2.9] for a favourable outcome (using the global outcome measure) at 3 months. The analysis was adjusted for centre, time to treatment, study part, age, baseline NIHSS, diabetes and pre-existing disability.

The NINDS study protocol used a dose of 0.9 mg/kg body weight (maximum 90mg) with 10% of the dose given as a bolus followed by the rest of the dose in a 60 minute infusion. No anticoagulants or antiplatelets were to be given for 24 hours after treatment and blood pressure was to be maintained within pre-specified values. The current EU SmPC uses the same dosing regimen, and also advises that heparin and aspirin should not be administered in the 24 hours following treatment. There is also a contraindication in place in patients with very high blood pressure.

No evidence was found for any of the adjusting variables modifying the rt-PA treatment effect. However, much of the controversy that prompted the re-analysis of NINDS related to the imbalance in baseline NIHSS scores.

The committee's evaluation of the imbalance of NIHSS at baseline (there were more rt-PA patients in category 0-5 than placebo) found no evidence that this had a

[3] The Chair had not participated in any of the studies or trials leading up to the NINDS-supported investigations regarding the use of t-PA as a therapy for acute ischemic stroke

statistically or clinically significant effect on the study results. The original models using both age and baseline NIHSS as continuous variables were considered to properly adjust for the role of these variables. Whilst both older age and higher NIHSS score are strongly negatively related to a favourable outcome, and there is a strong interaction between age and baseline NIHSS, there was no evidence of any age by baseline NIHSS subgroup responding differently to rt-PA treatment than the study group as a whole (see statistical assessment in the next section).

The committee also reviewed and concluded on a number of other specific issues in their evaluation:

Blood pressure, assessment and management: Non-compliance with the study protocol and issues with management and recording of blood-pressure were found. As a result it is not possible to assess the effect of hypertension management on clinical outcome in the NINDS study. The committee concluded that blood pressure variables should not be included in the statistical models (although their inclusion was not found to alter the rt-PA treatment effect).

Intracerebral haemorrhage: In NINDS, the overall risk of symptomatic intracerebral haemorrhage was 6.5% in rt-PA treated patients, vs 0.6% in patients receiving placebo. Symptomatic intracerebral haemorrhage in rt-PA patients had significant consequences (e.g. for the BI, favourable outcome was 10% vs. 55% for patients without intracerebral haemorrhage) and the three-month mortality was very high (75%). The trial was not powered to identify risk factors related to intracerebral haemorrhage or decreased likelihood of favourable outcome.

Onset to treatment time: The data from NINDS fail to support a conclusion that effect of rt-PA decreases with increasing time to treatment within the 3 hour time limit. However this does not meant that such a relationship does not exist.

Clinical centres: No significant difference in baseline characteristics between centres. Likelihood of a favourable outcome differed between centres, but between-centre variation of treatment effect of rt-PA was not statistically significant.

Stroke subtype: Data from the trial do not support any difference in rt-PA effect between stroke subtypes.

Pre-existing disability: Overall, patients with pre-existing disability had a significantly reduced chance of favourable outcome, but there was no evidence of a different response to rt-PA treatment compared with those without pre-existing disability.

Diabetes mellitus: The data indicated no benefit for patients with diabetes. However, there was also no statistical evidence of an rt-PA-diabetes interaction, no statistically significant evidence that diabetic and non-diabetic patients responded differently to rt-PA.

4.2.3. Statistical Assessor's assessment of the re-analysis of NINDS, also addressing concerns raised by Dr [redacted]

In response to concerns about the results of the NINDS rt-PA stroke study, an independent t-PA review committee was set up "*to address whether there is concern that eligible stroke patients may not benefit from rt-PA given according to the protocol used during the trials and whether the baseline imbalance (in baseline stroke severity) invalidates the entire trial as claimed by some of the critics.*"

The review committee concluded that "*despite an increased incidence of symptomatic intracerebral hemorrhage in t-PA treated patients and subgroup imbalances in baseline stroke severity, when t-PA was administered to acute ischemic stroke patients according to the study protocol, there was a statistically*

significant, and clinically important, benefit of t-PA treatment resulting in a higher likelihood of having a favourable clinical outcome at three months."

Therefore the review committee conclusions were in line with the original conclusions drawn from the study.

Despite the conclusion of the review committee being in line with the conclusions originally reached, based on the publication of the review online, Dr [redacted] considers that it identified issues that add to significant concerns over the quality of the data.

These concerns relate to:

- the difference in outcomes across centres
- concerns related to an imbalance in the time from onset of symptoms to randomisation
- the 'substantial' blood pressure protocol violations highlighted in the review which add to questions about study conduct.

Brief summary of the NINDS study

This was a randomised, placebo controlled study. The randomisation was stratified and balanced at each centre according to whether the patient was randomised within the first 90 minutes or in the 91-180 minute interval after stroke onset. Patients whose time since onset had exceeded 180 minutes were ineligible.

The NINDS investigators used four outcome measures. The Barthel index, modified Rankin scale, and Glasgow outcome scale are accepted as measures of functional status. The NIH Stroke Scale (NIHSS) is accepted as a measure of neurologic deficit. The primary response variable for each measure was a dichotomous indication of whether the outcome (at 90 days) was "favourable" or "not favourable." The definitions of "favourable" were: Barthel; 95 or 100, Rankin; 0 or 1, Glasgow; 1, and NIHSS; 0 or 1. For each measure death was treated as an unfavourable outcome.

The original study investigators used the principle of "intent-to-treat" in analysing all patients randomised in the study. Thus, they attempted formal follow up at 24 hours, 90 days and one year on all randomized patients. Patients "lost" in the sense that they were known to be alive but did not provide data permitting the determination of favourable/unfavourable status were assigned the least favourable known level for each index. With two exceptions, the review committee used the same approach. The two exceptions involved individuals mistakenly randomized into the study at a point more than 180 minutes after onset. Since their remit was to determine whether there were groups of patients who should not be treated with t-PA according to the study protocol, these two patients were excluded from the analyses.

There were two studies under the NINDS umbrella. Part 1 was a Phase 2 study to initially determine whether t-PA had activity with outcomes 24 hours after treatment as primary. Part 2 had the same design as part 1, except the 3 month time-point was primary. The review committee treated the two parts as a single study. This approach seems reasonable given the positive results in each part individually.

Assessment Instrument	Percentage with Favorable Outcome*		Odds Ratio (95% CI)†	Relative Risk (95% CI)†	P Value
	t-PA	Placebo			
Part 2, 0–180 min‡					
No. of patients	168	165			
Global test	—	—	1.7 (1.2–2.6)	—	0.008
Barthel index	50	38	1.6 (1.1–2.5)	1.3 (1.0–1.7)	0.026
Modified Rankin scale	39	26	1.7 (1.1–2.6)	1.5 (1.1–2.0)	0.019
Glasgow outcome scale	44	32	1.6 (1.1–2.5)	1.4 (1.0–1.8)	0.025
NIHSS	31	20	1.7 (1.0–2.8)	1.5 (1.0–2.2)	0.033
Part 1, 0–180 min‡§					
No. of patients	144	147			
Global test	—	—	2.1 (1.3–3.2)	—	0.001
Barthel index	54	39	1.8 (1.1–2.8)	1.4 (1.1–1.8)	0.012
Modified Rankin scale	47	27	2.3 (1.4–3.6)	1.7 (1.3–2.3)	<0.001
Glasgow outcome scale	47	31	2.0 (1.2–3.1)	1.5 (1.1–2.0)	0.005
NIHSS	38	21	2.2 (1.3–3.7)	1.8 (1.2–2.6)	0.002

The committee were given the full original data-set to work with. They initially ensured they could reproduce the tables from the original NINDS analyses, and then conducted further investigations to investigate the concerns relating to the study.

They primarily compared favourable outcome percentages between groups using odds ratios, choosing this statistic but this review generally presents differences in percentages to aid clinical interpretation.

Principal remit for the committee

1. Baseline NIHSS imbalance

As shown below the overall results for the trial were positive for all four endpoints. However there was some concern that this may be driven by an imbalance in baseline stroke severity between the treatment groups.

Overall results

Scale	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
Barthel Index	162/310 (52%)	119/312 (38%)	14.1%	(6.4, 21.9)
Modified Rankin Score	133/310 (43%)	83/312 (27%)	16.3%	(8.9. 23.7)
Glasgow outcome scale	141/310 (45%)	97/312 (31%)	14.4%	(6.8. 22.0)
NIHSS	106/310 (34%)	64/312 (21%)	13.7%	(6.7. 20.6)

The committee re-analysis confirmed the existence of an imbalance, with more t-PA patients with NIHSS scores 0-5, the patients with a better prognosis.

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			
	t-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			
	t-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			
	t-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			
	t-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 t-PA for all 4 endpoints in all of the sub-groups and statistical significance in favour of t-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 t-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.*"

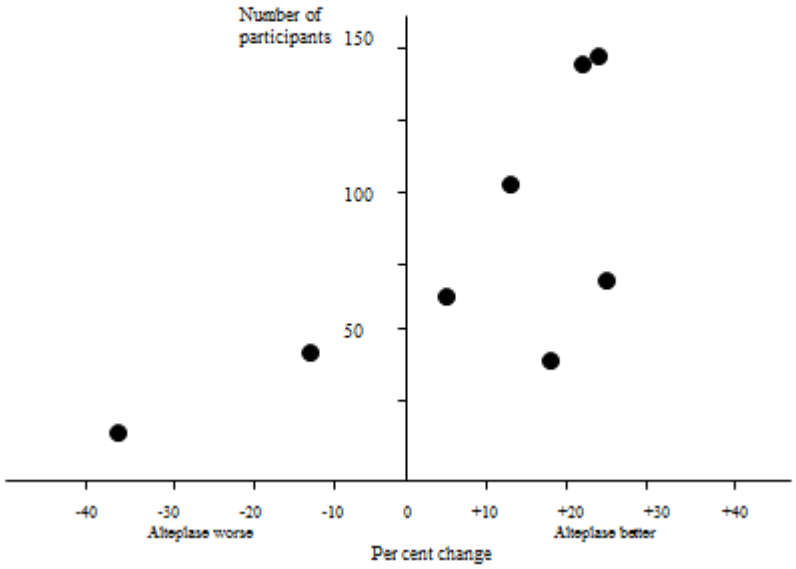
4.2.3.1 Concerns identified by Dr [redacted] based upon the review of the NINDS data

The committee produced many additional analyses as part their investigation into whether there was a subgroup of patients who received no benefit. Some of these have led Dr [redacted] to the conclusion that the review adds to concerns over the quality of the data.

Differences across centres

Dr [redacted] has concerns that the apparent success of rt-PA across the eight NINDS clinical trial centres 'differed considerably' and that plotting the per cent change in favourable outcome (modified Rankin score 0-1) with treatment, against the number of participants at each centre (figure 1), yields an unlikely spread of data, with smaller centres not underpinning the results of the larger centres.

Figure 1



Favourable outcomes by centre – results presented are raw percentages and associated unadjusted confidence intervals.

Barthel index by centre

Centre	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
5	45/74 (61%)	31/76 (41%)	20.0%	(4.2, 35.8)
4	43/74 (58%)	24/72 (33%)	24.8%	(9.0, 40.6)
8	25/51 (49%)	19/52 (37%)	12.5%	(-6.7, 31.7)
3	21/35 (60%)	15/36 (42%)	18.3%	(-4.9, 41.6)
2	8/29 (28%)	11/33 (33%)	-5.7%	(-29.1, 17.6)
1	11/20 (55%)	6/19 (32%)	23.4%	(-7.8, 54.6)
6&9	7/19 (37%)	10/18 (56%)	-18.7%	(-51.4, 14.0)
7	2/8 (25%)	3/6 (50%)	-25.0%	(-80.6, 30.6)

Modified Rankin score by centre

Centre	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
5	34/74 (46%)	17/76 (22%)	23.6%	(8.7, 38.4)
4	36/74 (49%)	19/72 (26%)	22.3%	(6.9, 37.7)
8	20/51 (39%)	13/52 (25%)	14.2%	(-3.8, 32.3)
3	21/35 (60%)	13/36 (36%)	23.9%	(0.9, 46.9)
2	8/29 (28%)	7/33 (21%)	6.4%	(-15.5, 28.2)
1	7/20 (35%)	3/19 (16%)	19.2%	(-8.3, 46.7)
6&9	6/19 (32%)	8/18 (44%)	-12.9%	(-45.0, 19.3)
7	1/8 (13%)	3/6 (50%)	-37.5%	(-88.8, 13.8)

Glasgow outcome scale by centre

Centre	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
5	33/74 (45%)	22/76 (29%)	15.6%	(0.3, 31.0)

4	39/74 (53%)	18/72 (25%)	27.7%	(12.4, 43.0)
8	20/51 (39%)	14/52 (27%)	12.3%	(-6.0, 30.5)
3	22/35 (63%)	17/36 (47%)	15.6%	(-7.6, 38.9)
2	9/29 (31%)	10/33 (30%)	0.7%	(-15.5, 28.2)
1	10/20 (50%)	4/19 (21%)	28.9%	(-22.8, 24.2)
6&9	7/19 (37%)	9/18 (50%)	-13.2%	(-46.0, 19.7)
7	1/8 (13%)	3/6 (50%)	-37.5%	(-88.8, 13.8)

NIHSS by centre

Centre	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
5	24/74 (32%)	16/76 (21%)	11.4%	(-2.8, 25.6)
4	26/74 (35%)	15/72 (21%)	14.3%	(-0.2, 28.8)
8	18/51 (35%)	9/52 (17%)	18.0%	(1.1, 34.9)
3	19/35 (54%)	9/36 (25%)	29.3%	(7.2, 51.4)
2	7/29 (24%)	6/33 (18%)	6.0%	(-14.9, 26.8)
1	5/20 (25%)	2/19 (11%)	14.5%	(-9.8, 38.7)
6&9	6/19 (32%)	6/18 (33%)	-1.8%	(-33.0, 29.5)
7	1/8 (13%)	1/6 (17%)	-4.2%	(-46.0, 37.6)

The review committee conclusion was as follows:

"*We found no significant difference between the centers in the baseline characteristics of the patients. The likelihood of having a favorable outcome differed considerably between the centers, those with fewer patients often having the worst outcome. However, the between-center variation in t-PA treatment effect for either the global outcome, or the individual outcome measures, was not statistically significant and did not invalidate the trial results. Nevertheless, it will be important in future studies to identify the factors that lead to good outcomes at institutions administering t-PA to treat acute ischemic stroke patients. This information will be very helpful to other institutions that are looking to develop the resources needed to administer t-PA safely to acute ischemic stroke patients.*"

Assessor's comment: 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 ranks 3rd, 2nd, 3rd and 5th of the 8 centres across the 4 endpoints, while centre 4 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.

Time from onset of symptoms

Dr [redacted] has concerns that the review noted irregularities in the recruitment process, with 50 per cent of all patients who were within the 0 - 90 min time window reportedly

treated between 89 and 90 min. In addition, the poor prognosis of placebo patients recruited between 91 and 133 min produced 'exceptionally high' odds for a favourable outcome with rt-PA within this time-frame.

Barthel index by time from onset of symptoms

Time	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
0-90	83/157 (53%)	55/145 (38%)	14.9%	(3.8, 26.1)
91-133	19/31 (61%)	13/50 (26%)	28.9%	(13.9, 56.6)
134-154	20/40 (50%)	16/39 (41%)	9.0%	(-13.2, 31.2)
155-173	18/42 (43%)	16/39 (41%)	1.8%	(-20.8, 23.7)
174-180	22/40 (55%)	19/39 (49%)	6.3%	(-16.1, 28.6)

Modified Rankin score by time from onset of symptoms

Time	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
0-90	63/157 (40%)	41/145 (28%)	11.9%	(1.2, 22.5)
91-133	16/31 (52%)	6/50 (12%)	39.6%	(19.5, 59.7)
134-154	20/40 (50%)	11/39 (28%)	21.8%	(0.5, 43.1)
155-173	16/42 (38%)	13/39 (33%)	4.8%	(-16.4, 25.9)
174-180	18/40 (45%)	12/39 (31%)	14.2%	(-7.3, 35.7)

Glasgow outcome scale by time from onset of symptoms

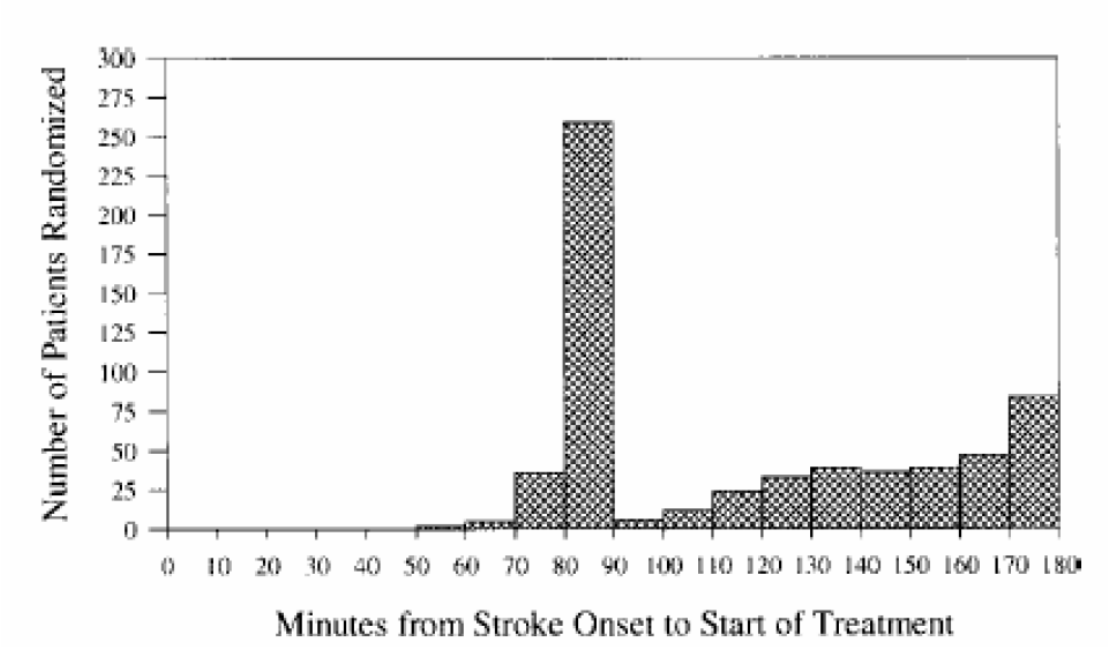
Time	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
0-90	68/157 (43%)	47/145 (32%)	10.9%	(0.0, 21.8)
91-133	17/31 (55%)	9/50 (18%)	36.8%	(16.0, 57.7)
134-154	18/40 (45%)	12/39 (31%)	14.2%	(-7.3, 35.7)
155-173	20/42 (48%)	13/39 (33%)	14.3%	(-7.2, 35.8)
174-180	18/40 (45%)	16/39 (41%)	4.0%	(-18.2, 26.1)

NIHSS by time from onset of symptoms

Time	n/N (%) Favourable outcome			
	t-PA	Placebo	Difference	95% CI
0-90	53/157 (34%)	29/145 (20%)	13.8%	(3.9, 23.7)
91-133	15/31 (48%)	7/50 (14%)	34.4%	(14.0, 54.7)
134-154	11/40 (28%)	10/39 (26%)	1.9%	(-17.9, 21.6)
155-173	13/42 (31%)	8/39 (21%)	10.4%	(-8.7, 29.6)
174-180	14/40 (35%)	10/39 (26%)	9.4%	(-11.1, 29.8)

The review committee conclusion regarding time from symptom onset to treatment was as follows:

"*Based on the substantially nonlinear nature of the distribution of time from symptom onset to treatment (OTT), and an idiosyncratic distribution of favorable response rates among the placebo patients, we conclude that the data provided by this study failed to support a conclusion that the effect of t-PA therapy diminished with increasing values of OTT within the protocol specified 3 hour time limit. However, this does not mean such a relationship does not exist, and further studies are needed to address the question of a differential t-PA treatment effect related to time from symptom onset to treatment. It is also important to recognize that the results from this study provide no data on the effectiveness of thrombolytic therapy administered to acute ischemic stroke patients more than 180 minutes after symptom onset.*"



Assessor's comment: 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 "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 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 assessment and management

The committee noted:

"Our analysis identified a number of problems regarding pre- and post-randomization blood pressure measurement and management:

- Non-compliance with the defined protocol was substantial, and persistent, throughout the study with regard to both the documentation of blood pressure readings, and adherence to the treatment regimen for hypertension.
- There was limited rigor with regard to the pharmacologic characteristics of antihypertensive regimens. In some instances pharmacologic monitoring was performed by representatives (nurses) of the sponsoring pharmaceutical firm. Medications employed were listed by date, but not by time, eliminating consequential interpretive utility.
- The exact number of patients who received medication to lower blood pressure either prior to, or after, receiving study treatment is unknown.
- The confusion regarding blood pressure documentation, and the lack of knowledge of treatment of hypertension either prior to, or after, receiving study treatment, could have led to an unknown number of patients receiving treatment in violation of the nominal study protocol.

Based on these observations, we reached the following conclusions:

- It was not possible to assess the effect of hypertension management on clinical outcome in acute ischemic stroke patients treated in the NINDS study.
- The blood pressure variables should not be included in the statistical models. However, we also found that inclusion of the blood pressure variables in the statistical models would have been inconsequential with regards to altering the t-PA treatment effect.

Finally, the inconsistent documentation of both blood pressure readings and hypertension management seriously undermines the NINDS investigators statement that blood pressure management was a significant part of the protocol that contributed to the success of the study. Nonetheless, we concur with the NINDS investigators premise that blood pressure management should be included in the protocol for treating acute ischemic stroke patients with t-PA. It is biologically plausible that hypertension management could affect clinical outcome in acute ischemic stroke patients treated with t-PA, and data from the cardiology literature has already demonstrated that in acute myocardial infarct patients, the risk of having an intracerebral hemorrhage is related to pre-treatment blood pressure. However, further clinical studies will be needed to assess whether blood pressure management is related to better clinical outcomes in acute ischemic stroke patients treated with t-PA."

Assessor's comment: 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 committee it is not considered that this calls into question the primary conclusions of the study regarding the efficacy of t-PA. It is not clear whether/how the poor monitoring and management of blood pressure extrapolates to other aspects of the trial conduct.

4.2.3.2 Statistical Assessor's Overall conclusion

The results of the committee's review were consistent with the original conclusions that a benefit had been shown for t-PA when given within 3 hours and that the baseline imbalance in stroke severity does not invalidate that. Dr [redacted] considers that some of the additional analyses performed by the committee cast further doubt on the trial data.

When assessing clinical trials, we are very concerned if an applicant is felt to be "cherry picking" the best results and ignoring poor findings, which is why protocols have to pre-specify the primary analyses and the conclusions should be primarily based upon those. In the same way when performing an assessment it would not be valid to pick out all the most negative findings and emphasise those. It might be felt that by looking at the many analyses performed by the committee and picking out those where there are imbalances, or surprising sub-group findings, we could be in danger of falling into this trap. If the data is broken down enough times into different sub-groups there will some findings that appear initially concerning or surprising, despite the pre-specified primary analyses being positive.

The sub-groups analyses by centre and time from symptoms to treatment do not suggest that the benefit of treatment is entirely driven by one centre or time from onset or an imbalance in any of these, and the committee's conclusion that the trial demonstrates a benefit of treatment on the primary scales is supported.

4.2.4. Other concerns regarding the NINDS trial raised by Dr [redacted]

Assessor's comments: [comment from GCP inspectorate] Trials conducted before May 2004 were prior to the clinical trials regulation and therefore MHRA does not have the legal right to inspect them.

[redacted]

2. "The current Cochrane review of stroke thrombolysis has, pertinently, indicated how flawed trial design might lead to bias through incomplete blinding. Firstly, reconstituted alteplase is initially frothy, unlike a clear saline

placebo. Secondly, treated participants frequently had visible external bleeding. Thirdly, 'observers' monitoring outcomes potentially could be influenced by knowledge of these events at randomisation. The NINDS review has clarified that local delivery of the trial was assisted by nurses employed by Genentech."

Assessor's comments: It should be noted that the Cochrane review comments regarding the issues with blinding of therapy in thrombolysis trials were made generally and not in specific reference to the NINDS trial.

1) The NINDS publication states that Genentech supplied both the rt-PA and the placebo, and there is no mention of saline. [redacted]

2&3) 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).

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.

3. "The Cochrane review, again, highlights 'administrative' problems with the NINDS randomisation procedures. The resulting imbalance between the two arms of the study between 91 and 180 min has not been disputed"

Assessor's comments: 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 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.

[redacted]

4.2.5. Discussion and conclusions on the NINDS re-analysis and regulatory implications

The NINDS re-analysis was commissioned to address concerns that had been raised about the results of the NINDS trials, in particular whether patients receiving rt-PA according to the trial protocol may not benefit, and whether imbalances in baseline stroke severity might invalidate the whole trial.

The committee concluded overall that there is a statistically significant and clinically important benefit of treatment compared with placebo at 3 months. The evaluation of imbalance of stroke severity at baseline found no evidence for a significant effect on the study results.

Although the conclusions of the committee were in line with the original study publication, Dr [redacted] has raised concerns relating to the difference in outcomes across centres and relating to an imbalance in the time from onset of symptoms to randomisation.

Upon review of the re-analysis report and discussed in the statistical assessor's comments above, it is concluded that none of these concerns are considered to cast doubt on the results of the NINDS trials.

[redacted]

Overall, the conclusions of the re-analysis committee are supported, that the trial demonstrates a benefit of treatment in the primary outcome. The re-analysis therefore is not considered to have any regulatory implications.

4.3 Impact of functional status at six months on long-term survival: prospective cohort studies[15] (see Annex 5 for publication)

This paper by Slot *et al* was published in 2008, however was not considered during the variation to extend the time-window for treatment with rt-PA to 4.5 hours. Whilst these prospective cohorts do not include any data on treatment with rt-PA, Dr [redacted] has highlighted this as an important piece of evidence.

4.3.1. Study description

The objective of the studies was to estimate the impact on long-term survival of functional status at six months after ischaemic stroke. Three cohorts were included, the Oxfordshire community stroke project, the Lothian stroke register and the first international stroke trial (in the UK).

A total of 7710 patients with stroke registered between 1981 and 2000 were included, with a maximum follow-up of 19 years. The main outcome was functional status at 6 months (mRS and 'two simple questions' – see Glossary) and mortality during follow up.

The Oxfordshire cohort registered patients from 1981-1986, with follow up at 1, 6 and 12 months and then annually for 5 years. A total of 539 patients were included.

The Lothian cohort registered patients from 1990-2000, with follow-up at 6, 12, 24, and 36 months. A total of 2054 patients were included.

The first international stroke trial was a randomised study of aspirin, s.c. heparin, both or neither started within 48 hours of ischaemic stroke. Patients were enrolled from 1991-1997, with follow up at 6 months. Assessment of independence was done by 'two simple questions'. A total of 5117 patients were included.

Survival data for all 3 cohorts was collected from the Office for National Statistics until 2000. Patients for whom no notification of death was received by this time were assumed to be alive.

4.3.2. Study findings

In a combined analysis of all three cohorts in patients who have survived to six months following ischaemic stroke, the median length of subsequent survival was 9.7 years, 95% CI [8.9-10.6] for patients who were independent in daily living; and 6.0 years, 95% CI [5.7-6.4] for those who were dependent.

In a combined analysis of the Oxfordshire and Lothian cohorts, subsequent median survival fell progressively from 12.9 years [10.0-15.9] for patients with a Rankin Score of 0-1 at 6 months following stroke to 2.5 years [1.4-3.5] for patients with a Rankin Score of 5.

The influence of functional outcome at 6 months on survival remained significant after adjustment for relevant covariates, e.g. age, presence of atrial fibrillation, visible infarct on CT scan, subtype of stroke.

Survival curves for the three cohorts are provided (see section 4.3.3. for the Oxfordshire cohort survival curve).

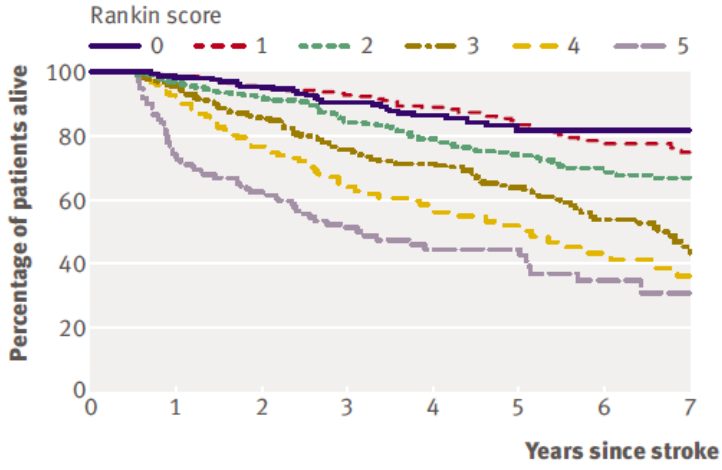


Figure 5: Survival curve for the Lothian cohort by mRS at 6 months

Figure 6: Survival curve for the international stroke trial cohort by independence at 6 months

4.3.3. Epidemiological Assessor's assessment of the Oxfordshire cohort

This project was a community based incidence study of stroke and transient ischaemic attacks. Patients were registered from 1981 to 1986.

Baseline characteristics were recorded in a standardised form. Trained study nurses followed up surviving patients at one, six, and 12 months from the date of stroke onset and then annually for up to five years. When possible, a study physician assessed survivors at the end of clinical follow-up.

Survival is shown below, by Rankin Score 6 months after stroke

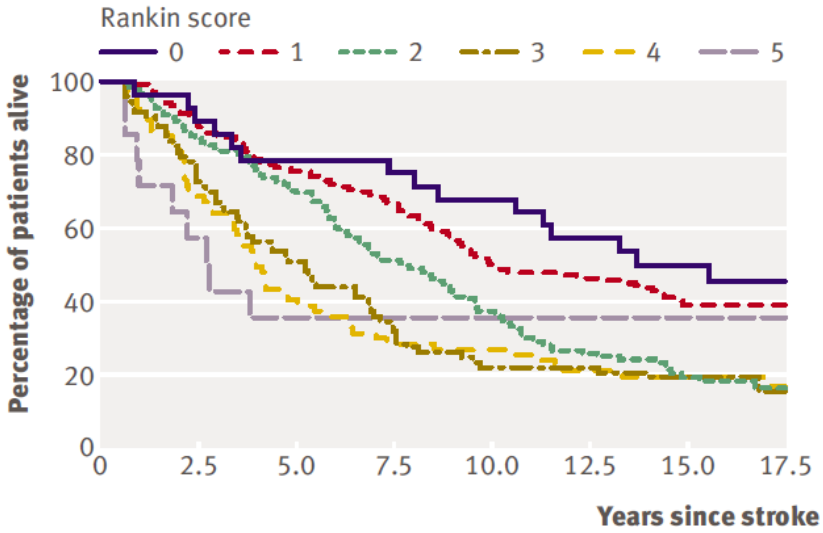


Figure 7: Survival curve for the Oxfordshire cohort by mRS at 6 months

Epidemiological Assessor's Comment:

The data clearly show that the lower the modified Rankin score at 6 months, the better the longer term outcome for the patient. This would provide some evidence that the 6 month Rankin score is an appropriate endpoint in a clinical trial as a surrogate for longer term outcomes. A score of 0 or 1 is clearly much better than 2 or above. The performance of those in Rankin Group 5 is clearly difficult to interpret as there are so few patients, resulting in a long horizontal KM curve from 5 years onwards, whereas all others show a decrease. However the data is very old, and it is

difficult to draw any further conclusions from this, for example on the expected survival rates of patients with stroke today.

4.3.4 Conclusions on the prospective cohort studies

These prospective cohort studies 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.

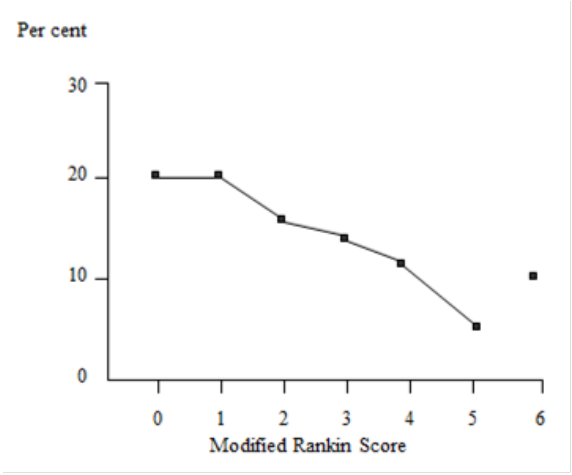
The authors of the paper also found a slight improvement in survival over the time period, and therefore these data are likely an underestimate of average survival following stroke under current medical care. This would reflect improvements in medical care, and perhaps in particular more intensive secondary prevention.

4.3.5 Specific concerns raised by Dr [redacted] regarding the prospective cohort studies

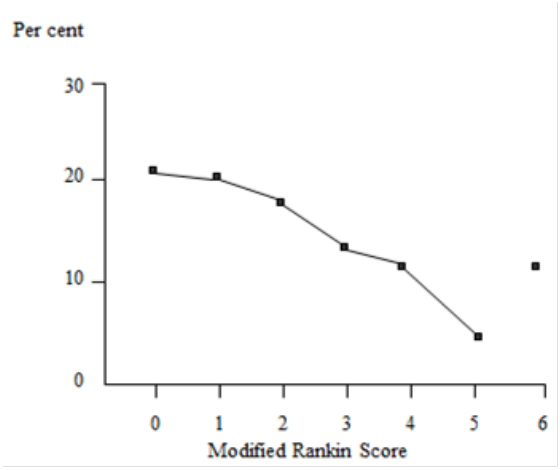
Dr [redacted] has compared the pattern of mRS observed in these cohorts at 6 months following stroke with that found in the combined analysis of clinical trials by Lees et al (ECASS, ATLANTIS, NINDS and EPITHET trials).

1. "…the patterns seen when the pooled outcome data 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'"

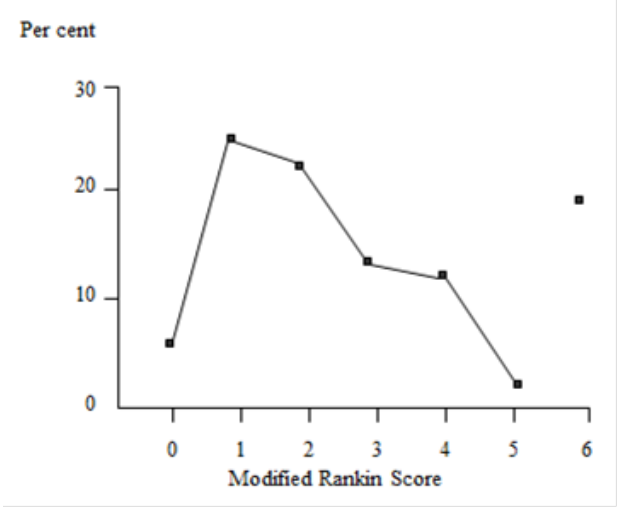
The following are the figures used by Dr [redacted] to illustrate these points:



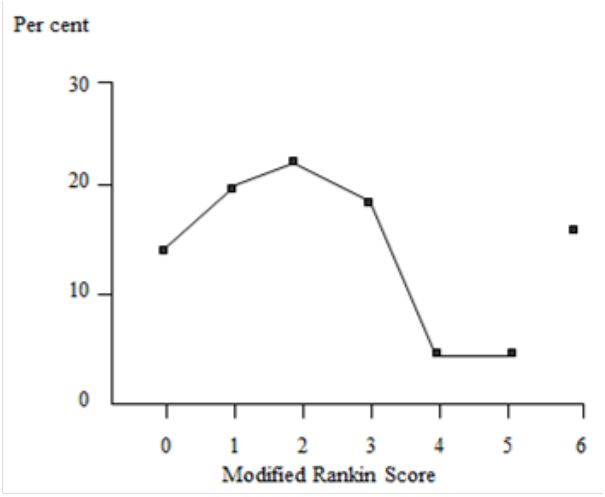
a) Cohort - 3 months after treatment with rt-PA within 180 min (SITS-ISTR, n=10 231)



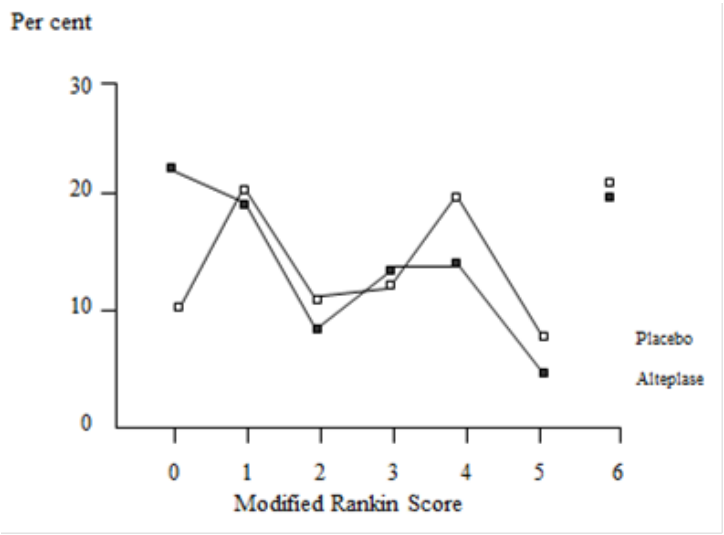
b) Cohort - 3 months after treatment with rt-PA within 181 to 270 min (SITS-ISTR, n=541)



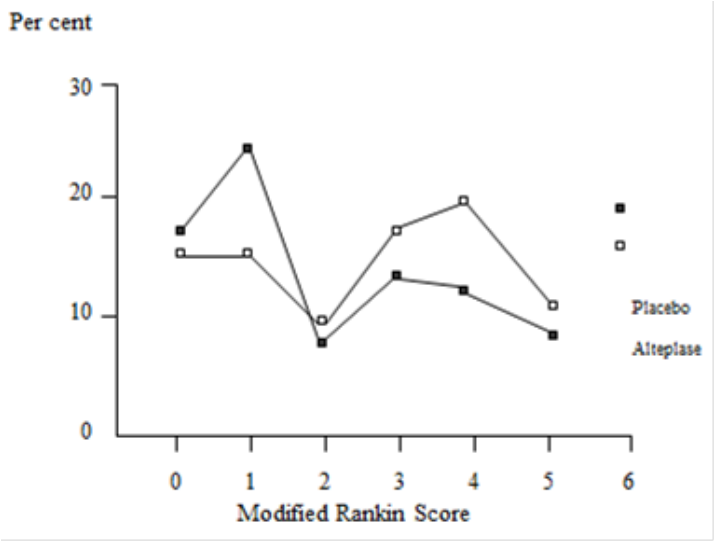
c) Cohort - 6 months following cerebral infarction in the Oxford Community Stroke Project (n=539)



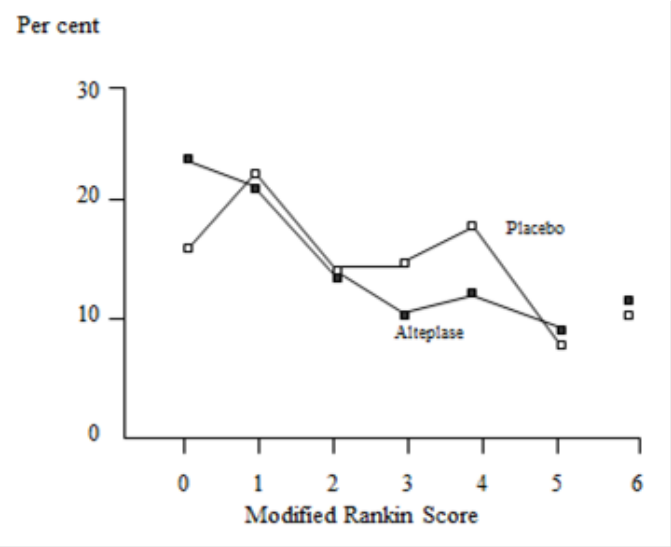
d) Cohort - 6 months following cerebral infarction in the Lothian cohort (n=2054)



e) Randomised trials of rt-PA, 3 month outcome - within 90 min (n=312)



f) Randomised trials of rt-PA, 3 month outcome - 91 to 180 min (n=618)



g) Randomised trials of rt-PA, 3 month outcome - 181 to 270 min (n=1620)

Statistical assessor's comments on mRS plots

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 and so would be expected to be smooth, while figure e) is based on only about 150 for each curve, 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, 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.

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.

Assessor's comments: As explained above, these plots do not raise concerns about the data from trials of rt-PA. The value of comparing these cohorts of data is questioned – untreated, unselected observational cohorts of patients compared with clinical trial populations with their many inclusion and exclusion criteria. As with all randomised clinical trials, the populations included in these studies is unlikely to be representative of all patients with ischaemic stroke.

As an aside, it is noted that overall the plots have generally higher percentages for mRS=0 for rt-PA treated patients compared with untreated patients whilst the opposite is generally found for mRS=6.

4.4 Impact of treatment delay, age and stroke severity on the effects of intravenous thrombolysis with rt-PA in acute ischaemic stroke: an individual-patient-data meta-analysis (unpublished)

4.4.1. Study description

This meta-analysis was presented recently at the American Stroke Association meeting in San Diego. The meta-analysis is yet to be published; however the presenter (Jonathan Emberson, The Stroke Thrombolysis Trialist's Collaborative Group) has shared their slides from the talk. The protocol for the meta-analysis has been published[18]. This meta-analysis includes data from 9 trials (ATLANTIS A/B, ECASS I/II/III, EPITHET, IST-3, NINDS part 1 and 2) and 6756 randomised patients.

The meta-analysis aims to assess:

1. the extent to which treatment delay modifies the effect of rt-PA on stroke outcome
2. the extent to which age or stroke severity modify these effects and
3. the effects of rt-PA on risk of symptomatic ICH and mortality.

The primary efficacy outcome is mRS 0-1 at 3-6 months post-stroke. Safety endpoints are 90 day mortality, symptomatic ICH, fatal ICH within 7 days.

4.4.2. Results

The results presented for the meta-analysis were as follows, for efficacy:

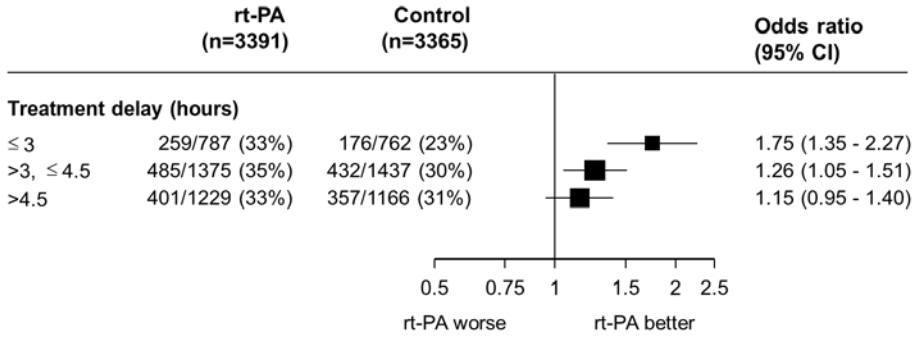


Figure 8: The effect on mRS 0-1 at 3-6 months by treatment delay

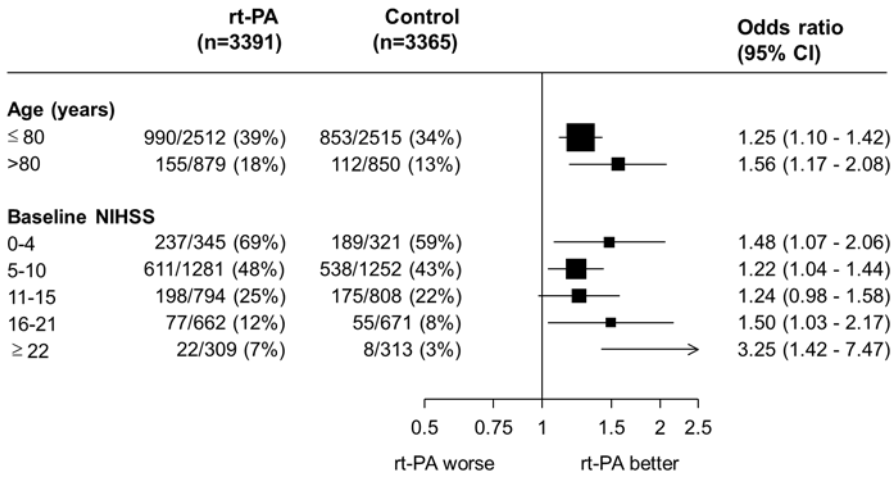


Figure 9: The effect on mRS 0-1 at 3-6 months by age and stroke severity

The risk of ICH was increased with rt-PA treatment compared with controls, both at 36 hours post-stroke and at 7 days. The risk of fatal ICH within 7 days was also

significantly raised, RR=7.14, 95% CI [3.98-12.8]. Death within 90 days was numerically increased but not statistically significantly greater in the rt-PA group compared with controls, RR=1.11, 95% CI [0.99-1.25].

The 90-day mortality risk was found to increase with increasing time to treatment, although no time period reached statistical significance:

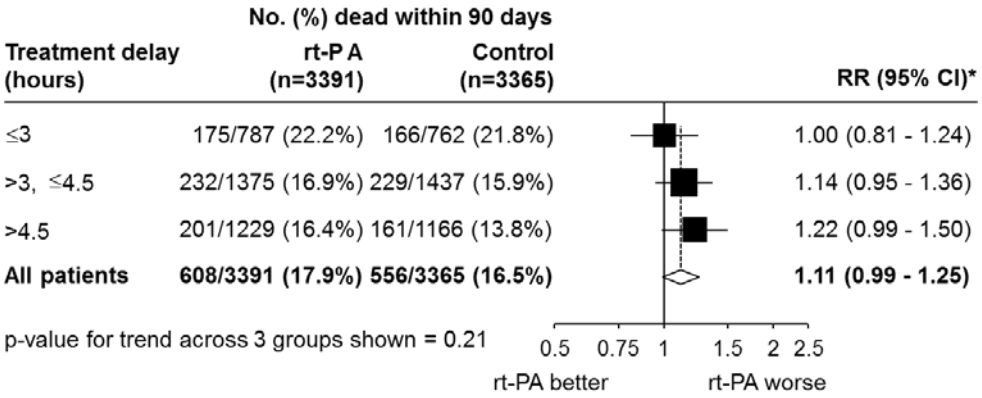


Figure 10: The effect of treatment delay on 90-day mortality

4.4.3. Epidemiological Assessor's assessment of the meta-analysis

It is extremely difficult to interpret these data without access to the full peer-reviewed details including a better understanding of what data has been included in the meta-analysis. The methods chosen to analyse the data have not been specified in the protocol, although it seems likely that standard fixed or random effects models have been used. The Forest plots do not show a huge amount of heterogeneity, so it seems unlikely that the statistical method chosen for the meta-analysis would alter the results.

What is perhaps most difficult to interpret is the effect in the 3 – 4.5 hour period. The data provided by ECASS III were modest, and only mildly positive. It is noted elsewhere in this assessment report that the point estimate of effect in this time period in IST-3 was substantially against rt-PA. Consequently it is surprising that the meta-analysis for this time period shows a statistically significant beneficial effect. Further information on the methodology and the effects of each trial for this time point are essential to help interpret the findings. It is possible there is substantial heterogeneity between trials, and understanding this further is important.

It is also unclear whether all surviving patients, or all patients randomised to treatment comprise the denominator in the individual studies that contribute to this analysis. As discussed in the review of the IST-3 study, this could have an impact on the interpretation. It is also possible that this is not consistent across studies.

The data as presented do suggest a beneficial effect in the currently indicated population, and an increased risk of ICH, particularly at 7 days, in line with what is currently known. Whether further understanding of the data that went into the meta-analysis and subsequent sensitivities analyses would change this is unknown.

4.4.4. Conclusions on unpublished meta-analysis

Without the full information on this meta-analysis, which will not be available until it is published, it is not possible to draw any final conclusion on these data. From the preliminary presentation at the American Stroke Association meeting, it would seem that the meta-analysis results support a beneficial effect of rt-PA up to 4.5 hours, whilst confirming an increased risk of ICH particularly within the first 7 days after treatment. This is in line with current understanding of rt-PA and would not impact on the licensed use.

5. Discussion and conclusion

The balance of benefits and risks of rt-PA in the indication of acute ischaemic stroke has prompted extensive discussion and analysis since initial licensing of the indication in 2002. These discussions have taken place both within the regulatory organisations at a national and European level and as debate in the scientific literature amongst the medical community, where viewpoints are often polarised on the benefit or otherwise of the use of rt-PA.

For several reasons, the balance of benefits and risks of rt-PA treatment for acute ischaemic stroke is perhaps particularly difficult to judge and achieve consensus. The efficacy of rt-PA in improving outcome has been demonstrated in some randomised controlled trials, however there are also RCTs that failed in their efficacy endpoints. Initial licensing discussions were not straightforward, having to consider the applicability of the NINDS results to the European stroke population, and the effects of the different criteria used in the different trials – varying enrolment time windows, doses, baseline severity of stroke. Treatment with rt-PA carries an increased risk of intracranial haemorrhage, with potentially devastating consequences. Balancing the evidence for efficacy against the risk of intracranial haemorrhage, whilst giving consideration to all of the other variables involved, is not a straightforward task. Furthermore, a patient's perspective should be borne in mind, if possible, when considering this balance.

The point has been made, initially by a UK clinician on behalf of the MAH and accepted by the MHRA clinical assessor at the time, that faced with a potentially life-changing ischaemic stroke, patients are likely to be willing to receive a treatment that provides the chance of complete recovery – despite the risk of serious intracranial bleeding, and may be expected to accept a higher risk than would be acceptable for a less severe condition. It is easy to understand why opinions on the overall balance of benefits and risks will vary substantially.

Throughout the initial licensing procedure for the indication in stroke, the UK held a negative position on benefit:risk balance – and maintained a negative stance at the final vote at CPMP. A major objection was also raised by the UK during the extension to the time-window of treatment to 4.5 hours. This resulted in extensive analyses of the data and discussion at national committees, and finally led the UK to agree to the extension.

Irrespective of the previous national position, the starting point for this new review is the current EU regulatory position. This makes it difficult to re-open any debate on the benefits and risks of rt-PA without important new data. However, it is essential that any such new data or major concerns about the data that underpin the current marketing authorisation are thoroughly investigated.

The current paper considers data that have become available since the grant of the extension to the time-window for treatment, or that have not previously been considered. The decision of which data are pertinent to the current review was taken in conjunction with Dr [redacted] and attempts have been made to evaluate the specific concerns raised by him.

The greatest amount of new raw data to have become available since the extension to the time-window is provided by the results of the IST-3 trial. This was a large, international, randomised, open-label trial, which enrolled a total of 3035 patients. The primary outcome was the proportion of patients alive and independent at 6 months following stroke, with a time-window for treatment of up to 6 hours after symptom onset.

Overall, the primary outcome for the trial failed, and there was a clear finding of excess mortality in the rt-PA treated group at 7 days. The mortality difference was

not observed at 6 months, but the primary outcome did not compensate for this early negative effect. Secondary endpoints provided some suggestion of a positive shift in OHS in favour of rt-PA treatment, although the shifts were relatively small and there is difficulty in interpretation of secondary endpoints when the primary outcome has failed. In addition, improvements in secondary endpoints did not result in a mortality benefit at 18 months of follow-up.

The subgroup treated within 3 hours of symptom onset did demonstrate some improvement in the proportion of patients alive and independent at 6 months, however this was lost when the data for the 3-4.5 hour group was added in. Therefore, whilst some support is provided for the use of rt-PA in ischaemic stroke up to 3 hours post symptom onset, there is no support from this trial for its use between 3 and 4.5 hours.

From a regulatory perspective, there are some additional considerations that should be taken into account:

- The IST-3 trial evaluated treatment with rt-PA up to 6 hours after the onset of symptoms, whilst the licence specifies treatment up to 4.5 hours. Therefore it could be argued that the failure of the primary outcome per se cannot be considered to be truly reflective of a negative balance of benefits and risks for use under the terms of the licence.
- A mortality benefit for rt-PA has never been claimed and therefore the lack of improvement in mortality cannot be considered to impact on the assessment of benefit and risk in the licenced use of rt-PA.
- The IST-3 trial purposefully enrolled patients for whom rt-PA was not specifically indicated, and who did not meet the prevailing licence criteria (95% of enrolled patients). Over half of patients were over the age of 80 years. Whilst a positive result in such a trial population might provide supporting evidence for a positive benefit-risk balance in the licenced population, the implications of a negative result are less clear – particularly as the trial population could be considered to be a higher risk group.

Debate surrounding the results of the pivotal NINDS trials eventually prompted an independent re-analysis of the trial data. This was carried out in 2004, however it was not considered when the time window for treatment was extended to 4.5 hours. Dr [redacted] considers that the re-analysis identifies issues that add to concerns over the quality of the data on which the existing licences were originally based.

The committee's findings from 2004 are supported, that the trial demonstrates a benefit of treatment in the primary outcome. Upon evaluation, the potential issues highlighted by Dr [redacted] regarding the data (centre effects, the pattern of time to treatment onset and the violations of the protocol with regards to blood pressure) are not considered to raise concerns. [redacted]

[redacted] the other points relating to e.g. the randomisation process and the composition of the placebo, are not considered to have a significant impact.

Additional concerns raised by Dr [redacted] relate to published pooled analyses, and in particular that the pattern of data generated from the mRS for the pooled analyses from RCTs is improbable when compared with observational data. However there are good reasons why we would not necessarily expect clinical trial data (with all their exclusion and inclusion criteria, and their relatively small size) to mirror larger unselected populations of patients. Comparing the two sets of data is made more difficult because the plots of mRS values do not show the variability around the curves, which is important when considering if the distributions are truly different.

The additional information that has most recently become available is the unpublished results of a meta-analysis. Without access to the full data it is not possible to draw any firm conclusions, however superficially the findings were supportive of the current understanding of rt-PA and therefore would not be expected to impact negatively on the licensed use.

In conclusion, after careful consideration, the new data and the issues highlighted by Dr [redacted] are not considered to impact on the balance of benefits and risks of rt-PA in the authorised indication of acute ischaemic stroke. Whilst the benefit-risk balance of treatment from 3-4.5 hours was negative in the IST-3 trial, it was conducted in a patient population generally not covered by the current EU licence.

6. Advice sought

Does the Commission consider that the data discussed have implications for the balance of benefits and risks of rt-PA as currently authorised in the EU?

The Commission is also asked whether they consider that the issue needs to be further explored by an ad-hoc expert working group.

7. References

1. Lenzer J, *Why we can't trust clinical guidelines*. BMJ 2013; 346:f3830
2. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. *Tissue plasminogen activator for acute ischaemic stroke* NEJM 1995; 333:1581–1587.
3. Hacke W, Kaste M, Fieschi C, Toni D, Lesaffre E, von Kummer R, Boysen G, Bluhmki E, Höxter G, Mahagne MH, et al. *Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS)* JAMA 1995; 274(13):1017-25.
4. 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 *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(9136):1245-51.
5. 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 investigators.* Stroke. 2000; 31(4):811-6.
6. 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. Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke.* JAMA. 1999; 282(21):2019-26.
7. Hacke W, Kaste M, Bluhmki E, Brozman M, Dávalos A, Guidetti D, Larrue V, Lees KR, Medeghri Z, Machnig T, Schneider D, von Kummer R, Wahlgren N, Toni D; ECASS Investigators. *Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke.* NEJM 2008; 359(13):1317-29
8. Wahlgren N, Ahmed N, Dávalos A, Hacke W, Millán M, Muir K, Roine RO, Toni D, Lees KR; SITS investigators. *Thrombolysis with alteplase 3-4.5 h after acute ischaemic stroke (SITS-ISTR): an observational study.* Lancet 2008; 372(9646):1303-9.
9. Koennecke HC, Nohr R, Leistner S, Marx P. *Intravenous tPA for ischemic stroke team performance over time, safety, and efficacy in a single-center, 2-year experience.* Stroke. 2001; 32(5):1074-8
10. Lopez-Yunez AM, Bruno A, Williams LS, Yilmaz E, Zurrú C, Biller J. *Protocol violations in community-based rTPA stroke treatment are associated with symptomatic intracerebral hemorrhage.* Stroke 2001; 32(1):12-6.
11. 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(9):1145-50.
12. 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.

13. 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.

14. 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

15. Slot KB, 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.

16. European Stroke Council, European Neurological Society and European Federation of Neurological Societies. *European Stroke Initiative recommendations for stroke management.* Cerebrovasc Dis. 2000; 10(5):335-51.

17. Sandercock *et al*.: *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.

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.* International Journal of Stroke 2013; 8: 278–283.

8. Glossary

1) **NIHSS: National Institute of Health Stroke Scale** [Brott et al, Stroke 1989;20:864-870]

A 15-item neurologic examination stroke scale used to evaluate level of consciousness, language, neglect/inattention, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss.

The scale runs from 0 – 42, with 42 being the worst outcome:

≤5 mild impairment

6-14 moderately severe impairment

15-24 severe impairment

≥25 very severe neurological impairment

A 2-point or greater increase on the NIHSS administered serially indicates stroke progression. A change from 0 to 1 may indicate a new deficit.

2) **mRS: modified Rankin Scale**

A seven point scale (0-6), as follows:

0 no symptoms

1 no significant disability

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

3) **OHS: Oxford Handicap Score**

The OHS is a modification of the Rankin Scale:

	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

4) **BI: the Barthel Index**

A 10-item examination that assesses feeding, chair/bed transfer, grooming, toileting, bathing, ambulation, stair climbing, dressing, bowel control and bladder control.

Scores range from 0 (dependent) to 100 (independent), although patients with a score of 100 are not necessarily able to live independently.

5) Glasgow outcome scale

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

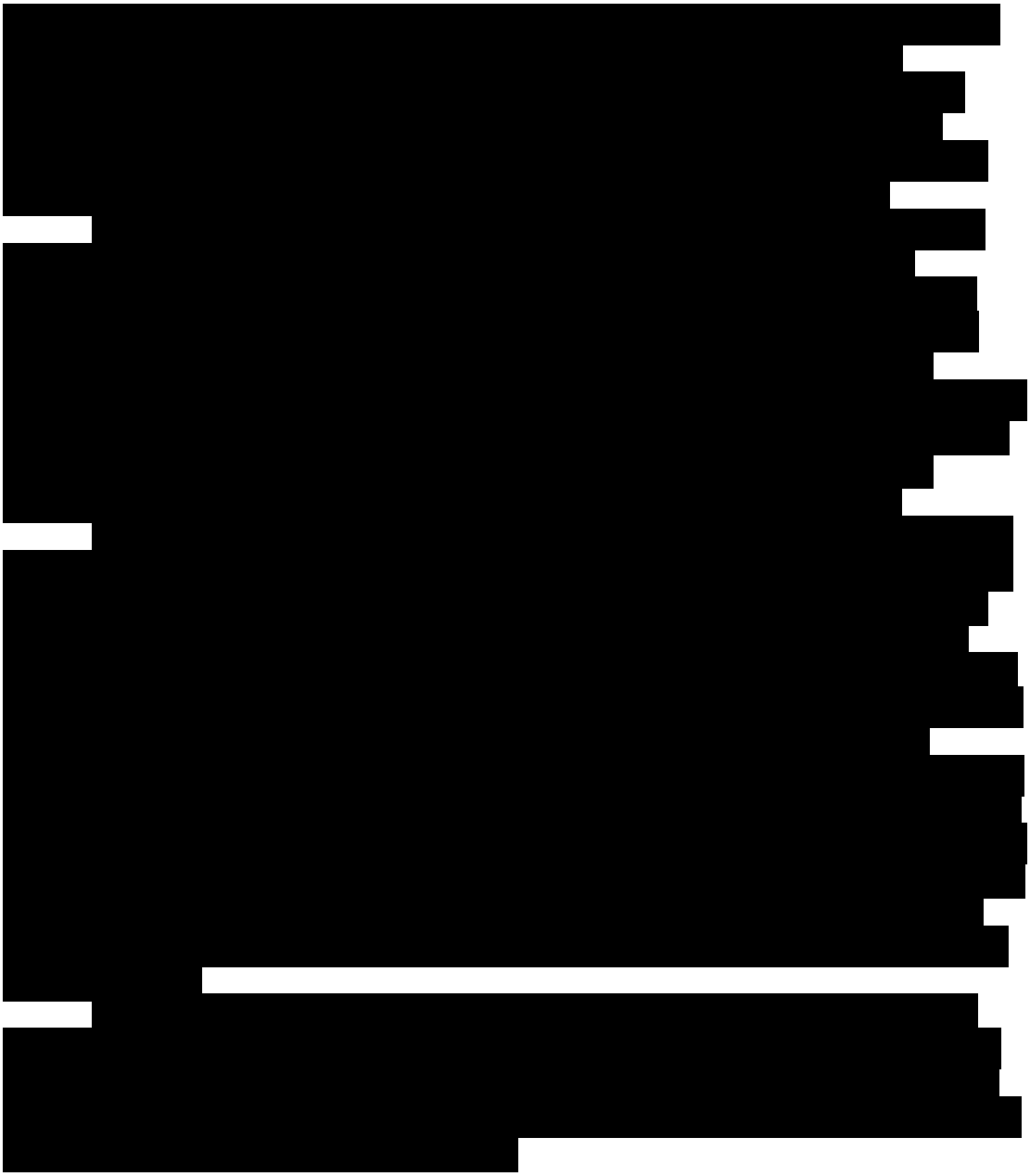
4 persistent vegetative state – patient exhibits no obvious cortical function

5 death

6) 'Two simple questions' (used in the first IST)

This was developed to assess functional outcome after stroke in large scale trials. Patients were asked if they had needed help from another person to perform everyday activities within the past two weeks (such as bathing, feeding, walking, dressing, or use of the toilet). Patients not requiring any help were classified as independent.

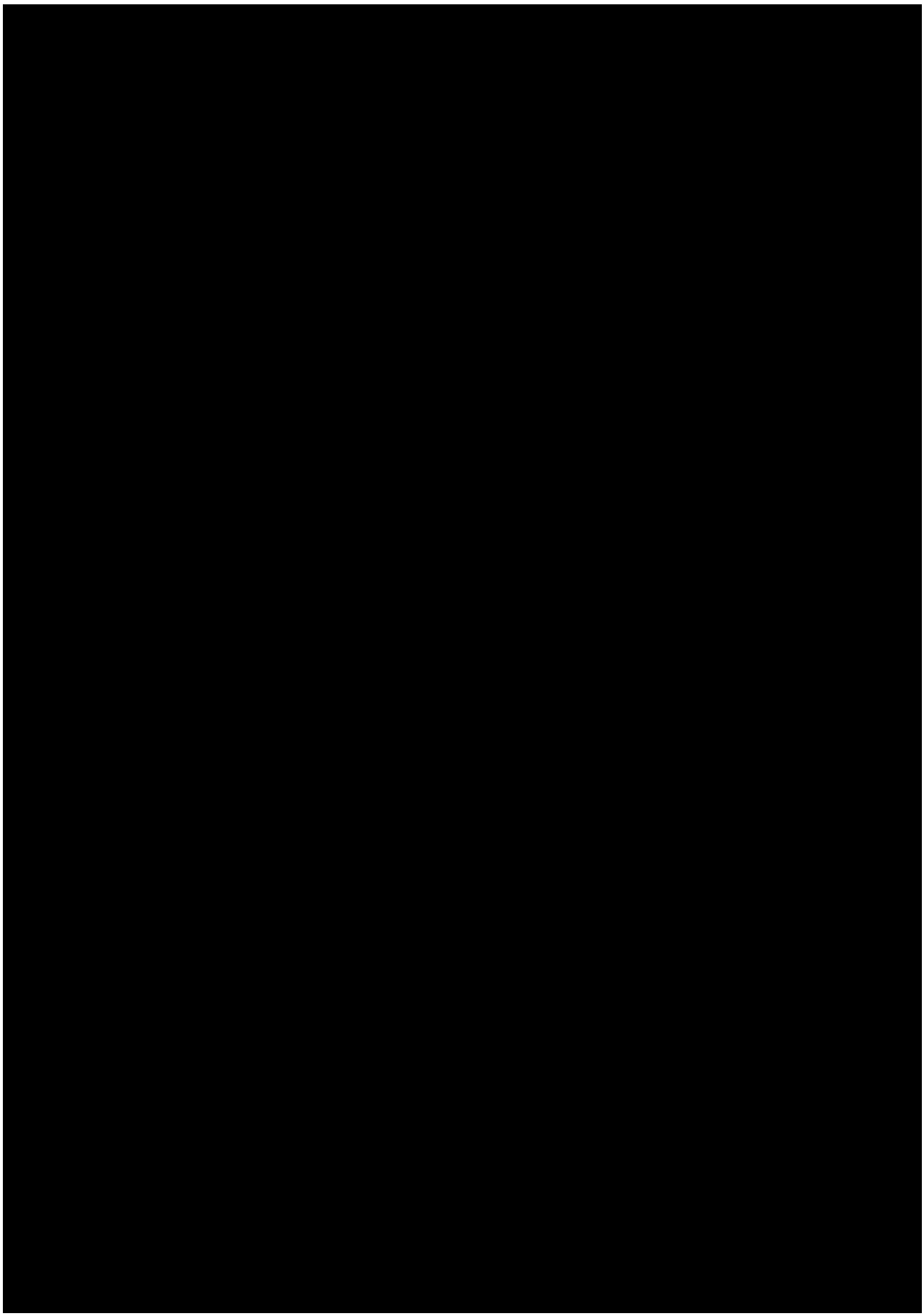
[redacted]



[redacted]

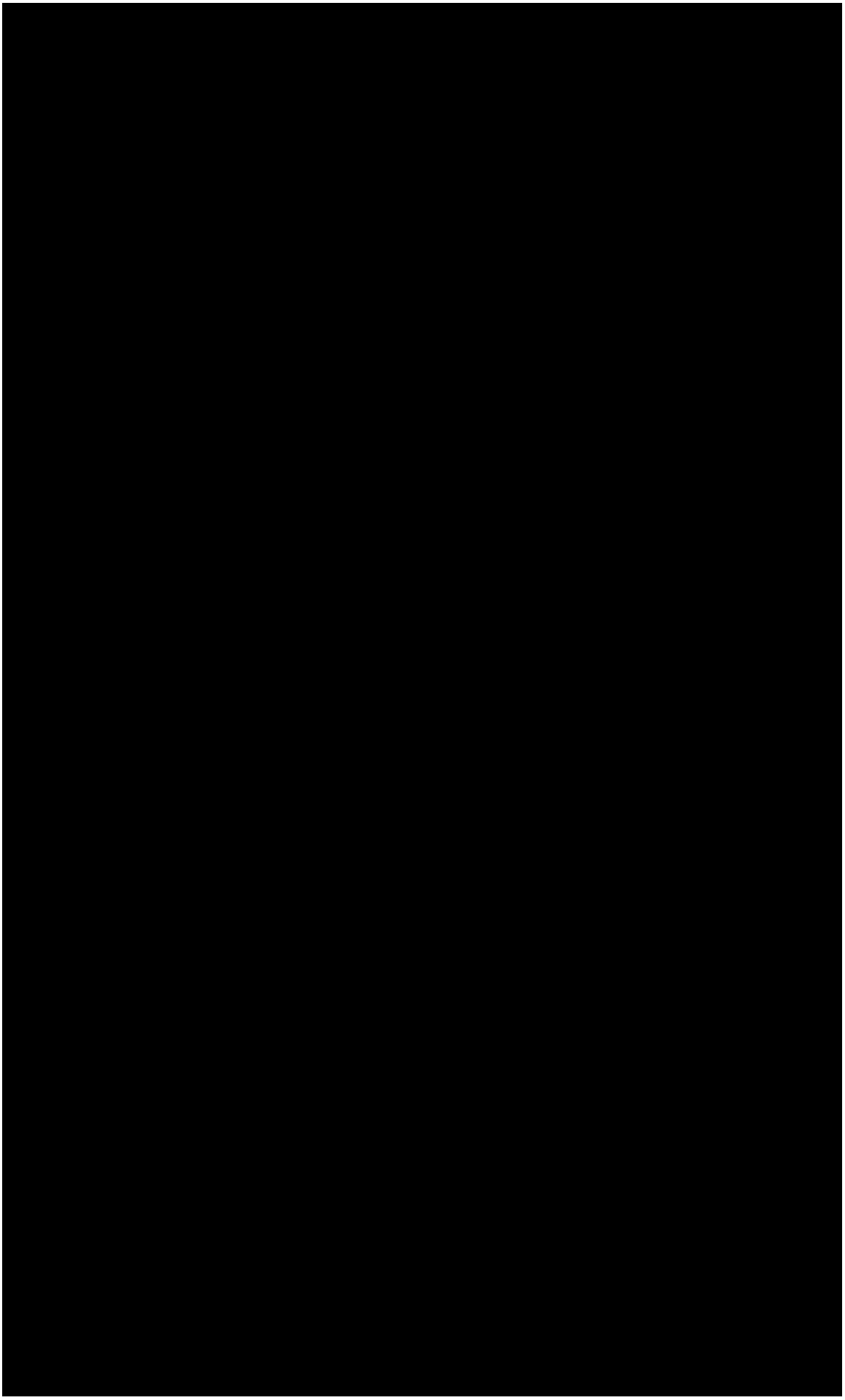
[redacted]

[redacted]



[redacted]

[redacted]



[redacted]

[redacted]

[redacted]

ANNEX 3: IST-3 publications and web appendices

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

*The IST-3 collaborative group**

Lancet 2012; 379: 2352–63
Published Online
May 23, 2012
DOI:10.1016/S0140-6736(12)60768-5
This online publication has been corrected. The corrected version first appeared at thelancet.com on August 24, 2012

See Comment page 2320
See Articles page 2364

*Members listed in the appendix
Correspondence to:
Prof Peter Sandercock, Division of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK
peter.sandercock@ed.ac.uk

Summary

Background 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. The third International Stroke Trial (IST-3) sought to determine whether a wider range of patients might benefit up to 6 h from stroke onset.

Methods In this international, multicentre, randomised, open-treatment trial, patients were allocated to 0·9 mg/kg intravenous recombinant tissue plasminogen activator (rt-PA) or to control. The primary analysis was of the proportion of patients alive and independent, as defined by an Oxford Handicap Score (OHS) of 0–2 at 6 months. The study is registered, ISRCTN25765518.

Findings 3035 patients were enrolled by 156 hospitals in 12 countries. All of these patients were included in the analyses (1515 in the rt-PA group *vs* 1520 in the control group), of whom 1617 (53%) were older than 80 years of age. At 6 months, 554 (37%) patients in the rt-PA group versus 534 (35%) in the control group were alive and independent (OHS 0–2; adjusted odds ratio [OR] 1·13, 95% CI 0·95–1·35, p=0·181; a non-significant absolute increase of 14/1000, 95% CI –20 to 48). An ordinal analysis showed a significant shift in OHS scores; common OR 1·27 (95% CI 1·10–1·47, p=0·001). Fatal or non-fatal symptomatic intracranial haemorrhage within 7 days occurred in 104 (7%) patients in the rt-PA group versus 16 (1%) in the control group (adjusted OR 6·94, 95% CI 4·07–11·8; absolute excess 58/1000, 95% CI 44–72). More deaths occurred within 7 days in the rt-PA group (163 [11%]) than in the control group (107 [7%], adjusted OR 1·60, 95% CI 1·22–2·08, p=0·001; absolute increase 37/1000, 95% CI 17–57), but between 7 days and 6 months there were fewer deaths in the rt-PA group than in the control group, so that by 6 months, similar numbers, in total, had died (408 [27%] in the rt-PA group *vs* 407 [27%] in the control group).

Interpretation For the types of patient recruited in IST-3, despite the early hazards, thrombolysis within 6 h improved functional outcome. Benefit did not seem to be diminished in elderly patients.

Funding UK Medical Research Council, Health Foundation UK, Stroke Association UK, Research Council of Norway, Arbetsmarknadens Partners Forsakringsbolag (AFA) Insurances Sweden, Swedish Heart Lung Fund, The Foundation of Marianne and Marcus Wallenberg, Polish Ministry of Science and Education, the Australian Heart Foundation, Australian National Health and Medical Research Council (NHMRC), Swiss National Research Foundation, Swiss Heart Foundation, Assessorato alla Sanita, Regione dell'Umbria, Italy, and Danube University.

Introduction

Each year, about 22 million people have a stroke worldwide,[1,2] of whom 4 million reside in high-income countries,[3,4] where thrombolytic therapy is affordable and feasible. The burden of ischaemic stroke among the elderly is large and increasing;[2,5] and we estimate that annually ischaemic stroke affects about a million people older than 80 years of age in high-income countries and about 3 million in low-income and middle-income countries.

Thrombolytic therapy with intravenous recombinant tissue plasminogen activator (rt-PA), when approved in Europe, was restricted to the treatment of patients younger than 80 years of age with acute ischaemic stroke who could be treated within 3 h. A Cochrane systematic review of the 11 completed trials of thrombolysis (including 3977 patients) with intravenous rt-PA for acute ischaemic stroke showed that treatment was associated with a significant increase in survival free of disability, despite an early 3% excess of fatal intracranial haemorrhage.[6] The review also suggested that treatment might be beneficial up to 6 h.[6] An individual patient data meta-analysis of a subset of intravenous rt-PA trials further showed that the earlier treatment was given, the greater the chance of a favourable outcome.[7] Older people have been under-represented in stroke trials in general,[8] and in stroke thrombolysis trials in particular (only 79 people aged older than 80 years had been included in trials of rt-PA).[6] As a result of the current European Union (EU) approval criteria, treatment is only applicable to a small proportion of patients with acute stroke.[9]

The Third International Stroke Trial (IST-3), therefore, had the following objectives: to establish the balance of benefits and harms of thrombolytic therapy with rt-PA in patients who did not exactly meet the licence criteria (especially elderly patients); determine whether a wider range of patients might benefit from this treatment; assess which categories of patients were most likely to benefit by investigating possible interactions between treatment effect and various factors (including age, stroke severity, and early brain imaging results); refine current estimates of the duration of the therapeutic time window; and to improve the external validity and precision of the existing estimates of the overall treatment effects (benefits and harms). The primary trial hypothesis was that 0·9 mg/kg rt-PA (maximum 90 mg) given to adult patients of all ages with acute ischaemic stroke, within 6 h of symptom onset, increased the proportion of people who were alive and independent at 6 months.

Methods

Study design and patients

IST-3 was a pragmatic[10] international, multicentre, randomised-controlled, open-treatment trial. The initial pilot phase was double-blinded and placebo-controlled. At the end of the pilot phase, since the main phase compared treatment with open control, several additional measures were introduced to minimise bias in the assessment of early and late outcomes.[11] We have published reports of the rationale for the trial,[12] the protocol,[13] an update on recruitment, amendments to the protocol and the baseline characteristics of the patients recruited,[11] and the statistical analysis plan.[14]

The eligibility criteria can be summarised in terms of the uncertainty principle.[15–17] Inclusion and exclusion criteria are listed in detail in the protocol.[13] Briefly, patients were eligible according to the following criteria: they had symptoms and signs of clinically definite acute stroke; the time of stroke onset was known; treatment could be started within 6 h of onset; and CT or MRI had reliably excluded both intracranial haemorrhage and structural brain lesions, which could mimic stroke (eg, cerebral tumour). Additionally, if the patient had a clear indication for intravenous thrombolysis with rt-PA, they were to be treated in accordance with local guidelines. Equally, if the patient had a clear contraindication to treatment they were not to be entered in the trial. Only if both the clinician and the patient (or a relevant proxy for the patient) felt that the treatment was promising but unproven, could the patient be included in the trial after appropriate informed consent from the patient or a valid proxy. The protocol was approved by the Multicentre Research Ethics Committees, Scotland (reference MREC/99/0/78), and by local ethical committees.

This study is registered, ISRCTN25765518.

Procedures

Clinicians entered baseline data via a telephone voice-activated or a secure web-based randomisation system. After the system had recorded and checked the data, patients were allocated either immediate thrombolysis with 0·9 mg/kg of intravenous rt-PA to a maximum of 90 mg (10% bolus with the remainder over 1 h) or control treatment. The system would not accept patients with blood pressure or glucose levels outside protocol-defined criteria (appendix pp 4–5) or other data inconsistencies. The system used a minimisation algorithm to achieve optimum balance for key prognostic factors (table 1), and from January, 2006, minimisation was additionally stratified by world region and then minimised on all the other key factors within regions.

See Online for appendix

To be eligible to join the trial, participating hospitals had to have an organised system of stroke care. Acute-care protocols were not specified by the trial, but had to include the components of effective stroke-unit care,[19] including, soon after admission, intravenous access, monitoring of physiological variables, correction of any abnormalities, and where clinically appropriate, intravenous-fluid therapy. All patients in the trial were to be treated within that organised system of stroke care, irrespective of treatment allocation. Patients allocated to the control group were to avoid treatment with rt-PA and received stroke care in the same clinical environment as those allocated to the rt-PA group. Both treatment groups had blood pressure monitored closely over the first 24 h. In the double-blinded phase, both groups were to avoid antiplatelet or anticoagulant therapy for 24 h. In the open phase, patients allocated to the control group were to start aspirin immediately. Blood pressure was managed in the same way in both treatment groups, according to local protocol. Additionally, all centres were asked for their pretrial experience of thrombolysis for treatment of stroke, and if the centre had, before joining the trial, a protocol for open-label use of rt-PA and had treated at least three people in the 12 months before joining the trial, the centre was classed as experienced.

For the **study protocol** see http://www.ist3.com

All patients had a CT or MRI brain scan before randomisation and a follow-up scan at 24–48 h. A repeat brain scan was required if the patient deteriorated neurologically or intracranial haemorrhage was suspected for any reason. Although CT scanning was preferred, MRI was allowed. All scans were sent to the trial centre in Edinburgh for masked central rating of any signs of visible early ischaemia (presence and extent of hypoattenuation, swelling, hyperattenuated artery), haemorrhage, and background brain changes (leukoaraiosis, atrophy, prior stroke lesions, non-stroke lesions) with validated rating methods.[20–25] Images were assessed with all original identifiers stripped from the record, and then viewed via a secure web-based image viewing system by an international panel of expert radiologists. All assessments were made masked to all patient details and treatment allocation.

The primary outcome specified in version 1·93 of the protocol and in the published statistical analysis plan[14] was the proportion of patients alive and independent as

measured by the Oxford Handicap Score (OHS),[26] a commonly used variant of the modified Rankin score.[27] Patients with an OHS of 0, 1, or 2 were classed as independent. The statistical analysis plan specified an ordinal analysis of the OHS score at 6 months. Additional secondary outcomes were to be reported separately.

Events occurring within 7 days of stroke were recorded by the local trial clinician on the 7-day form: deaths subdivided by cause (swelling of the initial infarct, intracranial haemorrhage, other deaths from the initial stroke, recurrent ischaemic stroke, recurrent stroke of unknown type, any other cause); symptomatic intracranial haemorrhage; recurrent ischaemic stroke; recurrent stroke of unknown type; neurological deterioration attributed to swelling of the initial ischaemic stroke; neurological deterioration not attributable to swelling of the initial ischaemic stroke or haemorrhage; and major extracranial haemorrhage (operational definitions of

	rt-PA (n=1515)	Control (n=1520)
Baseline variables collected before treatment allocation*		
Region†		
Northwest Europe (UK, Austria, Belgium, Switzerland)	792 (52%)	797 (52%)
Scandinavia (Norway, Sweden)	251 (17%)	250 (16%)
Australasia	89 (6%)	90 (6%)
Southern Europe (Italy, Portugal)	204 (13%)	204 (13%)
Eastern Europe (Poland)	174 (11%)	173 (11%)
Americas (Canada, Mexico)	5 (<1%)	6 (<1%)
Age (years)†		
18–50	59 (4%)	68 (4%)
51–60	98 (6%)	104 (7%)
61–70	188 (12%)	177 (12%)
71–80	353 (23%)	371 (24%)
81–90	706 (47%)	701 (46%)
>90	111 (7%)	99 (7%)
Sex†		
Female	782 (52%)	788 (52%)
NIHSS†		
0–5	304 (20%)	308 (20%)
6–10	422 (28%)	430 (28%)
11–15	306 (20%)	295 (19%)
16–20	270 (18%)	273 (18%)
>20	213 (14%)	214 (14%)
Delay in randomisation†‡		
0–3·0 h	431 (28%)	418 (28%)
3·0–4·5 h	577 (38%)	600 (39%)
4·5–6·0 h	507 (33%)	500 (33%)
>6·0 h	0 (0%)	2 (<1%)
Atrial fibrillation	473 (31%)	441 (29%)
Systolic blood pressure		
≤143 mm Hg	487 (32%)	492 (32%)
144–164 mm Hg	498 (33%)	518 (34%)
≥165 mm Hg	530 (35%)	510 (34%)
Diastolic blood pressure§		
≤74 mm Hg	462 (31%)	445 (29%)
75–89 mm Hg	541 (36%)	588 (39%)
≥90 mm Hg	500 (33%)	480 (32%)
Blood glucose¶		
≤5 mmol/L	254 (18%)	285 (21%)
6–7 mmol/L	664 (48%)	638 (46%)
≥8 mmol/L	455 (33%)	456 (33%)
(Continues in next column)		
(Continued from previous column)		
Treatment with antiplatelet drugs in previous 48 h†	775 (51%)	787 (52%)
Predicted probability of poor outcome at 6 months‖		
<40%	351 (23%)	378 (25%)
40–50%	169 (11%)	160 (11%)
50–75%	361 (24%)	357 (23%)
≥75%	634 (42%)	625 (41%)
Stroke clinical syndrome†**		
TACI	639 (42%)	666 (44%)
PACI	596 (39%)	551 (36%)
LACI	168 (11%)	164 (11%)
POCI	110 (7%)	136 (9%)
Other	2 (<1%)	3 (<1%)
Baseline variables collected from prerandomisation scan		
Expert reader's assessment of acute ischaemic change††	..	..
Scan completely normal	140 (9%)	129 (8%)
Scan not normal but no sign of acute ischaemic change	743 (49%)	781 (51%)
Signs of acute ischaemic change	624 (41%)	600 (40%)

Data are number (%). Percentages exclude missing values from denominators. rt-PA=recombinant tissue plasminogen activator. NIHSS=National Institutes of Health Stroke Scale. TACI=total anterior circulation infarct. PACI=partial anterior circulation infarct. LACI=lacunar infarct. POCI=posterior circulation infarct. *Data for these variables were gathered via the web-based or telephone randomisation system and had to be entered, complete, and have passed range and consistency checks before the system would issue a treatment allocation. †Variables were used in the minimisation algorithm. ‡Two patients in the control group were randomly assigned at more than 6 h (protocol violation). One of these was recorded as having severe swelling on the randomisation scan, because the stroke had in fact occurred about 24 h earlier. §Diastolic blood pressure missing for 12 patients in the rt-PA group and seven in the control group. ¶For the first 282 patients, glucose levels were not recorded. After patient 282, glucose levels were measured at randomisation. One further patient had a missing value. ‖Risk predicted by novel model designed by Konig and colleagues.[18] This model predicts outcome (death or Bartel Index <95) at 3 months. If we assume that those who die between 3 months and 6 months were dependent at 3 months, and those who do not die between 3 months and 6 months do not change their dependency status, then the risk estimates are likely to be quite accurate for death or dependency at 6 months. **Stroke clinical syndrome derived from baseline clinical features assigned by an algorithm (algorithm available on request). For the randomisation algorithm TACI, PACI, and POCI were combined as non-lacunar so the process ensured balance in the number of lacunar syndromes in each treatment group. ††Expert panel's masked assessment of prerandomisation scan. This assessment was done by members of the expert panel after randomisation and masked to treatment allocation and all clinical details. Prerandomisation scans were unavailable for eight patients in the rt-PA group and ten in the control group.

Table 1: **Baseline characteristics**

each of these events are provided in the published protocol[13] and statistical analysis plan[14]). Other fatal and non-fatal non-cerebral events were also recorded and coded. Data on potential reports of any of these events were extracted from the trial database and presented to the adjudication committee who were masked to treatment allocation.

Randomisation and masking

To avoid predictable alternation of treatment allocation, and thus potential loss of allocation concealment, patients were allocated with a probability of 0·80 to the treatment group that would minimise the difference between the groups on the key prognostic factors. Additional details of the procedures used in the double-blinded phase of the study are reported elsewhere.[11] The randomisation system informed local clinicians of the patients' unique trial identification number, and the weight-adjusted dose of drug or placebo in the double-blinded phase, or of the weight-adjusted drug dose among those allocated thrombolysis in the open phase, to be given as a 10% bolus with the remainder by an infusion over 1 h.

With the exception of the 276 patients treated in the double-blinded phase of the trial, treatment was given openly and neither the patient nor the treating clinicians were masked. Hospital staff completed an early outcome form at 7 days, death, or hospital discharge, whichever occured first, recording details of events occurring in hospital within 7 days, details of background treatments given and functional status. 6 months after randomisation, general practitioners (or hospital coordinators) were contacted by the IST-3 trial office staff to check that the patient was alive and inform them that they might be approached for follow-up. If appropriate, the IST-3 trial office masked staff then mailed a postal questionnaire to patients to assess outcome. Non-responders were contacted by telephone, and follow-up data was obtained by telephone interview. In Italy and Austria, all follow-ups were done as telephone interviews by a clinician, who was masked to treatment allocation and was highly experienced in outcome assessment. In Portugal, patients were followed up in clinic by clinicians not involved in the patients' initial treatment, again, masked to treatment allocation as far as possible. To assess the durability of any treatment benefit beyond 6 months, patients recruited in the UK (and in other countries where appropriate funding had been obtained) were also followed up at 18 months. All follow-up done by patient contact for these analyses ceased on March 31, 2012, but recording of deaths from national registries of deaths continues in UK, Norway, and Sweden.

Statistical analysis

At the outset of the trial in 2000, we estimated that, among the type of patients likely to be recruited at the time, to detect both an absolute difference of 10% in the proportion of patients alive and independent at 6 months after treatment and to have sufficient power to permit reliable analyses of the prespecified subgroups, a sample of 6000 patients would be needed. A trial of that size could detect a clinically worthwhile net benefit of as little as 3% absolute difference in the primary outcome (80% power, α=0·05). However, it was clear by 2007 that obtaining a sample of 6000 was no longer feasible, and the Steering Committee agreed a revised recruitment target.[11] The sample size, re-estimated in 2007 on the basis of event rates in both treatment groups combined, was 3100. This sample size gave 80% power to detect an absolute difference of 4·7% in the primary outcome.[11]

We monitored the quality and integrity of the accumulating clinical data according to a protocol agreed with the study sponsors, which involved central statistical monitoring according to the principles described by Buyse and colleagues,[28] supplemented by onsite monitoring and detailed source data verification in a random sample of 10% of records in centres that had recruited more than 30 patients, or when patterns in the data at a centre seemed anomalous. All IST-3 monitoring procedures were compliant with requirements of all study sponsors, the national ethics committees and regulatory agencies in the 12 participating countries, and they met all appropriate regulatory and Good Clinical Practice requirements. All baseline data, 7-day, and 6-month outcome data were subject to verification checks built into the randomisation and data management system. We monitored all baseline and postrandomisation imaging, which provided additional cross-checks on recruited patients and centre performance. An expert radiologist checked all scans, masked to clinical details and treatment allocation, immediately on receipt at the trial office, for evidence of adverse events and protocol deviations. The independent data monitoring committee met at least annually to review the unmasked data on major outcome events in the trial, on the background stroke-unit care received by trial patients (to ensure it was equal in both treatment groups), relevant external data (including updates of the Cochrane systematic review and reports from large-scale registries of rt-PA use) in strict confidence throughout the course of the trial. The committee judged these data never met the protocol-specified criteria to recommend modification of the protocol or halt recruitment to the study.

The statistical analysis plan was published[14] before unmasking of the authors to the data. All randomly assigned patients were included in the analysis. Masked analysis of the patients' baseline characteristics showed clear differences in key prognostic factors (age, stroke severity, degree of ischaemic change on baseline CT or MRI) in patients randomly assigned at different times after stroke onset, which might complicate the estimation of the effect of treatment overall and in subgroups.[11] Therefore, the primary analysis of the effect of treatment on the primary outcome was adjusted by logistic regression

for linear effects for the following covariates: age; National Institutes of Health stroke scale (NIHSS) score; time from onset of stroke symptoms to randomisation; and presence (*vs* absence) of ischaemic change on the prerandomisation brain scan according to expert assessment. An unadjusted analysis is also presented.

The trial did not meet its original target of 6000 patients, and so was no longer adequately powered to detect a 3% absolute difference in the primary outcome (with 80% power and α=0·05). The statistical-analysis-plan writing committee, which did not have access to the accumulating data, was therefore expanded to include an independent statistician (Gordon Murray, University of Edinburgh, Edinburgh, UK) to advise on the correct approach. The writing group was persuaded by the recent empirical evidence that the ordinal method was both statistically more efficient (effectively reducing the sample size required in stroke trials[29]) and robust against substantial deviations from the proportional assumption.[30] We therefore specified in the statistical analysis plan an ordinal logistic regression analysis, as a secondary outcome, in which the OHS as a dependent variable had 5 levels: levels 4, 5, and 6 were combined into a single level and levels 0, 1, 2, 3 were retained as distinct.

In this model the treatment odds ratios between one level and the next were assumed to be constant, so a single parameter summarises the shift in outcome distribution between treatment and control groups. For patients known to be alive at 6 months, but with an unknown OHS, we used the level of function recorded on the 7-day form (ie, measured at 7 days or before discharge from hospital) to impute 6-month functional status.[14] We chose this simple form of imputation because it effectively classified 6-month outcomes in patients for whom both 7-day and 6-month data were known (data not shown). Analyses were done with SAS (version 9.2).

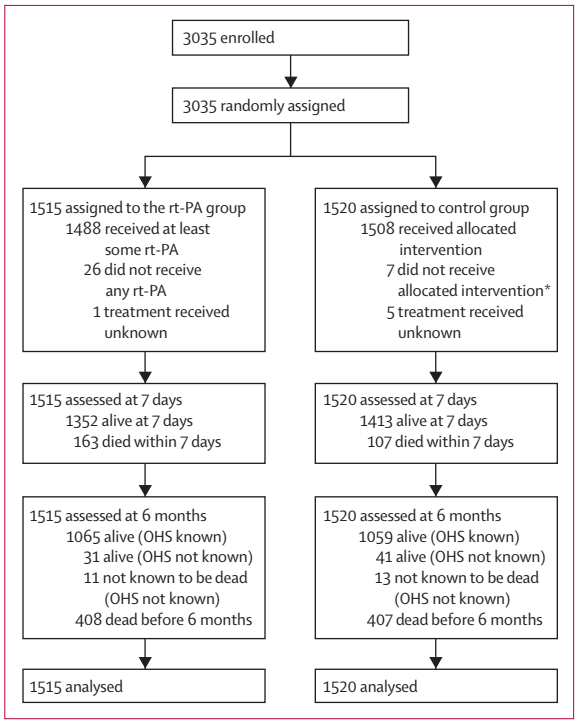


Figure 1: **Trial profile**
rt-PA=recombinant tissue plasminogen activator. OHS=Oxford Handicap Score. *Of the patients allocated to control, seven actually received some rt-PA. Appendix pp 4–5 gives more detail of treatment actually received and background care.

Role of the funding source

The sponsors of the study had no role in design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between May, 2000, and July, 2011, 3035 patients were enrolled in 156 centres in 12 countries. Baseline characteristics were well balanced between treatment groups (figure 1, table 1). 1617 (53%) patients were older than 80 years of age. Vital status at 6 months was known for 99% (3011 of 3035) of patients. Overall, 2581 (95%) of 2714 patients with data (data for some relevant variables were not collected in the initial phase) did not meet the prevailing EU-licence-approval criteria. Additional baseline characteristics are shown in appendix pp 2–3.

Of those assigned to the rt-PA group, 26 (2%) did not receive any rt-PA treatment, and of those assigned to the control group, seven (<1%) received at least some rt-PA. Among patients allocated to the rt-PA group, the mean time from randomisation to injection of the bolus was 18 min, the mean time from onset to treatment was 4·2 h (SD 1·2), median 4·2 h (IQR 3·2–5·2). Appendix pp 2–3 documents deviations from the protocol and the background treatments that were given during the first 7 days. Most patients were cared for in a stroke unit, and there was no evidence of a major imbalance in the use of background treatments or place of care (admissions ward, or stroke unit) over the first 7 days; an analysis of blood pressure in patients measured after randomisation showed no significant difference at each timepoint over the first 24 h in either systolic or diastolic blood pressures between the two treatment groups. However, the proportion of those who had spent at least 1 day in a high-dependency area was somewhat higher among patients assigned to the rt-PA group than in the control group (328 [24%] *vs* 237 [17%]), though in both groups, the median stay in such an area was just 1 day. 76 (49%) centres were classed as experienced in treating stroke with thrombolysis, and 1143 patients were recruited by these centres.

Patients recruited within 1–2 h of onset were significantly more likely to have a more severe neurological deficit did than those recruited at later timepoints after onset (test for linear trend p<0·0001). Similarly, patients

	rt-PA (n=1515)	Control (n=1520)	Adjusted analysis*		Unadjusted analysis†		
			Odds ratio (95% CI)	p value	Odds ratio (95% CI)	p value	Absolute difference per 1000 (95% CI)‡
Died within 7 days	163 (11%)	107 (7%)	1·60 (1·22 to 2·08)	0·001	1·59 (1·23 to 2·07)	0·0004	37 (17 to 57)
Died between 7 days and 6 months	245 (16%)	300 (20%)	0·73 (0·59 to 0·89)	0·002	0·78 (0·65 to 0·95)	0·011	–36 (–63 to –8)
Status at 6 months							
Vital status unknown, disability imputed	11	13	..	..	..	..	..
Alive at 6 months, disability imputed	31	41	..	..	..	..	..
Known 6 month vital and disability status	1473	1466	..	..	..	..	..
Number included in analysis (status known or imputed)	1515	1520	..	..	..	..	..
OHS at 6 months§							
0	138 (9%)	116 (8%)	..	..	..	..	..
1	225 (15%)	204 (13%)	..	..	..	..	..
2	191 (13%)	214 (14%)	..	..	..	..	..
3	235 (16%)	193 (13%)	..	..	..	..	..
4	115 (8%)	140 (9%)	..	..	..	..	..
5	203 (13%)	246 (16%)	..	..	..	..	..
Died before 6 months	408 (27%)	407 (27%)	0·96 (0·80 to 1·15)	0·672	1·01 (0·86 to1·19)	0·924	2 (–30 to 33)
Alive and favourable outcome (0+1)	363 (24%)	320 (21%)	1·26 (1·04 to 1·53)	0·018	1·18 (0·99 to 1·41)	0·055	29 (–1 to 59)
Alive and independent (0+1+2)¶	554 (37%)	534 (35%)	1·13 (0·95 to 1·35)	0·181	1·06 (0·92 to1·24)	0·409	14 (–20 to 48)

Data are number (%) unless otherwise stated. rt-PA=recombinant tissue plasminogen activator. OHS=Oxford Handicap Scale. *Odds ratios and p values were calculated by logistic regression after adjusting for age (linear), National Institutes of Health Stroke Scale (linear), time (linear), and presence or absence of visible acute ischaemic change on baseline scan as judged by the expert reader. †p value calculated from test of difference between percentages for rt-PA and control, using normal approximation. ‡Absolute difference calculated as rt-PA – control, so a positive number indicates this outcome was more frequent in the treatment group. §OHS: 0, no symptoms at all; 1, symptoms, but these do not interfere with everyday life; 2, symptoms that have caused some changes in lifestyle but patients are still able to look after themselves; 3, symptoms that have significantly changed lifestyle and patients need some help looking after themselves; 4, severe symptoms requiring help from other people but not so bad as to need attention day and night; 5, severe handicap needing constant attention day and night. ¶Primary outcomes.

Table 2: **Deaths by 6 months and functional outcome at 6 months**

recruited at earlier time points were significantly older than those recruited later (test for linear trend p<0·0001). The proportion of patients with a definitely visible ischaemic lesion (*vs* only possible or no early ischaemic change) on baseline imaging rose with time (test for linear trend p=0·0045).

At 6 months, 554 (37%) in the rt-PA group versus 534 (35%) in the control group were alive and independent in activities of daily living (OHS 0–2; table 2). A secondary ordinal analysis provided evidence of a favourable shift in the distribution of OHS scores at 6 months with treatment (p<0·001; figure 2). More patients died within 7 days in the rt-PA group than in the control group, but between 7 days and 6 months there were correspondingly fewer deaths in the rt-PA group.

Symptomatic intracranial haemorrhage and fatal or non-fatal deterioration due to swelling of the infarct within 7 days occurred in more patients in the rt-PA group than in the control group (table 3). rt-PA was associated with a significant increase in extracranial haemorrhages (table 3).

To assess the effect of treatment on the primary outcome, the statistical analysis plan predefined a small subset of key prognostic subgroups (figure 3). The

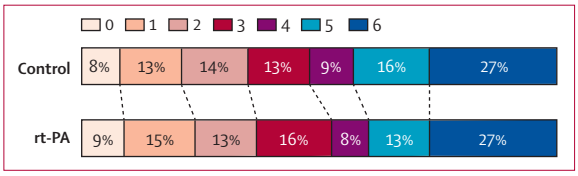


Figure 2: **Outcome at 6 months: Oxford Handicap Scale (OHS) by treatment group**
For the ordinal analysis, which was adjusted for age, National Institutes of Health Stroke Scale (NIHSS), delay (all linear), and and presence or absence of visible acute ischaemic change on baseline scan as judged by the expert reader, the statistical analysis plan prespecified that OHS levels 4, 5, and 6 were grouped and 0, 1, 2, 3 remained discrete. In that analysis, the common odds ratio was 1·27 (95% CI 1·10–1·47; p=0·001). An ordinal analysis with OHS levels 0, 1, 2, 3, 4, 5, and 6 all discrete, adjusted in the same way, gave an odds ratio of 1·17 (95% CI 1·03–1·33; p=0·016). rt-PA=recombinant tissue plasminogen activator.

subgroup analyses are of the adjusted effects and take account of the fact that, for a specific prognostic factor, the distribution of other factors might differ between subcategories. For example, in older patients the time to randomisation was shorter. The subgroup analyses for a specific factor provide estimated effects within subcategories that adjust for such imbalances. Overall, little variation occurred in the adjusted effects of treatment in different subgroups. However, a significant difference

	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)
Symptomatic intracranial haemorrhage¶					
Non-fatal	49 (3%)	9 (1%)	5·56 (2·72 to 11·4)	<0·0001	26 (17 to 36)
Fatal	55 (4%)	7 (<1%)	8·12 (3·68 to 17·9)	<0·0001	32 (22 to 42)
Total	104 (7%)	16 (1%)	6·94 (4·07 to 11·8)	<0·0001	58 (44 to 72)
Neurological deterioration not due to swelling or haemorrhage					
Non-fatal	107 (7%)	79 (5%)	1·37 (1·02 to 1·86)	0·038	19 (2 to 36)
Fatal	38 (3%)	49 (3%)	0·74 (0·48 to 1·14)	0·167	−7 (−19 to 5)
Total	145 (10%)	128 (8%)	1·14 (0·88 to 1·46)	0·320	11 (−9 to 32)
Recurrent ischaemic stroke					
Non-fatal	18 (1%)	15 (1%)	1·21 (0·61 to 2·42)	0·583	2 (−5 to 9)
Fatal	3 (0%)	5 (<1%)	0·61 (0·14 to 2·57)	0·499	−1 (−5 to 2)
Total	21 (1%)	20 (1%)	1·06 (0·57 to 1·97)	0·846	1 (−8 to 9)
Recurrent stroke of unknown type					
Non-fatal	1 (<1%)	2 (<1%)	0·50 (0·05 to 5·56)	0·574	−1 (−3 to 2)
Fatal	2 (<1%)	1 (<1%)	1·98 (0·18 to 22·3)	0·581	1 (−2 to 3)
Total	3 (<1%)	3 (<1%)	0·98 (0·20 to 4·89)	0·981	0 (−3 to 3)
Non-cerebral events‖					
Myocardial infarction					
Non-fatal	18 (1%)	19 (1%)	0·89 (0·46 to 1·71)	0·717	−1 (−8 to 7)
Fatal	5 (<1%)	4 (<1%)	1·25 (0·33 to 4·68)	0·738	1 (−3 to 5)
Total	23 (2%)	23 (2%)	0·95 (0·53 to 1·71)	0·859	0 (−9 to 9)
Extracranial bleed					
Non-fatal	14 (1%)	1 (<1%)	14·5 (1·90 to 110)	0·010	9 (4 to 14)
Fatal	2 (<1%)	2 (<1%)	0·99 (0·14 to 7·13)	0·995	0 (−3 to 3)
Total	16 (1%)	3 (<1%)	5·46 (1·59 to 18·8)	0·007	9 (3 to 14)
Allergic reaction					
Non-fatal	12 (1%)	0 (0%)	..	..	8 (3 to 12)
Fatal	0 (0%)	0 (0%)	..	..	0 (0 to 0)
Total	12 (1%)	0 (0%)	..	..	8 (3 to 12)
Total deaths from cerebral causes within 7 days	145 (10%)	87 (6%)	1·76 (1·32 to 2·34)	0·0001	38 (20 to 57)
Total deaths from non-cerebral causes within 7 days**	18 (1%)	20 (1%)	0·89 (0·47 to 1·69)	0·717	−1 (−9 to 7)
Total deaths within 7 days	163 (11%)	107 (7%)	1·60 (1·22 to 2·08)	0·001	37 (17 to 57)

Data are number (%) unless otherwise stated. rt-PA=recombinant tissue plasminogen activator. *One patient in the control group was missing a 7-day form but did return a 6-month form, so was known to be alive at 7 days. This case has been omitted from the analysis. †Odds ratio and p value calculated from logistic regression after adjusting for age (linear), National Institutes of Health Stroke Scale (linear), time (linear), and presence or absence of visible acute ischaemic change on baseline scan. When no events occurred in one treatment group the logistic model was not applied. ‡Absolute difference was calculated as rt-PA–control, so a positive number indicates this outcome was more frequent in the treatment group. §Symptomatic swelling of the original infarct was defined as significant neurological deterioration accompanied by evidence of significant brain swelling as determined by the independent masked expert assessment of the scan defined as: shift of the midline away from the side of the ventricle or effacement of the basal cisterns or uncal herniation on a postrandomisation scan (or autopsy if not rescanned before death). The presence of some degree of haemorrhagic transformation was permitted, provided it was not identified by the expert CT reader to be a major contributor to the mass effect. ¶Symptomatic intracranial haemorrhage was defined as significant neurological deterioration accompanied by clear evidence of significant intracranial haemorrhage on the postrandomisation scan (or autopsy if not rescanned and death occurs after 7 days). Significant haemorrhage was present on any postrandomisation scan if the expert reader both noted the presence of significant haemorrhagic transformation of the infarct or parenchymal haematoma and indicated that haemorrhage was a major component of the lesion (or was remote from the lesion and likely to have contributed significantly to the burden of brain damage). This event included clinical events described as a recurrent stroke within 7 days, in which the recurrent stroke was confirmed to be caused by an intracranial haemorrhage. ‖Non-fatal cerebral events are exclusive. However, non-fatal non-cerebral events are not exclusive. A given patient could have one or more non-fatal non-cerebral events and a non-fatal cerebral event. **The deaths in the fatal rows are exclusive (a patient can only contribute to one of the fatal rows). Total deaths from non-cerebral causes include deaths not attributed to myocardial infarction, extracranial bleed, or allergic reaction.

Table 3: **Fatal and non-fatal cerebral and non-cerebral events within 7 days of randomisation**

Subgroup	Events/number of patients		Adjusted odds ratio (99% CI)	Adjusted p value
	rt-PA	Control		
Age (years)				0·029
≤80	331/698 (47·4%)	346/719 (48·1%)	0·92 (0·67–1·26)	
>80	223/817 (27·3%)	188/799 (23·5%)	1·35 (0·97–1·88)	
NIHSS score				0·003
0–5	221/304 (72·7%)	232/308 (75·3%)	0·85 (0·52–1·38)	
6–14	276/728 (37·9%)	268/724 (37·0%)	1·08 (0·81–1·45)	
15–24	50/402 (12·4%)	33/421 (7·8%)	1·73 (0·93–3·20)	
≥25	7/81 (8·6%)	1/65 (1·5%)	7·43 (0·43–129·00)	
Predicted probability of poor outcome at 6 months				0·009
<0·4	256/351 (72·9%)	290/377 (76·9%)	0·81 (0·52–1·26)	
0·4–0·5	88/169 (52·1%)	76/160 (47·5%)	1·20 (0·68–2·13)	
0·5–0·75	127/361 (35·2%)	118/357 (33·1%)	1·10 (0·73–1·65)	
>0·75	83/634 (13·1%)	50/624 (8·0%)	1·73 (1·07–2·82)	
Time to randomisation (h)				0·613
0–3	132/431 (30·6%)	95/418 (22·7%)	1·64 (1·03–2·62)	
3–4·5	182/577 (31·5%)	226/600 (37·7%)	0·73 (0·50–1·07)	
>4·5	240/507 (47·3%)	213/500 (42·6%)	1·31 (0·89–1·93)	
Acute ischaemic change on randomisation scan according to expert panel				0·534
No	392/883 (44·4%)	379/910 (41·6%)	1·17 (0·88–1·56)	
Yes	158/624 (25·3%)	149/598 (24·9%)	1·05 (0·70–1·59)	
Sex				0·409
Female	239/782 (30·6%)	235/787 (29·9%)	1·21 (0·86–1·69)	
Male	315/733 (43·0%)	299/731 (40·9%)	1·04 (0·75–1·43)	
Stroke syndrome				0·465
TACI	106/639 (16·6%)	96/665 (14·4%)	1·36 (0·89–2·08)	
PACI	281/596 (47·1%)	254/550 (46·2%)	1·07 (0·76–1·51)	
LACI	100/168 (59·5%)	103/164 (62·8%)	0·91 (0·48–1·72)	
POCI	66/110 (60·0%)	79/136 (58·1%)	1·04 (0·49–2·22)	
Clinician's assessment of recent ischaemic change at randomisation				0·703
No evidence	381/894 (42·6%)	366/897 (40·8%)	1·13 (0·84–1·51)	
Possible evidence	105/361 (29·1%)	108/340 (31·8%)	0·92 (0·56–1·51)	
Definite evidence	68/260 (26·2%)	60/281 (21·4%)	1·39 (0·74–2·61)	
Atrial fibrillation				0·574
No	440/1042 (42·2%)	436/1078 (40·4%)	1·09 (0·83–1·43)	
Yes	114/473 (24·1%)	98/440 (22·3%)	1·20 (0·76–1·90)	
Systolic blood pressure (mm Hg)				0·737
≤143	172/487 (35·3%)	170/491 (34·6%)	1·18 (0·78–1·78)	
144–164	196/498 (39·4%)	196/518 (37·8%)	1·09 (0·74–1·62)	
≥165	186/530 (35·1%)	168/509 (33·0%)	1·11 (0·74–1·65)	
Diastolic blood pressure (mm Hg)				0·154
≤74	151/462 (32·7%)	133/445 (29·9%)	1·32 (0·86–2·01)	
75–89	204/541 (37·7%)	219/586 (37·4%)	1·08 (0·73–1·58)	
≥90	193/500 (38·6%)	178/480 (37·1%)	0·97 (0·64–1·46)	
Glucose (mmol/L)				0·444
≤5	109/254 (42·9%)	109/285 (38·2%)	1·23 (0·72–2·12)	
6–7	261/664 (39·3%)	242/636 (38·1%)	1·16 (0·82–1·66)	
≥8	143/455 (31·4%)	144/456 (31·6%)	1·03 (0·67–1·60)	
Treatment with antiplatelet drugs in previous 48 h				0·383
No	288/736 (39·1%)	282/725 (38·9%)	1·02 (0·73–1·43)	
Yes	265/775 (34·2%)	251/786 (31·9%)	1·20 (0·87–1·65)	
Trial phase				0·479
Blinded	34/136 (25·0%)	38/140 (27·1%)	0·91 (0·42–1·98)	
Open	520/1379 (37·7%)	496/1378 (36·0%)	1·14 (0·89–1·45)	
Centre with experience of thrombolysis				0·911
No	313/940 (33·3%)	309/950 (32·5%)	1·10 (0·82–1·48)	
Yes	241/575 (41·9%)	225/568 (39·6%)	1·14 (0·78–1·66)	
Total	**554/1515 (36·6%)**	**534/1518 (35·2%)**	**1·12 (0·89–1·41)**	

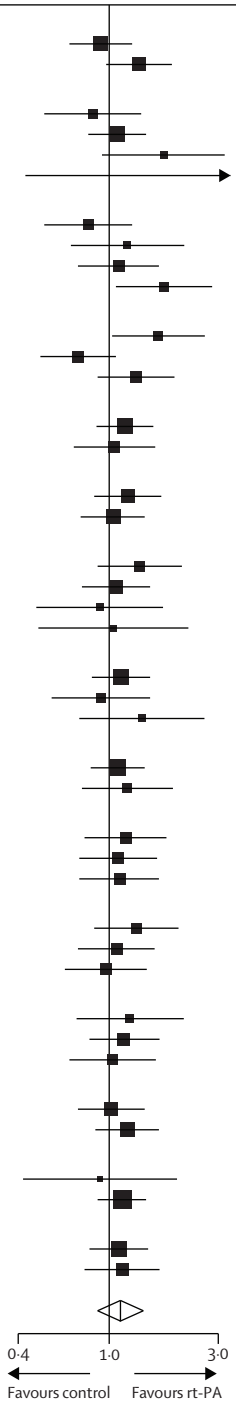


Figure 3: **Adjusted effect of treatment on the primary outcome (alive and independent, Oxford Handicap Score 0, 1, or 2) in subgroups**
The key predefined subgroups were age 80 years or younger, age older than 80 years, time from stroke onset to randomisation (0–3·0 h, 3·0–4·5 h, 4·5–6·0 h), initial stroke severity as measured by National Institutes of Health stroke score, and the appearance of the baseline brain scan on expert read for each subgroup (whether ischaemic change is visible or not). The treatment odds ratio in each subgroup has been adjusted for the linear effects of the other key variables (age, NIHSS, and delay) but not for the presence or absence visible ischaemic change. It is for this reason that the adjusted odds ratio in the "Total" row at the bottom of the table does not exactly agree with the odds ratio in table 2. The choice of cut-points to define certain subgroups is slightly different to those given in table 1.[14] On the graph, for each subgroup, the horizontal line represents the 99% CI, the diamond is centred on the overall estimate and it represents the 95% CI. The graph was generated with R (version 2.11.1). rt-PA=recombinant tissue plasminogen activator. NIHSS=National Institutes of Health Stroke Scale. TACI=total anterior circulation infarct. PACI=partial anterior circulation infarct. LACI=lacunar infarct. POCI=posterior circulation infarct.

did occur in the adjusted effect of treatment between patients older than 80 years and in patients 80 years or younger (p=0·027), suggesting greater benefit in those older than 80 years of age; contrary to expectations.[14] Treatment appeared at least as effective in this age group as in younger patients. Significant trends towards larger effects of treatment in more severe strokes were also seen (as assessed by the NIHSS and by the predicted probability of a poor outcome[18]). Benefit was greatest in patients treated within 3 h, but there was insufficient power to examine decay of benefit with time. An analysis of the treatment effect in each of three equal-sized cohorts of patients (ie, those recruited in 2000–06, 2007–08, 2009–11) did not provide any evidence of period effects (data not shown). We also undertook a sensitivity analysis restricted to the 2939 (96%) patients with known 6-month vital and disability status (appendix pp 4–5), and the results were not qualitatively different from those in table 2.

Discussion

Although the increase in the number of patients treated with rt-PA who were alive and independent at 6 months was smaller than originally anticipated and was not significant, the secondary analysis provides supportive evidence of benefit. The ordinal analysis provided evidence that on average, patients treated with intravenous thrombolysis up to 6 h after stroke survived with less disability. At 6 months, vital status was known for most patients and there was no evidence of any difference in the number of deaths, despite the excess of deaths within 7 days of stroke (mainly due to intracranial haemorrhage). Since mortality at 6 months was equal in the two groups, and in view of the evidence that the lower the patients' degree of disability at 6 months, the greater their subsequent survival,[31] long-term follow-up beyond 6 months is important. Follow-up for survival, therefore, continues in the UK, Norway, and Sweden to assess whether an overall survival advantage from rt-PA after 6 months emerges.

Since we sought to recruit older patients and patients who did not strictly meet prevailing licence criteria for thrombolytic therapy with rt-PA, we anticipated a higher risk of adverse events, chiefly symptomatic intracranial haemorrhage. The patient information leaflet stated that rt-PA treatment might be associated with an increased risk of fatal intracranial haemorrhage of 4%, which indeed was the rate reported in the trial. Furthermore, applying a similar definition of symptomatic intracerebral haemorrhage as in the Cochrane systematic review, the frequency of this disorder within 7 days in IST-3 patients treated with rt-PA (6·8%) was comparable with the 7·3% reported in the Safe Implementation of Thrombolysis in Stroke (SITS) registry of 6483 patients treated within licence in routine clinical practice.[32] We also expected a higher risk of death in the control group, and a smaller proportion alive and independent than in previous trials. Reassuringly, despite the different event rates in the control group, for most of the outcomes, there was no clear evidence that the effects of treatment were qualitatively different in IST-3 to those seen in earlier randomised trials, with two exceptions. We identified significant trends towards larger effects of treatment in patients with more severe strokes. We also anticipated a reduction in fatal and non-fatal neurological deterioration due to swelling of the initial infarct,[6] so the clear 17 per 1000 excess was unexpected, and inconsistent with data from previous trials.[6]

As proposed by Kent and colleagues,[33] we reported the effect of treatment on the primary outcome in several prespecified subgroups and included the effects subdivided by the result of a prognostic score. Benefit with treatment was greatest within 3 h, but the analyses did not have sufficient power to define the shape of the relation between benefit and time beyond 3 h. The effect of treatment in patients older than 80 years of age was at least as large as in patients younger than 80 years of age. A formal test for trend showed a significant difference for greater benefit of rt-PA in patients with increasingly severe strokes. However, in view of the overall non-significant benefit for the primary outcome, the significant interactions across subgroups in these analyses should be interpreted with caution. As specified in the statistical analysis plan, we planned additional secondary analyses to explore these apparent effects on the primary outcome (and on other outcomes, such as symptomatic intracranial haemorrhage) and to decide if these effects were due to chance.

Lyden[34] has identified limitations in these data, chiefly that IST-3 recruited only half the number of patients originally intended and so was underpowered for the primary outcome (and more so for the subgroup analyses). The many changes in the regulatory environment over the course of the trial delayed the approval of the trial in many centres and precluded the participation of several countries and hence was a significant factor in our failing to achieve our original target.[11] Nonetheless, the trial was the largest-ever trial of thrombolysis therapy for stroke[34] (over three times larger than any previous trial) and included more patients treated within 3 h of stroke (n=849) than were included in the National Institute of Neurological Disorders and Stroke (NINDS) trial (n=624), the only previous trial examining specifically treatment within 3 h (panel). The fact that most of the IST-3 patients treated within 3 h were older than 80 years of age (n=726), yet achieved similar benefit to younger patients in NINDS trial, adds to the NINDS trial.

The absence of masking is most relevant for the assessment of the events within 7 days. However, every possible precaution was taken to ensure masking of the expert panel assessing the scans, and the adjudication committee, who also assessed clinical data on all potential cerebral events. The proportional effect of

treatment on fatal and non-fatal events within 7 days was very similar, which perhaps suggest that masking of the assessors was successful. The self-assessment at 6 months by patients or their carer by postal questionnaire or masked telephone interview was unmasked and so could be subject to reporting bias.[34] However, self-reported outcome by patients is necessarily subjective and affected by many things besides knowledge of treatment allocation. The subgroup analysis subdivided by trial phase provides some reassurance in that no significant difference was seen in the effect of treatment on the primary outcome in the double-blind phase and the open phase (figure 3). The measurement of outcome with OHS at 6 months is different from previous trials that measured the modified Rankin score at 3 months. When we planned IST-3 in 1998, the modified Rankin score and OHS were judged to be equivalent. Both are derivatives of the original Rankin scale, developed by members of our group. While the proportion of patients recorded as dependent might be slightly different with each scale, the choice of outcome scale would not bias the assessment of treatment effect between treatment and control groups.

The outcome was recorded at 6 months and 18 months, to assess the effects on survival free of disability after a few months and also in the long term (the longer the benefit persists, the greater the cost-effectiveness). The longer time to follow-up allowed any differential effect of rt-PA on early and late death to become clearer. Outcome (other than survival) was not recorded at 3 months, although the proportional effects on death and disability seen at 6 months in IST-3 are comparable with those seen at 3 months in previous trials.

Lyden also comments that the sampling approach to monitoring in IST-3 was less intense than in many commercial studies, and is a potential concern, but also states: "many clinical trialists believe that source verification of some clinical trial data assures safety, accuracy, and validity of the trial data. Authorities do not agree on the minimum quantity of verified data to assure validity (100%, half, 10% sample)...but there is no evidence to suggest any problems with the [IST-3] data set due to limited monitoring."[34]

When the results of IST-3 are incorporated into an updated systematic review,[35] the estimates of relative treatment effect are broadly compatible with the previous rt-PA trials for each of the main outcomes: alive and independent; death at final follow-up; and fatal intracranial haemorrhage.

Our trial was underpowered to reliably detect important subgroup effects, and so a collaborative individual patient data meta-analysis (the Stroke Thrombolysis Trialists Collaboration [STTC]) has been established, which will include data from all the completed intravenous rt-PA trials and will update the previous pooled analysis.[7] The meta-analysis will explore which baseline factors, other than time, might modify the effects of treatment on major outcomes (such as death, functional outcome, and intracerebral haemorrhage), and so provide better guidance for clinicians and patients to apply this treatment as effectively as possible in routine practice.

For the types of patient recruited in IST-3 (about three quarters of whom were randomised after 3 h, and half of all patients were older than 80 years of age), by 6 months there was evidence that rt-PA improved functional outcome. The data add weight to the policy of treating patients as soon as possible, and also justify extending treatment to patients older than 80 years of age. The data do not support any restriction of treatment on the basis of stroke severity or the presence of early ischaemic change on the baseline brain scan. The data support the need for randomised trials of thrombolysis in selected patients more than 4·5 h after stroke.

Panel: Research in context

Systematic review

To update the published systematic review of randomised-controlled trials of recombinant tissue plasminogen activator (rt-PA) in patients with acute ischaemic stroke and incorporate the third International Stroke Trial (IST-3) results,[6] we searched for additional randomised trials of intravenous rt-PA versus control within 6 h of onset of acute ischaemic stroke up to March 30, 2012, in the Cochrane Stroke Trials Registry (November, 2011), Internet Stroke Trials Centre (March, 2011), Medline and Embase (search strategy available on request), and references lists in review articles and conference abstracts. The primary analysis was for all patients treated up to 6 h after stroke. Data were available for 7012 patients in 12 trials. We tested for heterogeneity between the estimates of effect for key outcomes from two strata: all trials before IST-3 and IST-3. The tests for heterogeneity in the proportional effects of treatment across these two strata were not significant for symptomatic intracranial haemorrhage (χ^2 2·13, p=0·1), deaths within 7 days (χ^2 1·44, p=0·2), deaths by the end of follow-up (χ^2 1·0, p=0·3) and, the proportion alive and independent (modified Rankin score 0–2: χ^2 3·08, p=0·08). Similarly, no heterogeneity occurred across the two strata for patients of all ages treated within 3 h (χ^2 0·25, p=0·6). The review established that the effects of treatment reported in IST-3—in this wider range of patients (generally outside the current approvals)—were consistent with those seen in previous trials.

Interpretation

By providing estimates on the benefits and harms of treating patients with acute ischaemic stroke outside the current approvals, IST-3 enables clinicians to consider thrombolytic treatment for a wider range of patients, especially those older than 80 years of age. The data reinforce the need for further efforts to increase the proportion of all ischaemic strokes treated within 3 h. The additional data from IST-3 give greater confidence that mortality is not increased by treatment. The implications for ongoing research are that the data strengthen the rationale for the ongoing trials of thrombolysis in patients presenting more than 4·5 h after onset of stroke, and suggest that the imposition of upper age limits on future trials in acute stroke will become harder to justify.

Contributors

The study was conceived by the co-chief investigators, PS, RIL, and JMW. JMW led the development of all of the imaging aspects of the study. The study was designed by PS, RIL, and JMW, with input from all the other listed contributors who act as coordinators of the trial in their own country. PS, RIL, JMW, MD, and KI designed the study and wrote the protocol. KI is the study coordinator. GC is the study statistician who prepared the analyses for this paper. PS, RIL, MD, GV, AC, AK, EB, KBS, VM, AP, GJH, KM, MB, SR, GG, SJP, AA, MC, and PL recruited patients

to the study. GV, AC, AK, EB, KBS, VM, AP, GJH, KM, MB, SR, GG, SJP, AA, MC, and PL acted as National Coordinators. PS drafted the Article and all authors commented on drafts and approved the final version.

IST-3 collaborative group

The members of the collaborative group are listed in full in the appendix.

Writing committee

Peter Sandercock (University of Edinburgh, Edinburgh, Scotland), Joanna M Wardlaw (University of Edinburgh, Edinburgh, Scotland), Richard I Lindley (Sydney Medical School – Westmead Hospital and The George Institute for Global Health, University of Sydney, Australia), Martin Dennis (University of Edinburgh, Edinburgh, Scotland), Geoff Cohen, Gordon Murray, Karen Innes (University of Edinburgh, Edinburgh, Scotland), Graham Venables (Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK), Anna Czlonkowska (Institute of Psychiatry and Neurology, Warsaw, Poland, and Medical University of Warsaw, Warsaw, Poland), Adam Kobayashi (Institute of Psychiatry and Neurology, Warsaw, Poland), Stefano Ricci (Department of Neurology ASL1, Ospedale, Citta' di Castello, Italy), Veronica Murray (Karolinska Institutet, Stockholm, Sweden), Eivind Berge (Oslo University Hospital, Oslo, Norway), Karsten Bruins Slot (Oslo University Hospital, Oslo, Norway), Graeme J Hankey (Royal Perth Hospital, Perth, Australia), Manuel Correia (Hospital Geral de Santo Antonio, Porto, Portugal), Andre Peeters (Cliniques Universitaires Saint-Luc, Bruxelles, Belgium), Karl Matz (Landesklinikum Donauregion Tulln, Tulln, Austria), Phillippe Lyrer (University Hospital Basel, Basel, Switzerland), Gord Gubitz (Dalhousie University and Queen Elizabeth II Health Sciences Centre, Halifax, Canada), Stephen J Phillips (Dalhousie University and Queen Elizabeth II Health Sciences Centre, Halifax, Canada), Antonio Arauz (Instituto Nacional de Neurologia, Mexico City, Mexico).

Trial steering committee

Independent chairmen: Colin Baigent (University of Oxford, Oxford UK); David Chadwick (University of Liverpool, Liverpool UK). *Independent member*: Pippa Tyrrell (University of Manchester, Manchester, UK); Gordon Lowe (University of Glasgow, Glasgow, UK). *Co-principal investigators*: PS; RIL. *Chief investigator for neuroradiology*: JMW; MD. *Statistician*: GC. *Trial Co-ordinator*: KI. *Lay representative*: Heather Goodare.

CT and MRI reading panel

JMW, Andrew Farrall, Rüdiger von Kummer, Lesley Cala, Anders von Heijne, Zoe Morris, Alessandro Adami, AP, Gillian Potter, Nick Brady.

Data monitoring committee

Rory Collins (Oxford University, Oxford, UK; Chairman), Philip Bath (Nottingham University, Nottingham, UK), Jan van Gijn (University of Utrecht, Utrecht, Netherlands), Richard Gray (University of Oxford, Oxford, UK), Robert Hart (McMaster University, ON, Canada), Salim Yusuf (McMaster University, ON, Canada).

Event adjudication committee

Keith Muir (Institute of Neurological Science, University of Glasgow, Glasgow, UK), PS, RIL.

National coordinators and associate national coordinators

Australia: RIL, GJH. *Austria:* KM, Michael Brainin. *Belgium*: AP. *Canada*: GG, SJP. *Italy*: SR. *Mexico*: AA. *Norway*: EB, KBS. *Poland*: AC, AK. *Portugal*: MC. *Switzerland*: PL; Stefan Engelter. *Sweden*: VM, Andreas Terent, Bo Norrving, Per Wester: UK: GV.

Trial coordinating centres

Division of Clinical Neuroscience, University of Edinburgh, Edinburgh, Scotland: KI, Alison Clark, David Perry, Vera Soosay, David Buchanan, Sheila Grant, Eleni Sakka, Jonathan Drever, Pauli Walker, Indee Herath, Ann Leigh Brown, Paul Chmielnik, Christopher Armit, Andrea Walton, Mischa Hautvast, Steff Lewis, Graeme Heron, Sylvia Odusanya, Pam Linksted, Ingrid Kane, Will Whiteley, Robin Sellar, Philip White, Peter Keston, Andrew Farrell, Zoe Morris, Hector Miranda. *Clinical Trials Service Unit, Oxford, UK*: Lisa Blackwell.

National coordinating centres

Italy (up to Sept, 2008): Maria Grazia Celani; Enrico Righetti. Italy (after Sept, 2008): Silvia Cenciarelli; Tatiana Mazzoli. *Central follow-up for Italy*: Teresa Anna Cantisani. *Poland:* Jan Bembenek. *Sweden:* Eva Isaakson. *Norway*: EB, KBS. *Australia*: Genevieve Freys. The list of participating hospitals in each country is in the appendix.

Conflicts of interest

EB has received honoraria for lectures at meetings arranged by Boehringer Ingelheim, and reimbursement for costs for attending these meetings. AC has received lecture fees and conference travel costs from Boehringer Ingelheim. GB has received honoraria and speaker fees from Boehringer Ingelheim, Sanofi Synthlabo Aventis, Hoffman La Roche, and Novo Nordisk. AK has received lecture fees and conference travel costs from Boehringer Ingelheim. RIL has received payment in his role as conference scientific committee member and for occasional lectures from Boehringer Ingelheim; has attended national stroke meetings organised and funded by Boehringer Ingelheim; and is not a member of any industry advisory boards. PS has received lecture fees (paid to the Division of Clinical Neurosciences, University of Edinburgh) and travel expenses from Boehringer Ingelheim for occasional lectures given at international conferences; and was a member of the Independent Data and Safety Monitoring Board (DSMB) of the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) trial funded by Boehringer Ingelheim and received attendance fees and travel expenses for attending DSMB meetings (paid to the Division of Clinical Neurosciences, University of Edinburgh). KBS has received an honorarium for a lecture from Boehringer Ingelheim and had costs for participating in scientific meetings reimbursed; is a member of the European Medicines Agency's Committee for Medicinal Products for Human Use (CHMP) and the Cardiovascular Working Party. The views expressed in this article are the personal views of KBS and should not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties. VM has received an unrestricted educational grant for a meeting on thrombolysis in stroke at which IST-3 was discussed. JMW received reimbursement for reading CT scans for European Cooperative Acute Stroke Study III (ECASS III) from Boehringer Ingelheim in the form of funding to her department, the Division of Clinical Neurosciences, University of Edinburgh; is the contact reviewer for the Cochrane systematic reviews of thrombolytic treatment for acute stroke; has attended meetings held by Boehringer Ingelheim as an unpaid independent external adviser during the licensing of rt-PA, but was refunded her travel expenses and the time away from work; has attended and spoken at national and international stroke meetings organised and funded by Boehringer Ingelheim for which she received honoraria and travel expenses; and is director of the Brain Research Imaging Centre for Scotland, which is located within the Department of Clinical Neurosciences at the University of Edinburgh, Edinburgh, Scotland and houses a research MRI scanner, which was funded by the UK Research Councils Joint Research Equipment Initiative, supplemented by grants and donations from various other sources including Novartis, Schering, General Electric, and Boehringer Ingelheim. These commercial sources contributed to the purchase of the scanner, but not the running costs or any individual studies. All other members of the writing committee declare that they have no conflicts of interest.

Acknowledgments

The IST-3 collaborative group thanks all the patients who participated in the study, and the many individuals not specifically mentioned in the paper who have supported the study. IST-3 is an investigator led trial. The University of Edinburgh and the Lothian Health Board are cosponsors. The start-up phase was supported by a grant from the Stroke Association, UK. The expansion phase was funded by The Health Foundation UK. The main phase of the trial is funded by the following organisations: UK MRC (grant numbers G0400069 and EME 09-800-15) and managed by NIHR on behalf of the MRC-NIHR partnership; The Research Council of Norway; AFA Insurances (Sweden); the Swedish Heart Lung Fund; The Foundation of Marianne and Marcus Wallenberg; Stockholm County Council and Karolinska Institute Joint ALF-project grants (Sweden); the Government of Poland (grant number 2PO5B10928); the Australian Heart Foundation (grant number G 04S 1638); Australian NHMRC (grant number 457343); the Swiss National Research Foundation; the Swiss Heart Foundation; Foundation for health and cardio-/neurovascular research, Basel, Switzerland; the Assessorato alla Sanita, Regione

dell'Umbria; Danube University, Krems, Austria. Drug and placebo for the 300 patients in the double-blind component of the start-up phase were supplied by Boehringer Ingelheim. We thank the NIHR Stroke Research Network, NHS Research Scotland, through the Scottish Stroke Research Network, and the National Institute for Social Care and Health Research Clinical Research Centre for their support. The imaging work was undertaken at the Brain Imaging Research Centre, a member of the SINAPSE collaboration, at the Division of Clinical Neurosciences, University of Edinburgh. SINAPSE is funded by the Scottish Funding Council and the Chief Scientist Office of the Scottish Executive. Additional support was received from Chest Heart and Stroke Scotland, DesAcc, University of Edinburgh, Danderyd Hospital R&D Department, Karolinska Institutet, Oslo University Hospital, and the Dalhousie University Internal Medicine Research Fund. This report presents independent research supported by the NIHR through the UK Stroke Research Network. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

References

1 Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. *Lancet Neurol* 2007; **6:** 182–87.

2 Rothwell PM, Coull AJ, Giles MF, et al, and the Oxford Vascular Study. Change in stroke incidence, mortality, case-fatality, severity, and risk factors in Oxfordshire, UK from 1981 to 2004 (Oxford Vascular Study). *Lancet* 2004; **363:** 1925–33.

3 Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol* 2009; **8:** 355–69.

4 Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; **367:** 1747–57.

5 Islam MS, Anderson CS, Hankey GJ, et al. Trends in incidence and outcome of stroke in Perth, Western Australia during 1989 to 2001: the Perth Community Stroke Study. *Stroke* 2008; **39:** 776–82.

6 Wardlaw JM, Murray V, Berge E, Del Zoppo GJ. Thrombolysis for acute ischaemic stroke. *Cochrane Database Syst Rev* 2009; **4:** CD000213.

7 Lees KR, Bluhmki E, von Kummer R, et al, and the ECASS, ATLANTIS, NINDS and EPITHET rt-PA Study Group. Time to treatment with intravenous alteplase and outcome in stroke: an updated pooled analysis of ECASS, ATLANTIS, NINDS, and EPITHET trials. *Lancet* 2010; **375:** 1695–703.

8 Sanossian N, Ovbiagele B. Prevention and management of stroke in very elderly patients. *Lancet Neurol* 2009; **8:** 1031–41.

9 Bembenek J, Kobayashi A, Sandercock P, Czlonkowska A. How many patients might receive thrombolytic therapy in the light of the ECASS-3 and IST-3 data? *Int J Stroke* 2010; **5:** 430–31.

10 Ware JH, Hamel MB. Pragmatic trials—guides to better patient care? *N Engl J Med* 2011; **364:** 1685–87.

11 Sandercock P, Lindley R, Wardlaw J, et al, 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.

12 Whiteley W, Lindley R, Wardlaw J, Sandercock P, and the IST-3 Collaborative Group. Third international stroke trial. *Int J Stroke* 2006; **1:** 172–76.

13 Sandercock P, Lindley R, Wardlaw J, et al. Third international stroke trial (IST-3) of thrombolysis for acute ischaemic stroke. *Trials* 2008; **9:** 37.

14 Sandercock P, Lindley R, Wardlaw J, Whiteley W, Murray G, and the ST-3 collaborative group. Statistical analysis plan for the third International Stroke Trial (IST-3); part of a 'thread' of reports of the trial. *Int J Stroke* 2012; **7:** 186–87.

15 Warlow C, D'Hooghe M, Joos F, et al. MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70–99%) or with mild (0–29%) carotid stenosis. *Lancet* 1991; **337:** 1235–43.

16 Peto R, Baigent C. Trials: the next 50 years. Large scale randomised evidence of moderate benefits. *BMJ* 1998; **317:** 1170–71.

17 Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *Lancet* 2001; **357:** 373–80.

18 König IR, Ziegler A, Bluhmki E, et al, and the Virtual International Stroke Trials Archive (VISTA) Investigators. Predicting long-term outcome after acute ischemic stroke: a simple index works in patients from controlled clinical trials. *Stroke* 2008; **39:** 1821–26.

19 Langhorne P, Pollock A, and the Stroke Unit Trialists' Collaboration. What are the components of effective stroke unit care? *Age Ageing* 2002; **31:** 365–71.

20 Barber PA, Demchuk AM, Zhang J, Buchan AM. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. *Lancet* 2000; **355:** 1670–74.

21 Wardlaw JM, Sellar RJ. A simple practical classification of cerebral infarcts on CT and its interobserver reliability. *AJNR Am J Neuroradiol* 1994; **15:** 1933–39.

22 Wardlaw JM, Farrall AJ, Perry D, et al, and the Acute Cerebral CT Evaluation of Stroke Study (ACCESS) Study Group. Factors influencing the detection of early CT signs of cerebral ischemia: an internet-based, international multiobserver study. *Stroke* 2007; **38:** 1250–56.

23 van Swieten JC, Hijdra A, Koudstaal PJ, van Gijn J. Grading white matter lesions on CT and MRI: a simple scale. *J Neurol Neurosurg Psychiatry* 1990; **53:** 1080–83.

24 Cordonnier C, Potter GM, Jackson CA, et al. Improving inter-rater agreement about brain microbleeds. Development of the Brain Observer MicroBleed Scales (BOMBS). *Stroke* 2009; **49:** 94–99.

25 Farrell C, Chappell F, Armitage PA, et al. Development and initial testing of normal reference MR images for the brain at ages 65–70 and 75–80 years. *Eur Radiol* 2009; **19:** 177–83.

26 Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. A prospective study of acute cerebrovascular disease in the community: the Oxfordshire Community Stroke Project—1981–86. 2. Incidence, case fatality rates and overall outcome at one year of cerebral infarction, primary intracerebral and subarachnoid haemorrhage. *J Neurol Neurosurg Psychiatry* 1990; **53:** 16–22.

27 van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; **19:** 604–07.

28 Buyse M, George SL, Evans S, et al. The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Stat Med* 1999; **18:** 3435–51.

29 Bath PM, Gray LJ, Collier T, Pocock S, Carpenter J, and the 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:** 1911–15.

30 McHugh GS, Butcher I, Steyerberg EW, et al. A simulation study evaluating approaches to the analysis of ordinal outcome data in randomized controlled trials in traumatic brain injury: results from the IMPACT Project. *Clin Trials* 2010; **7:** 44–57.

31 Slot KB, Berge E, Dorman P, Lewis S, Dennis M, 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:** 376–79.

32 Wahlgren N, Ahmed N, Dávalos A, et al, and the 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.

33 Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials* 2010; **11:** 85.

34 Lyden PD. In Anticipation of International Stroke Trial-3 (IST-3). *Stroke* 2012; published online May 3. DOI:10.1161/STROKEAHA.112.656876.

35 Wardlaw J M, Murray V, Berge E, et al. Recombinant tissue plasminogen activator for acute ischaemic stroke: an updated systematic review and meta-analysis. *Lancet* 2012; published online May 23. DOI:10.1016/S0140-6736(12)60738-7.

THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: 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; published online May 23. DOI:10.1016/S0140-6736(12)60768-5.

Web appendix contents:

Web table 1 Additional baseline data

	rt-PA Con		trol	
	No. (%	) N	o.	(%)
Number randomised	1515		1520	
Baseline variables collected before treatment allocation[1]				
Clinician's assessment of pre-randomisation scan				
No evidence of recent ischaemic change	894	(59%)	898	(59%)
Possible evidence of recent ischaemic change	361	(24%)	340	(22%)
Definite evidence of recent ischaemic change	260	(17%)	282	(19%)
Baseline variables collected from pre-randomisation scan				
Lesion territory				
MCA or ACA or Borderzone	589	(39%)	555	(37%)
Posterior	22 (	1%) 3	6	(2%)
Lacunar	11 (1%)		5	(<1%)
Indeterminate[3] 8	85	(59%)	914	(61%)
Lesion size				
None	885 (59%	)	914	(61%)
Small	110 (	7%)	97	(6%)
Medium	250 (17%	)	250	(17%)
Large	124 (8%)		137	(9%)
Very large	138 (9%)		112	(7%)
Depth of tissue damage				
None	892 (59%	)	922	(61%)
Mild	503 (33%	)	492	(33%)
Severe	112 (	7%)	96	(6%)
Degree of swelling				
None	1152 (76%	) 1171		(78%)
Mild Sulcal	283 (19%	)	265	(18%)
Mild Ventricular	71 (5%)	73		(5%)
Moderate	1 (<	1%)	0	(0%)
Severe[4]	0 (0%)	1		(<1%)
Location of hyperdense arteries				
None	1131 (75%	) 1151		(76%)
Anterior	360 (24%	)	342	(23%)
Posterior	16 (	1%) 1	7	(1%)
Evidence of atrophy	1161 (77%	) 1166		(77%)
Evidence of periventricular lucencies	765	(51%)	782	(52%)
Evidence of old lesions	685	(45%)	651	(43%)
Evidence of non-stroke lesions	73	(5%)	77	(5%)
Baseline variables collected from seven-day form				
Pre-trial history of stroke	354	(23%)	345	(23%)
Pre-trial treatment with aspirin	639	(47%)	667	(49%)
Pre-trial treatment with dipyridamole	66	(5%)	59	(4%)
Pre-trial treatment with clopidogrel	69	(5%)	77	(6%)
Pre-trial treatment with anticoagulants				
Warfarin or other oral anticoagulant	61	(4%)	57	(4%)

	rt-PA Con		trol	
	No. (%	) N	o.	(%)
Heparin[5] (low dose)	15	(1%)	5	(0%)
None of the above	1292	(94%)	1309	(95%)
Pre-trial treatment for hypertension	975	(64%)	979	(65%)
Pre-trial treatment for diabetes	184	(12%)	204	(13%)
Phase of trial in which patient recruited				
Blinded	136 (9%)		140	(9%)
Open	1379 (91%	) 1380		(91%)
Patients recruited in centre with pre-trial experience of thrombolysis[6] 5	75	(38%)	568	(37%)

NIH = National Institutes of Health, TACI= Total Anterior Circulation Infarct, PACI = Partial Anterior Circulation Infarct, LACI = Lacunar Infarct, POCI = Posterior Circulation Infarct, MCA = middle Cerebral Artery, ACA = Anterior Cerebral Artery

1. These variables were collected via the web-based or telephone randomisation system and had to be entered, complete and passed range and consistency checks before the system would issue a treatment allocation. Variables marked with an asterisk* were employed in the minimisation algorithm.
2. Expert panel's blinded assessment of pre-randomisation scan. This assessment was performed by the expert panel members after randomisation & blinded to treatment allocation and all clinical details.
3. Indeterminate because no infarct was visible.
4. Two patients in Control group were randomised at more than 6 hours (protocol violation). One of these was recorded as having severe swelling on the randomisation scan, it was later discovered that the stroke had occurred about 24 hours earlier.
5. Heparin: unfractionated or low-molecular weight heparin.
6. Pre-trial experience of thrombolysis is defined as the centre had, before joining the trial, a protocol for open label rtPA and had treated at least 3 people in the 12 months before joining the trial; 76 (49%) centres met this criterion.

Web table 2 Adherence to treatment protocol and background stroke care

	rt-PA (n= 1515)		Control (n=1520)	
	No· (%	) N	o·	(%)
Eligibility deviations				
Dependent pre-stroke[1]	8 (0	·5%) 9		(0·6%)
Haemorrhage on pre-randomisation scan	1	(0·1%)	0	(0·0%)
Advanced ischaemic change on pre-randomisation scan[2]	0 (0	·0%) 1		(0·1%)
Tumour or non-stroke lesion on pre-randomisation CT[3]	0 (0	·0%) 0		(0·0%)
Pre-randomisation low dose heparin	16	(1·1%)	6	(0·4%)
Systolic BP <90 or >220 mmHg or diastolic BP <40 or >130 mmHg	0	(0·0%)	0	(0·0%)
Glucose outside allowable limits (3·0 to 20 mmol/l)	0	(0·0%)	0	(0·0%)
Thrombolysis for stroke within previous 14 days	0	(0·0%)	1	(0·1%)
Infusion compliance among those allocated rt-PA[4]				
Did not get bolus	26	(1·7%)	·	
Got bolus, but did not start infusion	4	(0·3%)	·	
Got bolus, started infusion, but halted, wrong total dose[5] 29		(1·9%)	·	
Got bolus, started infusion, but halted, right total dose	62	(4·1%)	·	
Got bolus and infusion, wrong total dose[5]	45	(3·0%)	·	
Got bolus and infusion, right dose	1348	(89·0%)	·	
Infusion compliance among those allocated placebo or open control[6]				
Blinded phase : got bolus, but did not start infusion	·		1	(0·1%)
Blinded phase : got bolus, started infusion, but halted	·		6	(0·4%)
Blinded phase : got bolus and planned infusion	·		133	(8·8%)
Open phase : did not receive rtPA	·		1368	(90·3%)
Open phase : received at least some rtPA	·		7	(0·5%)
Treatments given within 24 h				
Double-blind phase				
Aspirin given	12 (8	·8%) 10		(7·1%)
Other antiplatelet given	1	(0·7%)	0	(0·0%)
No antiplatelet given	123	(90·4%)	130	(92·9%)
Low dose heparin for DVT prophylaxis given	6	(4·4%)	4	(2·9%)
Full dose heparin given[7]	1 (0	·7%) 0		(0·0%)
Intravenous fluids given[8] 11		(73·3%)	8	(47·1%)
Insulin given[8]	0 (0	·0%) 1		(5·9%)
Open phase[9]				
Aspirin given	183	(13·3%)	1044	(75·8%)
Other antiplatelet given	53	(3·8%)	218	(15·8%)
No antiplatelet given	1167	(84·8%)	271	(19·7%)
Low dose heparin for DVT prophylaxis given	46	(3·3%)	223	(16·2%)
Full dose heparin given	17	(1·2%)	52	(3·8%)
Intravenous fluids given	838	(62·1%)	804	(59·4%)
Insulin given	96 (7	·1%) 99		(7·3%)
Other treatments given between 24 h and 7 days				
Aspirin given	1114 (73	·8%) 1284		(84·7%)
Other antiplatelet given	318	(21·1%)	401	(26·5%)
Low dose heparin or LMWH for DVT prophylaxis given	315	(20·9%)	406	(26·8%)
Full anti-coagulation[10] 117		(7·7%)	122	(8·1%)
Any treatment to lower blood pressure	890	(58·9%)	889	(58·7%)
Any non-trial thrombolysis	3	(0·2%)	0	(0·0%)

	rt-PA (n= 1515)		Control (n=1520)	
	No· (%	) N	o·	(%)
Antibiotics 378		(27·7%)	361	(26·4%)
Feeding via nasogastric tube or percutaneous gastronomy	256	(18·8%)	304	(22·2%)
Place of treatment in 7 days since randomisation	**N [11] (%)**	**Median stay[12]**	**N [11] (%)**	**Median stay[12]**
	1392		1402	
Admissions area[13]	30 (2·2%)	2 40	(2·9%)	1
High dependency ward, intensive care ward or critical care area	328 (23·6%)	1 237	(16·9%)	1
Stroke unit or stroke rehabilitation unit	1248 (89·7%)	6 1252	(89·3%)	6
General Ward[14] 215	(15·4%)	4 219	(15·6%)	5

1. In the early part of the trial, patients with a minimal degree of pre-stroke dependency could be included. After a protocol amendment to change eligibility, the randomisation programme was changed in September 2004 and such patients could not be included in the remainder of the trial.
2. Marked degree of ischaemic change on pre-randomisation CT or MR incompatible with onset less than 6 hours previously.
3. Tumour or non-stroke lesion sufficient to account for symptoms leading to randomisation.
4. Base of percentages is number with infusion record (1514).
5. Dose violations occur when dose given is greater than 10% above or below the prescribed dose, or when a Control patient in the Open phase received any dose of rt-PA.
6. Base of percentages is number with infusion record (1515). .
7. Full-dose unfractionated heparin or high-dose low molecular weight heparin.
8. Questions on intravenous fluids and insulin were only added in 2004. Hence few participants in the blinded phase were asked these questions (15 rt-PA, 17 Control).
9. Patients in the control arm of the open phase who receive these drugs are not protocol violators, but are shown here for information. Base of percentages is number with valid seven-day follow-up in given trial phase and treatment group.
10. Full-dose unfractionated heparin, high-dose low molecular weight heparin or oral anticoagulants.
11. N is the number of patients who spent at least one night on the particular type of ward. The base of percentages is the number of patients who spent at least one night in any of these ward types This question was not asked in the early part of the trial (pre 2003).
12. Median number of nights spent among patients who stayed at least one night in given type of ward.
13. Accident and Emergency Department or Medical admissions unit.
14. General Ward: Neurology Ward, Geriatric Medicine Ward, General Internal Medicine Ward, Neurosurgical Ward, Geriatric Ward, Rehabilitation Ward or Other Ward.

Webtable 3: Outcomes at six months for patients with known disability status

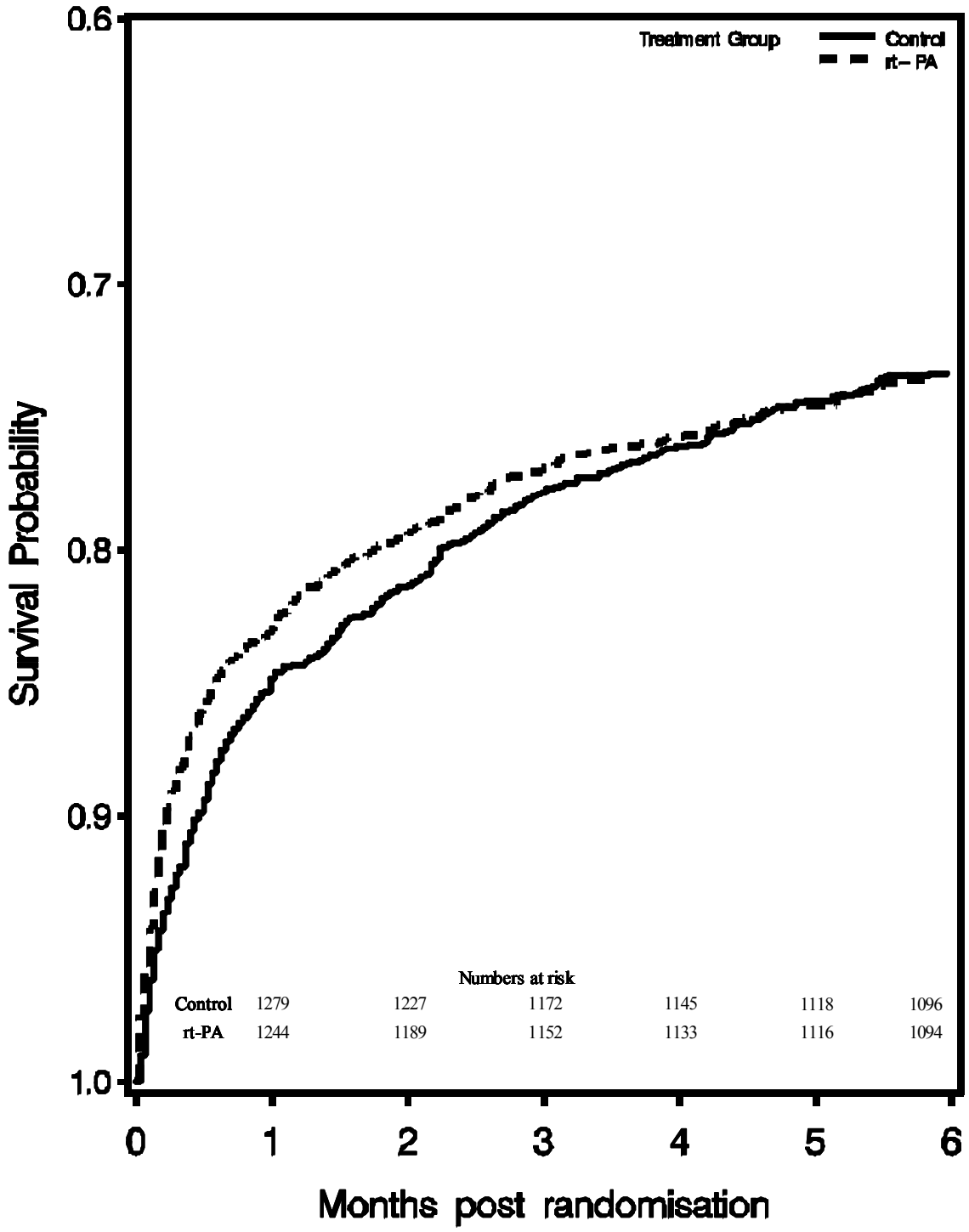
r	t-PA		Placebo		Adjusted analysis[1] Un		adjusted analysis[2]		
	N	(%)	N	(%)	OR (95% CI)	P	OR (95% CI)	P	Difference per 1000 (95% CI)
No. randomised	1515		1520						
No. with known 6 month disability status	1473		1466						
Oxford Handicap Score									
0	137	(9%)	116	(8%)					
1	225	(15%)	204	(13%)					
2	183	(12%)	200	(13%)					
3	234	(15%)	192	(13%)					
4	115	(8%)	140	(9%)					
5	171	(11%)	207	(14%)					
6 (Died before 6 months)	408	(27%)	407	(27%)	0·94 (0·79 , 1·13)	0·533	1.00 (0·85 , 1·18)	0·969	1 (-31, 32)
Alive and favourable outcome (0+1)	362	(25%)	320	(22%)	1·25 (1·03 , 1·51)	0·026	1·17 (0·98 , 1·39)	0·078	28 (1, 58)
Alive and independent (0+1+2)	545	(37%)	520	(35%)	1.14 (0.95 , 1.36)	0.160	1.07 (0.92 , 1.25)	0.389	15 (-19, 50)
Total deaths < 7 days	163	(11%)	107	(7%)	1·60 (1·22 , 2·08)	0·001	1·59 (1·22 , 2·07)	<0·001	-37 (-57, -17)

.

Notes:

1. OR = Odds Ratio. Odds ratio and p value calculated from logistic regression after adjusting for age (linear), NIHSS (linear), time (linear) and presence/absence of visible acute ischaemic change on baseline scan as judged by the expert reader. The two cases with delay time greater than 6 hours were omitted from the adjusted analysis.
2. Significance p value calculated from test of difference between percentages for rt-PA and Control, using normal approximation.
3. Oxford Handicap Scale:
0. No symptoms at all. 1. Symptoms, but these do not interfere with everyday life. 2. Symptoms which have caused some changes in lifestyle but still able to look after oneself. 3. Symptoms which have significantly changed lifestyle and need some help in looking after oneself. 4. Severe symptoms requiring help from other people but not so bad as to need attention day and night. 5. Severe handicap needing constant attention day and night.
4. Primary outcome shown in bold

Web Figure 1



Web Figure. Kaplan Meier plot of survival to six months. The dotted line is the treatment group and the solid line the control. Deaths at day 0 excluded from at risk at day 0. Logrank test of difference in survival curves: P=0.83

List of participating hospitals in each country.

Figures in parentheses are the number of patients recruited in the country or by the centre. ***UK*** *(1 447)* Royal Hallamshire Hospital (118): G Venables, C Blank, H Bowler, C Doyle, K Endean, K Harkness, E Parker, M Randall. University Hospital of North Staffordshire (97): C Roffe, N Ahmad, A Arora, S Brammer, J Chembala, B Davies, S Ellis, E Epstein, K Finney, C Jackson, C Jadun, R Kinston, H Maguire, I Memon, I Natarajan, M Poulson, R Sanyal, S Sills, A Vreeburg, E Ward. Western General Hospital (95): P Sandercock, R Al-Shahi Salman, R Davenport, M Dennis, P Hand, S Hart, I Kane, S Keir, M MacLeod, L McKinlay, H Milligan, E Sandeman, J Stone, C Sudlow, P Taylor, J Wardlaw, C Warlow, W Whiteley, A Williams. The National Hospital for Neurology & Neurosurgery (84): M Brown, B Athwal, V Bassan, N Bhupathiraju, J Bowler, C Davie, D Doig, R Erande, S Gilbert, L Ginsberg, R Greenwood, S Gregoire, N Harding, N Losseff, R Luder, N Passeron, R Perry, P Rayson, R Simister, S Stone, D Werring. Arrowe Park Hospital (83): J Barrett, H Aitken, S Cherian, R Davis, S Downham, L Godd, V Gott, D Jose, V Little, D Lowe, L Luxford, M McGrory, P Owings, N Price, J Richards, G Sangster, J Sherlock, S Vargese, I Wakefield, P Weir. Southend University Hospital (77): P Guyler, T Attygale, S Chandler, L Coward, S Feasey, C Khuoge, T Loganathan, S Martin, A O'Brien, D Sinha, V Thompson, S Tysoe, R Walsh. Norfolk and Norwich University Hospital (67): K Metcalf, J Cochius, R Fulcher, N Gange, C Green, J Jagger, M Lee, P Myint, J Potter, G Ravenhill, S Shields, N Shinh, T Staunton, E Thomas, W Woodward, P Worth, N Wyatt. Nottingham City Hospital (63): W Sunman, P Bath, P Berman, J Clarke, C Gaynor, F Hammonds, R Harwood, K Mitchell, S Munshi, S Pacey, A Shetty, N Sprigg, H Stear, G Subramanian, A Wills. Guy's & St.Thomas Hospital (60): A Rudd, H Audebert, A Bhalla, J Birns, R Chowdhury, G Cluckie, I Davies, C Gibbs, P Holmes, N Mitchell, F Schiavone, E White, M Yeung. Darlington and Bishop Auckland Hospitals (56): A Mehrzad, V Baliga, E Brown, L Burnside, B Esisi, J Kent, P Orr, D Stead, E Wayman. University Hospital Aintree (46): R Durairaj, C Cullen, R Kumar, H Martin, D McDowell, A Sharma, V Sutton, R White. University Hospital of Wales (46): T Hughes, K Ali, J Anderson, K Baker, K Bethune, K Bethune, M Booth, M Cossburn, S Halpin, M Hourihan, E Marsh, K Peall, R Powell, H Shetty, M Wardle, M Williams. Derby Royal Hospital (37): K Muhiddin, J Beavan, M Clarke, R Donneley, S Elliott, P Fox, P Gorman, M Harper, M Mangoyana, I Memon, L Mills, L Wright. Addenbrookes Hospital (34): L Warburton, J Baron, P Barry, D Day, T Harold, P Martin, J Mitchell, E O'Brien, J Rycarte, M Turnham. St George's Healthcare NHS Trust (34): G Cloud, L Choy, B Clarke, C Griffin, O Halse, I Jones, F Kennedy, U Khan, R Lewis, A Loosemore, C Lovelock, H Markus, B Moynihan, J O'Reilly, O Paul, A Pereira, M Punter, P Rich, D Rolfe, F Schiavone. Royal Devon & Exeter Hospital (Wonford) (30): M James, J Bell, A Bowring, L Boxall, J Cageao, H Eastwood, S Elyas, F Hall, S Harries, A Hemsley, S Jackson, S Keenan, P Mudd, A Sekhar, D Strain, J Sword, N Wedge. Aberdeen Royal Infirmary (26): M MacLeod, M Bruce, A Joyson, M Kemp, K McMullan, J Reid, O Robb, J Webster, S Wilkinson. Hammersmith Hospitals & Imperial College (24): P Sharma, P Bentley, H Jenkins, A Kar, T Sachs. Northwick Park Hospital (20): D Cohen, R Bathula, J Devine, M Mpelembue. William Harvey Hospital (20): D Hargroves, I Balogun, L Cowie, A Maidment, D Rand, J Rowe, H Rudenko, D Smithard, L Wray. Scarborough Hospital (17): J Paterson, J Brown, J Hampton, S Jamieson, R Rose, A Volans. Countess of Chester Hospital NHS Foundation Trust (17): K Chatterjee, G Abbott, R Brookes, C Castle, C Kelly, S Leason, A Nallasivan, A Sen. Watford General Hospital (17): D Collas, M Cottle, N Damani, P Jacob, D Oza, D Werring. University Hospitals Coventry & Warwickshire NHS Trust (15): A Kenton, N Adab, L Aldridge, H Allroggen, Y Brown, R Cross, L Galvin, K Ghosh, A Grubneac, A Lindahl, H Mehta, M Pritchard, C Randall, P Ray, A Shehu, S Thelwell. Royal Bournemouth & Christchurch NHS Trust (12): D Jenkinson, J Bell, T Black, O David, J Kwan, A Orpen, C Ovington, D Tiwari, Z ud Din Babar. Leeds General Infirmary (12): A Hassan, A Bailey, J Bamford, C Bedford, R Bellfield, J Cooper, L Dunsmure, J Greig, M Keeling, L Mandizvidza, J Rankine, E Roberts, P Wanklyn, T Webb, S Williamson. York Health Services NHS Trust (12): J Coyle, S Crane, C Croser, P Duffey, R Evans, E Iveson, M Keeling, G Kitching, M Porte, C Rhymes. Queen Elizabeth Hospital (Gateshead) (12): D Barer, M Armstrong, M Bokhari, T Cassidy, B McClelland. Queen Elizabeth The Queen Mother Hospital (10): G Gunathilagan, P DOLKE, S Jain, S Jones, A Maidment, L Rosser, G Thomas, C White. Worcestershire Royal Hospital (10): P Sanmuganathan, C Scholtz, E Stratford. Blackpool Victoria Hospital (10): M O'Donnell, H Goddard, G Hoadley, J Howard, S Leach, J McIlmoyle, A Stewart, A Strain. Basildon & Thurrock University Hospitals NHS FT (9): F Huwez, P Croot, N Gadi, N Mguni, U Umasankar. Royal Infirmary of Edinburgh (8): G Mead, B Chapman, A Coull, S Hart, A Kinnear, B Morrow, F Morrow. St Mary's Hospital (8): D Ames, J Ball, S Bannerjee, J Chataway. Yeovil District Hospital (8): K Rashed, C Buckley, D Donaldson, D Hayward, C Lawson. Luton and Dunstable Hospital (8): L Sekaran, K Bharaj, F Justin, G Jutlla, D Phiri, S Sethuraman, M Tate. Solihull Hospital, Heart of England NHS Trust (8): D Sandler, P Carr, G Jones, J Lyons, K Warren. King's College Hospital (7): L Kalra, A Davis, J Jarosz, D Manawadu, L Sztriha. Doncaster Royal Infirmary (7): D Chadha, A Holford, P Willcoxson. Royal United Hospital Bath (7): L Shaw, D Button, A Cunningham, L Dow, J Dutson, T Hall, C Hardy, N Jakeman, P Kaye, B Madigan, K O'Brien, D Pressdee, M Price, L Robinson, C Taylor, D Williamson. Birmingham Heartlands Hospital (6): D Sandler, P Carr, J Lyons, J McCormack, C Stretton. University Hospital North Durham (6): P Earnshaw, E Brown, S Bruce, C Church, S Desai, B Esisi, M Myint, N Watt. Wansbeck General Hospital (6): C Price, S Elliott, H Graham, R Lakey, K Mitchelson. Bristol Royal Infirmary (6): P Murphy, L Ball, S Caine, J Dovey, J Hughes, A Steele. Stepping Hill Hospital (6): K Dizayee, A Brown, T Chattopadhyay, J Cheetham, H Cochrane, A Datta, M Datta-chaudhuri, C Fox, D Kilroy, S Krishnamoorthy, F Levy, S Metha, P Ngoma, B Venkatesh. Princess Royal Hospital Brighton & Sussex University Hospitals Trust (5): K Ali, R Gautam, N Henderson, M Jones, S Murphy, G Spurling. Belfast City Hospital (5): I Wiggam, C Boyd, K Fullerton, P Gray, M Kinnaird, S MacNair, C Morgan, M Reid, S Tauro. Royal Liverpool University Hospital (5): S Loharuka, D Balmforth,

P Cox, G Fletcher, A Ledger, A Manoj, M Wilkinson. City Hospital, Sandwell & West Birmingham Hospital (5): D Nicholl, S Clegg, S Hurdowar, S Kausar, K Law, A Singal, S Sturman. Royal London Hospital (4): P Gompertz, J Evanson, A Farrell, A Petrou, K Saastamoinen, T Sachs, A Salek-Haddadi, R Yadava. Sunderland Royal Hospital (4): J O'Connell, H Brew, S Butler, S Crawford, C Gray, D Gulliver, N Majmu dar, R O'Brien. Morriston Hospital (4): M Wani, L Dacey, L Davies, R Evans, D Harris, T Jones, S Storton. Royal Preston Hospital (4): S Punekar, A Ashton, S Duberley, H Emsley, C Gilmour, B Gregary, L Hough, S Philip, S Wuppalapati. The Royal Wolverhampton Hospitals NHS Trust (4): K Fotherby, P Bourke, D D'Costa, K Kauldhar, D Leung, R Lodwick, S McBride, D Morgan, M Qaiyum, G Sahota, M Srinvasan. Royal West Sussex NHS Trust, St Richard's Hospital (4): I Kane, N Chuter, L Garrad, M Hookway, S Ivatts, G Kennedy. Queen's Hospital Romford (4): K Darawil, L Al Dhahirl, S Andole, M Baig, P Dugh, K Dunne, H Kariuki, M Khan, S Rathnayaka. Ulster Hospital (3): M Power, K Dynan, J Finnerty, A Heaney, C Leonard, K McKnight, J Turkington, B Wroath. Great Western Hospital (3): B Dewan, S Cotton, M Gardiner, T Saunders, B Vincent. The Queen Elizabeth Hospital Birmingham (3): D Sims, P Guest, E Jones, J McCormack, D Nicholl, J Savanhu, R Tongue, M Willmot. Leicester General Hospital (3): D Eveson, S Dawson, M Dickens, M Fotherby, R Hunt, S Khan, T Kumar, R Marsh, A Mistri, T Robinson, J Thompson. Darent Valley Hospital, Dartford & Gravesham NHS Trust (3): P Aghoram, T Daniel, M Gatehouse, S Hussein, A Jackson, T Shanganya, E Strachan, G Tan. Nevill Hall Hospital, Aneurin Bevan Local Health (3): B Richard, S Elaine, S Hanson, S Mosely, H Reed, M Williams. Colchester Hospital University Foundation Trust (3): R Saksena, S Cook, D Demuran, M Keating, R Needle, V Paramsothy, A Sebastian, R Sivakumar, A Wright. Salford Royal Hospital Foundation NHS Trust (2): R Grue, E Barberan, C Dickson, C Douglas, J Jellicoe, T Marsden, J Priestley, E Quick, C Sherrington, A Singh, C Smith, J Stevens, P Tyrell, J Wainwright. Leicester Royal Infirmary (2): M Ardron, J Birchall. Queen Elizabeth Hospital (Kings Lynn) (2): R Shekhar, C Barsted, S Coleman, S Fletcher, J Graham. John Radcliffe Hospital (2): A Buchan, J Hinkle, J Kennedy, A Manoj, M Westwood. Derriford Hospital (2): A Mohd Nor, S Allder, B Hyams, A Pace. West Cumberland Hospital (1): E Orugun, C Brewer, L Huntley, R Jolly, C Summers. Sandwell General Hospital (1): K Sharobeem, J Khaira, J Leahy, E Linehan, G Moore, J Rizkalla, J Wilkinson. Torbay Hospital (1): D Kelly, C Hilaire. Warrington & Halton Hospitals NHS Foundation Trust (1): O Otaiku, L Connell, G Delaney-Sagar, G James, L Lomax, D Matthew, J Simpson, H Whittle. Medway Maritime Hospital (1): S Sanmuganathan, S Burrows, A Mahmood. Southampton General Hospital (1): G Durward, S Barker, J Cantle, P Crawford, S Evans, V Pressly, N Weir. Victoria Hospital (1): V Cvoro, K McCormick. ***Poland*** *(347)* 2nd Department of Neurology, Institute of Psychiatry & Neurology (190): A Czlonkowska, J Bembenek, M Bilik, G Chabik, W Czepiel, J Dzierka, M Gluszkiewicz, K Grabska, B Janus-Laszuk, J Jedrzejewska, A Kobayashi, T Litwin, A Oskedra, A Piorkowska, M Skowronska, A Sliwinska, U Stepien. SPZZOZ w Sandomierzu (43): P Sobolewski, A Gajewska, M Grzesik, R Hatalska-Zerebiec, I Labudzka, B Loch, A Medrykowska, M Sledzinska, A Sobota, W Szczuchniak, G Wolak, I Zdyb. Medical University of Gdansk (35): W Nyka, D Gasecki, K Chwojnicki, A Gojska, B Karaszewski, G Kozera, M Kwarciany, M Nowak, M Swierkocka-Miastkowska, S Szczyrba, M Wisniewska, E Wnorowska. 1st Department of Neurologyt, Institute of Psychiatry & Neurology (25): P Richter, A Bochynska, M Chahwan, A Graban, R Rola. Military Medical Institute (24): A Stepien, B Brodacki, M Grotowska, J Kotowicz, J Staszewski, J Swistak, S Zaloga. Szpital Powiatowy (14): J Stoiñski, K Czajkowaka-Fornal, P Czubak, A Kaczor, J Kraska, E Nowakowska-Sledz, J Ozdoba-Rot, E Zawadzka. Szpital Specjalistyczny w Konskich (8): M Fudala, D Adamczyk, W Brola, I Guldzinska, K Kaluzny, M Kucharska-Lipowska, M Mosiolek, M Polewczyk, M Ziomek. Central University Hospital (7): G Opala, M Arkuszewski, M Kudlacik, P Malgorzata, M Swiat. SPSK im. Prof. W. Orlowskiego CMKP (1): U Fiszer, M Lenska-Mieciek. ***Italy*** *(326)* Ospedale Di Citta' di Castello (62): S Cenciarelli, A Barilaro, R Condurso, F Coppola, S Dioguardi, E Gallinella, A Mattioni, C Menichetti, S Ricci. Nuovo Ospedale Civile "S.Agostino-Estense" (40): F Casoni, M Bacchelli, M Cavazzuti, M Malagoli, A Zini. Ospedale Beato Giacomo Villa - Citta' della Pieve (36): G Benemio, M Celani, R Allegrucci, V Bondo, S Cupella, L Guerra, S Guerrieri, C Ottaviani, E Righetti, C Rossi, N Sacchi, M Scucchi, V Stefanini. Ospedale Di Branca (36): T Mazzoli, A Bigaroni, L Greco, R Paris, P Parise, S Ricci. A.O. Niguarda Ca'Granda (25): A Ciccone, L Basso, R Causarano, P Doneda, E Ferrante, A Gatti, A Guccione, A Gullo, F Imbesi, S Jann, R Marazzi, E Moro, C Motto, D Parodi, A Protti, M Riva, A Rosiello, I Santilli, R Sterzi, P Tiraboschi, G Venturelli. R.Guzzardi Hospital - Vittoria (RG) (19): F Iemolo, R Campagna, G Campagnolo, A Carnemolla, N D' Apico, G D'Asta, S Giannarita, A Giordano, E San zar. Ospedale Regionale di Aosta (17): E Bottacchi, S Cordera, G Corso, M Di Giovanni, G Giardini, C Lia, T Meloni, M Pesenti Campagnoni, P Tosi. Ospedale Sacro Cuore - Negrar Verona (12): A Adami, G Rossato, T Zuppini. Ospedale Maggiore - Bologna (12): G Procaccianti, T Sacquegna. Universita degli Studi di Genova (11): C Gandolfo, M Balestrino, C Bruno, L Castellan, M Del Sette, A Ferrari, C Finocchi, N Reale, D Rizzi. Ospedale Civlle S.Andrea (9): M Del Sette, L Benedetti, C Capellini, E Carabelli, E Cibei, M Godani, G Guariglia, E Landini, A Mannironi, B Nucciarone, S Parodi, S Tonelli, E Traverso, D Zito. Ospedale S. Giovanni Battista - Foligno (8): P Brustenghi, F Corea, O Flamini, S Lolli, G Pelliccia, R Ricci, S Stefanucci, M Zampolini. Ospedale a Vibo Valentia (8): D Consoli, F Galati, P Postorino. Azienda Ospedaliero-Universitaria " Ospedali Riuniti" di Foggia (8): G Rinaldi, E Carapelle, G Grilli, M Guido, L Specchio. Ospedale Valduce di Como (8): N Checcarelli, G Borin, L Chiveri, R Clerici, E Corengia, L Gandola, P Garavaglia, M Guidotti, A Martegani, M Mauri, F Muscia, F Raudino. Ospedale di Cattinara - Trieste (4): F Chiodo Grandi, A Brat ina, N Carraro, M Gaio, A Granato, N Koscica, M Naccarato, V Sarra, P Schincariol, C Vil otti, Z Zugna. Clinica Dr Pederzoli Spa (4): D Idone, C Bonato, E De Angelis, A Forgione, M Gambera, F Recchia, S Tamburin, P Tinazzi Martini, G Zanette. Ospedale Civile San Matteo Degli Infermi - Spoleto (2): S Grasselli. Ospedale Silvestrini - Perugia (2): G Agnelli, A Andrea, A Billecia, V Caso, V Casso, R Fabiola, P Fanelli, M Paciaroni, B Sergio, M Vemti. Mater Salutis Hospital, Legnago VR (2): M Silvestri, L Altarini, A Bonfante, M Bonornetti, B Costa, N D'Attoma, N Deluca, F Frattini, R Niego, D Rafaele, V Ravenna, M Turazzini. Ospedale Guglielmo da Saliceto - Piacenza (1): S Cammarata. ***Sweden*** *(297)* Uppsala University Hospital (100): E Lundström, L Jonsson, U Söderström,

A Terént. Danderyd Hospital (46): V Murray, A Alvelius, M Arbin von, I Dalenbring, Å Doverhall, Å Franzén-Dahlin, N Greilert, M Hallberg, A Heijne von, E Isaksson, H Kumpulainen, A Laska, A Lundström, C Martin, J Muhrbeck, E Näslund, N Ringart, E Roth, R Undén, P Waldenström. Hassleholm Hospital (29): M Esbjornsson, M Petranek. Karnsjukhuset (25): B Cederin, E Bertholds, A Elgåsen, T Johansson, B Witteborn. Koping Hospital (20): M Kwiatkowska, E Gustafsson, T Noren, J Saaf. Mora Hospital (17): J Teichert, M Bertilsson, S Nilsson, S Oestberg. LidkopingHospital (11): L Welin, K Fredricson, L Pehn. Falu Hospital (11): J Hambraeus, I Lonn. Capio S:tGoran Hospital (9): B Hojeberg, A Adolfsson, M Anzen. Vastervik Hospital (5): T Wallen, R Schloenzig, P Söderström, A Wennerberg. University Hospital MAS (5): F Buchwald, K Abul-Kasim, A Berkeskold, J Petersson, E Poromaa. University Hospital of Northern Sweden (4): P Wester, R Backlund, A Sjöström. Helsingborgs lasarett (4): B Hedström, E Campbell, K Johnsson, B Karlsson, N Lekokotla, C Lundahl, A Risedal, P Sandgren, A Svensson. Visby Hospital (4): S Byselll, E Smedberg, A Vestberg Bysell. Sundsvall Hospital (3): V Sjögren, B Högvall. University Hospital Lund (2): G Andsberg, T Cronberg, A Lindgren. Vasteras Hospital (1): H Wannberg, F Ax, L Nyren. Karlstad Central Hospital (1): J Sanner, H Andersson, F Andler, S Holmgård, R Johansson, I Magnussan, K Nilsson, J Rådberg. ***Norway*** *(204)* Trondheim University Hospital (69): B Indredavik, H Ellekjær, A Østvik, G Rohweder, D Steckhan, J Storvold. Oslo University Hospital Sykehus (66): E Berge, Y Rønning, R Aakvik, K Bruins Slot, G Knutsen, M Moxness, R Pettersen, T Wyller. University Hospital of North Norway (23): C Wahl, O Iversen, S Johnsen, B Norderhus, L Steffensen, E Stensland. Kongsvinger Hospital (13): T Asak, J Aaseth, T Rotnes, J Sparby, S Wetterhus. Levanger Hospital (9): H Hallan, A Aardal, T Graven, H Hansbakk Skjetne, B Klykken, K Lindqvist, A Tommy. University Hospital of North Norway (8): T Engstad, M Antonsen, R Bajic, W Fønnebø, S Hykkerud, I Lyngmo, A Nyrnes, S Rogne, S Sparr. Harstad Hospital (7): O Kildahl-Andersen, K Pedersen, H Ulrichsen. Ålesund Hospital (4): O Skogen, I Alnes, R Hukari, Y Seljeseth, P Vadset. Asker and Bærum Hospital (2): G Knutsen, B Fure, H Ihle-Hansen, N Johnsen, L Kornberg. Namsos Hospital (2): S Schuler, M Heibert. Volda Hospital (1): M Lillebø, O Aasen, I Eskeland, T Hamre, S Hareide, H Helset, K Kolnes, B Lødemel, H Osevelle, S Reite, E Velle. ***Australia*** *(179)* Nambour General Hospital (51): R Grimley, E Ahern, C Cocks, M Courtney, R Devin, J Endacott, C Fawcett, V Harrington, C Johnston, M Koltermann, S Murray, K Ng, G Styles, A Tampiyappa. John Hunter Hospital (29): C Levi, K Chung, L Dark, M Evans, Y Gawarikar, E Kerr, A Loiselle, F Miteff, A Moore, W O'Brien, M Parsons, D Quain, A Royan, M Russell, N Spratt. Gosford Hospital (24): J Sturm, D Crimmins, D Griffiths, P Kavelieros, J Kinsella, A Malhotra, B O 'Brien, A Schutz, M Webb, S Whyte, V Zenteno. Westmead Hospital (16): R Lindley, A Bleasel, N Cordato, A Duggins, V Fung, L Gomes, N Ingham, J Ip, P Landau, J Morris, S Vucic. Royal Perth Hospital (15): G Hankey, A Claxton, N Lillywhite. The Canberra Hospital (12): C Lueck, C Andrews, G Danta, C Das, I Harvey, A Hughes, C McColl, A Oon, R Tuck. Royal Brisbane and Women's Hospital (10): S Read, M Badve, M Broad, G Cadigan, H Cavanagh, J Chalk, D Copsinis, K Etherington, R Henderson, R Hull, J O'Sullivan, J Pandian, L Ross-Lee, M Roxas, N Sheikh, G Skinner, A Wong. Austin Health - Repatriation Campus (8): H Dewey, A Brodtmann, G Donnan, A Hughes, M Karonen, H Ma, T Mulcahy, S Petrolo, L Walker, D Young, J Zavala. Nepean Hospital (7): M Thieben, C Harris, M Krause, S Lane, H Park, M Shaffi, J Wood. Box Hill Hospital (7): C Bladin, A Buckland, K Coughlan, B Coulton, A Gilligan, P Lee, S Mullen, Z Ross, P Sien Loh, C Szoeke. ***Portugal*** *(82)* UAVC. Centro Hospitalar de Trás-os-Montes e Alto Douro (42): M Silva, F Afonso, J Gabriel, P Guimarães, A Velon. Hospital Pero da Covilhã (19): M Castelo- Branco, F Alvarez, V Branco, C Coxo, P Goulao, D Leal, S Morgado, R Oliveira, F Paiva, A Rodrigues, M Simoes. Hospital de Santo António (12): G Lopes, T Almeida, M Cardoso, J Chaves, C Correia, M Correia, J Damásio, R Felgueiras, J Pereira, A Tuna. Hospital S.Marcos (9): C Ferreira, E Lourenco, A Machado, R Mare, J Rocha. ***Belgium*** *(73)* Cliniques Universitaires St Luc (73): A Peeters. ***Austria*** *(46)* Landesklinikum Donauregion Tulln (34): K Matz, M Brainin, G Funk, V Reiner-Deitemyer. Krankenhaus Der Barmherzigen Bruder Wien (9): J Ferrari, A Flamm-Horak, G Gruber, R Rattinger. Krankenhaus Göttlicher Heiland (3): W Muellbacher, D Doppelbauer, R Kalchmayr, W Schima, T Wieser, M Zart. ***Switzerland*** *(23)* Universitatsspital Basel (22): P Lyrer, L Bonati, S Engelter, F Fluri, S Muller, E Radue, A Tiemessen, L Walz, F Weisskopf, S Wetzel. Universitätsspital Zürich (1): A Luft, D Fetz, B Hertler, A Pangalu. ***Canada*** *(8)* QEII Health Sciences Centre (8): G Gubitz, P Boulton, J Jarrett, J Moeller, S Phillips. ***Mexico*** *(3)* Instituto Nacional de Neurologia y Neurojrugia MVS (3): A Arauz, L Bermudez, J Calleja, R Garcia.

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

*The IST-3 collaborative group**

Lancet Neurol 2013; 12: 768–76

Published Online
June 21, 2013
http://dx.doi.org/10.1016/S1474-4422(13)70130-3

See Comment page 731

*Members listed in the appendix

Correspondence to:
Prof Peter Sandercock, Division of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK
peter.sandercock@ed.ac.uk

See Online for appendix

Summary

Background Few data are available from randomised trials about the effect of thrombolysis with alteplase on long-term functional outcome in patients who have had acute ischaemic stroke and no trial has reported effects on health-related quality of life. A secondary objective of the third International Stroke Trial (IST-3) was to assess the effect of thrombolysis on such outcomes at 18 months.

Methods In this open-label, international, multicentre, randomised, controlled trial, 3035 patients with ischaemic stroke from 12 countries were randomly allocated within 6 h of onset via a secure central system to either intravenous alteplase (0·9 mg/kg; n=1515) plus standard care or standard care alone (control; n=1520). 2348 patients were scheduled for 18-month follow-up. For our main analysis, survivors were assessed at 18 months with the Oxford handicap scale (OHS; the primary outcome was the adjusted odds of OHS score 0–2). We also used the EuroQoL (EQ) instrument and asked questions about overall functioning and living circumstances. We analysed the OHS and the five EQ domains by ordinal logistic regression and calculated the mean difference between treatment groups in EQ utility index and visual analogue scale score. Analyses were adjusted for key baseline prognostic factors. This study is registered with controlled-trials.com, number ISRCTN25765518.

Findings At 18 months, 408 (34·9%) of 1169 patients in the alteplase group versus 414 (35·1%) of 1179 in the control group had died (p=0·85). 391 (35·0%) of 1117 patients versus 352 (31·4%) of 1122 had an OHS score of 0–2 (adjusted odds ratio [OR] 1·28, 95% CI 1·03–1·57; p=0·024). Treatment was associated with a favourable shift in the distribution of OHS grades (adjusted common OR 1·30, 95% CI 1·10–1·55; p=0·002). Alteplase treatment was associated with significantly higher overall self-reported health (adjusted mean difference in EQ utility index 0·060; p=0·019). The differences between the groups in visual analogue scale score and the proportion living at home were not significant.

Interpretation IST-3 provides evidence that thrombolysis with intravenous alteplase for acute ischaemic stroke does not affect survival, but does lead to statistically significant, clinically relevant improvements in functional outcome and health-related quality of life that are sustained for at least 18 months.

Funding UK Medical Research Council, Health Foundation UK, Stroke Association UK, Research Council of Norway, AFA Insurances Sweden, Swedish Heart Lung Fund, The Foundation of Marianne and Marcus Wallenberg, Polish Ministry of Science and Education, the Australian Heart Foundation, Australian National Health and Medical Research Council, Swiss National Research Foundation, Swiss Heart Foundation, Assessorato alla Sanita (Regione dell'Umbria, Italy), and Danube University.

Introduction

Intravenous alteplase has been approved for treatment of acute ischaemic stroke in Europe for patients who are younger than 80 years and can be treated within 4·5 h. Such use is associated with improved functional outcome at 3 months after stroke,[1] but whether treatment improves survival and sustains functional recovery in the long term is unclear. Of the 12 completed randomised controlled trials, ten reported outcomes at 90 days or less,[1] two reported outcomes at 6 months,[2,3] and one reported outcomes at 12 months,[3] but none have reported effects at more than 1 year after stroke. Furthermore, the effect of thrombolysis on health-related quality of life—an important measure of the clinical and economic value of treatment—has not been reported to our knowledge.

The third International Stroke Trial (IST-3)[2] recruited 3035 patients—half of whom were older than 80 years—to assess the effect of thrombolytic treatment with intravenous alteplase within 6 h of onset of acute ischaemic stroke. The results showed that although thrombolytic treatment was not associated with a significant difference in the proportion of patients who were alive and independent at 6 months, treatment did seem to improve functional outcome. A prespecified secondary ordinal analysis of Oxford handicap scale scores showed that treatment was associated with a favourable shift in the distribution of Oxford handicap

scale scores (odds ratio [OR] 1·27, 95% CI 1·10–1·47; p=0·001).[2] A secondary aim of IST-3 was to assess whether thrombolytic treatment improved outcomes more than 1 year after stroke, and sought to assess survival, functional outcome, health-related quality of life, overall functioning, and living circumstances at 18 months.[4,5]

Methods

Study design and participants

The methods of the trial have been described in full previously.[2,4–6] IST-3 was a randomised, open-label trial of intravenous alteplase (0·9 mg/kg) plus standard care compared with standard care alone (control). Eligibility criteria were: symptoms and signs of clinically definite acute stroke, known time of stroke onset, treatment could be started within 6 h of onset, and exclusion by CT or MRI of intracranial haemorrhage and structural brain lesions that could mimic stroke (eg, cerebral tumour). A patient could only be included in the trial if both they (or a proxy) and their clinician believed that the treatment was promising but unproven—ie, there was neither a clear indication for treatment, nor a clear contraindication against treatment. The effect that using this uncertainty principle approach as a key eligibility criterion had on the type of patients included and excluded from the trial has been described in detail elsewhere.[2,6] Generally, patients who could be treated within licence were rarely enrolled, unless there was a specific reason that led the clinician or patient to be uncertain about whether to treat or not; as a result, 95% of enrolled patients did not meet the terms of the prevailing EU approval for treatment. All participants or proxies gave informed consent. The protocol was approved by the Multi-Centre Research Ethics Committee (Scotland) and by local ethics committees.

For the analysis presented here, we planned to assess outcome in patients who had follow-up at 6 months and 18 months. In seven countries (Austria, Belgium, Canada, Italy, Mexico, Poland, and UK) follow-up had to cease on Jan 30, 2012; therefore, we excluded any patients from these countries who were recruited after June 30, 2010, because they would not reach the 18-month follow-up point. In three countries (Australia, Norway, and Sweden), all recruited patients were to be followed up to 18 months, as part of a sub-study. Two countries (Portugal and Switzerland) followed up patients to 6 months only and were not included in this analysis.

Randomisation

After enrolment, patients were randomly assigned by a secure central telephone or web-based computer system, which recorded baseline data and generated the treatment allocation only after the baseline data had been checked for range and consistency. The system used a minimisation algorithm to balance for key prognostic factors: geographic region, age, National Institutes of Health stroke scale score, sex, time since onset of stroke, stroke clinical syndrome, and presence or absence of visible ischaemic change on the pre-enrolment brain scan.[4,5] To avoid predictable alternation of treatment allocation, and thus potential loss of allocation concealment, patients were allocated with a probability of 0·80 to the treatment group that would minimise the difference between the groups for the key prognostic factors. Recruitment in the small double-blind phase (n=276) began in May, 2000, continued without interruption into the open-treatment phase (n=2759), and was completed in July, 2011.

Procedures

In the ten countries participating in follow-up at 6 months and 18 months after enrolment (Australia, Austria, Belgium, Canada, Italy, Mexico, Norway, Poland, Sweden, and UK), if the patient was not known to have died, staff at each national coordinating centre contacted the patient's doctor (or hospital coordinator) to confirm that the patient was alive and that they might be approached for follow-up. In Austria and Italy, experienced stroke physicians, masked to treatment allocation, contacted all patients by telephone. In the other eight countries, IST-3 trial office staff posted a questionnaire to patients to assess outcome. Non-responders were sent a second questionnaire. If no questionnaire was returned, an experienced, masked clinician or stroke nurse assessed the patient by telephone interview. Telephone assessment of disability in stroke survivors is as valid as face-to-face interviews[7] and postal questionnaires.[8]

The primary outcome of the trial was the proportion of patients alive and independent with an Oxford handicap scale[9] score of 0–2 at 6 months (this outcome was chosen

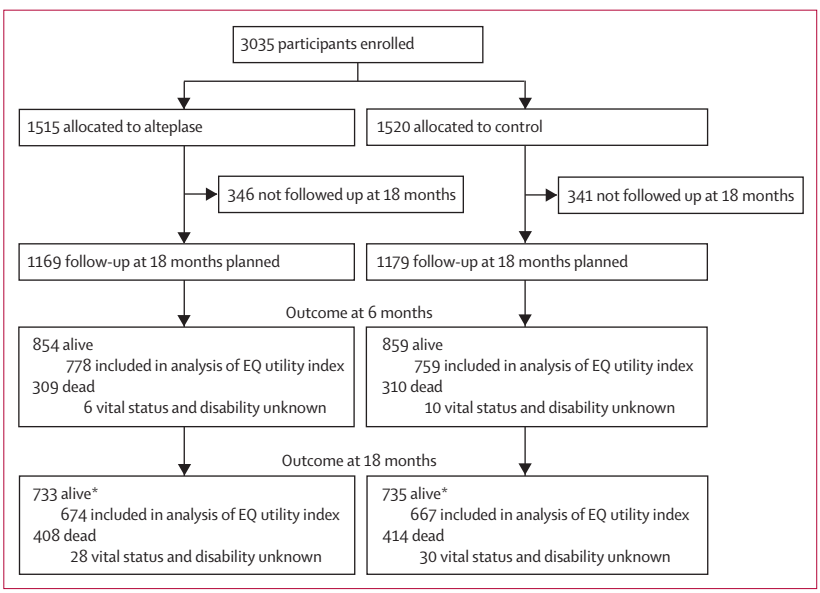


Figure 1: **Trial profile**
EQ=EuroQoL. *Of the patients who were known to be alive at 18 months, 24 in the alteplase group versus 27 in the control group had a known date of death more than 18 months after enrolment, but their disability status at 18 months was unknown.

instead of survival alone because many people regard survival after a stroke in a disabled or dependent state as worse than death). The secondary endpoints at 18 months were: survival, Oxford handicap scale score, health-related quality of life, overall functioning, and living circumstances. The Oxford handicap scale is a six-point scale almost identical to the modified Rankin scale.[10] In emergency care of acute ischaemic stroke, recording quality of life at baseline before randomisation was not possible; instead, quality of life was measured at 6 months and 18 months with the EuroQoL instrument,[11] which assesses current self-rated health by a combination of questions about wellbeing and a visual analogue scale score. The questions are about the five dimensions of mobility, self-care, activity, pain or discomfort, and anxiety (the EQ-5D). Each dimension has three levels (no problems, some problems, severe problems), which can be presented individually. A unique health state is defined by combining one level from each of the five dimensions. Patients' responses can then be combined into an EQ utility index with scores ranging from –1 to +1 (where +1 represents perfect health, 0 represents a state equivalent to death, and –1 represents a state worse than death). Calculation of the EQ utility index requires valuations for all health states, and these have been estimated for the UK and other European populations.[12] For the visual analogue scale, 100 represents the best imaginable health and 0 the worst imaginable health. We used the EuroQoL instrument because it is short and simple, and in patients with stroke it has been validated,[13–17] is responsive to change,[18] and is associated with higher response rates and fewer missing data than more complex instruments.[16] Many patients who have had severe strokes might not be able to complete the questionnaire themselves and because responses from a proxy have reasonable validity,[15,19] we therefore accepted responses submitted by a spouse, partner, close relative, or carer.

	Alteplase group (n=1169)	Control group (n=1179)
Region		
Americas (Canada, Mexico)	5 (<1%)	6 (1%)
Australia	89 (8%)	90 (8%)
Eastern Europe (Poland)	158 (14%)	159 (13%)
Northwest Europe (UK, Austria, Belgium)	550 (47%)	556 (47%)
Scandinavia (Norway, Sweden)	251 (21%)	250 (21%)
Southern Europe (Italy)	116 (10%)	118 (10%)
Age		
18–50 years	49 (4%)	57 (5%)
51–60 years	83 (7%)	81 (7%)
61–70 years	153 (13%)	158 (13%)
71–80 years	291 (25%)	304 (26%)
81–90 years	523 (45%)	512 (43%)
>90 years	70 (6%)	67 (6%)
Women	592 (51%)	596 (51%)
National Institutes of Health stroke scale score		
0–5	235 (20%)	236 (20%)
6–10	323 (28%)	330 (28%)
11–15	244 (21%)	235 (20%)
16–20	207 (18%)	219 (19%)
>20	160 (14%)	159 (13%)
Delay in enrolment		
≤3·0 h	320 (27%)	307 (26%)
>3·0–4·5 h	471 (40%)	481 (41%)
>4·5–6·0 h	378 (32%)	389 (33%)
>6·0 h	0 (0%)	2 (<1%)
Atrial fibrillation	347 (30%)	331 (28%)
Systolic blood pressure		
≤143 mm Hg	380 (33%)	380 (32%)
144–164 mm Hg	379 (32%)	405 (34%)
≥165 mm Hg	410 (35%)	394 (33%)
Diastolic blood pressure		
≤74 mm Hg	342 (29%)	343 (29%)
75–89 mm Hg	409 (35%)	448 (38%)
≥90 mm Hg	406 (35%)	381 (32%)
(Continues in next column)		
(Continued from previous column)		
Blood glucose concentration*		
≤5 mmol/L	202 (20%)	207 (20%)
6–7 mmol/L	501 (49%)	485 (47%)
≥8 mmol/L	324 (32%)	347 (33%)
Treatment with antiplatelet drugs in previous 48 h	599 (51%)	610 (52%)
Assessment of acute ischaemic change		
Scan normal	99 (8%)	102 (9%)
Scan not normal but no sign of acute change	551 (47%)	579 (49%)
Signs of acute change	511 (44%)	490 (42%)
Predicted probability of poor outcome at 6 months†		
<40%	633 (54%)	640 (54%)
≥40–<50%	130 (11%)	113 (10%)
≥50–<75%	275 (24%)	304 (26%)
≥75%	131 (11%)	122 (10%)
Stroke syndrome		
TACI	491 (42%)	509 (43%)
PACI	460 (39%)	430 (36%)
LACI	137 (12%)	133 (11%)
POCI	79 (7%)	104 (9%)
Other	2 (<1%)	3 (<1%)

Data are n (%). TACI=total anterior circulation infarct. PACI=partial anterior circulation infarct. LACI=lacunar infarct. POCI=posterior circulation infarct. *Baseline glucose concentration was not recorded for the first 282 patients recruited; thus, glucose measurements were available for 2066 of 2348 participants (88%; 1027 allocated to alteplase and 1039 allocated to control). †Calculated from a model based on age and baseline National Institutes of Health stroke scale score.[22]

Table 1: **Baseline characteristics of patients included in 18-month follow-up**

We also assessed binary (yes or no) answers to two questions, about global functioning: "Has the stroke left you with any problems?" and activities of daily living: "Do you need help from anybody with everyday activities (in washing, dressing, feeding, and going to the toilet)?" These questions have been validated[17] and were used previously in a large trial.[20] We also asked whether patients were living in their own home, a relative's home, a residential home, a nursing home, or were still in hospital. Finally, the questionnaire asked patients enrolled in the open-label treatment phase what treatments they recalled being given in hospital, including thrombolysis with alteplase. If the patient or proxy did not complete a specific item on a postal questionnaire, we did not re-contact them.

Statistical analysis

All randomly assigned patients who were due to be followed up at 18 months were included in the analysis of survival. We constructed Kaplan-Meier survival curves, and compared treatment groups with the log-rank test. Survival times were censored at 548 days after enrolment if patients died at a later date or returned an 18-month form at a later date. For patients from the Australia, Norway, Sweden, and UK, where reporting of deaths was prompt, if there was no known death date and no return of an 18-month form, patients were censored at 548 days. For patients from other countries who had no reported death date and no 18-month form, survival was censored at the date of return of the 6-month form or at the last date of contact, whichever was later. The justification for, and the methods for statistical adjustment of, the outcomes and the ordinal analyses of the Oxford handicap scale score at 18 months were specified in the statistical analysis plan and also described in the report of the primary outcomes.[2,5] We divided the Oxford handicap scale into five levels: 0, 1, 2, and 3 were retained and 4, 5, and 6 were combined into a single level. The treatment OR between one level and the next was assumed to be constant, so a single parameter (a common OR) summarises the shift in outcome distribution between treatment and control groups.

In the main analysis, we report results without imputing missing data. In the sensitivity analysis, for patients with an unknown Oxford handicap scale score at 18 months, we imputed the value from their 6-month assessment (last observation carried forward). For the EuroQoL instrument, we analysed the three levels of each EQ-5D domain as ordered categories by ordinal logistic regression, calculated the mean overall difference in visual analogue scale score between treatment groups, and estimated the EQ-5D index—calculated with a set of valuations derived from a sample of the UK population with the time trade-off method and also the UK visual analogue scale and European visual analogue scale valuations.[12] Analyses were adjusted for baseline prognostic factors (age, National Institutes of Health stroke scale score, delay between onset and enrolment, and presence of acute ischaemic change on the baseline scan). We did several sensitivity analyses to assess the effect of missing data for Oxford handicap scale score and EQ-5D, and we assessed the effect of setting utility to zero for patients who had died. We did subgroup analyses of the effect of treatment on Oxford handicap scale score (ordinal logistic regression, as in the study by Frank and colleagues[21]) and utility subdivided by age (>80 *vs* ≤80 years), time to randomisation (≤3·0, >3·0–4·5, >4·5–6·0 h), baseline National Institutes of Health stroke scale score (0–5, 6–15, 16–25, >25), phase of the trial (masked *vs* open label), and by the person completing the form (patient *vs* proxy). For National Institutes of Health stroke scale score, we also fitted a model with baseline severity as a linear regressor with treatment-specific slopes. Analyses were done with SAS (version 9.3).

This study is registered with controlled-trials.com, number ISRCTN25765518.

Role of the funding source

The sponsors had no role in data collection, data storage, data analysis, preparation of this report, or the decision to publish. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Of the 3035 patients enrolled by 156 hospitals in 12 countries, 2348 (77·4%) met the criteria for inclusion in the 18-month follow-up study—1169 assigned to alteplase, 1179 assigned to control (figure 1). The baseline characteristics of this subset were well balanced between groups (table 1) and were not much different from those who were ineligible for the 18-month follow-up analysis (appendix).

Of the 2348 patients scheduled for 18-month follow-up, vital status and Oxford handicap scale score at 18 months were known for 2290 (97·5%). Survival at 18 months did

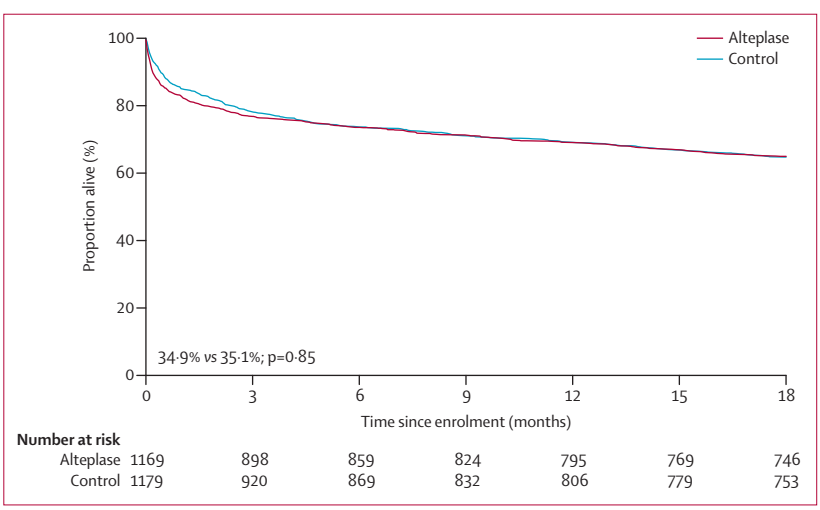


Figure 2: **Kaplan-Meier survival curves**

not differ significantly between groups: 408 of 1169 (34·9%) participants allocated to alteplase versus 414 of 1179 (35·1%) allocated to control died (log-rank p=0·85; figure 2).

At 18 months, of 2348 participants, vital status and disability were known for 2239 (95·3 %), vital status only was known for 51 (2·2%), and vital status and disability were unknown for 58 (2·5%). Oxford handicap scale scores were available for 1117 participants assigned to alteplase versus 1122 assigned to control. 391 (35·0%) patients allocated to alteplase versus 352 (31·4%) allocated to control were alive and independent (Oxford handicap scale score 0–2) at 18 months (adjusted odds ratio 1·28, 95% CI 1·03–1·57; p=0·024; unadjusted OR 1·18, 95% CI 0·99–1·40; p=0·068; table 2), with a favourable shift in Oxford handicap scale score (adjusted common OR 1·30, 95% CI 1·10–1·55; p=0·002). The size and statistical significance of the effect on Oxford handicap scale score at 18 months was robust to sensitivity analyses for missing data (data not shown). The appendix shows Oxford handicap scale score at 6 months in patients scheduled for 18-month follow-up who had data available at 6 months.

The EQ utility index could be calculated for 1341 (91·3%) of the 1468 patients who were alive at 18 months. 591 (44%) of these assessments were completed by patients themselves, 724 (54%) by a valid proxy, and 25 (2%) by a doctor. Treatment was associated with significant improvements in mobility, self-care, ability to do usual activities, and pain or discomfort, with no evidence of an effect on anxiety or depression (table 3). At 18 months, alteplase was associated with significantly fewer patients reporting being left with problems and needing help with everyday activities (table 3).

Although treatment with alteplase was associated with a significantly higher EQ utility index in survivors (p=0·028; table 4), the mean adjusted difference in visual analogue scale score was not significant (p=0·072; table 4). These findings were robust in the sensitivity analyses (data not shown). The appendix shows EQ-5D, EQ utility index, and visual analogue scale score at 6 months and 18 months using different valuations. Of the participants who were still alive, the proportion who were resident at home did not differ significantly between groups (appendix).

For the ordinal subgroup analysis of Oxford handicap scale score at 18 months, significant interactions existed between baseline variables and treatment effect. Greater differences in favour of alteplase were reported for age older than 80 years (p=0·032) and high National Institutes of Health stroke scale score (p=0·021), but not for time to treatment, respondent (patient *vs* proxy), or masking of assessment of outcome (double blind *vs* open label; appendix). When age, delay, and National Institutes of Health stroke scale score were treated as continuous variables, the interaction of ordinal Oxford handicap scale score with age became non-significant, delay remained non-significant, and for National Institutes of Health stroke scale score the p value for a trend was 0·004 (appendix). For EQ utility index, when

	Alteplase group	Control group	Adjusted anaylsis*		Unadjusted analysis†		Difference per 1000 patients† (95% CI)
			OR (95% CI)	p value	OR (95% CI)	p value	
Planned 18-month follow-up	1169	1179	..	..	..	..	..
Missing OHS data at 18 months‡	52 (4%)	57 (5%)	..	..	..	..	..
Number analysed (both vital and OHS status known)	1117 (96%)	1122 (95%)	..	..	..	..	..
OHS score at 18 months§							
0	119 (11%)	83 (7%)	..	..	..	..	..
1	135 (12%)	141 (13%)	..	..	..	..	..
2	137 (12%)	128 (11%)	..	..	..	..	..
3	132 (12%)	138 (12%)	..	..	..	..	..
4	81 (7%)	107 (10%)	..	..	..	..	..
5	105 (9%)	111 (10%)	..	..	..	..	..
Died before 18 months§¶	408 (37%)	414 (37%)	0·95 (0·78 to 1·16)	0·628	0·98 (0·83 to 1·17)	0·855	4 (–36 to 44)
Alive and independent (OHS score 0–2)§	391 (35%)	352 (31%)	1·28 (1·03 to 1·57)	0·024	1·18 (0·99 to 1·40)	0·068	–36 (–75 to 3)
Alive and had favourable outcome (OHS score 0 or 1)§	254 (23%)	224 (20%)	1·23 (0·98 to 1·55)	0·076	1·18 (0·96 to 1·44)	0·109	–28 (–62 to 6)

Data are n (%) unless stated otherwise. OHS=Oxford handicap score. *Logistic regression of outcome on treatment group, adjusted for age, National Institutes of Health stroke scale score, and delay (all linear) and visible infarct on baseline scan. †Standard binomial test with normal approximation. ‡Includes patients who did not return an 18-month form but died more than 18 months after enrolment (figure 1). §Percentages based on number analysed for OHS. For one participant, OHS was imputed on the basis of responses to EQ-5D. ¶If all patients known to be alive are included in the denominators, the percentage dead at 18 months are 35·8% in the alteplase group and 36·0% in the control group.

Table 2: **Oxford handicap scale scores at 18 months**

	Alteplase group	Control group	Odds ratio (95% CI)*	p value	Difference per 1000 patients† (95% CI)
EQ-5D					
Mobility	702	692	..	..	..
No problems walking	283 (40%)	259 (37%)	1·30 (1·05 to 1·61)	0·017	–29 (–80 to 22)
Some problems walking	343 (49%)	346 (50%)	..	..	11 (–41 to 64)
Confined to bed	76 (11%)	87 (13%)	..	..	17 (–16 to 51)
Self-care	695	689	..	..	..
No problems with self-care	372 (54%)	328 (48%)	1·43 (1·16 to 1·78)	0·001	–59 (–112 to –7)
Some problems washing or dressing	176 (25%)	191 (28%)	..	..	24 (–23 to 70)
Unable to wash or dress	147 (21%)	170 (25%)	..	..	35 (–9 to 79)
Usual activities	699	694	..	..	..
No problems with usual activities	235 (34%)	209 (30%)	1·32 (1·07 to 1·62)	0·008	–35 (–84 to 14)
Some problems with usual activities	258 (37%)	256 (37%)	..	..	0 (–51 to 50)
Unable to do usual activities	206 (29%)	229 (33%)	..	..	35 (–13 to 84)
Pain or discomfort	698	694	..	..	..
No pain or discomfort	344 (49%)	304 (44%)	1·26 (1·02 to 1·56)	0·029	–55 (–107 to –2)
Moderate pain or discomfort	316 (45%)	355 (51%)	..	..	59 (6 to 111)
Extreme pain or discomfort	38 (5%)	35 (5%)	..	..	–4 (–27 to 19)
Anxiety or depression	693	690	..	..	..
Not anxious or depressed	353 (51%)	349 (51%)	1·05 (0·85 to 1·29)	0·668	–4 (–56 to 49)
Moderately anxious or depressed	292 (42%)	290 (42%)	..	..	–1 (–53 to 51)
Extremely anxious or depressed	48 (7%)	51 (7%)	..	..	5 (–23 to 32)
Additional questions about overall function					
Stroke left patient with problems	484/700 (69%)	542/699 (78%)	1·67 (1·30 to 2·17)	<0·0001	84 (38 to 130)
Needs help with everyday activities	298/696 (43%)	350/692 (51%)	1·59 (1·25 to 2·00)	<0·0001	78 (25 to 130)

Data are n (%) unless stated otherwise.*Logistic regression of outcome on treatment group, adjusted for age, National Institutes of Health stroke scale score, and delay (all linear) and visible infarct on baseline scan. †Standard binomial test with normal approximation.

Table 3: **EQ-5D and other assessments of function at 18 months**

	Alteplase group		Control group		Adjusted analysis*		Unadjusted analysis†	
	n	Mean (SE)	n	Mean (SE)	Mean difference (SE)	p value	Mean difference (SE)	p value
Visual analogue scale score	653	62·07 (0·90)	648	60·57 (0·91)	2·18 (1·21)	0·072	1·49 (1·28)	0·244
EQ utility index	674	0·550 (0·015)	667	0·502 (0·016)	0·062 (0·020)	0·002	0·049 (0·022)	0·028

*Adjusted for age, National Institutes of Health Stroke Scale score, delay from onset to enrolment, and presence of visible ischaemia on the baseline scan. †Significance based on *t* test. Utility based on UK time trade-off valuations on a scale of –1 to +1.

Table 4: **EQ utility index and visual analogue scale score assessment of overall health at 18 months**

subgroups were in discrete categories, none of the interactions were statistically significant (appendix). However, when the National Institutes of Health stroke scale score was treated as continuous, every five-point increase in score reduced the EQ utility index by 0·12 in the alteplase group versus 0·15 in the control group (adjusted estimates; p=0·008 for difference in slopes). For delay in enrolment time and age there was no trend in EQ utility index, irrespective of whether the variables were grouped or entered into models as a linear trend (data not shown).

Of the 1468 patients who were alive at 18 months, 1260 were asked to recall if they had been given thrombolytic treatment (appendix); 273 in the alteplase group versus 156 in the control group correctly recalled whether or not they had received thrombolytic treatment. In both treatment groups, the ability to recall treatment correctly was associated with better outcome; patients with correct recall were more likely to have an Oxford handicap scale score of 0–2 than were those who remembered incorrectly or did not know (62·5% *vs* 49·3%; 0·0001). Of patients with correct recall, those treated with alteplase were more likely to have an Oxford handicap scale score of 0–2 than were those in the control group (66·7% *vs* 55·1%; 0·018), whereas of those who did not remember correctly, outcomes did not differ significantly between groups (OHS 0–2 48·6% *vs* 49·9%; 0·714); a significant interaction existed between recall status and treatment (p<0·0001).

Discussion

We have shown that, for treatment of acute ischaemic stroke, thrombolysis with intravenous alteplase seems to provide a benefit at 18 months. Treatment had no effect on survival, but was associated with a significant increase in the likelihood of being alive and independent. However, the unadjusted absolute difference in the number of patients alive and independent at 18 months was not significant, so judgment on whether or not the results are clinically significant rests on the quality of the data and the overall patterns of effect seen across all measures. The ordinal estimates of effect at 6 months and 18 months were similar and significant. Treatment was also associated with a gain in health-related quality of life that was significant for four of the five dimensions of the EQ-5D and the overall EQ utility index (though not for visual analogue scale score). Living circumstances did not differ significantly between groups.

Strengths of this study are the large number of patients and the completeness of follow-up. Of the patients scheduled for 18-month follow-up, a small proportion were missing data for both vital and functional outcome status.

***Panel:* Research in context**

Systematic review

The primary results of IST-3[2] included a systematic review of randomised controlled trials of alteplase in acute stroke.[1] To accompany this review we searched up to April 30, 2013, for additional randomised trials of intravenous alteplase versus control within 6 h of onset of acute stroke in the Cochrane Stroke Trials Registry, Internet Stroke Trials Centre, and reference lists in review articles and conference abstracts. For the Cochrane Stroke Trials Registry we searched for interventions with thrombolytic drugs in acute ischaemic stroke added since the last update of the Cochrane review. For the Internet Stroke Center, we searched for "acute ischemic stroke", "acute ischaemic stroke", "thrombolysis", "thrombolytic therapy", "alteplase", and "recombinant tissue plasminogen activator". For each trial, we checked the primary trial publication, and when available, the trial protocol, to determine if it was planned to collect long-term clinical outcome data (ie, more than 90 days after enrolment) or health-related quality-of-life data, as assessed by a valid instrument such as EQ-5D or Short Form 36.

Of the 12 completed randomised controlled trials, ten reported outcome at 90 days or less,[1] two reported clinical outcome at 6 months[2,3] and one at 12 months,[3] but none reported effects more than 12 months after stroke. The Second European Collaborative Acute Stroke Study collected data on health-related quality of life at 90 days with the SF-36, but has yet to report those data. In the NINDS Trial,[3] mortality at 12 months did not differ significantly between alteplase and placebo groups (24% *vs* 28%; p=0·29). The primary outcome was favourable outcome, defined as minimal or no disability as measured by the Barthel index, the modified Rankin scale, and the Glasgow outcome scale, and the treatment effect was assessed with a global statistic. The global statistic favoured the alteplase group at 6 months (OR for a favourable outcome 1·7, 95% CI 1·3–2·3) and at 12 months (1·7, 1·2–2·3).

Interpretation

IST-3 confirms the evidence from previous trials on the neutral effect of thrombolysis with alteplase on survival after stroke in a much larger sample, and adds to the evidence that improvements in function reported at earlier timepoints are evident at 18 months. IST-3 also provides the first validated estimates of the effect of thrombolysis with alteplase on health-related quality of life.

We estimated the EQ utility index in more than 91% of survivors (a similar proportion to that in a trial[23] of younger and less impaired patients with coronary artery disease) and our sensitivity analyses also showed that the estimates of overall health-related quality of life with the EQ utility index were robust to various assumptions about missing data. Although thrombolytic treatment was associated in survivors with less functional impairment, better health-related quality of life, and less likelihood of being left with problems and needing help with daily activities after stroke, it did not translate into a higher proportion of patients living at home at 18 months, perhaps because living circumstances are affected by social and financial factors that are not influenced by treatment. We believe that the direction and size of the effects are clinically significant and will inform health economic assessments of thrombolytic treatment. For example, in 2002, the estimated cost of long-term care of an independent survivor of stroke was £876 per year and that of a dependent survivor was £11 292 per year,[24] so even a small difference in the proportion of patients who survive and are independent will have substantial economic impact.

Lyden[25] has identified limitations of IST-3, chiefly that treatment was not masked. Patient-reported outcomes—eg, health-related quality of life—are subjective,[26] and recall of thrombolytic treatment could affect patient responses. Only 30% of survivors correctly recalled whether or not they had received thrombolytic treatment. As expected, accurate recall was associated with better outcome in both treatment groups. Thus, recall bias might have affected our findings. However, the analysis of recall was based on a variable measured in a subset of survivors after randomisation and so could itself be biased. The effects of treatment on the Oxford handicap scale score and EQ utility index were much the same in the masked and open-label parts of the study (appendix). Assessment of health-related quality of life is limited because many patients who have had a stroke are unable to complete the form themselves. The high proportion of forms completed by a proxy in IST-3 is a result of the severity of stroke in the patients included in the trial. Although the use of surrogates is a potential weakness, it did enable us to achieve satisfactory response rates; however, because proxies tend to assign worse health status than do patients,[15] we were reassured that there was no interaction between the person responding and the effect of treatment on utility or Oxford handicap scale score. Not all enrolled patients were scheduled to be followed up for 18 months, but the selection criteria for the longer follow-up cohort did not seem to introduce relevant imbalances at baseline, nor were the characteristics of the cohort substantially different from those not included in long-term follow-up. We therefore believe the 18-month follow-up cohort is representative of the trial as a whole.

Another weakness is that the trial was under-powered, so the subgroup analyses of the effects of baseline age, stroke severity, and delay to enrolment on the Oxford handicap

scale score and health-related quality of life should be treated with caution. These are secondary analyses of a secondary outcome, and the apparent lack of effect of time to treatment might be due to chance. Furthermore, a more appropriate assessment of the complex interactions between age, stroke severity, and time to treatment will be available from a meta-analysis of individual patient data by the Stroke Thrombolysis Trialists.[27]

In conclusion, IST-3 adds to the evidence from previous trials (panel) and shows that although thrombolysis for acute ischaemic stroke with intravenous alteplase does not improve survival, there is evidence of improvement in several measures of function and quality of life in survivors of all ages for up to 18 months after treatment.

Contributors

The study was conceived by the cochief investigators—PS, RIL, and JW. JW led the imaging. The study was designed by PS, RIL, and JW, with input from all others who were coordinators of the trial in their own country. PS, RIL, JW, MD, and KI wrote the protocol. KI is the study coordinator. GC is the study statistician who prepared the analyses. GM advised on statistical aspects. PS, RIL, MD, WW, GV, AC, AK, EB, KBS, VM, AP, GH, KM, SR, GG, SP, AA, MC, and PL recruited patients. GV, AC, AK, EB, KBS, VM, AP,GH, KM, SR, GG, SP, AA, MC, and PL were national coordinators. PS wrote the first draft and all authors commented on subsequent drafts and approved the final version.

Writing committee

University of Edinburgh (Edinburgh, UK)—Peter Sandercock, Joanna M Wardlaw, Martin Dennis, Geoff Cohen, Gordon Murray, Karen Innes, Will Whiteley; *Sydney Medical School—Westmead Hospital and The George Institute for Global Health (University of Sydney, Sydney, Australia)*—Richard I Lindley; *Sheffield Teaching Hospitals NHS Foundation Trust (Sheffield, UK)*—Graham Venables; *Institute of Psychiatry and Neurology and Medical University of Warsaw (Warsaw, Poland)*—Anna Czlonkowska; *Institute of Psychiatry and Neurology (Warsaw, Poland)*—Adam Kobayashi; *Department of Neurology (Ospedale, Citta' di Castello, Italy)*—Stefano Ricci; *Karolinska Institutet (Stockholm, Sweden)*—Veronica Murray; *Oslo University Hospital (Oslo, Norway)*—Eivind Berge, Karsten Bruins Slot; *School of Medicine and Pharmacology (The University of Western Australia, Perth, Australia) and Royal Perth Hospital (Perth, Australia)*—Graeme J Hankey; *Hospital Geral de Santo Antonio (Porto, Portugal)*—Manuel Correia; *Cliniques Universitaires Saint-Luc (Bruxelles, Belgium)*—Andre Peeters; *Landesklinikum Donauregion Tulln (Tulln, Austria)*—Karl Matz; *University Hospital Basel (Basel, Switzerland)*—Phillippe Lyrer; *Dalhousie University and Queen Elizabeth II Health Sciences Centre (Halifax, Canada)*—Gord Gubitz, Stephen J Phillips; *and Instituto Nacional de Neurologia (Mexico City, Mexico)*—Antonio Arauz.

Conflicts of interest

EB and AC have received honoraria and travel costs from Boehringer Ingelheim. GB has received honoraria and speaker fees from Boehringer Ingelheim, Sanofi Synthelabo Aventis, Hoffman La Roche, and Novo Nordisk. AK has received lecture fees and conference travel costs from Boehringer Ingelheim. RIL has been paid for his role as a member of a conference scientific committee and for lectures by Boehringer Ingelheim and has attended national stroke meetings organised and funded by Boehringer Ingelheim. PS has received lecture fees (paid to the Division of Clinical Neurosciences, University of Edinburgh) and travel expenses from Boehringer Ingelheim, was a member of the independent data and safety monitoring board of the RE-LY trial funded by Boehringer Ingelheim for which attendance fees and travel expenses were paid (to the Division of Clinical Neurosciences, University of Edinburgh). KBS has received an honorarium for a lecture from Boehringer Ingelheim and had costs for participating in scientific meetings reimbursed; is a member of the European Medicines Agency's Committee for Medicinal Products for Human Use and the Cardiovascular Working Party. The views expressed in this Article are the personal views of KBS and should not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties. VM has received an unrestricted educational grant from Boehringer Ingelheim for a meeting on thrombolysis in stroke at which IST-3 was discussed. JMW received funding to the Division of Clinical Neurosciences, University of Edinburgh for reading CT scans for ECASS III from Boehringer Ingelheim, is the contact reviewer for Cochrane systematic reviews of thrombolytic treatment for acute stroke, has attended meetings held by Boehringer Ingelheim as an unpaid independent adviser during the licensing of alteplase, but was refunded her travel expenses and the time away from work, has attended and spoken at meetings organised and funded by Boehringer Ingelheim for which she received honoraria and travel expenses, and is director of the Brain Research Imaging Centre for Scotland, which has received some funding supplemented by grants and donations from Novartis, Schering, General Electric, and Boehringer Ingelheim. All other members of the writing committee declare that they have no conflicts of interest.

Acknowledgments

We thank all the patients who participated in the study, and the many individuals not specifically mentioned in the report who have supported the study. We also thank NIHR Stroke Research Network, NHS Research Scotland, and the National Institute for Social Care and Health Research Clinical Research Centre. IST-3 is an investigator-led trial. The University of Edinburgh and the NHS Lothian Health Board are cosponsors. The double-blind start-up phase was supported by the Stroke Association (UK). The expansion phase was funded by the Health Foundation UK. The main phase of the trial was funded by: UK Medical Research Council and managed by the National Institutes of Health Research on behalf of the MRC-NIHR partnership, the Research Council of Norway, Arbetsmarknadens Partners Forsakringsbolag Insurances Sweden, the Swedish Heart Lung Fund, The Foundation of Marianne and Marcus Wallenberg, Stockholm County Council, Karolinska Institutet, Polish Ministry of Science and Education, Australian Heart Foundation, Australian National Health and Medical Research Council, Swiss National Research Foundation, Swiss Heart Foundation, Foundation for Health and Cardio-/Neurovascular Research (Basel, Switzerland), Assessorato alla Sanita (Regione dell'Umbria, Italy), and Danube University (Krems, Austria). Boehringer Ingelheim GmbH donated drug and placebo for the double-blind phase, but thereafter had no role in the trial. The UK Stroke Research Network adopted the trial on May 1, 2006, supported the initiation of new UK sites, and in some centres, data were collected by staff funded by the Network or working for associated NHS organisations. Central imaging was done at the Brain Imaging Research Centre (University of Edinburgh), partly funded by the Scottish Funding Council and the Chief Scientist Office of the Scottish Executive. Additional support was received from Chest Heart and Stroke Scotland, DesAcc, University of Edinburgh, Danderyd Hospital R&D Department, Karolinska Institutet, Oslo University Hospital, and the Dalhousie University Internal Medicine Research Fund.

References

1 Wardlaw JM, Murray V, Berge E, et al. Recombinant tissue plasminogen activator for acute ischaemic stroke: an updated systematic review and meta-analysis. *Lancet* 2012; **379:** 2364–72.

2 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.

3 Kwiatkowski TG, Libman RB, Frankel M, et al. Effects of tissue plasminogen activator for acute ischemic stroke at one year. National Institute of Neurological Disorders and Stroke Recombinant Tissue Plasminogen Activator Stroke Study Group. *N Engl J Med* 1999; **340:** 1781–87.

4 Sandercock P, Lindley R, Wardlaw J, et al. The third international stroke trial (IST-3) of thrombolysis for acute ischaemic stroke. *Trials* 2008; **9:** 37.

5 Sandercock P, Lindley R, Wardlaw J, Whiteley W, Murray G, IST3 Collaborative Group. Statistical analysis plan for the third International Stroke Trial (IST-3); part of a 'thread' of reports of the trial. *Int J Stroke* 2012; **7:** 186–87.

6 Sandercock P, Lindley R, Wardlaw J, et al. 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.

7 Bruno A, Akinwuntan AE, Lin C, et al. Simplified modified Rankin scale questionnaire: reproducibility over the telephone and validation with quality of life. *Stroke* 2011; **42:** 2276–79.

8 Dennis M, Mead G, Doubal F, Graham C. Determining the modified Rankin score after stroke by postal and telephone questionnaires. *Stroke* 2012; **43:** 851–53.

9 Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. A prospective study of acute cerebrovascular disease in the community: the Oxfordshire Community Stroke Project—1981–86. 2. Incidence, case fatality rates and overall outcome at one year of cerebral infarction, primary intracerebral and subarachnoid haemorrhage. *J Neurol Neurosurg Psychiatry* 1990; **53:** 16–22.

10 van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; **19:** 604–07.

11 EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16:** 199–208.

12 EuroQol Group Monographs; EQ5D value sets; inventory, comparative review and user guide. Szende A, Oppe M, Devlin N, eds. Dordrecht: Springer; 2007.

13 Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ. A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health Qual Life Outcomes* 2010; **8:** 13.

14 Dorman P, Slattery J, Farrell B, Dennis M, Sandercock P, the United Kingdom collaborators in the International Stroke Trial. A randomised comparison of the EuroQol and Short Form-36 after stroke. *BMJ* 1997; **315:** 461.

15 Dorman P, Waddell F, Slattery J, Dennis M, Sandercock P. Are proxy assessments of health status after stroke with the EuroQol questionnaire feasible, accurate and unbiased? *Stroke* 1997; **28:** 1883–87.

16 Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? *Stroke* 1997; **28:** 1876–82.

17 Dorman P, Dennis MS, Sandercock P. Are the modified "simple questions" a valid and reliable measure of health related quality of life after stroke? *J Neurol Neurosurg Psychiatry* 2000; **69:** 487–93.

18 Pickard AS, Johnson J, Feeny D. Responsiveness of generic health-related quality of life measures in stroke. *Qual Life Res* 2005; **14:** 207–19.

19 Pickard AS, Johnson JA, Feeny DH, Shuaib A, Carriere KC, Nasser AM. Agreement between patient and proxy assessments of health-related quality of life after stroke using the EQ-5D and Health Utilities Index. *Stroke* 2004; **35:** 607–12.

20 Sandercock P, Collins R, Counsell C, et al. The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19 435 patients with acute ischaemic stroke. *Lancet* 1997; **349:** 1569–81.

21 Frank B, Grotta JC, Alexandrov AV, et al. Thrombolysis in stroke despite contraindications or warnings? *Stroke* 2013; **44:** 727–33.

22 Konig IR, Ziegler A, Bluhmki E, et al. Predicting long-term outcome after acute ischemic stroke: a simple index works in patients from controlled clinical trials. *Stroke* 2008; **39:** 1821–26.

23 Cohen DJ, Van Hout B, Serruys PW, et al. Quality of life after PCI with drug-eluting stents or coronary-artery bypass surgery. *N Engl J Med* 2011; **364:** 1016–26.

24 Sandercock P, Berge E, Dennis M, et al. A systematic review of the effectiveness, cost-effectiveness and barriers to the implementation of thrombolytic and neuroprotective treatment in the NHS. Health Technology Assessment 6[26]. Southampton: NHS Health Technology Assessment Programme; 2002.

25 Lyden PD. In anticipation of International Stroke Trial-3 (IST-3). *Stroke* 2012; **43:** 1691–94.

26 Calvert M. Reporting of patient-reported outcomes in randomized trials: the consort PRO extension. *JAMA* 2013; **309:** 814–22.

27 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 LANCET Neurology

Supplementary webappendix

This webappendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: The 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; published online June 21. http://dx.doi.org/10.1016/S1474-4422(13)70130-3.

Supplementary appendix to
Effect of thrombolysis with recombinant tissue plasminogen activator for acute ischaemic stroke on long-term outcomes in the third international stroke trial (IST-3) : a randomised controlled trial

The IST-3 Collaborative Group

Contents of appendix

Supplementary Tables

Web Table 1 Comparison of baseline characteristics of patients scheduled for 18-month follow –up versus those not scheduled for such follow-up

	Planned 18-month follow-up		No planned 18-month follow-up	
	No.	**(%)**	**No.**	**(%)**
Age in years				
18-50	106	(5%)	21	(3%)
51-60	164	(7%)	38	(6%)
61-70	311	(13%)	54	(8%)
71-80	595	(25%)	129	(19%)
81-90	1035	(44%)	372	(54%)
over 90	137	(6%)	73	(11%)
Female	1188	(51%)	382	(56%)
NIHSS				
0 to 5	471	(20%)	141	(21%)
6 to 10	653	(28%)	199	(29%)
11 to 15	479	(20%)	122	(18%)
16 to 20	426	(18%)	117	(17%)
> 20	319	(14%)	108	(16%)
Delay in randomisation				
0-3 hours	627	(27%)	222	(32%)
3-4.5 hours	952	(41%)	225	(33%)
4.5-6 hours	767	(33%)	240	(35%)
>6 hours	2	(0%)	0	(.%)
Atrial Fibrillation	678	(29%)	236	(34%)
Systolic BP				
<= 143 mmHg	760	(32%)	219	(32%)
144 - 164 mmHg	784	(33%)	232	(34%)
>= 165 mmHg	804	(34%)	236	(34%)
Diastolic BP				
<= 74 mmHg	685	(29%)	222	(32%)
75 - 89 mmHg	857	(37%)	272	(40%)
>= 90 mmHg	787	(34%)	193	(28%)
Blood glucose[1]				
<= 5 mmol/l	409	(20%)	130	(19%)
6 - 7 mmol/l	986	(48%)	316	(46%)
>= 8 mmol/l	671	(32%)	240	(35%)
Treatment with antiplatelet drugs in previous 48hrs	1209	(51%)	353	(51%)
Expert reader's assessment of acute ischaemic change				
Scan completely normal	201	(9%)	69	(10%)
Scan not normal but no sign of acute change	1130	(48%)	400	(58%)
Signs of acute change	1001	(43%)	216	(32%)
Predicted probability of poor outcome at 6 months[2]				
< 40%	1273	(54%)	378	(55%)
40% - 50%	243	(10%)	83	(12%)
50% - 75%%	579	(25%)	163	(24%)
>= 75%	253	(11%)	63	(9%)
Stroke syndrome				
TACI	1000	(43%)	306	(45%)
PACI	890	(38%)	256	(37%)
LACI	270	(11%)	62	(9%)
POCI	183	(8%)	63	(9%)
OTHER	5	(0%)	0	(.%)

NIHSS = National Institutes of Health Stroke Scale, TACI= Total Anterior Circulation Infarct, PACI = Partial Anterior Circulation Infarct, LACI = Lacunar Infarct, POCI = Posterior Circulation Infarct

1. For the first 282 patients recruited in the study, pre-randomisation glucose was not recorded; as a result, glucose measurements were available for 2066/2348 (88%).
2. Probability of a poor outcome calculated from a model based on age and baseline NIHSS[27]

Web Table 2 OHS at six months for patients with planned 18-month follow-up and known disability status

r	t-PA		Control		Adjusted analysis[1]		Unadjusted analysis[2]		(Control - rt-PA) difference per 1000[3] (95% CI)
	No. (%)	No	. (%)		OR rt-PA:Control (95% CI) P		OR rt-PA:Control (95% CI) P		
No. planned 18 month follow-up	1169		1179						
Missing 6 month status[4]	29		41						
No. for 6 month analysis[5]	1140		1138						
0 115		(10%)	89	(8%)					
1 170		(15%)	155	(14%)					
2 157		(14%)	173	(15%)					
3 184		(16%)	154	(14%)					
4 82		(7%)	105	(9%)					
5 152		(13%)	193	(17%)					
Died before 6 months[6]	309	(27%)	310	(27%)	0.94 (0.76 , 1.16)	0.554	0.99 (0.83 , 1.19)	0.942	1 (-35, 38)
Alive and independent (0+1+2)	437	(38%)	409	(36%)	1.18 (0.97 , 1.45)	0.101	1.11 (0.93 , 1.31)	0.237	-24 (-64, 16)
Alive and favourable outcome (0+1)	284	(25%)	244	(21%)	1.29 (1.04 , 1.61)	0.022	1.22 (1.00 , 1.48)	0.050	-35 (-69, -0)

1 Logistic regression of outcome on treatment group, adjusted for age, NIHSS and delay (all linear) and visible infarct on baseline scan. For OHS an ordinal analysis was also performed (with levels 0, 1, 2, 3 discrete and 4+5+6 grouped): adjusted common OR = 1·35 (95%CI 1·15 -1·59, p=0·0004)

2 Logistic regression of outcome on treatment group

3 Standard binomial test with normal approximation

4 Patients who did not return a 6 month form· Vital status at 6 months was known for all members of 18 month follow-up cohort·

5 The percentages by OHS category are based on these totals·

6 If all patients known to be alive (see note 4) are included in denominators percents dying by 6 months are : rt-PA 26·4%, Control 26·3%

Web Table 3 : EQ-5D and additional questions at six months (in planned 18-month follow-up population)

	rt-PA	Control			(Control - rt-PA) difference per 1000 (95% CI) [2]
	No (%)	No· (%)	Odds Ratio (95% CI)[1]	p	
Mobility (N= 815 rt-PA, 814 Control)					
No problems walking	320 (39%)	313 (38%)	1·14 (0·94 - 1·38)	0·189	-8 (-55, 39)
Some problems walking	393 (48%)	386 (47%)			-8 (-57, 41)
Confined to bed	102 (13%)	115 (14%)			16 (-17, 49)
Self-care (N= 819 rt-PA, 813 Control)					
No problems with self care	429 (52%)	394 (48%)	1·37 (1·13 - 1·67)	0·002	-39 (-88, 9)
Some problems washing or dressing	233 (28%)	211 (26%)			-25 (-68, 18)
Unable to wash or dress self	157 (19%)	208 (26%)			64 (24, 104)
Usual activities (N= 816 rt-PA, 804 Control)					
No problems with usual activities	271 (33%)	231 (29%)	1·39 (1·15 - 1·68)	0·001	-45 (-90, 0)
Some problems with usual activities	299 (37%)	288 (36%)			-8 (-55, 39)
Unable to perform usual activities	246 (30%)	285 (35%)			53 (7, 99)
Pain or discomfort (N= 806 rt-PA, 802 Control)					
No pain or discomfort	389 (48%)	381 (48%)	1·03 (0·85 - 1·25)	0·74	-8 (-56, 41)
Moderate pain or discomfort	366 (45%)	374 (47%)			12 (-36, 61)
Extreme pain or discomfort	51 (6%)	47 (6%)			-5 (-28, 19)
Anxiety or depression (N= 808 rt-PA, 801 Control)					
Not anxious or depressed	395 (49%)	377 (47%)	1·10 (0·91 - 1·33)	0·330	-18 (-67, 31)
Moderately anxious or depressed	344 (43%)	350 (44%)			11 (-37, 60)
Extremely anxious or depressed	69 (9%)	74 (9%)			7 (-21, 35)
Additional questions on overall function					
Stroke left patient with problems (N= 827 rt-PA, 821 Control)	596 (72%)	648 (79%)	1·54[3](1·20 – 1·92)	<0·0001	69 (27, 110)
Needs help with everyday activities (N= 827 rt-PA, 820 control)	387 (47%)	423 (52%)	1·32[3] (1·06 - 1·64)	0·011	48 (0, 96)

1 Logistic regression of outcome on treatment group, adjusted for age, NIHSS and delay (all linear) and visible infarct on baseline scan; odds ratio > 1 indicates treatment associated with greater odds of a lesser degree of problems
2 Standard binomial test with normal approximation
3. Odds ratio > 1 indicates increases odds of not being left with problems or needing help

Web Table 4 Treatment group difference in EQ utility and VAS index at six and 18 months: effects of different valuations

	Adjusted differences						Unadjusted differences					
	UK TTO		UK VAS		European VAS		UK TTO		UK VAS		European VAS	
	Mean (SE)	P	Mean (SE)	P	Mean (SE)	P	Mean (SE)	P	Mean (SE)	P	Mean (SE)	P
EQ-5D Utility at Six months	0·040 (0·019)	0·033	0·031 (0·014)	0·025	2·9 (1·3)	0·024	0·031 (0·021)	0·150	0·024 (0·016)	0·129	2·2 (1·4)	0·123
EQ-5D Utility at Six months (survivors to 18 months)	0·044 (0·019)	0·022	0·033 (0·015)	0·024	3·0 (1·3)	0·024	0·032 (0·022)	0·131	0·024 (0·017)	0·147	2·2 (1·5)	0·142
EQ-5D Utility at Eighteen months	0·060 (0·019)	0·002	0·047 (0·015)	0·002	4·3 (1·3)	0·001	0·049 (0·022)	0·028	0·038 (0·017)	0·025	3·5 (1·5)	0·024
Change : Eighteen minus Six months	0·007 (0·015)	0·650	0·008 (0·011)	0·478	0·7 (1·0)	0·462	0·005 (0·015)	0·735	0·007 (0·011)	0·551	0·6 (1·0)	0·535

TTO= Time trade-off, VAS= Visual analogue scale· The valuations are described and reported in EQ-5D value sets: Inventory, comparative review and user guide[12]

Web Table 5 Living circumstances at six and 18 months

	rt-PA		Control	
	N	%	N	%
6 months				
Ow n home	639	76·9	626	75·6
Relative's home	49	5·9	56	6·8
Residential home	45	5·4	41	5·0
Nursing home	93	11·2	96	11·6
Still in hospital	5	0·6	9	1·1
All 8	31	100·0	828	100·0
18 months				
Ow n home	574	81·0	553	78·2
Relative's home	29	4·1	40	5·7
Residential home	45	6·3	42	5·9
Nursing home	60	8·5	69	9·8
Still in hospital	1	0·1	3	0·4
All 7	09	100·0	707	100·0

Web Table 6. Adjusted subgroup effects on OHS at 18 months: Ordinal analysis

	O		rdinal analysis			
	Alive and independent / Total		Adjusted [1] Una		djusted	
rt-	PA	Control	OR (99% CI) [2] P	[3] OR	(99% CI) [2]	P[3]
Age [4]			·			·
<= 80 years	243/547 (44%)	249/569 (44%)	1·10 (0·82, 1·47)	0·032	1·07 (0·81, 1·42)	0·123
> 80 years	148/570 (26%)	103/553 (19%)	1·57 (1·11, 2·23)	·	1·38 (1·00, 1·90)	·
Delay [4]			·			·
<= 3	92/308 (30%)	63/300 (21%)	1·54 (0·96, 2·47)	0·492	1·47 (0·95, 2·26)	0·266
>3 - =<4·5	144/450 (32%)	145/449 (32%)	1·16 (0·81, 1·66)	·	1·06 (0·77, 1·48)	·
>4·5 - =<6	155/359 (43%)	144/371 (39%)	1·32 (0·91, 1·92)	·	1·15 (0·81, 1·63)	·
NIHSS [4]			·			·
0 to 5	157/224 (70%)	152/224 (68%)	1·23 (0·80, 1·90)	0·021	1·20 (0·78, 1·85)	0·025
6 to 15	194/539 (36%)	185/535 (35%)	1·14 (0·84, 1·54)	·	1·07 (0·80, 1·44)	·
16 to 24	35/293 (12%)	15/315 (5%)	2·33 (1·28, 4·25)	·	1·99 (1·13, 3·49)	·
>= 25 [5]	5/61 (8%)	0/48 (0%)	7·81 (0·82, 74·1)	·	5·21 (0·67, 40·8)	·
Respondent			·			·
Patient	263/318 (83%)	240/308 (78%)	1·36 (0·94, 1·97)	0·551	1·26 (0·87, 1·82)	0·631
Proxy	128/390 (33%)	111/401 (28%)	1·59 (1·12, 2·27)	·	1·32 (0·94, 1·85)	·
Trial phase			·			·
Blinded	34/131 (26%)	31/136 (23%)	1·29 (0·64, 2·58)	0·955	1·17 (0·61, 2·22)	0·994
Open	357/986 (36%)	321/986 (33%)	1·30 (1·02, 1·65)	·	1·17 (0·94, 1·46)	·

1 Analysis by age adjusted for linear effects of delay and NIHSS; analysis by delay adjusted for linear effects of age and NIHSS; analysis by NIHSS adjusted for linear effects of age and delay; analysis by proxy group and trial phase adjusted for linear effects of age, NIHSS and delay. Additionally all analyses adjusted for visible ischaemia on baseline scan.

2 Odds ratio for effect of rtPA on an improvement shift of one level in OHS where this scale has been grouped into 5 levels (see text).

3 Wald test of interaction between treatment and subgroup factor.

4 Trend tests of treatment interaction give P=0·24, 0·004 and 0·41 for age, NIHSS and delay respectively, in each case adjusting for the other variables· Unadjusted trend tests give P=0·67, 0·006 and 0·34 respectively.

5 Logistic model for this subgroup has quasi-complete separation, hence OR estimates are unreliable.

Web Table 7. Adjusted subgroup effects on EQ utility index at 18 months

	Mean Utility (SE) [1] Adj		usted Analysis [2] Unadj			usted Analysis [3]	
rt-PA		**Control**	**Adjusted difference (SE)**	**P**	**Interaction test [4]**	**Unadjusted difference (SE)**	**P**
Age [5]					·		
<= 80 years	0·599 (0·020)	0·558 (0·020)	0·045 (0·025)	0.071	0·277	0·041 (0·028)	0·141
> 80 years	0·480 (0·024)	0·409 (0·026)	0·087 (0·033)	0.009	·	0·071 (0·035)	0·043
Delay [5]					·		
<= 3	0·475 (0·033)	0·423 (0·036)	0·081 (0·046)	0.080	0·717	0·051 (0·049)	0·297
>3 - =<4·5	0·536 (0·025)	0·502 (0·026)	0·033 (0·032)	0.293	·	0·034 (0·036)	0·344
>4·5 - =<6	0·617 (0·023)	0·546 (0·023)	0·081 (0·030)	0.008	·	0·071 (0·033)	0·030
NIHSS [5]					·		
0 to 5	0·726 (0·024)	0·715 (0·020)	0·017 (0·031)	0.580	0·111	0·012 (0·031)	0·707
6 to 15	0·537 (0·021)	0·499 (0·022)	0·055 (0·030)	0.070	·	0·039 (0·030)	0·198
16 to 24	0·342 (0·037)	0·200 (0·035)	0·146 (0·050)	0.004	·	0·142 (0·050)	0·005
>= 25	0·212 (0·089)	0·018 (0·137)	0·251 (0·164)	0.140	·	0·194 (0·168)	0·258
Respondent					·		
Patient	0·759 (0·015)	0·737 (0·016)	0·026 (0·021)	0.217	0·403	0·023 (0·022)	0·307
Proxy	0·378 (0·021)	0·324 (0·021)	0·072 (0·028)	0.011	·	0·054 (0·030)	0·072
Trial phase					·		
Blinded	0·469 (0·051)	0·386 (0·052)	0·087 (0·068)	0.200	0·487	0·084 (0·073)	0·253
Open	0·560 (0·016)	0·516 (0·016)	0·059 (0·021)	0.005	·	0·043 (0·023)	0·060

1. Utility determined from EQ5D responses using UK valuations on time trade-off basis.
2. Linear regression of utility on treatment plus adjustment factors. Analysis by age adjusted for linear effects of delay and NIHSS; analysis by delay adjusted for linear effects of age and NIHSS; analysis by NIHSS adjusted for linear effects of age and delay; analysis by proxy group and trial phase adjusted for linear effects of age, NIHSS and delay. Additionally, all analyses adjusted for visible ischaemia on baseline scan.
3. t-tests of difference in mean utility between treatment groups within each subgroup.
4. F test of interaction between treatment and subgroup factor in linear model with adjustment variables.
5. Trend tests of treatment interaction give P=0·21, 0·008 and 0·56 for age, NIHSS and delay respectively.

Web Table 8 Recall at 18 months of thrombolytic drug administration on day of hospital admission among the cohort scheduled for 18-month follow-up

	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

1. The question about recall of treatment was first asked towards the end of the double blind phase of the trial. Some patients who were recruited in the open phase were not sent the correct questionnaire, and so were inadvertently not asked the question

Committees and List of members of IST-3 collaborative group

Trial Steering Committee:
Independent Chairman, Professor Colin Baigent (University of Oxford), Professor David Chadwick (Former Chairman, Retired)): University of Liverpool; Independent members: Dr Pippa Tyrrell (Manchester University), Professor Gordon Lowe (Glasgow University); Co-principal Investigators: Professor Peter Sandercock (University of Edinburgh); Professor Richard Lindley (Sydney Medical School – Westmead Hospital, and the George Institute for Global Health, University of Sydney); Chief Investigator for Neuroradiology: Professor Joanna Wardlaw (University of Edinburgh); Professor Martin Dennis (University of Edinburgh); Statistician: Geoff Cohen; Trial Co-ordinator: Karen Innes. Lay representative: Heather Goodare.

CT and MR reading panel
Joanna Wardlaw, Andrew Farrall, Rüdiger von Kummer, Lesley Cala, Anders von Heijne, Zoe Morris, Alessandro Adami, Andre Peeters, Gillian Potter, Nick Brady.

Data Monitoring Committee
Professor Rory Collins, Oxford University, UK (Chairman), Professor Philip Bath (Nottingham University), Professor Jan van Gijn (former member, now retired) (University of Utrecht, The Netherlands), Professor Richard Gray (University of Oxford), Professor Robert Hart (McMaster University, Canada), Professor Salim Yusuf (McMaster University, Canada).

Event Adjudication Committee
Keith Muir, Peter Sandercock, Richard Lindley.

National Co-ordinators and Associate National Co-ordinators
Australia: Richard Lindley, Graeme J. Hankey. Austria: Karl Matz, Michael Brainin. Belgium: Andre Peeters. Canada: Gord Gubitz, Stephen Phillips. Italy: Stefano Ricci. Mexico: Antonio Arauz. Norway: Eivind Berge, Karsten Bruins Slot. Poland: Anna Czlonkowska, Adam Kobayashi. Portugal: Manuel Correia. Switzerland: Phillippe Lyrer, Stefan Engelter. Sweden: Veronica Murray, Bo Norrving, Andreas Terént, Per Wester: UK: Graham Venables.

Trial Coordinating Centre, Edinburgh. Karen Innes, Alison Clark, David Perry, Vera Soosay, David Buchanan, Sheila Grant, Eleni Sakka, Jonathan Drever, Pauli Walker, Indee Herath, Ann Leigh Brown, Paul Chmielnik, Christopher Armit, Andrea Walton, Mischa Hautvast, Steff Lewis, Graeme Heron, Sylvia Odusanya, Pam Linksted, Ingrid Kane, Will Whiteley, Robin Sellar, Philip White, Peter Keston, Andrew Farrell, Zoe Morris, Hector Miranda. **CTSU, Oxford**, Lisa Blackwell. **National Co-ordinating centres**: ***Italy*** up to September 2008: Maria Grazia Celani; Enrico Righetti. After September 2008: Silvia Cenciarelli; Tatiana Mazzoli. Central Follow up for Italy: Teresa Anna Cantisani. ***Poland*** Jan Bembenek. ***Sweden*** Eva Isaksson. ***Norway*** Eivind Berge, Karsten Bruins Slot.

List of members of IST-3 collaborative group and participating hospitals in each country.
Figures in parentheses are the number of patients recruited in the country or by the centre. ***UK*** *(1447)* Royal Hallamshire Hospital (118): G Venables, C Blank, H Bowler, C Doyle, K Endean, K Harkness, E Parker, M Randall. University Hospital of North Staffordshire (97): C Roffe, N Ahmad, A Arora, S Brammer, J Chembala, B Davies, S Ellis, E Epstein, K Finney, C Jackson, C Jadun, R Kinston, H Maguire, I Memon, I Natarajan, M Poulson, R Sanyal, S Sills, A Vreeburg, E Ward. Western General Hospital (95): P Sandercock, R Al-Shahi Salman, R Davenport, M Dennis, P Hand, S Hart, I Kane, S Keir, M MacLeod, L McKinlay, H Milligan, E Sandeman, J Stone, C Sudlow, P Taylor, J Wardlaw, C Warlow, W Whiteley, A Williams. The National Hospital for Neurology & Neurosurgery (84): M Brown, B Athwal, V Bassan, N Bhupathiraju, J Bowler, C Davie, D Doig, R Erande, S Gilbert, L Ginsberg, R Greenwood, S Gregoire, N Harding, N Losseff, R Luder, N Passeron, R Perry, P Rayson, R Simister, S Stone, D Werring. Arrowe Park Hospital (83): J Barrett, H Aitken, S Cherian, R Davis, S Downham, L Godd, V Gott, D Jose, V Little, D Lowe, L Luxford, M McGrory, P Owings, N Price, J Richards, G Sangster, J Sherlock, S Vargese, I Wakefield, P Weir. Southend University Hospital (77): P Guyler, T Attygale, S Chandler, L Coward, S Feasey, C Khuoge, T Loganathan, S Martin, A O'Brien, D Sinha, V Thompson, S Tysoe, R Walsh. Norfolk and Norwich University Hospital (67): K Metcalf, J Cochius, R Fulcher, N Gange, C Green, J Jagger, M Lee, P Myint, J Potter, G Ravenhill, S Shields, N Shinh, T Staunton, E Thomas, W Woodward, P Worth, N Wyatt. Nottingham City Hospital (63): W Sunman, P Bath, P Berman, J Clarke, C Gaynor, F Hammonds, R Harwood, K Mitchell, S Munshi, S Pacey, A Shetty, N Sprigg, H Stear, G Subramanian, A Wills. Guy's & St.Thomas Hospital (60): A Rudd, H Audebert, A Bhalla, J Birns, R Chowdhury, G Cluckie, I Davies, C Gibbs, P Holmes, N Mitchell, F Schiavone, E White, M Yeung. Darlington

and Bishop Auckland Hospitals (56): A Mehrzad, V Baliga, E Brown, L Burnside, B Esisi, J Kent, P Orr, D Stead, E Wayman. University Hospital Aintree (46): R Durairaj, C Cullen, R Kumar, H Martin, D McDowell, A Sharma, V Sutton, R White. University Hospital of Wales (46): T Hughes, K Ali, J Anderson, K Baker, K Bethune, K Bethune, M Booth, M Cossburn, S Halpin, M Hourihan, E Marsh, K Peall, R Powell, H Shetty, M Wardle, M Williams. Derby Royal Hospital (37): K Muhiddin, J Beavan, M Clarke, R Donneley, S Elliott, P Fox, P Gorman, M Harper, M Mangoyana, I Memon, L Mills, L Wright. Addenbrookes Hospital (34): L Warburton, J Baron, P Barry, D Day, T Harold, P Martin, J Mitchell, E O'Brien, J Rycarte, M Turnham. St George's Healthcare NHS Trust (34): G Cloud, L Choy, B Clarke, C Griffin, O Halse, I Jones, F Kennedy, U Khan, R Lewis, A Loosemore, C Lovelock, H Markus, B Moynihan, J O'Reilly, O Paul, A Pereira, M Punter, P Rich, D Rolfe, F Schiavone. Royal Devon & Exeter Hospital (Wonford) (30): M James, J Bell, A Bowring, L Boxall, J Cageao, H Eastwood, S Elyas, F Hall, S Harries, A Hemsley, S Jackson, S Keenan, P Mudd, A Sekhar, D Strain, J Sword, N Wedge. Aberdeen Royal Infirmary (26): M MacLeod, M Bruce, A Joyson, M Kemp, K McMullan, J Reid, O Robb, J Webster, S Wilkinson. Hammersmith Hospitals & Imperial College (24): P Sharma, P Bentley, H Jenkins, A Kar, T Sachs. Northwick Park Hospital (20): D Cohen, R Bathula, J Devine, M Mpelembue. William Harvey Hospital (20): D Hargroves, I Balogun, L Cowie, A Maidment, D Rand, J Rowe, H Rudenko, D Smithard, L Wray. Scarborough Hospital (17): J Paterson, J Brown, J Hampton, S Jamieson, R Rose, A Volans. Countess of Chester Hospital NHS Foundation Trust (17): K Chatterjee, G Abbott, R Brookes, C Castle, C Kelly, S Leason, A Nallasivan, A Sen. Watford General Hospital (17): D Collas, M Cottle, N Damani, P Jacob, D Oza, D Werring. University Hospitals Coventry & Warwickshire NHS Trust (15): A Kenton, N Adab, L Aldridge, H Allroggen, Y Brown, R Cross, L Galvin, K Ghosh, A Grubneac, A Lindahl, H Mehta, M Pritchard, C Randall, P Ray, A Shehu, S Thelwell. Royal Bournemouth & Christchurch NHS Trust (12): D Jenkinson, J Bell, T Black, O David, J Kwan, A Orpen, C Ovington, D Tiwari, Z ud Din Babar. Leeds General Infirmary (12): A Hassan, A Bailey, J Bamford, C Bedford, R Bellfield, J Cooper, L Dunsmure, J Greig, M Keeling, L Mandizvidza, J Rankine, E Roberts, P Wankyn, T Webb, S Williamson. York Health Services NHS Trust (12): J Coyle, S Crane, C Croser, P Duffey, R Evans, E Iveson, M Keeling, G Kitching, M Porte, C Rhymes. Queen Elizabeth Hospital (Gateshead) (12): D Barer, M Armstrong, M Bokhari, T Cassidy, B McClelland. Queen Elizabeth The Queen Mother Hospital (10): G Gunathilagan, P DOLKE, S Jain, S Jones, A Maidment, L Rosser, G Thomas, C White. Worcestershire Royal Hospital (10): P Sanmuganathan, C Scholtz, E Stratford. Blackpool Victoria Hospital (10): M O'Donnell, H Goddard, G Hoadley, J Howard, S Leach, J McIlmoyle, A Stewart, A Strain. Basildon & Thurrock University Hospitals NHS FT (9): F Huwez, P Croot, N Gadi, N Mguni, U Umasankar. Royal Infirmary of Edinburgh (8): G Mead, B Chapman, A Coull, S Hart, A Kinnear, B Morrow, F Morrow. St Mary's Hospital (8): D Ames, J Ball, S Bannerjee, J Chataway. Yeovil District Hospital (8): K Rashed, C Buckley, D Donaldson, D Hayward, C Lawson. Luton and Dunstable Hospital (8): L Sekaran, K Bharaj, F Justin, G Jutlla, D Phiri, S Sethuraman, M Tate. Solihull Hospital, Heart of England NHS Trust (8): D Sandler, P Carr, G Jones, J Lyons, K Warren. King's College Hospital (7): L Kalra, A Davis, J Jarosz, D Manawadu, L Sztriha. Doncaster Royal Infirmary (7): D Chadha, A Holford, P Willcoxson. Royal United Hospital Bath (7): L Shaw, D Button, A Cunningham, L Dow, J Dutson, T Hall, C Hardy, N Jakeman, P Kaye, B Madigan, K O'Brien, D Pressdee, M Price, L Robinson, C Taylor, D Williamson. Birmingham Heartlands Hospital (6): D Sandler, P Carr, J Lyons, J McCormack, C Stretton. University Hospital North Durham (6): P Earnshaw, E Brown, S Bruce, C Church, S Desai, B Esisi, M Myint, N Watt. Wansbeck General Hospital (6): C Price, S Elliott, H Graham, R Lakey, K Mitchelson. Bristol Royal Infirmary (6): P Murphy, L Ball, S Caine, J Dovey, J Hughes, A Steele. Stepping Hill Hospital (6): K Dizayee, A Brown, T Chattopadhyay, J Cheetham, H Cochrane, A Datta, M Datta-chaudhuri, C Fox, D Kilroy, S Krishnamoorthy, F Levy, S Metha, P Ngoma, B Venkatesh. Princess Royal Hospital Brighton & Sussex University Hospitals Trust (5): K Ali, R Gautam, N Henderson, M Jones, S Murphy, G Spurling. Belfast City Hospital (5): I Wiggam, C Boyd, K Fullerton, P Gray, M Kinnaird, S MacNair, C Morgan, M Reid, S Tauro. Royal Liverpool University Hospital (5): S Loharuka, D Balmforth, P Cox, G Fletcher, A Ledger, A Manoj, M Wilkinson. City Hospital, Sandwell & West Birmingham Hospital (5): D Nicholl, S Clegg, S Hurdowar, S Kausar, K Law, A Singal, S Sturman. Royal London Hospital (4): P Gompertz, J Evanson, A Farrell, A Petrou, K Saastamoinen, T Sachs, A Salek-Haddadi, R Yadava. Sunderland Royal Hospital (4): J O'Connell, H Brew, S Butler, S Crawford, C Gray, D Gulliver, N Majmudar, R O'Brien. Morriston Hospital (4): M Wani, L Dacey, L Davies, R Evans, D Harris, T Jones, S Storton. Royal Preston Hospital (4): S Punekar, A Ashton, S Duberley, H Emsley, C Gilmour, B Gregary, L Hough, S Philip, S Wuppalapati. The Royal Wolverhampton Hospitals NHS Trust (4): K Fotherby, P Bourke, D D'Costa, K Kauldhar, D Leung, R Lodwick, S McBride, D Morgan, M Qaiyum, G Sahota, M Srinivasan. Royal West Sussex NHS Trust, St Richard's Hospital (4): I Kane, N Chuter, L Garrad, M Hookway, S Ivatts, G Kennedy. Queen's Hospital Romford (4): K Darawil, L Al Dhahirl, S A ndole, M Baig, P Dugh, K Dunne, H Kariuki, M Khan, S Rathnayaka. Ulster Hospital (3): M Power, K Dynan, J Finnerty, A Heaney, C Leonard, K McKnight, J Turkington, B Wroath. Great Western Hospital (3): B Dewan, S Cotton, M Gardiner, T Saunders, B Vincent. The Queen Elizabeth Hospital Birmingham (3): D Sims, P Guest, E Jones, J McCormack, D Nicholl, J Savanhu, R Tongue,

M Willmot. Leicester General Hospital (3): D Eveson, S Dawson, M Dickens, M Fotherby, R Hunt, S Khan, T Kumar, R Marsh, A Mistri, T Robinson, J Thompson. Darent Valley Hospital, Dartford & Gravesham NHS Trust (3): P Aghoram, T Daniel, M Gatehouse, S Hussein, A Jackson, T Shangany a, E Strachan, G Tan. Nevill Hall Hospital, Aneurin Bevan Local Health (3): B Richard, S Elaine, S Hanson, S Mosely, H Reed, M Williams. Colchester Hospital University Foundation Trust (3): R Saksena, S Cook, D Demuran, M Keating, R Needle, V Paramsothy, A Sebastian, R Sivakumar, A Wright. Salford Royal Hospital Foundation NHS Trust (2): R Grue, E Barberan, C Dickson, C Douglas, J Jellicoe, T Marsden, J Priestley, E Quick, C Sherrington, A Singh, C Smith, J Stevens, P Tyrell, J Wainwright. Leicester Royal Infirmary (2): M Ardron, J Birchall. Queen Elizabeth Hospital (Kings Lynn) (2): R Shekhar, C Barsted, S Coleman, S Fletcher, J Graham. John Radcliffe Hospital (2): A Buchan, J Hinkle, J Kennedy, A Manoj, M Westwood. Derriford Hospital (2): A Mohd Nor, S Allder, B Hyams, A Pace. West Cumberland Hospital (1): E Orugun, C Brewer, L Huntley, R Jolly, C Summers. Sandwell General Hospital (1): K Sharobeem, J Khaira, J Leahy, E Linehan, G Moore, J Rizkalla, J Wilkinson. Torbay Hospital (1): D Kelly, C Hilaire. Warrington & Halton Hospitals NHS Foundation Trust (1): O Otaiku, L Connell, G D elaney-Sagar, G James, L Lomax, D Matthew, J Simpson, H Whittle. Medway Maritime Hospital (1): S Sanmuganathan, S Burrows, A Mahmood. Southampton General Hospital (1): G Durward, S Barker, J Cantle, P Crawford, S Evans, V Pressly, N Weir. Victoria Hospital (1): V Cvoro, K McCormick. ***Poland*** *(347)* 2nd Department of Neurology, Institute of Psychiatry & Neurology (190): A Czlonkowska, J Bembenek, M Bilik, G Chabik, W Czepiel, J Dzierka, M Gluszkiewicz, K Grabska, B Janus-Laszuk, J Jedrzejewska, A Kobayashi, T Litwin, A Oskedra, A Piorkowska, M Skowronska, A Sliwinska, U Stepien. SPZZOZ w Sandomierzu (43): P Sobolewski, A Gajewska, M Grzesik, R Hatalska- Zerebiec, I Labudzka, B Loch, A Medrykowska, M Sledzinska, A Sobota, W Szczuchnik, G Wolak, I Zdyb. Medical University of Gdansk (35): W Nyka, D Gasecki, K Chwojnicki, A Gojska, B Karaszewski, G Kozera, M Kwarciany, M Nowak, M Swierkocka-Miastkowska, S Szczyrba, M Wisniewska, E Wnorowska. 1st Department of Neurologyt, Institute of Psychiatry & Neurology (25): P Richter, A Bochynska, M Chahwan, A Graban, R Rola. Military Medical Institute (24): A Stepien, B Brodacki, M Grotowska, J K otowicz, J Staszewski, J Swistak, S Zaloga. Szpital Powiatowy (14): J Stoiñski, K Czajkowska-Fornal, P Czubak, A Kaczor, J Kraska, E Nowakowska-Sledz, J Ozdoba-Rot, E Zawadzka. Szpital Specjalistyczny w Konskich (8): M Fudala, D Adamczyk, W Brola, I Guldzinska, K Kaluzny, M Kucharska-Lipowska, M Mosiolek, M Polewczyk, M Ziomek. Central University Hospital (7): G Opala, M Arkuszewski, M Kudlacik, P Malgorzata, M Swiat. SPSK im. Prof. W. Orlowskiego CMKP (1): U Fiszer, M Lenska-Mieciek. ***Italy*** *(326)* Ospedale Di Citta' di Castello (62): S Cenciarelli, A Barilaro, R Condurso, F Coppola, S Dioguardi, E Gallinella, A Mattioni, C Menichetti, S Ricci. Nuovo Ospedale Civile "S.Agostino-Estense" (40): F Casoni, M Bacchelli, M Cavazzuti, M Malagoli, A Zini. Ospedale Beato Giacomo Villa - Citta' della Pieve (36): G Benemio, M Celani, R Allegrucci, V Bondo, S Cupella, L Guerra, S Guerrieri, C Ottaviani, E Righetti, C Rossi, N Sacchi, M Scucchi, V Stefanini. Ospedale Di Branca (36): T Mazzoli, A Bigaroni, L Greco, R Paris, P Parise, S Ricci. A.O. Niguarda Ca'Granda (25): A Ciccone, L Basso, R Causarano, P Doneda, E Ferrante, A Gatti, A Guccione, A Gullo, F Imbesi, S Jann, R Marazzi, E Moro, C Motto, D Parodi, A Protti, M Riva, A Rosiello, I Santilli, R Sterzi, P Tiraboschi, G Venturelli. R.Guzzardi Hospital - Vittoria (RG) (19): F Iemolo, R Campagna, G Campagnolo, A Carnemolla, N D'Apico, G D'Asta, S Giannarita, A Giordano, E Sanzar. Ospedale Regionale di Aosta (17): E Bottacchi, S Cordera, G Corso, M Di Giovanni, G Giardini, C Lia, T Meloni, M Pesenti Campagnoni, P Tosi. Ospedale Sacro Cuore - Negrar Verona (12): A Adami, G Rossato, T Zuppini. Ospedale Maggiore - Bologna (12): G Procaccianti, T Sacquegna. Universita degli Studi di Genova (11): C Gandolfo, M Balestrino, C Bruno, L Castellan, M Del Sette, A Ferrari, C Finocchi, N Reale, D Rizzi. Ospedale Civile S.Andrea (9): M Del Sette, L Benedetti, C Capellini, E Carabelli, E Cibei, M Godani, G Guariglia, E Landini, A Mannironi, B Nucciarone, S Parodi, S Tonelli, E Traverso, D Zito. Ospedale S. Giovanni Battista - Foligno (8): P Brustenghi, F Corea, O Flamini, S Lolli, G Pelliccia, R Ricci, S Stefanucci, M Zampolini. Ospedale a Vibo Valentia (8): D Consoli, F Galati, P Postorino. Azienda Ospedaliero-Universitaria "Ospedali Riuniti" di Foggia (8): G Rinaldi, E Carapelle, G Grilli, M Guido, L Specchio. Ospedale Valduce di Como (8): N Checcarelli, G Borin, L Chiveri, R Clerici, E Coreng ia, L Gandola, P Garavaglia, M Guidotti, A Martegani, M Mauri, F Muscia, F Raudino. Ospedale di Cattinara - Trieste (4): F Chiodo Grandi, A Bratina, N Carraro, M Gaio, A Granato, N Koscica, M Naccarato, V Sarra, P Schincariol, C Vilotti, Z Zugna. Clinica Dr Pederzoli Spa (4): D Idone, C Bonato, E De Angelis, A Forgione, M Gambera, F Recchia, S Tamburin, P Tinazzi Martini, G Zanette. Ospedale Civile San Matteo Degli Infermi - Spoleto (2): S Grasselli. Ospedale Silvestrini - Perugia (2): G Agnelli, A Andrea, A Billecia, V Caso, V Casso , R Fabiola, P Fanelli, M Paciaroni, B Sergio, M Vemti. Mater Salutis Hospital, Legnago VR (2): M Silvestri, L Altarini, A Bonfante, M Bonornetti, B Costa, N D'Attoma, N Deluca, F Frattini, R Niego, D Rafaele, V Ravenna, M Turazzini. Ospedale Guglielmo da Saliceto - Piacenza (1): S Cammarata. ***Sweden*** *(297)* Uppsala University Hospital (100): E Lundström, L Jonsson, U Söderström, A Terént. Danderyd Hospital (46): V Murray, A Alvelius, M Arbin von, I Dalenbring, Å Doverhall, Å Franzén-Dahlin, N Greilert, M Hallberg, A Heijne von, E Isaksson, H Kumpulainen, A Laska, A Lundström, C Martin, J Muhrbeck, E Näslund, N Ringart, E Rooth, R Undén, P Waldenström. Hassleholm Hospital (29): M

Esbjornsson, M Petr anek. Karnsjukhuset (25): B Ce derin, E Ber tholds, A Elg åsen, T Johansson, B Witteborn. Koping Hospital (20): M Kwi atkowska, E Gustafsson, T Noren, J Saaf . Mora Hos pital (17): J Teichert, M B ertilsson, S Nilsson, S Oestber g. LidkopingHospital (11): L Welin, K Fredr icson, L Pehn. Falu Hospital (11): J Ha mbraeus, I Lonn. Capio S: tGor an Hospita l (9): B Hojebe rg, A Adol fsson, M Anzen. Vastervik Hos pital (5): T Wallen, R Schloenzig , P Sö derström, A We nnerberg. University Hospital MAS (5): F Buchwald, K Abul-Kasim, A Berkeskold, J Petersson, E Poromaa. University Hospital of Northern Sweden (4): P Wester , R Backlun d, A Sj öström. Helsingborgs lasar ett (4): B Hedst röm , E Ca mpbell, K Johnsson, B Karlsson, N Lekoko tla, C Lu ndahl, A Rise dal, P Sandgr en, A Svensson . Visby Hospit al (4): S Bysell, E Smedberg, A Vestber g By sell. Sundsvall Hosp ital (3): V Sjögr en, B Hö gvall. University Hospit al Lund (2): G A ndsberg, T Cronber g, A Lin dgren. Vasteras Hospital (1): H Wannber g, F Ax, L Nyr en. Karlstad Central Hospital (1): J San ner, H An dersson, F An dler, S Hol mgård, R Johansson, I Magnussan, K Nilsson, J Rådberg. ***Norway*** *(204)* Trondheim University Hospital (69): B Indredavik, H Ellekjær, A Østvik, G Rohweder, D Steckhan, J Storvold. Oslo University Hospital Sykehus (66): E Berge, Y Rønning, R Aakvik, K Bruins Slot, G Knutsen, M Moxness, R Petter sen, T Wyller. University Hospital of North Norway (23): C Wahl, O I versen, S Johnsen, B Norderhus, L Steffensen, E Stensland. Kongsvinger Hospital (13): T As ak, J Aaseth, T Rotnes, J Sparby, S Wetterhus. Levanger Hospital (9): H Hallan, A Aardal, T Graven, H Hansbakk Skjetne, B Klykken, K Lindqvist, A Tommy. University Hospital of North Norway (8): T Engstad, M A ntonsen, R Baj ic, W Fønnebø, S Hykkerud, I Lyngmo, A Nyrnes, S Rogn e, S Sparr. Harstad Hospital (7): O Kildahl-Andersen, K Pedersen, H Ulrichsen. Ålesund Hosp ital (4): O Skogen, I Alnes, R Hukar i, Y Seljeseth, P Vadset. Asker an d Bær um Hospital (2): G Knutse n, B Fur e, H Ihle-Hansen, N John sen, L Kor nberg. Namsos Hospital (2): S Schuler , M Heibert. Volda Hospi tal (1): M Lillebø , O Aasen, I Eskeland, T Ha mre, S Hareide, H Helset, K Kolnes, B Lødemel, H Ose Velle, S Rei te, E Velle. ***Australia*** *(179)* Nambour General Hospital (51): R Grimley, E Ahern, C Cocks, M Cour tney, R Devin, J Endacott , C Fawce tt , V Harrington, C Johnston, M Koltermann, S Mur ray, K Ng, G Sty les, A Tampiyappa. John Hunter Hospital (29): C Levi, K Ch ung, L Dar k, M Evans, Y Gawar ikar, E Ker r, A Loiselle, F Mit eff, A Moor e, W O'Br ien, M Par sons, D Quain, A Royan, M Russell, N Spratt. Gosford Hospital (24): J St urm, D Cri mmins, D Gri ffiths, P Kavelieros, J Kinsella, A M alhotra, B O'Brien, A Schutz, M Webb, S Why te, V Z enteno. Westmead Hospital (16): R Lindley , A Bleasel, N Cordato, A Dugg ins, V Fung, L Go mes, N I ngham, J I p, P Landa u, J Mor ris, S Vucic. Royal Per th Hospital (15): G Hankey, A Claxton, N Lillywhite. The Canberra Hospital (12): C Lueck, C Andrews, G Danta, C Das, I Harvey, A Hughes, C McColl, A Oo n, R Tuck. Royal Br isbane and Wo men's Hospital (10): S Read, M B adve , M Broad, G Cad igan, H Ca vanagh, J Chal k, D Cops inis, K Ether ington, R Hender son, R Hul l, J O'S ullivan, J Pandian, L Ross- Lee, M Roxas, N Sheikh, G Skinner , A Wong. Austin Health - Repatr iation Ca mpus (8): H Dewey, A Brodtmann, G Donnan, A Hughes, M Karonen, H Ma, T Mulcahy, S Petrolo, L Walker, D Y oung, J Zavala. Nepean Hos pital (7): M Th ieben, C Harris, M Krause, S Lane, H Park, M Shaffi, J Wo od. Box H ill Hospital (7): C Bladin, A Buckland, K Cou ghlan, B Coulton, A Gilligan, P Lee, S Mullen, Z Ross, P Sien Loh, C Szoeke. ***Portugal*** *(82)* UAVC. Centro Hospitalar de Trás-os-Montes e Alto D ouro (42): M Silva, F A fonso, J Gabriel, P Gui marães, A Vel on. Hospital Per o da Covi lhã (19): M Castelo- Branco , F Alvarez, V Branco, C Coxo, P Go ulao, D Leal, S Mor gado, R O liveira, F Paiv a, A Rodr igues, M Si moes. Hospital de Santo António (12): G Lopes, T Al meida, M Car doso, J Ch aves, C Cor reia, M Correia, J Damásio, R Felgueiras, J Pereira, A Tuna. Hospital S.Marcos (9): C Fe rreira, E Lourenco, A Machado, R Mare, J Rocha. ***Belgium*** *(73)* Cliniques Universitaires St Lu c (73): A Peeters. ***Austria*** *(46)* Landesklin ikum Donauregion Tulln (34): K Matz, M Brainin, G Funk, V Reiner -Deitemyer. Krankenhaus Der Barmherzigen Br uder Wien (9): J Ferrari, A Flamm-Horak, G Gruber, R Ratti nger. Krankenhaus Gö ttlicher Heiland (3): W Muell bacher, D Doppe lbauer, R Kalchmayr, W Schi ma, T Wieser, M Zart. ***Switzerland*** *(23)* Universitatsspital Basel (2 2): P Ly rer, L Bonati, S Engelter, F Flur i, S Muller , E Radue, A Ti emessen, L Wa lz, F Weisskop f, S Wetzel. Universitätsspital Zürich (1): A Luft, D Fetz, B Hertler, A Pangalu. ***Canada*** *(8)* QEII Health Sciences Centre (8): G Gubitz, P Boulton, J Jarrett, J Moeller, S Phillips. ***Mexico*** *(3)* Instituto Nacional de Neurolo gia y Neurojrugia MVS (3): A Arauz, L Bermudez, J Calleja, R Garcia.

ANNEX 4: Re-analysis of NINDS study

Report of the t-PA Review Committee

W. Michael O'Fallon, Ph.D., Chair

Kjell Asplund, Ph.D., M.D.

Lewis R. Goldfrank, M.D.

Vicki Stover Hertzberg, Ph.D.

Timothy J. Ingall, MB, BS, Ph.D.

Thomas A. Louis, Ph.D.

August 25, 2004

t-PA Review Committee Roster

Chair

W. Michael O'Fallon, PhD
Professor of Biostatistics
Division of Biostatistics
Mayo Clinic Rochester
Rochester, Minnesota

Committee Members

Kjell Asplund, MD, PhD
Professor of Medicine
Head, Department of Medicine
University Hospital
Umeå, Sweden

Lewis R. Goldfrank, MD
Director, Emergency Medicine
Bellevue Hospital Center
Professor, Clinical Medicine and Surgery
New York University School of Medicine
New York, New York

Vicki Stover Hertzberg, PhD
Associate Professor
Department of Biostatistics
Emory University
Atlanta, Georgia

Timothy J Ingall MB BS, PhD
Associate Professor of Neurology
Cerebrovascular Diseases Center
Mayo Clinic Scottsdale
Scottsdale, Arizona

Thomas A. Louis, PhD
Professor, Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

- **Data Analyst**

  Teresa J. H. Christianson, BS
  Data Analyst II
  Division of Biostatistics
  Mayo Clinic Rochester
  Rochester, Minnesota

Table of Contents:

1. EXECUTIVE SUMMARY

1.1 The NINDS Charge

In May 2002, in response to concerns about the results of the NINDS rt-PA Stroke Study, the independent t-PA Review Committee was established at the request of NINDS. 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."

The committee was also asked, as a secondary issue, to explore if "pharmaceutical company participation biased the results of the trial". The committee declined to consider this charge 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 t-PA

1.2 Principal Findings

The principal findings of the Review Committee are as follows:

1. Using the global statistic devised by the NINDS investigators and the GEE, 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. The analysis was adjusted for center, time to treatment (0-90 minutes and 91-180 minutes), study part, age, baseline NIHSS, diabetes, and preexisting disability.
2. We examined all of the adjusting variables to determine if they modified the treatment effect of t-PA as measured by the adjusted t-PA to placebo OR. Our analyses found no evidence that any variable modified the t-PA treatment effect. In particular, neither baseline NIHSS, nor time from symptom onset to treatment, modified the t-PA treatment effect. Baseline NIHSS was analyzed both as a continuous and categorical variable, while time from symptom onset to treatment was analyzed as a dichotomous variable (0-90 minutes and 91-180 minutes) reflecting its role as a stratification factor in the design of the study.

1.3 Secondary Analyses

The Review Committee considered the following issues in their evaluation of the NINDS t-PA study:

1.3.1 Blood Pressure Assessment and Management

Our analysis identified a number of problems regarding pre- and post-randomization blood pressure measurement and management:

- Non-compliance with the defined protocol was substantial, and persistent, throughout the study with regard to both the documentation of blood pressure readings, and adherence to the treatment regimen for hypertension.
- There was limited rigor with regard to the pharmacologic characteristics of antihypertensive regimens. In some instances pharmacologic monitoring was performed

by representatives (nurses) of the sponsoring pharmaceutical firm. Medications employed were listed by date, but not by time, eliminating consequential interpretive utility.
- The exact number of patients who received medication to lower blood pressure either prior to, or after, receiving study treatment is unknown.
- The confusion regarding blood pressure documentation, and the lack of knowledge of treatment of hypertension either prior to, or after, receiving study treatment, could have led to an unknown number of patients receiving treatment in violation of the nominal study protocol.

Based on these observations, we reached the following conclusions:

- It was not possible to assess the effect of hypertension management on clinical outcome in acute ischemic stroke patients treated in the NINDS study.
- The blood pressure variables should not be included in the statistical models. However, we also found that inclusion of the blood pressure variables in the statistical models would have been inconsequential with regards to altering the t-PA treatment effect.

Finally, the inconsistent documentation of both blood pressure readings and hypertension management seriously undermines the NINDS investigators statement that blood pressure management was a significant part of the protocol that contributed to the success of the study. Nonetheless, we concur with the NINDS investigators premise that blood pressure management should be included in the protocol for treating acute ischemic stroke patients with t-PA. It is biologically plausible that hypertension management could affect clinical outcome in acute ischemic stroke patients treated with t-PA, and data from the cardiology literature has already demonstrated that in acute myocardial infarct patients, the risk of having an intracerebral hemorrhage is related to pre-treatment blood pressure. However, further clinical studies will be needed to assess whether blood pressure management is related to better clinical outcomes in acute ischemic stroke patients treated with t-PA.

1.3.2 Intracerebral Hemorrhage

In the NINDS trial, the overall risk of symptomatic ICH was 6.5% in t-PA treated patients vs. 0.6% in patients receiving placebo. When a symptomatic ICH occurred after treatment with t-PA, there were significant clinical consequences. Only a small minority had a favorable outcome (e.g., for the Barthel index, the favorable outcome in patients with symptomatic ICH was 10% vs. 55% in patients without ICH) and the three month mortality rate was very high (75%).

A number of putative risk factors for ICH were identified, with many of them being interrelated. Our exploratory analysis found four risk factors, age >70 years, baseline NIHSS >20 points, plasma/serum glucose >300 mg/L and edema and/or mass effect on the initial CT scan, that were associated with both an increased risk of having an SICH and a lower likelihood of having a favorable outcome. For patients with either no risk factors or only one risk factor, the likelihood of having a favorable outcome favored the t-PA treatment group, while for the group at highest risk (> 1 risk factor), there was essentially no difference between the t-PA and placebo groups with regards to the likelihood of having a favorable outcome. However, the analysis also found that the adjusted t-PA to placebo odds ratios for favorable outcome in the

three subgroups with different numbers of risk factors were not significantly different, and were consistently in favor of the t-PA treatment group.

We conclude that there was no statistically significant evidence of the existence of any subgroup of acute ischemic stroke patients in whom the risk, and consequences, of having a symptomatic ICH clearly outweighed the beneficial effects of t-PA. However, it is important to keep in mind that because of the study design and the small number of patients who had an SICH, this trial was not powered to identify risk factors related to having either an SICH or a decreased likelihood of a favorable outcome. Risk factors for ICH acute ischemic stroke patients treated with t-PA should be evaluated in future studies that are designed, and powered, to evaluate this question.

How the findings of this exploratory analysis are used in the management of the individual patient with acute ischemic stroke, balancing risks and benefits based on very limited scientific information, is for the patient and the attending physician to decide.

1.3.3 Baseline NIHSS Imbalance

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 a either a statistically or clinically significant effect on the study results, We further believe that the original models using both Age and baseline NIHSS as continuous variables properly adjust for the complex role played by these two variables, both strongly (negatively) related to the likelihood of a favorable outcome. There was a strong interaction between age and baseline NIHSS with respect to both the global analysis and the analysis 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 score above 20. However, there was no evidence of any Age by baseline NIHSS subgroup responding significantly differently to t-PA treatment than the study group at large.

1.3.4 Baseline Stroke Severity and Age

This analysis found evidence that age, baseline stroke severity as assessed by the baseline NIHSS score, and the interaction between age and baseline NIHSS, were related significantly in a negative manner to the likelihood of a favorable outcome. We believe that the original models using both Age and baseline NIHSS as continuous variables properly adjust for the complex role played by these two variables. There was a strong interaction between age and baseline NIHSS with respect to both the global analysis and the analysis 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 above 20. Patients with minor symptoms at baseline (NIHSS 0-5) had similar high odds for favorable outcome whether or not they were treated with t-PA. However, there was no statistical evidence of any Age by baseline NIHSS subgroup responding significantly differently to t-PA treatment than the study group at large.

1.3.5 Onset to Treatment Time

Based on the substantially nonlinear nature of the distribution of time from symptom onset to treatment (OTT), and an idiosyncratic distribution of favorable response rates among the placebo patients, we conclude that the data provided by this study failed to support a conclusion that the effect of t-PA therapy diminished with increasing values of OTT within the protocol specified 3 hour time limit. However, this does not mean such a relationship does not exist, and further studies are needed to address the question of a differential t-PA treatment

effect related to time from symptom onset to treatment. It is also important to recognize that the results from this study provide no data on the effectiveness of thrombolytic therapy administered to acute ischemic stroke patients more than 180 minutes after symptom onset.

1.3.6 Clinical Centers

We found no significant difference between the centers in the baseline characteristics of the patients. The likelihood of having a favorable outcome differed considerably between the centers, those with fewer patients often having the worst outcome. However, the between-center variation in t-PA treatment effect for either the global outcome, or the individual outcome measures, was not statistically significant and did not invalidate the trial results. Nevertheless, it will be important in future studies to identify the factors that lead to good outcomes at institutions administering t-PA to treat acute ischemic stroke patients. This information will be very helpful to other institutions that are looking to develop the resources needed to administer t-PA safely to acute ischemic stroke patients.

1.3.7. Stroke Subtype

We conclude that it was appropriate that stroke subtype was not included as a covariate in the analytic models. Further, we conclude that the data of this trial do not support any claim regarding either the presence, or absence, of a differential t-PA treatment effect within stroke subtype.

1.3.8 Preexisting Disability

Although patients with a preexisting disability had a significantly reduced chance of experiencing a favorable outcome, there was no evidence that they responded any differently to t-PA therapy than those without a preexisting disability.

1.3.9 Diabetes Mellitus

The observed data, and the adjusted estimated t-PA effects, indicated a strong benefit for patients without diabetes mellitus (DM), but no benefit among patients with DM. However, this comparison must be treated cautiously because there was no statistical evidence of a t-PA*DM interaction. The trial found no statistically significant evidence that diabetic and non-diabetic acute ischemic stroke patients responded differently to t-PA therapy.

1.4 Issues in Need of Further Investigation

The NINDS t-PA trial was a prototype study of acute ischemic stroke that demonstrated a beneficial effect of thrombolytic treatment with t-PA when administered within three hours of the onset of stroke symptoms. The study was designed to show differences in the entire group of eligible patients and not in subgroups. The exploratory analyses conducted previously by the trial investigators, and now by us, found a number of issues that need to be explored further so that t-PA can be used confidently by a broad range of practitioners in routine clinical practice. Analysis of these issues could be done by either conducting new large-scale clinical trials, or combining primary data from all the t-PA in ischemic stroke trials that have already been conducted. Both strategies are ongoing.

Based on the findings of this review committee, some of the most critical questions that need to be addressed are:

- Is there a subgroup of patients with ischemic stroke in whom the risk for intracerebral hemorrhage is so high that the group as a whole has no benefit from t-PA treatment? Candidate high risk factors are; age > 70 years, baseline NIHSS > 20, high glucose levels, and signs of edema or mass effect on CT.

- Is there a subgroup of patients with only mild symptoms in whom t-PA provides no net benefit?

- Within the time frame of the NINDS trial (treatment within 180 min), is there evidence of a differential t-PA treatment effect related to time from symptom onset to treatment?

- What is the impact of elevated blood pressure, and its management, before and after t-PA treatment on clinical outcome?

- Can data from other trials be used to validate the cut-off for t-PA treatment used by the NINDS investigators (blood pressure <=185/110)?

- Can the exploratory analysis finding in the NINDS trial that stroke patients with diabetes do not benefit from t-PA treatment be confirmed?

1.5 Conclusion

The committee concluded that, despite an increased incidence of symptomatic intracerebral hemorrhage in t-PA treated patients and subgroup imbalances in baseline stroke severity, when t-PA was administered to acute ischemic stroke patients according to the study protocol, there was a statistically significant, and clinically important, benefit of t-PA treatment resulting in a higher likelihood of having a favorable clinical outcome at three months.

2. INTRODUCTION

2.1 Background: In 1995, a group of investigators, the NINDS rt-PA Stroke Study Group, published a seminal manuscript summarizing the results of two studies of t-PA as a therapy for acute ischemic stroke[1]. Prior to the publication of this manuscript, the study group had conducted several investigations in preparation for the performance of the two pivotal studies[2-4]. These investigations involved pilot studies of the use of t-PA, studies of the reliability of the NIH Stroke Scale[5] and the Barthel scale[6] in the setting of a clinical trial, and a study of the factors related to the risk of intracranial hematoma formation in patients being treated for ischemic stroke with t-PA[7].

Subsequent to the 1995 publication[1] the FDA considered and approved an application from Genentech for the approval of t-PA as a therapy for acute ischemic stroke when administered according to the NINDS protocol. In the meantime, the NINDS rt-PA Stroke Study Group published a series of manuscripts designed to: i) elucidate their methods[8], ii) refine their message regarding the therapeutic efficacy of t-PA[9], iii) examine the long-term consequences of therapy[10], v) consider the factors affecting the risk of ICH[11], and, v) describe the frequency of pre- and post-treatment hypertension and the effect of its management[12].

As t-PA was used in emergency departments around the country, results were not as universally successful as anticipated, and doubts began to arise. Eventually, these doubts were expressed in the form of publications[13-17], and a commentary[18]. As a result of the concerns raised by these publications, NINDS appointed this t-PA Review Committee.

2.2 Announcement: The following announcement of the creation of this Review Committee appeared in the October 2003, *NINDS Notes*

> "The NINDS recently invited an independent committee to review and consider the data from the five-year, multi-site "Tissue Plasminogen Activator for Acute Ischemic Stroke," published by the NINDS r-TPA Stroke Study Group in the New England Journal of Medicine, December 14, 1995. The study represents the first treatment for acute ischemic stroke, and the therapeutic agent t-PA was approved by the FDA for this usage in June of 1996.
>
> In recent months, public debate about the study findings has resulted in some discussion within the medical community about the appropriate use of this treatment for stroke. In answer to this, the NINDS asked that the committee "address whether there is concern that eligible stroke patients may not benefit from rt-PA given according to the protocol used in the (NINDS) trials, and whether any subgroup imbalances invalidate the trial as claimed by some of the critics."
>
> The committee is chaired by Dr. W. Michael O'Fallon, Ph.D., Professor of Biostatistics and former Chair of the Department of Health Sciences Research at Mayo Clinic, Rochester, Minnesota. Dr. O'Fallon chose the members of the committee, who represent an international cadre of physician-scientists with expertise in biostatistics, clinical medicine, cerebrovascular disease, neurology, and emergency medicine. None of the committee members has a connection with the previous published study or with the manufacturer of t-PA. (See attached sheet for a roster of the committee members.)
>
> The committee has full access to the study data, will re-analyze the study, and hopes to report its findings by early spring. The NINDS looks forward to the group's findings and the presentation of the data analysis at professional meetings and in the scientific literature."

2.3 Charge: The actual charge to the Review Committee, delivered by Dr. John Marler on May 24, 2002, read as follows:

> "As the effort to implement the acute stroke care guidelines resulting from the publication of the results of the NINDS rt-PA Stroke Study has proceeded, increasing scrutiny of the results has occurred. One group in particular has recently raised concerns about the implications of an imbalance in the severity of the baseline stroke between different subgroups for the two treatment arms.
>
> I would like the committee 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 invalidates the entire trial as claimed by some of the critics. The issue of whether pharmaceutical company participation biased the results of the trial is an important, but secondary issue for the group.
>
> The committee will have full cooperation and access to the data in any manner that they wish for their own independent analysis or for analysis by the statistician from the trial."
>
> John Marler
> NINDS/National Institutes of Health

2.4 Membership: As described above, in May of 2002, Dr. Marler of NINDS invited Dr. O'Fallon to appoint a committee that would be viewed as fair and objective to both advocates and critics of the NINDS t-PA trials. The NINDS did not participate in the selection of any member other than Dr. O'Fallon and did not review the credentials of the members he selected. Furthermore, neither NINDS staff nor original NINDS rt-PA Stroke Study Group investigators participated in any of the meetings of the committee. They communicated with the committee and/or Dr. O'Fallon only rarely and then at the committee's invitation. The committee members, who represent an international cadre of physician-scientists with expertise in biostatistics, clinical medicine, cerebrovascular disease, neurology, and emergency medicine, were paid as hired consultants to an independent contractor to NINDS. None of the committee members has a connection with the previous published study or with the manufacturer of t-PA.

The committee consists of three clinicians (Drs. Kjell Asplund, Lewis Goldfrank and Timothy Ingall) and three statisticians (Dr. O'Fallon and Drs. Vicki Hertzberg and Thomas Louis). Full titles and affiliations are listed on the Committee Roster, a component of the Title Page of this report. The three statisticians are well acquainted, but had not previously collaborated on research projects or worked at the same institutions. Dr. O'Fallon recruited Dr. Ingall with whom he had worked when Dr. Ingall was a Neurology Fellow at Mayo Clinic in Rochester, Minnesota, but otherwise the three statisticians did not know the clinicians. Drs. Ingall and Asplund, are acquainted, having collaborated in the analysis of the WHO MONICA project, but have not been colleagues and Dr. Goldfrank, was not known to any of the other committee members.

Although Dr. O'Fallon had investigated the epidemiology of stroke, he had not participated in any of the studies or trials leading up to the NINDS-supported investigations regarding the use of t-PA as a therapy for acute ischemic stroke. Dr. Hertzberg, has participated in stroke related research, but none involving the investigations being appraised. Dr. Louis, had no experience in stroke research but has an extensive background in clinical trials, most recently in HIV-Aids. The three MDs have active research careers and, importantly, are practicing physicians whose professional responsibilities necessitate an intimate understanding of

appropriate assessment and management workup and therapy for individuals experiencing an ischemic stroke.

3. COMMITTEE PROCESSES

3.1 NINDS' Independent Contractor: The NINDS has a contract with an independent contractor in the Washington DC area, to maintain data sets, monitor the activities of investigators and establish contractual relationships with ad hoc groups engaged in small studies. The review committee acted independently of NINDS through this contractor, which provided the financial support required by the committee. This independent organization has been responsible for archiving data from completed studies sponsored by NINDS and thus provide the data from the t-PA studies to Dr. O'Fallon at Mayo for the committee's analysis.

3.2 Communications: Except for one in-person meeting (March 22, 2003), the considerable communication necessary among committee members was conducted via telephone and e-mail. Regularly scheduled conference calls were established and documented by approved minutes. The first conference call was held on June 4, 2002 and calls were held every two or three weeks until late fall, 2002, since which time weekly calls were held. A member of the contractor's staff joins these calls, records them and prepares draft minutes. The minutes are circulated electronically, reviewed and approved at subsequent calls. The contractor maintains and archives the conference call minutes as well as the exchanges of analyses among the reviewers.

3.3 Guiding Principles: The Committee was established within the framework of the following principles:

- In all interactions, openness and candor were encouraged and respected.
- The committee's work must be performed independently of NINDS, Genentech and the investigators involved in the original studies.
- The committee must have unhindered and complete access to the original data upon which the published manuscripts and the FDA approval were based. The evaluation required analysis of the original data; it could not depend solely on reading the literature and arriving at a conclusion.
- The committee must be in control of its data analyses. To this end it was necessary to arrange for the data to be made available to a data analyst from the Division of Biostatistics at Mayo Rochester, who worked under the guidance of Dr. O'Fallon.
- The committee requested that the scientific community be made aware of its existence and charge.
- The committee declined to consider the "secondary issue" in the charge 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 t-PA

3.4 Timeline:

~May 1, 2002	O'Fallon appointed as Chair & asked to form a committee
May 20, 2002	Committee of 5 completed
May 24, 2002	Charge to Committee issued by Dr. Marler of NINDS

June 4, 2002	First Conference call
June 18, 2002	Committee budget proposed to NINDS
August 6, 2002	Final member joins the committee
August 27, 2002	Discussions regarding an alternative analytic support plan
Sept. 5, 2002	Arrangements for Mayo to provide analytic support
Sept. 15, 2002	Data made available to Mayo
Sept. 26, 2002	First Mayo replication of results
Oct. 25, 2002	NINDS Notes announces the committee's charge and composition
Nov. 15, 2002	Committee decides not to participate in NINDS sponsored "Stroke Symposium" (12/12/02)
Nov. 16, 2002	BMJ Reporter & Article
Nov. 21, 2002	NINDS Investigators Tilley & Brott join committee conference call regarding blood pressure management issues and data
Jan. 2003	Additional blood pressure data made available
Jan. 2003	Exchanges with Drs. Marler & Penn regarding publications and presentations
Feb. 6, 2003	NINDS investigators participate in another conference call regarding blood pressure management issues and data
~March 1, 2003	Committee sends abstracts to European Stroke Conference (ESC) and Society for Academic Emergency Medicine (SAEM)
March 22, 2003	Committee meets to complete the framing of its final report.
May 24, 2003	Presentation at ESC, Valencia, Spain
May 29, 2003	Presentation and Panel discussion, SAEM, Boston, Massachusetts
June 10, 2003	BMJ Reports on SAEM presentation
July 31, 2003	Final Report submitted to Dr. Marler
Subsequently	Manuscript prepared for publication and presentations made in Europe and Australia

4. METHODS

4.1 Data Management

4.1.1 Data from NINDS' Independent Contractor: In September 2002, the committee received a CD from the Institute's independent contractor labeled "NINDS t-PA Stroke Study Data Collection." It contained a 43 page descriptive document, a main directory with 114 SAS datasets and 81 SAS programs. There was also an ancillary directory with 112 additional files. We examined all 114 SAS datasets containing a total of 4,795 variables (with many variables being in multiple datasets) to discern the variables of interest. Many of these datasets contained 624 observations, but several had more than 10,000 observations. All the variables we used in these analyses were found in one of five main datasets. In January 2003, we received 8 additional SAS datasets upon request pertaining to blood pressure. These datasets were not used in our analyses.

4.1.2 Variable Identification and Definition: As described in 4.1.1, data were obtained from an independent contractor. Definitions of the primary variables were obtained and are presented in Table 4.1. In general, this was a straightforward process, but where issues arose we contacted the contractor and, on occasion, the original study statisticians, programmers, or investigators for clarification.

In Table 4.2, which will be used to assess balance between the t-PA and Placebo groups in Section 5.1, we summarize the 64 variables assessed at or prior to the time of randomization. We have adopted a standard nomenclature that is as transparent as possible. In addition to the variable name, we distinguish timing by the three prefixes: i) "Pr" to indicate a determination (often a diagnosis) made prior to the stroke, ii) "Ad" to indicate measurements/determinations made at admission to the Emergency Department for treatment of the stroke, and iii) "Bs" for baseline to indicate measurements/observations made between admission and randomization. Time constant variables (e.g., age, sex, race) do not require a prefix and some prefixes are somewhat arbitrary. For example, "AdAspirin" indicates whether or not the patient had been taking aspirin as a regimen up to the time of the stroke. A person who had discontinued such a regimen before stroke would be coded "no." We coded diabetes "PrDM," indicating a prior diagnosis of DM rather than an indicator of elevated glucose at arrival to the ED.

Many variables are dichotomous, indicating the presence or absence of some characteristic. Usually, we use "1" for presence and 0 for absence. If the coding might be unclear, we indicate the coding rule in ( ) after the variable name. Thus, sex (male) means that we coded males as 1. All dichotomous or polychotomous variables are summarized as percents rounded to the nearest one decimal. If a variable is continuous, we indicate the units of the measurement and in Table 4.2 the variable is summarized by its median.

4.1.3 Result Replication: We undertook to replicate results reported in several of the published manuscripts. Results of this replication process will be discussed in Section 5.1, Specifically, we replicated:

1995[1]: Tables 1, 2, 3, & 4

1997[11]: Table 1

1999[10]: Table 3

2000[19]: Table 3

4.2 Study Design

4.2.1 Stratification Factors: Evaluation of design, conduct, analysis, and interpretation issues was restricted to the committee's charge. The original investigators published multiple manuscripts and it was not the purview of the committee to assess and judge all aspects of each of these manuscripts.

The primary manuscript[1] describes two studies, referred to as Parts 1 and 2. The investigators describe Part 1 essentially as a Phase 2 study that seeks to determine whether the agent had activity. However, it was a randomized, placebo controlled, study designed to address the evidence for t-PA activity with respect to outcomes assessed 24 hours after stroke onset. Part 2, designed to assess results 90 days after stroke onset, was designed exactly as Part 1. The investigators, essentially the same as for Part 1, were blinded as to the results of Part 1 until Part 2 was complete. The studies were placebo controlled, randomized clinical trials (RCTs). Both were conducted at approximately 8 clinical centers with independent randomization at each. Randomization was stratified and balanced at each center according to whether the patient was randomized within the first 90 minutes or in the 91-180 minute interval after stroke onset. Patients whose time since onset had exceeded 180 minutes were ineligible. Furthermore, the investigators conducted assessments in both studies at 24 hours, 90 days and 1 year after stroke onset.

Consequently, as proposed by the investigators, Parts I and II can be treated as independent, replicate studies. For analytic purposes, we treated the two studies as a single, large RCT with three stratification factors, Part (1 or 2), Center, and onset to treatment time, (OTT: 0-90 or 91-180 minutes). After detailed examination of outcome data at 90 days and 1 year, the review committee decided to restrict its analysis to the outcome assessment at 90 days after stroke onset of all the patients in the two studies.

4.2.2 Outcome Measures: The NINDS investigators used four outcome measures. The Barthel index, modified Rankin scale, and Glasgow outcome scale are accepted as measures of functional status. The NIH Stroke Scale (NIHSS) is accepted as a measure of neurologic deficit. The primary response variable for each measure was a dichotomous indication of whether the outcome (at 90 days) was "favorable" or "not favorable." The definitions of "favorable" were: Barthel; 95 or 100, Rankin; 0 or 1, Glasgow; 1, and NIHSS; 0 or 1. For each measure death was treated as an unfavorable outcome. Since the measures assess different aspects of the consequences of stroke, they are neither completely congruent nor statistically independent. Therefore, the committee evaluated comparisons of the Placebo and t-PA treatments for each of the measures individually.

In addition, the NINDS investigators constructed what they refer to as a "Global" indicator of a favorable response[1, 20]. This Global indicator is a 4-dimensional vector of the favorable/unfavorable indicators, for each of the 4 indices. Thus, each patient had a global response vector consisting of 4 elements, each either zero or one, with zero indicating an unfavorable outcome and one a favorable outcome. Those dead by 90 days had a global

response vector of the form (0, 0, 0, 0) while those whose outcome was favorable on all the measures had a global response vector of the form (1, 1, 1, 1). The review committee replicated the results of the global analyses reported by the investigators and carried out additional analyses deemed appropriate.

4.2.3 Intent to Treat: The study investigators used the principle of "intent-to-treat" in analyzing all patients randomized in the study. Thus, they attempted formal follow up at 24 hours, 90 days and one year on all randomized patients. Patients "lost" in the sense that they were known to be alive but did not provide data permitting the determination of favorable/unfavorable status were assigned the least favorable known level for each index (4, 11) with its consequent favorability status. With two exceptions, the review committee used the same approach. The two exceptions involved individuals mistakenly randomized into the study at a point more than 180 minutes after onset. Since an essential component of our charge was to determine whether there were groups of patients who should not be treated with t-PA according to the study protocol, we excluded these two patients from subsequent analyses.

4.3 Analytic Methods

4.3.1 Treatment Group Balance: In theory, the process of the completely random assignment of patients into one of two treatments within strata should produce nearly equivalent distributions of observed covariables (not treatment effect variables) in the two treatment groups. We will examine whether or not that happened for the variables listed in Table 2 using Chi-Squared tests for the dichotomous and polychotomous variables and rank sum tests for the continuous variables.

4.3.2 Missing Data Imputation: As is seen in Table 4.2, there were patients whose values of several of the variables were missing. Before any in depth analyses could be undertaken we elected to take the following actions:

1. We eliminated 5 variables from all subsequent analyses because they were each missing for more than 40 patients. From Table 4.2 it is seen that these five variables are: BMI, Prior Atherosclerosis, Prior Hyperlipidemia, Baseline fibrinogen and Prior TIA.
2. We imputed the other missing values essentially by sampling at random from the existing data. In this imputation, if a variable was categorical with some categories being observed less than 10% of the time we used 10% as the probability for that category and adjusted the most common category appropriately to assure that the sum of the several percents was 100.

For the logistic regression analyses, we performed "best case-worst case" imputation and replicated the random imputation process several independent times, running the regression models for each resulting data set. The distributions of parameter estimates were then examined to determine if any aberrations were observed. Seeing none considered critical, we ran one final random imputation thus creating an analysis data set including 622 patients each with a complete set of values for the variables to be considered in future analyses. It should be noted that this use of a single imputation sample will result in underestimated standard errors. However, the number of patients and variables for which imputation was necessary was small so the bias should be negligible. There was no specific evidence in the published material as to what, if any, actions were taken by the NINDS investigators in reaction to the

missing values, although there is an allusion in the FDA application in which the sponsor states: “An Intent-to-treat analysis was the performed, and the data imputation plan for missing values as devised by the NINDS Investigator group would be utilized”[21]. It is possible that our analyses might differ from theirs in minor ways as a consequence of the use of different imputation strategies.

4.3.3 Analytic Models: With the primary analyses focused on the “favorability” outcome, statistical models must be appropriate to the analysis of proportions (equivalently, probabilities or percents). Such data can be analyzed on one of three scales: the original scale examining the difference between two proportions, the ratio (log) scale analyzing the ratio (Relative Risk) of two proportions, or the odds (logit) scale analyzing the ratio of two odds (Odds Ratio). Each measurement scale has its advantages and disadvantages. The original scale is most clinically relevant, but the log and logit scales generally produce more parsimonious models with better understood statistical properties. The investigators reported relative risks (RR) and odds ratios (OR) when possible, but performed their most extensive analyses in the logit scale using univariate and multiple logistic regression models and reported the resulting odds ratio estimates.

Validated approaches and software are available to implement each approach. A Generalized Linear Model (GLiM) with the identity, log or logit “links” and “binomial” variation unified the approach. The GLiM can be used to compare two treatments with respect to an outcome measured on the probability scale while adjusting for stratification factors, confounding factors and even effect modifying factors, with model specification being essentially identical to that for standard, linear regression.

In analyzing the Global outcome measure, the investigators used the Generalized Estimating Equation model (an extension of a GLiM) with the logit link function and with the correlation structure estimated by the empirical, observed, correlations among the four indices[1, 20]. This analysis yields a general odds ratio estimate comparing the odds of a (global) favorable outcome in the t-PA treated group to that in the placebo group while adjusting for stratification and baseline factors. After determining that this was appropriate the review committee used these same models in its analyses.

4.3.4 Subgroup Analysis and Interaction Detection: In addition to evaluating overall results, the t-PA Review Committee was charged with considering the question of whether subgroups of patients might actually be harmed by the use of t-PA therapy. The investigators addressed that issue[9]. Possibly the FDA Advisory Committee considered the subgroup issue, however FDA approval was without conditions other than the restriction that the therapy be administered according to the protocol.

The sample sizes in Parts I and II were determined so that the study would have sufficient statistical power to detect a clinically relevant difference between t-PA and placebo. Neither study was powered to detect clinically important subgroup effects or treatment interaction effects. The combined studies still have low power for these investigations. Even though the power is low, a large number of evaluations are likely to generate some statistically (and apparently clinically) significant results even when the underlying truth is that no such treatment/subgroup relations are operating. Consequently, subgroup analyses and evaluations of interactions operate in a low power, exploratory context.

The NINDS investigators attempted to address the low power issue by performing the tests at a very generous p-value (0.2 & 0.1)[9]. While this certainly increases the power (decreases the chances of a Type II error), it does so at the price of an increase in the chances of a Type I error and may result in spurious "findings." The investigators quite correctly point out (and we concur) that such findings are best used as motivation for further studies designed specifically to address the issue raised by the identification of these interesting groups. The review committee has examined some of the potentially interesting subgroups in considerable detail, reports results and emphasizes the caveats and cautions.

4.3.5 Logistic Models:

The term "Odds" refers to the ratio of a probability to it's complement (e.g., P/[1-P]). In this report the term "odds ratio (OR)" always refers to the ratios of the odds of a favorable outcome in one group of patients to the odds of a favorable outcome in another group. Since favorable outcome is defined on 4 scales, it is essential that the appropriate scale be kept in mind, but do not incorporate references to Barthel, Rankin, Glasgow, or NIHSS in the "OR" notation.

Define

P[F | T] = probability of a favorable outcome on the treatment
and P[F | P] = probability of a favorable outcome on the placebo.

In this notation the odds ratio is:

$$\mathrm{OR} = \frac{\mathrm{P[F|T]/\{1\text{-}P[F|T]\}}}{\mathrm{P[F|P]/\{1\text{-}P[F|T]\}}}.$$

With "log" indicating the natural logarithm, the basic logistic model takes the form,

$$\text{logit} = \log(OR) = \beta_0 + \beta_1 X + \underline{\gamma}^t \underline{Z},$$

where X is a 0/1 indicator with "0" the comparison group, $\underline{Z}$ is a vector of covariates with the corresponding coefficient vector $\underline{\gamma}$. The OR adjusted for the covariates $\underline{Z}$ is estimated by $e^{\hat{\beta}_1}$. We use SAS Proc Logit to estimate the parameters β_0, β_1, and $\underline{\gamma}$.

Effect modification, for example by component Z_1 of $\underline{Z}$ (i.e., the influence of X on the OR is dependent on Z_1) is modeled by an interaction term:

$$\text{logit (favorable outcome)} = \beta_0 + \beta_1 X + \beta_2 X * Z_1 + \underline{\gamma}^t \underline{Z}.$$

We test the null hypothesis that $\beta_2 = 0$ to asses whether or not there is sufficient evidence to declare Z_1 an effect modifier. This type of question and test is critical to the question of whether baseline imbalances have influenced results.

In our logistic models, the vector $\underline{Z}$ of covariates must include all stratification variables and will include variables that are statistically significantly related to the likelihood of a favorable outcome. Note that imbalance of a variable between the t-PA and placebo groups does not necessarily imply that it must (will) be included in the model.

4.3.6 Global Analysis

The Global analysis, described in detail by the NINDS investigators[20] provides a powerful, multi-outcome approach to assessing the relation of baseline variables and treatment to the probability of a favorable outcome. This approach treats the four binary outcomes as a four-dimensional outcome vector which is then related to covariates much as in the basic logistic regression models. Correlation among the outcome measures must be taken into account. Software for estimating the parameters in such a comprehensive model was limited, when the original investigators conducted their analyses. Now, SAS Genmod and other implementations of the Generalized Estimating Equation (GEE) approach facilitate such analyses.

4.3.7 Analyses in the Probability Scale

As mentioned earlier, logistic regression methods are sometimes criticized because they are based on the odds scale. The fundamental results of logistic regression on a clinical trial analysis is an estimate of an odds ratio. This odds ratio (OR) relates the odds of a "success" (in this study, the odds of a favorable outcome at 90 days) among those on the active therapy (t-PA in this study) to the odds of a success in the comparison (placebo) group. Estimates of the <u>difference</u> between the probability of success in two groups rather than the odds ratio of probabilities does provide a more clinically relevant comparison.

We estimate the difference $\Delta = P[F \mid T] - P[F \mid P]$ rather than the logistic regression based odds ratio. Using the odds of a favorable outcome among the placebo treated patients to "represent" the status in the general stroke patient population, we estimate the odds ratio as described and then can estimate the difference between the two percents as follows: Define

K = odds of favorable outcome among placebo treated patients

OR = estimated t-PA to placebo odds ratio

Then the estimate of Δ is

$$\hat{\Delta} = \frac{K}{1+K}\left[\frac{OR-1}{1+ORgK}\right] \qquad \text{(Equation 1)}$$

In a more general context if patients are stratified into M groups, G_1, …, G_M, we can represent the data by the following table.

	G_1		G_2		…	G_M		
	F	UF	F	UF		F	UF	
Placebo								n_1
t-PA								n_2

In such a context, it is possible to extend the above formula to estimate the difference between the two proportions (probabilities) taking into account this sub-grouping and all other covariates. This formula is based on a logistic regression model which includes M-1 indicator variables distinguishing the M groups as well as all other covariates and, if necessary, M-1 variables accounting for the differential effects of t-PA in the M groups.

Define

1) $P[G_i] = n_i / n \quad i = 1, ..., M$

where n = total number of randomized patients

n_i = total number of randomized patients in group G_i.

2) Define the variable

$$T = \begin{cases} 1 \text{ for those randomized to t-PA} \\ 0 \text{ for those randomized to placebo} \end{cases}$$

3) For i=1, …, M-1, define variables:

$$X_i = \begin{cases} 1 \text{ for all patients in } G_i \\ 0 \text{ for all patients not in } G_i \end{cases}$$

Note: This designation of M-1 variables, X_i, specifies group G_M as the comparison group. That is completely arbitrary and in practice we will tend to use one of the groups containing a large number of patients.

4) θ = odds of a favorable outcome in the placebo treated patients from the group chosen to be the comparison group.

5) $\underline{Z}$ a vector of covariates with corresponding vector $\underline{\delta}$ of coefficients.

The most general logistic regression model takes the form:
logit (Favorable Outcome)

$$= \beta_0 + \beta_1 T + \sum_{j=1}^{M-1} \gamma_i G_i + \sum_{j=1}^{M-1} \gamma_{M+j} T * G_i + \underline{\delta^t Z}$$

and the generalized formula for the weighted difference in the likelihood of a favorable outcome between the t-PA and placebo groups is

$$\hat{\Delta} = \theta\left[\exp(\beta_1) - 1\right] \sum_{j=1}^{M} \left\{ \frac{\exp(\gamma_j)}{\left[1 + \theta \exp(\beta_1 + \gamma_j + \gamma_{M+j})\right]\left[1 + \theta \exp(\gamma_j)\right]} \right\} P(G_j) \qquad \text{(Equation 2)}$$

The $\gamma_M = \gamma_{2M} = 0$ and if the interaction terms are not included
$\gamma_{M+1} = \gamma_{M+2} = ... = \gamma_{2M} = 0$.

The standard errors of the difference estimators defined by equations 1 and 2 above were estimated using the Jackknife method described by Efron and Gong[22].

4.3.8 Covariate Determination (Stepwise Models): All baseline covariates available to the NINDS investigators (with the exception of baseline/admission blood pressure measurements, as explained in Section 6) were considered for initial inclusion in the models A forward stepwise selection process ($p < 0.05$ to enter and remain) was used to derive the final covariates for inclusion, after constraining the model to include the design stratification variables, CENTER, OTT, and PART. This modeling process was performed for each of the

outcome measures to derive a candidate list of covariates for consideration. A covariate was considered to be in the candidate list if it entered the stepwise process for at least one of the four outcome measures. In addition, all covariates in the candidate list were reviewed for clinical relevance, i.e., did the relationship make sense biologically. After arriving at this candidate list of covariates, these covariates were then screened for pairwise interactions, again using the forward stepwise selection process. From this process, any covariate or interaction (and any contributing lower order effects) was included in the final list of covariates if it remained in the model after this second stepwise screening process for any of the four outcome measures. A similar process was employed for the analyses described in Section 7 where the occurrence of an intracerebral hemorrhage was the endpoint.

5. RESULTS

5.1 Replication of Published Results: As indicated in Section 4.1.3, we selected tables from several of the NINDS investigators' publications[1, 10, 11, 19] and attempted to replicate them. In Tables 5.1 through 5.7 with the matched Tables 5.1a through 5.7a, respectively, we summarize our replication. There are nothing but trivial differences between our results (Tables 5.1 through 5.7) and the corresponding published results (Tables 5.1a through 5.7a), respectively. Thus, we concluded that we had access to the correct data and had defined the variables correctly so we continued with our planned analyses.

5.2 Baseline Balance Between t-PA and Placebo Groups: Table 4.2 was constructed to facilitate an examination of the balance at randomization of the distribution of the 64 variables which may be used in the analysis but were not specifically balanced by the randomization process. In general, randomization should yield a balance-on-average between the t-PA and Placebo groups. However, when many variables are assessed, chance alone will result in a statistically significant imbalance in some variables. The primary questions faced by the original investigators and the review committee must be whether any imbalances noted represent some excess above that expected by chance alone, whether such imbalances suggest an inherent flaw in the randomization process and whether observed imbalances confound treatment comparisons.

Many of the 64 variables included in Tables 4.1 and 4.2 are constructed from a common set of inputs and, consequently, the multiple "significant" p-values observed need to be considered with some care. We observe, as did the original investigators[1, 19], that there were imbalances in the following areas:

Age:	- Placebo group somewhat younger than the t-PA group
Weight:	- Placebo group somewhat heavier than the t-PA group
Aspirin	- Fewer in the Placebo group on a daily regimen of aspirin than in the t-PA group
Baseline NIHSS	- While the Placebo and t-PA group medians of baseline NIHSS (BsNIHSS) were not significantly different; when BsNIHSS was categorized as: 0-5, 6-10, 11-15, 16-20, and > 20, a significant imbalance was identified. Primarily, among patients in the 0-5 range, there was a greater proportion of patients randomized to t-PA than to placebo. It is with respect to this latter imbalance that much controversy regarding the study results has arisen.

Of course, the most critical question is whether or not an imbalance is so severe that any observed treatment effect could be explained by the imbalance (false positive effect) or any lack of observed effect could have been the result of the imbalance obscuring the effect (false negative). The NINDS investigators concluded that, regarding the above noted imbalances, neither instance seemed likely. We describe our investigations of this issue in the following sections.

5.3 Observed Outcomes: Table 5.8 contains the observed data regarding favorable outcomes for each of the 4 outcome measurements. The 622 patients are divided into the 310

treated with t-PA and the 312 treated with the Placebo, and are classified further as to whether they had had a favorable 90-day outcome according to each of the outcome measures. Results are summarized in 3 ways. For each outcome measure the difference between the percents favorable for the t-PA and Placebo groups, the ratio of these percents, and the corresponding odds ratios are all presented. For all of these comparison scales for each of the outcome measures, these data summaries show that the t-PA treatment is significantly more likely to produce a favorable 90-day outcome than the Placebo and the estimated treatment effect is clinically important.

In evaluating our covariate adjusted analyses (Section 5.5) it will be important to refer back to these unadjusted results. Assuming that randomization was properly conducted, these results are valid. For the four outcome measures, the proportion expressing a favorable outcome in the t-PA treated group exceeds that proportion in the Placebo group by between 13.7% and 16.3%. These differences indicate that if 1000 ischemic stroke patients received t-PA therapy according to the NINDS protocol, about 150 more of them would experience a favorable outcome than if t-PA had not been available or used. The four odds ratios, ranging from 1.78 to 2.07, are all significantly different from one, again suggesting that t-PA is more likely to produce a positive outcome than Placebo. It will be informative to consider the effect of the adjustments relative to these basic estimates.

Table 5.9 contains three sub-tables showing the joint distribution of the 4 dichotomous outcome variables so that their interrelationships can be examined. The first subtable shows the entire cohort of 622 randomized patients classified into the 16 possible categories. Here we see that 325 (52%) of the patients failed to have a favorable outcome on any of the 4 outcome measures which means that 48% had a favorable outcome on at least one of the outcome measures. At the other extreme, there were 151 (24%) of the randomized patients who had a favorable outcome on all of the 4 measures.

These two extremes suggest two straightforward ways to combine the 4 outcome measures into two consolidated outcome scales. In the lower two tables we see that among the patients treated with t-PA, 169 (54.5%) had at least one favorable outcome while among the Placebo patients 128 (41.0%) had at least one. So, in the “at least one” scale, t-PA is better than Placebo by 13.5% with an OR of 1.72. For the more rigorous condition of having a favorable response on all 4 of the outcome measures, 98 (31.6%) of the t-PA treated patients achieved that level while only 53 (17.0%) of the Placebo patients did. This represents a difference of 14.6% in favor of the t-PA treated patients with a corresponding OR of 2.26.

The global analysis described by the NINDS investigators[20] is a more sophisticated way of combining the four outcome measures. Because the four measures are correlated, combining them is not equivalent to simply increasing the sample size by a factor of 4. However, because they are not perfectly correlated, combining them brings more information through the global analysis than is contained in any analysis of an individual outcome measure. As a consequence, the global analysis is more powerful than the individual analyses, as emphasized by the NINDS investigators[20].

5.4 Covariate Selection: In this section we describe the process of developing the models that account for the study design and covariates used to adjust estimates of the t-PA to

Placebo odds ratios and the corresponding differences in the probability of a favorable outcome. In subsequent sections we use these tools to address several critical issues among which will be the following.

1. Did the BsNIHSS imbalance bias the treatment comparison in a critical way?
2. Does the increased risk of ICH among t-PA treated patients put in question the value of t-PA as therapy for acute ischemic stroke patients? In particular, are there subsets of patients in whom the risk and consequences of ICH outweigh the benefits of t-PA?
3. Do the data support an informative analysis of the impact of the time from onset to treatment on the efficacy of t-PA therapy?
4. Is the t-PA benefit consistent among the several centers involved in the study?

5.4.1 Logistic Analysis of Favorable Individual Outcome: The results of the first stage in the process of selecting the covariates to be included in the outcome models are summarized in Table 5.10. All of the variables in Table 4.2, except for those with a large number of missing values (Section 4.3.2) and for blood pressure measurements reported as made at admission or baseline (for reasons described in Section 6), were considered as potential covariates. For each of the four outcome measures, each variable was considered separately in a logistic model of favorable outcome constrained to include the stratification variables of PART, CENTER, and OTT. The top part of Table 5.10 lists the stratification variables and those variables that had a p-value <0.20 for association with a favorable outcome for at least one of the outcome measures. The variables are ranked in order of their level of significance within the Barthel model. Thus, in these analyses of one potential covariate at a time, 18 of the potential covariates have p-values <0.20 for at least one of the outcome measures and 9 have p <0.20 for all four of the outcome measures. Not surprisingly, baseline NIHSS (BsNIHSS) in either of two constructs, AGE, and evidence of a preexisting disability (PrDISAB) are all highly (negatively) related to a favorable outcome for all 4 outcome measures.

The lower half of Table 5.10 illustrates the results of a forward stepwise process (p <0.05 to enter and remain). The three variables, BsNIHSS (as a continuous variable) AGE and PrDISAB enter, in that order, for all four of the outcome measures. Seven other variables enter for at least one but not all of the four models. Some of these variables – most notably weight – entered these models even though their univariate p-values (at the onset of the stepwise process) were not <0.20.

The next stage of the process of identifying covariates to be included in the outcome models is illustrated in Table 5.11. Here, the top part of the table illustrates the four separate models, including all of the variables that entered in the aforementioned stepwise process. All of the potential interactions among those variables included in the models were made available as candidates to enter the model in another stepwise process (with p <0.05 to enter and remain). For each of the outcome variables the interaction between AGE and BsNIHSS (AGE*BsNIHSS) was highly significant and was included in all subsequent models. The role of this interaction between stroke severity and age will be discussed in Section 8.1.5.

Two other interactions entered with the Rankin score and one of them also with the Glasgow score. The resulting models are presented in Table 5.12. The most interesting aspect of Table 5.12 is the different impact of the inclusion of the AGE*BsNIHSS interaction within each model. Understandably, with the inclusion of AGE*BsNIHSS in a model containing AGE and BsNIHSS some impact on the two "main effect" terms is expected. What is seen is that for the Barthel index, and to a lesser degree for NIHSS, nearly all of the impact of AGE and BsNIHSS is found in the interaction term. In contrast, for the Rankin and Glasgow scores both of these main effect terms retain significance in the presence of the interaction term. Although not visible in this table (see Tables 5.17. through 5.21), another interesting aspect of this interaction term is that its coefficient is negative. Since increasing values of the two variables decreases the chance of a favorable outcome, this negative coefficient indicates that they are synergistic in their interaction with each other. This is actually a somewhat uncommon phenomenon since advancing age often overwhelms other factors regarding the effect of a disease. This will be discussed further in Section 8.1.

Following an argument described in Section 5.5.2, we decided that all of our models of treatment effect will include as covariates the three stratification factors, four main effects (BsNIHSS, AGE, PrDISAB and PrDM) and the AGE*BsNIHSS interaction. However, the next stage of the process is to use the information gleaned from the analyses of the individual outcome measures to develop a Global model.

5.4.2 Global Model of Favorable Outcome: The first stages of the process of building a Global model are illustrated in Table 5.13. To the left, with OTT, PART and CENTER fixed in all models, is a summary of the independent impacts in a global model of each of the variables that either had a p-value less than 0.20 or were potentially interesting for other reasons. The stepwise process, which is performed automatically for the logistic models, is, of necessity, performed one-variable-at-a-time in the Global model. The top part of the right side of Table 5.13 summarizes the order in which seven variables "entered" the model in this process. The first three of these seven variables are the same as the three that entered all of the models for the individual outcome measures. The bottom part of the right side of Table 5.13 shows what would be the final model if no interactions entered this global model.

Since the AGE*BsNIHSS interaction seemed certain to enter the global model, we began the process of looking for interactions among the covariates in the global model by entering that interaction into the model. The results are described in Table 5.14. Here, the introduction of that interaction term changes the p-value of one of the existing variables, BsED/ME, to be greater than 0.05. We thus applied a backwards removal process, removing that variable and, subsequently, two more of the original seven variables. This left a Global covariate model – Table 5.14 – with three stratification factors, four main effects (BsNIHSS, AGE, PrDISAB and PrDM) and the AGE*BsNIHSS interaction. The coefficient estimates and their standard errors for these covariates in the global model are also included in a small subtable of Table 5.14.

5.4.3 Primary Covariate Model: For the sake of uniformity, we declared the covariates described above to be the covariates to be included in all models used in subsequent treatment comparisons, those for each of the outcome measures as well as for the Global analysis. Having so declared our "final" covariate model, we again examined all interactions among them for each outcome measure, as well as for the Global model. We found no other

interactions of consequence so all further analyses are based on the comprehensive covariate model described above. The first of these analyses are discussed in Section (5.5).

5.5 Model-Based Treatment Comparisons: In Section 5.3 (Tables 5.8 & 5.9), we described the data regarding the comparison of t-PA therapy to Placebo in the most fundamental terms. For the 4 outcome measures of Barthel, Rankin, Glasgow and NIHSS, the actual counts of patients experiencing a favorable outcome resulted in odds ratio estimates of 1.78, 2.07, 1.85 & 2.01 respectively. The corresponding, unadjusted Global estimate is 1.88. These were all highly significantly different than 1, ($p < 0.0001$) indicating that t-PA is superior to Placebo insofar as the likelihood of a 90-day favorable outcome was concerned.

5.5.1 Logistic and Global Model Results: Subsequent to our examination of the fundamental data, we examined variables potentially related, either positively or negatively, to the prospects of a favorable outcome. As discussed in Section 5.4, we have identified several variables that significantly influence outcome and, consequently, should be included, along with the stratification variables, as covariates in any model-based estimates of the OR.

Tables 5.15 & 5.16 summarize the evolution of the process of estimating a t-PA to Placebo odds ratio (OR) for each outcome measure (Table 5.15) as well as for the global analysis (Table 5.16). As the estimation process became more sophisticated and complete through the use of models that "adjusted" the OR estimates first for the stratification factors alone and ultimately for the stratification factors and the covariates of BsNIHSS, Age, PrDiabetes Mellitus, and PrDisability, the adjusted OR estimates became 2.19, 2.43, 2.13, 2.19, & 2.13, respectively. These adjusted OR estimates are numerically larger and statistically more significant than their unadjusted counterparts.

On the basis of similar analyses, the NINDS investigators concluded that t-PA, when administered according to the NINDS protocol is significantly superior to Placebo[1, 11]. **The review committee concurs with this conclusion.**

5.5.2 Treatment by Covariate Interactions: Before the general conclusion stated above can be considered valid, we must examine for each of the outcome measures as well as in the global analysis whether any of the covariates in the model directly moderated the effect of t-PA. Such moderation could be synergistic (enhancing the t-PA effect) or antagonistic (depressing the effect of t-PA). As indicated in Section 4.3.3, such "effect modification" is assessed by the inclusion of appropriate interaction terms in the logistic and Global models. Only in the absence of terms that are large relative to the main effects, are we in a position to report, without qualification, a universal statement of the evidence regarding the effectiveness of t-PA after "adjusting" for the presence of a number of covariates. If the interaction effects are such that all estimates of treatment comparisons are in the same direction a general statement might be possible, but, even then, care in interpretation is essential (Section 4.3.6).

5.5.2.1 Results of t-PA by Covariate Interaction Tests: In a series of five tables (Tables 5.17, through 5.21),for the analyses of the four outcome measures; Barthel, Rankin, Glasgow, NIHSS and the Global analysis respectively, we provide extensive summaries of the analytic models. In each of these tables, the first two results columns, labeled "Estimate" & "std. Error", provide the estimates of the coefficients of each of the covariates within a strictly covariate model. Here we see the negative coefficient on the Age*BsNIHSS interaction term alluded to

earlier. No p-values are provided here because they have been presented in Tables 5.12 & 5.14.

In the next set of three columns we see the primary adjusting models leading to the adjusted Odds Ratio estimates seen in Tables 5.15 and 5.16 which is obtained by inserting the treatment variable (t-PA) into the covariate model. These estimates are obtained by exponentiating the t-PA coefficient (e.g. for the Barthel model OR = exp(.7837)). In these models we also note that the coefficients of the covariates are changed only a little by the addition of the treatment variable t-PA into the models.

The remainder of the sets of columns summarizes the testing of interactions between t-PA & Covariates (including the stratification factors) within each model. There are 4 dichotomous covariates, each with a single degree of freedom, and 2 polychotomous covariates with multiple degrees of freedom. The interaction between t-PA and the dichotomous covariate have a very direct interpretation so we elect to discuss them in more detail and, specifically, to examine the power of the tests that are performed within the models to determine if the interactions are significant and need to be included in the models.

If there is a dichotomous variable that interacts with the treatment variable, the treatment by placebo odds ratio, our basic indicator of a treatment effect, is different depending on whether the covariate is absent (coded 0) or present (coded 1). In such a situation it is not possible to refer simply to "a treatment effect" because there are two different ones. In the modeling process, the interaction between t-PA & a Covariate is tested by inserting the product of the t-PA indicator and the covariate, with regression slope θ, into the model and testing whether $\theta = 0$. This test involves estimating θ and its standard error. In Tables 5.17 through 5.21, we summarize 20 such tests by reporting the estimates of θ, the standard errors and the corresponding p-values. None of the p-values are <0.05, so we report that these interactions are not significant and do not retain them in the models when we summarize the treatment effect. However, tests of no interaction have notoriously low power, a point we will examine in a moment.

The interpretation of θ is summarized by:

$$e^{\theta} = \frac{OR(\text{cov} = 1)}{OR(\text{cov} = 0)},$$

where OR(cov = 1) is the t-PA versus placebo OR in the presence of the covariate, and OR(cov = 0) is the OR in the absence of the covariate.

Only when $\theta = 0$ is the ratio of ORs equal to 1, indicating that the effect of t-PA is the same whether the covariate is present or not.

5.5.2.2 Power of Interaction Tests: As mentioned, we report all 20 tests of no interaction of t-PA with dichotomous covariates as being not significant ($p > 0.05$). But, this study, as is the case for most clinical trials, was not designed to have much power to assess such interactions. In the tables below, we report how large the ratio of the two ORs would have to be before our tests would have had an 80% chance of being significant.

These “minimally detectable ORs” are based on the level of significance being set at 0.05 (2-sided in the first table and 1-sided in the second), the power set at 0.80 and using the empirically observed estimates of the standard error of the various estimates of θ.

	Barthel	Rankin	Glasgow	NIHSS	GLOBAL
OTT	3.6	3.8	3.6	4.0	2.9
PART	3.6	3.7	3.6	3.9	2.9
PrDisab	28.1	33.4	33.1	79.7	16.5
PrDM	4.8	5.2	5.1	6,4	3.7
OTT	2.7	2.8	2.7	2.9	2.3
PART	2.7	2.7	2.7	2.9	2.3
PrDisab	12.9	14.8	14.7	28.8	8.6
PrDM	3.3	3.6	3.5	4.1	2.8

Table entries are the ratio of odds ratios that would have to actually exist for the interaction tests just performed to have an 80% probability of being statistically significant. For example, the t-PA effect, as measured by the t-PA vs. Placebo OR, would have to be about 4 times higher (lower) in those with DM than in those without DM in order for these tests to have a reasonable likelihood of detecting the interaction.

The interactions involving polychotomous covariates with multiple degrees of freedom have even less power than indicated by these tables because they involve the spreading of patients over multiple classes, with smaller numbers per class.

Thus, while we, and the NINDS investigators, examined these interactions and report that they are not statistically significant, lack of significance does not imply the absence of interactions. Indeed, lack of evidence of an effect is not equivalent to evidence of the lack of an effect. Caution is needed in evaluating subgroup effects.

5.6 Absolute Risk Differences: Sensitive to the several concerns raised by many subsequent to the NINDS publications, the FDA approval, and the American Heart Association endorsement, the committee continued with further analyses of the data. Some of those analyses will be discussed in the subsequent sections; here we will discuss estimating the difference between the probabilities that a t-PA treated patient and a Placebo treated patient will experience a favorable outcome.

5.6.1 Differences Between Success Rates: Recall that in Table 5.8 the observed differences between success (favorable outcome) rates (%) for the t-PA and Placebo treatment groups were: 14.1%, 16.3%, 14.4%, & 13.7% for the Barthel, Rankin, Glasgow, and NIHSS outcome measures respectively. These estimates translate more directly than odds ratios into interpretations of the impact of treating a population of acute ischemic stroke

patients. If 1000 patients were treated according to the NINDS protocol, these numbers suggest that between 140 and 160 more patients would experience a favorable outcome than if t-PA therapy was not available or was not used on the whole population.

As outlined in Section 4.3.7, these differences between the t–PA and Placebo success rates can be estimated using the adjusted OR estimates. The resulting estimated differences and their 95% Confidence Intervals are: 19.3% (9.6-29.0%), 20.2% (10.6-29.8%), 17.9% (8.3-27.5%), and 15.6% (6.8-24.4%) respectively for the Barthel, Rankin, Glasgow and NIHSS outcome measures. Thus, after taking into account the modeling process, which led to slightly larger OR estimates, the estimated differences are also greater than the actual observed differences. In light of these differences, it seems reasonable to suggest that between 150 & 200 more of the hypothetical population of 1000 acute ischemic stroke patients would experience a favorable outcome if all 1000 are treated with t-PA according to the NINDS protocol than if none are.

5.6.2 Attributable Fraction: Computation of "attributable risk" or "attributable fraction" sheds additional light on the role of t-PA as a treatment for acute ischemic stroke. We can use this concept, which originated in the field of epidemiology to estimate what fraction of a disease in a population might be reasonably "attributed" to the presence of a risk factor in the population, to estimate the proportion of the unfavorable outcomes that can be "attributed" to exposure to the Placebo. Such estimates which are based on our OR estimates and the fraction of placebo patients with an unfavorable outcome can be interpreted as that fraction of the unfavorable outcomes that could be eliminated if all Placebo treatment could be eliminated in favor of the t-PA treatment. From the raw data, the estimates of the "attributable fractions are; 24.8%, 20.8%, 25.7% and 27.6% respectively for the Barthel, Rankin, Glasgow, and NIHSS outcome measures. The corresponding numbers using the model based OR estimates instead of the raw data are: 30.8%, 26.1%, 29.7%, and 29.8%. Thus, if we take the Barthel index as an example, 61.9% of the placebo treated patients had an unfavorable outcome. In our hypothetical 1000 acute ischemic stroke patients we thus expect ~620 to have an unfavorable outcome if all were treated with placebo. If, in contrast, all were treated with t-PA we expect a reduction in this number of unfavorable outcomes by between 25% (unadjusted) and 30% (adjusted). That is we expect between 435 and 465 unfavorable outcomes rather than 620 or a reduction of between 155 and 185 unfavorable outcomes achieved through the use of t-PA therapy. These numbers are obviously very similar to the figures quoted above corresponding to the increase in the number experiencing a favorable outcome.

5.7 Public Health Consequences and Exploratory Subgroup Analyses: The results of a clinical trial lead to population-based decisions rather than to patient-specific decisions. In this public health context, we (and the NINDS investigators) conclude that the use of t-PA in accord with the NINDS protocol will result in an increase in the total number of favorable responses among those acute ischemic stroke patients who satisfy the conditions of the protocol. However, physicians and patients face patient-specific decisions, even among patients who meet the conditions of the protocol, and further refinement of the results would be helpful in making these decisions. In subsequent chapters we examine subgroups to determine if there are groups of acute ischemic stroke patients who satisfy the conditions of the protocol but might not fare as well on t-PA therapy as the overall evidence suggests. In interpreting these subgroup analyses it is important to keep in mind both that the study was not designed to have substantial power to assess subgroup differences so these tests may fail to detect real

differences (see Sections 4.3.6 and 5.5.2) and that performing many exploratory analyses may deliver spuriously “significant” findings.

6. Blood Pressure Assessment and Management

6.1 Stated Methodology

6.1.1 Original Protocol: The NINDS investigators, in their first publication[1], made the following statements regarding patient eligibility for the clinical trial; (a) patients did not undergo randomization if they had "... a systolic blood pressure above 185 mmHg or diastolic blood pressure above 110 mmHg;...." (p 1582 Col 1 Para 3), (b) patients were also excluded if aggressive treatment was required to reduce their blood pressure to the specified limits. (p 1582 Col 1 Para 3), and (c) the protocol required that ... "blood pressure be maintained within prespecified values." (p 1582 Col 1 Para 5)

6.1.2 Blood Pressure Manuscript: In a subsequent publication[12], the NINDS investigators stated:" All patients had BP measurements at the time of admission to the emergency department and at the time of randomization (equivalent to the time of study-drug initiation) those with a systolic BP of $\leq$185 mm Hg and a diastolic BP of $\leq$110 mm/Hg were eligible for randomization." Patients with higher BP readings at the time of admission but who met BP criteria by the time of randomization were defined as hypertensive before randomization.

Between admission and randomization, aggressive antihypertensive therapy, defined as use of intravenous nitroprusside or repeated intravenous infusions of other medications, could not be used to meet eligibility criteria. After randomization, BP measurements were collected prospectively on a scheduled basis (Appendix 2)[12]. Patients with elevations of systolic BP >180 mm Hg or of diastolic BP >105 mm Hg in the 24 hours after randomization were defined as hypertensive after randomization. For such elevations, repeat BP determinations were recommended every 5 to 10 minutes but were not recorded in the trial. Prespecified antihypertensive treatment guidelines were given (Appendix 2)[12]. The date of administration of any antihypertensive treatment was recorded but not the time of administration. Acute antihypertensive therapy was defined as administration of intravenous nitroprusside, nicardipine, labetalol, or hydralazine; sublingual nifedipine; and sustained-released or topical nitroglycerin.

"To explore the relationship among BP reduction, thrombolytic therapy, and antihypertensive therapy, the severity of hypertension and declines in BPs were calculated at various time frames from randomization. To evaluate severity of hypertension, for each patient in the study the maximum mean arterial pressure during the first 24 hours after baseline was calculated. " To identify precipitous drops in BP soon after initiation of placebo or t-PA, the maximum abrupt decline, defined as the maximum decline between two consecutive mean arterial pressures during the first 8 hours, was calculated (measurements were hourly after the first 8 hours)."

6.1.3 Systems Approach: In another manuscript[8], the investigators offer additional guidelines (ibid. Table 5, p. 1539); (a) Patient Selection: contraindications: "On repeated measurements, systolic BP>185 mm Hg or diastolic BP>110 mm Hg at the time treatment is to begin, and patient requires aggressive treatment to reduce BP to within these limits:" and (b) BP control: pretreatment " Monitor BP every 15 minutes (should be <185/110 mm Hg), if >185/110, BP may be treated with one to two 10- to 20-mg doses of labetalol given IV push within 1 hour and/or nitroglycerin paste. If these measures do not reduce BP < 185/110 and keep it down, the patient should not be treated with rt-PA."

6.1.4 FDA Submission: PLA supplement 96-0350 Submitted by Genentech to the FDA

3/19/96[21]: (a.) (p23-24) It is stated that 17,367 patients with strokes were screened, but not enrolled; but none appear to have been excluded for reasons due to blood pressure. Unless BP is included in other serious illness 490/17367 this certainly would imply rapid spontaneous or therapeutic control of blood pressure. On further review the reasons for exclusion on p24 of the Genentech submission document only 95% of all patients. Possibly the investigators did not include patients excluded because of elevated BP. If so, then the 5% would be a reasonable estimate of the number potentially excluded for BP.

Exclusion reason	Number	%
Time from onset too long	8708	51.6
Symptoms rapidly improving	1749	10.4
Intracranial Hemorrhage	1306	7.8
Symptoms too minor	1106	6.6
Outside age range	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[1]	267	1.6
Recent prior stroke	219	1.3
Oral anticoagulants	210	1.2
Subarachnoid Hemorrhage	169	1.0

6.1.5 Exclusion Characteristics: In an attempt to define the exclusion characteristics of the study these data were compared with the investigators' report of 17,324 patients in the their 1997 manuscript[23]. The data sets in these two documents provide inconsistent tallies. The allocated percentages are inconsistent. The investigators suggest that these discrepancies are primary exclusionary criteria, but these also do not achieve 100% of the population. The graphic represents 93% of the population. Figure 3 in the manuscript rounds up/down inconsistently with relation to the Table above.

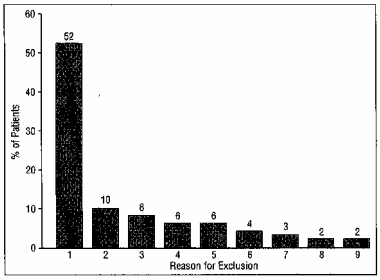


Figure 3. *Primary reasons for exclusion from the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial (17,324 patients screened; 624 patients randomized). Reasons are coded as follows: 1, time; 2, improving rapidly; 3, intracranial hemorrhage on baseline by computed tomographic scan; 4, minor symptoms; 5, age; 6, serious coexisting illness; 7, seizure at stroke onset; 8, stroke ruled out; and 9, other reason.*

On p 18 of the clinical review it is stated that the most common protocol violation involved blood pressure criteria represented by 29/54 patients with violations of the 624 study cases. No details are offered with regard to the 29 patients who had a BP eligibility violation. There are no details in the FDA submission, or details in any of the manuscripts.

Via teleconference calls and email communications with Drs. Tilley and Brott, the following information was obtained related to BP and its management:

a) Figure 3 from the original manuscript displayed causes for exclusion of the 16,741 patients out of the total 17,363 who were screened[1]. The Table showed that 2 percent were excluded for "other reasons," and of that group, 162 were excluded because of high blood pressure. Seven of those patients were trial patients. Patients whose primary reason for exclusion was something other than blood pressure might have had blood pressure issues as well.
b) The exact number of patients given BP lowering medication prior to receiving treatment with the study medication was unknown. No information was available on patients who were treated by non-study physicians before the study physician's arrival in the ED.
c) While information was available as to which cardioactive drugs were given to the study patients, no information was available regarding the indications for giving the medications. Thus, it was not known if the cardioactive drugs were given for reasons other than lowering BP such as the treatment of chest pain or managing the ventricular response rate in patients with atrial fibrillation. The investigators reported that they reviewed the medication list (prior to identifying any patient characteristics) to make suggestions as to which medications their reviewer should consider as antihypertensive therapy. Genentech nurses determined which specific medications recorded on forms were to be considered antihypertensive therapies and so coded the agent.

6.2 Stated Data Sets

6.2.1: From the Original Publication[1]

Table 1. The Medical Histories of the Patients in the Study. (p 1582)

Variable	**Part 1**		**Part 2**	
	t-PA N=144	Placebo N=147	t-PA N=168	Placebo N=165
	percent			
Stroke	17	17	12	9
Transient ischemic Attack	22	14	13	19
Aspirin therapy	41	31	40	26
Diabetes	24	21	20	20
Hypertension	66	64	67	67
Myocardial infarction	25	21	22	20
Atrial fibrillation	18	20	20	16
Angina pectoris	18	22	24	24
Congestive heart failure	14	17	16	19
Valvular heart disease	11	7	6	6
Smoking in year before stroke	43	37	27	35
No preexisting disability	90	91	95	93

Table 2. Base-line characteristics of the patients in the two parts of the study, according to treatment group. (p1583)[1]

Characteristic	Part 1		Part 2	
	t-PA N=144	Placebo N=147	t-PA N=168	Placebo N=165
Blood Pressure (mm Hg)				
Systolic	155±22	153±20	153±22	152±21
Diastolic	85±12	85±13	85±14	86±15

6.2.2 From the Manuscript on BP[12]: "Hypertension was present on admission for 121 (19%) of the 624 patients eventually randomized into the NINDS rt-PA Stroke Trial; 65 were placebo-treated patients and 56 were t-PA-treated patients (Table 1). Postrandomization hypertension was detected during the first 24 hours in 372 patients (60%); 195 were placebo-treated and 177 were t-PA patients. For all patients, the frequency of antihypertensive therapy was similar for both the placebo- and t-PA-randomized patients. Before randomization, 28 (9%) of the 312 placebo patients and 28 (9%) of the 312 t-PA patients received antihypertensive treatment, whether or not they were hypertensive as defined above; 1 patient in the t-PA treated group, included in our analysis, received aggressive antihypertensive therapy (i.e., intravenous nitroprusside, a protocol violation). After randomization, 92 placebo patients (29%) and 75 t-PA patients (24%) received antihypertensive therapy. Antihypertensive therapy was administered either before or after randomization to 110 placebo patients (35%) and 96 t-PA (31%) patients." (p 1506)

6.2.3 Hypertension on Admission and Antihypertensive Therapy Received Before Randomization (p1506)[12]: Of the 121 patients who were hypertensive on admission, slightly more placebo patients received antihypertensive therapy (22 of 65, 34%) before randomization than did t-PA patients (11 of 56, 20%), but the difference was not significant (Table 1). The effects of antihypertensive therapy before randomization were similar in the groups randomized to t-PA and placebo for all clinical outcomes except death at 3 months (Table 2).

Table 1. Antihypertensive Therapy by t-PA-Treated and Placebo-Treated Groups

Hypertension Recorded	Received Anti-Hypertensive Therapy Placebo n	Placebo %	t-PA n	t-PA %	P*
Admission **, }	65	34	56	20	0.17
Within 24 hours after randomization **, ω	195	41	177	37	0.33
Not hypertensive by definitions	109	9	127	11	0.81

*Mantel-Haenszel test adjusting for centers and time strata.

**Groups are not mutually exclusive.

6.3 Stated Goals, Discussion And Conclusions: In the first t-PA manuscript[1], the investigators stated; "In our trial treating physicians used an algorithm to manage blood pressure after treatment began." (p 1586. Col 2 Para 2). In the hypertension manuscript[12] they stated;

"BP eligibility criteria more stringent than those used for t-PA-treatment of acute myocardial infarction were instituted, but aggressive measures to lower BP to allow enrollment were prohibited to prevent precipitous falls in BP. After initiation of t-PA therapy, a BP management

algorithm was followed, adapted from a similar algorithm designed for treatment of stroke patients in general (9). Recommended drugs were selected because of their rapid onset of action and because of their predictable effects with low potential for overshoot. Adjustments in the algorithm were made in response to experience during the course of the trial."

In the NINDS t-PA Stroke Trial[1] the investigators chose BP eligibility criteria similar to those used in the dose-finding trial (p 1505 Col 1 Para 2). The authors focus on a systolic BP of > 185 mmHg and a diastolic BP of > 110 mmHg at admission. Tables 4 and 5, which focus on blood pressure deal solely with severity and rate of reduction of the mean arterial blood pressure. Abrupt decline is analyzed in two ways as per Appendix 2; q 15 min in first 2 hours and q 30 min in hrs 2-8 following randomization.

The authors emphasized “gentle management” (p 1504) in those patients “who were hypertensive”[12]. In the last paragraph the authors state that "after initiation of t-PA therapy a BP management algorithm adapted from an American Academy of Neurology guideline[24] was followed.

In Subjects and Methods, aggressive therapy was defined as intravenous Nitroprusside or *repeated* infusions of other medications. Their chosen antihypertensive intravenous medications were stated to be nicardipine, labetalol, or hydralazine or sublingual nifedipine and sustained release or topical nitroglycerin. Based on our other data set, furosemide and diltiazem were also utilized as therapeutic agents for reasons determined by study monitors.

It is not clear how the Appendix 2[12] relates to the Subjects and Methods section. As the authors did not initially use the mean blood pressures for study entry their emphasis on mean vs. systolic or diastolic does not describe individual abnormalities.

In their discussion they state; “The antihypertensive therapy used in the NINDS study was modest in its effects and had little potential for overshoot. Hypertensive placebo patients who received the antihypertensive therapy after randomization did not have a greater maximum decline in mean arterial pressure over the first 24 hours compared with hypertensive patients who did not receive antihypertensive therapy. In addition, abrupt declines in BP were not more pronounced among placebo patients who were treated with antihypertensive therapy compared with those who were not, reflecting the careful use and gentle effects of the antihypertensive therapy administered in this study (Appendix 2)[12].”

The interaction of antihypertensive therapy with intravenous t-PA in this exploratory analysis is intriguing, but interpretations should be cautious. For the patients randomized to receive t-PA, antihypertensive therapy administered before t-PA was not associated with differences in early or late outcomes. However, hypertensive t-PA patients who received antihypertensive therapy had a more pronounced abrupt decline in mean arterial BP. Hypertensive t-PA patients who received antihypertensive therapy after randomization were less likely to have a favorable outcome at 3 months than hypertensive t-PA patients who did not. One possible explanation is the nonrandomized administration of antihypertensive therapy at the bedside. Investigators could have been more likely to treat hypertensive patients they judged to be sicker. “ (p 1508 Col 1 Para 2, 3)[12]

“A randomized trial would be necessary to address adequately the effects of antihypertensive therapy on BP and on clinical outcome.” (p1508 Col 2 Para 2)

"In summary, hypertension was a common phenomenon in the NINDS trial. BP eligibility criteria were applied in a balanced fashion. The antihypertensive therapy was designed for, and resulted in, modest effects on BP with low potential for overshoot. The results do not suggest that use of antihypertensive therapy adversely affected BPs or clinical outcomes of placebo-randomized patients. The effects of antihypertensive therapy following treatment with t-PA are complex and merit further study. Careful attention to BP and gentle management remain warranted for stroke patients treated with t-PA." (p 1508, Col 2, Para 2, 3)

In the investigators' manuscript on ICH[11] we find (p.2111, last Para.) under the heading Baseline and Time Dependent Covariates a first citation for admission diastolic blood pressure > 100 mmHg as associated with increased risk of symptomatic ICH. Later, p. 2114, Para 1, the authors suggest high correlations between systolic BP and mean BP and between systolic BP and pulse pressure. On several occasions such as the next to last paragraph of the Methods in the last sentence the authors state "prerandomization and postrandomization antihypertensive therapies were evaluated with patients who were hypertensive." Under Results: In the next paragraph the authors state: "patients received antihypertensive treatment whether or not they were hypertensive as defined above." In the last paragraph under Results (maximum BPs and declines in BPs) they state that the "more severe BPs were more likely to be treated". The last sentence in that same paragraph with regard to BP decline states that "abrupt declines were noted more frequently in treated patients."

6.4 The NINDS t-PA Review Committee's Evaluation of the BP issue

6.4.1 Review Data sets: As described in Section 4.1 of this report, the review committee had access to extensive data and we sought strict definitions of the following variables, their names, and their locations. When necessary, more information was requested and some clarification was obtained.

(i) Hypertension: Prestroke, Post stroke – Prerandomization, and Post Randomization

(ii) Hypertension Therapy: Prestroke, Post stroke – Prerandomization, and Post Randomization

(iii) Blood Pressure: At Admission, At Randomization, and Subsequent to Randomization

6.4.2 Investigator Queries

(i) When comparing the mean and SD of baseline systolic (BsSYS) and diastolic (BsDIA) blood pressures with admission: systolic (AdSYS) and diastolic (AdDIA) blood pressures, the admission values were higher.

	Mean	Std Dev	Max
BsSYS:	153.12	21.27	227
AdSYS:	158.92	21.33	254
BsDIA:	85.32	13.53	134
AdDIA:	89.24	15.81	180

(ii) The investigators' study form 10 section C item 2 asks if the patient has a history of hypertension with the options of answering yes, no, or unknown. Item 2b followed up by asking "If yes was medication prescribed?" and again the answers are yes, no or unknown.

(iii) A review of the descriptive characteristics of the submitted variables and, the range of blood pressures in the dataset indicates that some of these readings would have placed a substantial number of the patients in an exclusionary status. Inclusion would have been in violation of the upper systolic and/or diastolic blood pressures limits established.

(iv) Drs. Tilley and Brott stated (personal communication) that blood pressure variables for readings at admission (on arrival at the ED) and baseline (time of randomization) should be available.

The investigators stated (personal communication) that they used the randomization blood pressure variables in their analyses, but they stated that they may have used the terms baseline and randomization interchangeably.

Although patients may have had additional blood pressure readings prior to randomization and after randomization only the randomization blood pressure was recorded.

The investigators stated that admission blood pressure could be quite high, but if an antihypertensive regimen could lower the blood pressure to 185/110 or below at any time before randomization, the patient could be randomized into the trial. Post-admission prerandomization medication information was not collected.

It was uncertain whether the authors restricted lowering BP to those who were hypertensive by their exclusion criteria or at any other specific levels. The authors do not define precisely their therapeutic goals: How far below the cut off post randomization values of 180/105 did they wish to go?

The investigators did not offer information defining specific data as to what antihypertensive therapy was employed. The protocol reviewer utilized the list of medications given during hospitalization with the times administered to determine what prerandomization drugs had been used to treatment hypertension. Labetalol was considered antihypertensive every time, whereas calcium channel blockers and diuretics were reviewed and judged to be antihypertensive or not depending on a retrospective chart analysis. This could have been a source of error, but the investigators believe that it was an error that affected t-PA and control patients uniformly. The investigators' goal was to look at the effects of antihypertensive therapy given after admission to hospital.

The investigators' determination of pre-randomization treatment referred to the history of hypertension question ("yes/no/unknown" from form 10, item 2.) The assessment of post randomization hypertensive therapy and its influence on outcome was based on post-hoc evaluations of the medications given and an ad hoc decision determining whether this represented antihypertensive treatment or not.

Dr. Brott stated that in Cincinnati they did not treat high blood pressure to permit entry into the

trial, but that other centers engaged in this practice at the time of the study. Throughout the documents it is suggested that it is acceptable to treat hypertension so that patients can be treated with t-PA providing that the treatment is not 'aggressive'.

6.4.3 Review of Study Datasets: Our examination of the blood pressure data led to the following observations:

(i) Nineteen individuals were found with abnormally elevated (BP >185 mmHg or >110 mmHg) at both admission and baseline.
(ii) Ten individuals were found to be hypertensive at baseline who had not been so at admission. Although this table states >185/110 it actually means either >185 mmHg or >110 mmHg.

	Baseline BP	
≤185/110	572	95.17%
>185/110	29	4.83%

Twenty-one patients were missing baseline BP.

	Admission BP	
≤185/110	501	80.68%
>185/110	120	19.32%

One patient was missing admission BP.

		Admission BP	
		≤185/110	>185/110
Baseline BP	≤185/110	474 (79.00)	10 (1.67)
	>185/110	97 (16.17)	19 (3.17)

Twenty-two patients were missing admission or baseline BP.

(iii) Admission BP readings were missing in 1 patient, and randomization BP readings were missing in 21 patients.

(iv) A pair wise comparison of the recorded admission and baseline blood pressures for the entire cohort was performed. Restricting our attention to the 622 patients who were randomized into the study within 180 minutes of onset it was noted that 112 (18%) of them had identical admission and baseline blood pressures.

6.5 Review Committee's Areas of Concern:

6.5.1 Definitions The NINDS trial[1] had no specific definitions for the 'Prior Medical History' conditions, including a 'History of hypertension'. It was left to the discretion of the investigators at each site to determine if a patient had one of these conditions. With regard to the history of hypertension, we were unable to determine those patients who had a previous history, a

current history, whether treatment was current or whether treatment had occurred in the ambulance prior to hospitalization.

6.5.2 Protocol Applications There appeared to be some patients included in the study whose blood pressures at randomization exceeded either the systolic or diastolic maximum permissible values.

6.5.3 Therapeutic Interventions Dr. Tilley stated (11.11.02) that the data set included blood pressure at admission and blood pressure at baseline plus post randomization blood pressures.

6.6 Review Committee's Findings:

6.6.1 Definition of Hypertension: Publications from the t-PA studies and written and oral communication with Drs. Tilley and Brott document confusing and inconsistent information with respect to nominal and actual procedures for BP recording and management, and confusing and inconsistent nominal and actual eligibility and exclusion criteria. It was never defined as to what was precisely meant when the term hypertensive was used. Was it always based on their exclusion values, standard terms, or history of treatment?

6.6.2 History of Hypertension: Item 6.7.1 creates confusion with regard to the numbers of patients considered to have a history of hypertension throughout diverse comments and manuscripts. This resultant variability stems from the uncertainty of the definitions of hypertension and of the term history of hypertension. Although an analysis of current, recent or past use of hypertensive medications could be of interest, the data definitions are not sufficiently precise to support these exploratory analyses.

6.6.3 Blood Pressure on Admission and at Baseline: There appears to be a persistent uncertainty of the investigators in their written and stated use of terms. We confirmed that the terms baseline and at randomization values have been used interchangeably. In various manuscripts this confusion seems to be represented periodically. We demonstrated that admission and baseline blood pressure were identical across all centers 18% of the time which suggests that the interpretation for each term was confused at various times. At one center 50% of the blood pressure values were identical at admission and baseline.

Teleconference and email communications with Dr.'s Brott and Tilley revealed that there was variability between centers in the interpretation of the definition of admission BP. This led to some centers using the BP reading taken at the time of randomization as both the admission and randomization BP measurements. There were also some patients where the admission BP reading was the BP reading taken at the time of admission into the ICU after receiving the study drug.

6.6.4 Blood Pressure Exclusion Criteria: There was substantial inconsistency in the presentation of the exclusionary blood pressure criteria for entry into this study. It appears that the investigator's intent was to exclude patients whose blood pressure at the time of randomization:

(1) Exceeded either 185 mm Hg systolic or 110 mmHg diastolic on repeated measures.

Or

(2) received “aggressive antihypertensive therapy” to fall within these limits.

Drs. Tilley and Brott explained that a patient whose pre-randomization BP readings were persistently ≤ 185/110, could be included in the study even if the BP reading at the time of randomization exceeded 185/110.

While the Stroke Trial Guide manuscript stated clearly that the exclusionary BP criteria were based on repeated BP readings, throughout the relevant papers these criteria are written as

(1) Exceed 185 mm Hg systolic and 110 mm Hg diastolic.

Or

(2) > 185/110

These are not equivalent exclusion criteria, confuse the reader, and may have confused the investigators at various sites as 29 patients may have been included in the study with randomization blood pressures that would have required exclusion.

6.6.5 Antihypertensive Therapy Before Randomization: The concept of aggressive therapy is uncertain in written terms to the investigators and probably to the site practitioners. Prehospital therapy by EMS and other pre-randomization therapy could have included numerous diverse exceptionally efficacious rapid acting agents without being termed aggressive. The effects of these interventions in addition to all else that was done between admission and baseline makes the blood pressure determinations of limited value from an analytic perspective.

The caveats with regard to what was or was not considered antihypertensive treatment remains a concern. What agents? What doses? At what time? To whom? What results? Patient charts were evaluated retrospectively. In addition some patients would have been treated pre-randomization with delayed effects post-randomization.

Did giving an antihypertensive medication result in a lowering of BP? The investigators theoretically had a large number of patients who became hypertensive (again) post randomization. This is an interesting question; however the quality of the existing dataset may not allow for a proper analysis.

It is stated that 9% of all patients enrolled in the study received prerandomization treatment[12]. It would appear that the investigators substantially underestimated the number of patients who were treated with antihypertensive medication prior to randomization in view of the neglected or unidentified treatment regimens utilized outside the study protocol.

Other patients would have had blood pressure return to normalcy because stress, pain, hypoxia and other clinical issues were treated. Although managing BP was an important part of the protocol, the q15 minute BP readings that were required prior to giving t-PA, were not recorded neither was it recorded whether medication was given specifically for treating elevated BP. There was no formal protocol of sequential BP measurement that might allow for analysis of peak effect or duration of drug effect.

6.7 Use of Blood Pressure Data in Review Committee Outcome Models: Because of the concerns expressed above, the review committee decided not to incorporate information regarding blood measurement or management obtained during the prerandomization workup in our principal models of treatment effect. However, some have questioned this decision and so we include a summary of the assessment of baseline (randomization) blood pressures as predictors of favorable outcome using the same methods as were used to derive the best covariate model as described in the analysis section of this report. Recall that the variables fixed in all models – for each of the four outcome measures as well as for the global analysis – were the three stratification variables (Center, Part, and OTT ($\pm$90 mm)) and, ultimately, the covariates BsNIHSS, Age, BsDisab, BsDM, and the interaction between Age and BsNIHSS that were selected as described. No blood pressure variables were included in the process of determining which covariates to include in the adjusting models

6.7.1 Blood Pressure Variables as Favorable Outcome Predictors There were seven baseline blood pressure variables as described and defined in Tables 4.1 and 4.2. As seen in Table 4.2, there were 22 patients with missing values for these measurements. Values were imputed for these patients as described in Section 4.3.2. In the table below for each of the seven variables and each of the four outcome measures as well as for the global analysis we provide the chi-square and p-values which these variables would have carried into the stepwise process had they been included.

	Barthel		Rankin		Glasgow		NIHSS		Global	
	χ^2	p-v	χ^2	p-v	χ^2	p-v	χ^2	p-v	χ^2	p-v
BsSYSbp	2.36	0.12	2.51	0.11	2.20	0.13	0.43	0.50	1.53	0.21
BsSYSbp>190	0.25	0.61	1.96	0.16	0.26	0.60	0.91	0.33	0.94	0.33
BsDIAbp	0.07	0.77	3.05	0.08	1.30	0.25	0.93	0.33	0.36	0.54
BsDIAbp>100	1.94	0.16	0.69	0.40	0.43	0.50	0.26	0.60	1.53	0.21
BsMBP	0.94	0.32	3.74	0.05	2.23	0.13	0.92	0.33	1.04	0.30
BsMBP>130	0.34	0.55	0.74	0.38	0.12	0.72	0.35	0.55	0.27	0.60
BsPulseP	2.40	0.12	0.32	0.56	0.78	0.37	0.00	0.94	0.96	0.32

These χ^2 and p-values should be compared to the values for the variables that are summarized in Table 5.10 of Section 5.4. Note particularly the p-values for BsNIHSS, Age, and Pr Disability, all of which were <.0001. While some of these BP variables were "borderline significant", none were even remotely as important as predictors of favorable outcome as the variables ultimately included in the models.

The next table shows, for the four outcome measures and the global analysis, the χ^2 and p-values corresponding to each of the seven BP variables if they were each individually added

to the models including the 3 stratification factors and the 5 covariates selected for our analyses.

	Barthel		Rankin		Glasgow		NIHSS		Global	
	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
BsSYSbp	1.97	0.16	4.39	0.03	3.92	0.04	0.64	0.42	1.74	0.18
BsSYSbp>190	1.15	0.28	4.26	0.03	1.53	0.21	1.93	0.16	2.68	0.10
BsDIAbp	1.78	0.18	7.00	0.00	3.97	0.04	2.57	0.10	3.38	0.06
BsDIAbp>100	0.18	0.66	0.00	0.94	0.08	0.77	0.12	0.72	0.04	0.84
BsMBP	2.45	0.11	7.43	0.00	5.11	0.02	2.00	0.15	3.53	0.06
BsMBP>130	1.25	0.26	1.91	0.16	0.68	0.40	0.97	0.32	1.38	0.24
BsPulseP	0.45	0.49	0.30	0.57	0.78	0.37	0.04	0.82	0.05	0.81

From this table we see that only in a few instances would any of these variables be selected for inclusion in the models predicting favorable outcome. For all but one of the outcome variables, the blood pressure variable with the smallest p-value and therefore at the top of the list to be added was BsMBP. So, we entered that variable into each model and the following table illustrates the impact that had on the remaining six BP variables.

	Barthel		Rankin		Glasgow		NIHSS		Global	
	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
BsSYSbp	0.02	0.87	0.19	0.65	0.01	0.91	0.58	0.44	0.14	0.70
BsSYSbp>190	0.34	0.55	1.90	0.16	0.25	0.61	1.08	0.29	1.06	0.30
BsDIAbp	0.02	0.87	0.19	0.65	0.01	0.91	0.58	0.44	0.14	0.70
BsDIAbp>100	2.29	0.13	3.09	0.07	1.33	0.24	0.27	0.60	2.10	0.14
BsMBP>130	0.21	0.64	0.03	0.86	0.07	0.80	0.17	0.68	0.10	0.75
BsPulseP	0.03	0.87	0.19	0.66	0.01	0.91	0.59	0.44	0.14	0.70

Clearly, only if we were very generous regarding the qualifications necessary for a variable to enter the model as a covariate would any of these variables enter.

6.7.2 Influence of BP Variables on OR Estimates We now assess the impact that the addition of these variables on the estimate of the t-PA effect. To set the stage, recall, (Table 5.15) that for the Barthel, Rankin, Glasgow and NIHSS outcome variables, the raw (unadjusted) odds ratio estimates were, respectively, 1.78, 2.07, 1.85, and 2.01. Following adjustment by the extremely significant covariates included in the model, these estimates

became, respectively, 2.19, 2.43, 2.13, and 2.19. In other words, adjusting for extremely significant covariates increased the odds ratio estimate by a relatively small amount.

Following the addition of BsMBP and any others with a p <0.10, the corresponding odds ratio estimates became: 2.20, 2.51, 2.16, and 2.20, respectively. Thus, inclusion of the blood pressure variables, which were only marginally related to the outcome, had, predictably, almost no influence on the odds ratio measure of treatment effect.

In the global analysis, the BsMBP variable had a p-value to enter the model of 0.06. After it was allowed to enter, the next most "significant" variable was BsDIAbp>100 with a p-value of 0.15 so no other blood pressure variable other than BsMBP was entered into the global model. The Global odds ratio estimates were:

Unadjusted:	1.88
Adjusted (w.o. bp variables):	2.13
Adjusted including BsMBP:	2.14

The addition of the blood pressure variables had no impact on the estimate of the t-PA to Placebo odds ratio estimate.

6.8 Summary and Conclusions

Our analysis identified a number of problems regarding pre- and post-randomization blood pressure measurement and management:

- Non-compliance with the defined protocol was substantial, and persistent, throughout the study with regard to both the documentation of blood pressure readings, and adherence to the treatment regimen for hypertension.

- There was limited rigor with regard to the pharmacologic characteristics of antihypertensive regimens. In some instances pharmacologic monitoring was performed by representatives (nurses) of the sponsoring pharmaceutical firm. Medications employed were listed by date, but not by time, eliminating consequential interpretive utility.

- The exact number of patients who received medication to lower blood pressure either prior to, or after, receiving study treatment is unknown.

- The confusion regarding blood pressure documentation, and the lack of knowledge of treatment of hypertension either prior to, or after, receiving study treatment, could have led to an unknown number of patients receiving treatment in violation of the nominal study protocol.

Based on these observations, we reached the following conclusions:

- It was not possible to assess the effect of hypertension management on clinical outcome in acute ischemic stroke patients treated in the NINDS study.

- The blood pressure variables should not be included in the statistical models. However, we also found that inclusion of the blood pressure variables in the statistical models would have been inconsequential with regards to altering the t-PA treatment effect.

Finally, the inconsistent documentation of both blood pressure readings and hypertension management seriously undermines the NINDS investigators statement that blood pressure management was a significant part of the protocol that contributed to the success of the study. Nonetheless, we concur with the NINDS investigators premise that blood pressure management should be included in the protocol for treating acute ischemic stroke patients with t-PA. It is biologically plausible that hypertension management could affect clinical outcome in acute ischemic stroke patients treated with t-PA, and data from the cardiology literature has already demonstrated that in acute myocardial infarct patients, the risk of having an intracerebral hemorrhage is related to pre-treatment blood pressure[25-27]. However, further clinical studies will be needed to assess whether blood pressure management is related to better clinical outcomes in acute ischemic stroke patients treated with t-PA.

7. INTRACEREBRAL HEMORRHAGE

7.1 Introduction: Prior to the initiation of the t-PA trials, there was concern that t-PA therapy for ischemic stroke might increase the risk of an intracerebral hemorrhage (ICH) to an unacceptably high level. Indeed, the NINDS investigators specifically stated, "… the use of rt-PA for cerebral arterial thrombolysis requires a careful evaluation of both the risks and potential benefits" (p. 1581)[1]. In each of the two primary studies, ICH was considered a serious adverse event and, consequently, the protocols required that a CT scan be performed at 24 hours and between 7 to 10 days after randomization and whenever symptoms suggested an intracerebral hemorrhage. A symptomatic intracerebral hemorrhage was defined as "a CT-documented hemorrhage that was temporally related to deterioration in the patient's clinical condition in the judgment of the clinical investigator"[1]. Asymptomatic hemorrhages were defined as those confirmed by the protocol designated CT, in the absence of symptoms[1].

The investigators' protocol stated that, "Interim analyses were required after every three symptomatic ICHs and after every 10 deaths" and that the rate of occurrence of symptomatic ICH among t-PA treated patients was "compared with the rate of 8% estimated from pilot studies using similar doses and times of treatment" (p. 1584)[1]. The NINDS investigators reported a total of 22 symptomatic and 23 asymptomatic ICHs within 36 hours of treatment (p.1586)[1]. Of the symptomatic ICHs, 20 occurring among patients treated with t-PA and two among those receiving placebo ($p < 0.001$) whereas, of the 23 asymptomatic ICHs, 14 occurred in patients treated with t-PA and 9 in those receiving placebo ($p = 0.23$). These data are summarized in the following table.

ICH	rt-PA	Placebo	Total
Symptomatic	20	2	22
Asymptomatic	14	9	23
None	278	301	579
	312	312	624

Subsequent to the publication of the primary analyses, the NINDS investigators published a manuscript focused on ICH[11]. The stated purpose of that manuscript was the identification of "variables associated with intracerebral hemorrhage in patients with acute stroke who receive t-PA". In this manuscript they again report 22 symptomatic ICHs (20 from the t-PA treated group and two from the placebo group) but report only 21 asymptomatic ICHs (13 from the t-PA treated group and 8 from the placebo group) in contrast to the 23 reported in the first manuscript[1, 11]. One of these exclusions is explained as a post-surgical ICH and the other apparently occurred more than 36 hours after treatment. Hemorrhages occurring more than 36 hours after t-PA therapy (there were 5 symptomatic) were deemed unrelated to therapy.

The manuscript describes complex statistical analyses designed to identify patients at a high risk of experiencing an ICH. These analyses utilized both prerandomization (baseline) and time dependent data collected during the 36 hours subsequent to the initiation of therapy. The investigators started with 45 variables with a stated goal of identifying risk factors for ICH in four different scenarios.

1. Symptomatic ICH; t-PA treated only (n=312, ICH=20)
2. Symptomatic ICH; t-PA & placebo patients (n=624, ICH=22)
3. Symptomatic & asymptomatic ICH; t-PA treated only (n=312, ICH=33)
4. Symptomatic & asymptomatic ICH; t-PA & placebo (n=624, ICH=43)

The variables remaining in their "final model" for each of the above scenarios were:

Scenario 1
- Baseline NIHSS Score (categorized into 5 levels)
- Edema/mass effect on baseline CT (yes/no)
- No time dependent covariates

Scenario 2
- Baseline NIHSS Score (categorized into 5 levels)
- Edema/mass effect on baseline CT (yes/no)
- A treatment indicator variable (t-PA/Placebo)
- No interaction of treatment with the covariates

Scenario 3
- Baseline NIHSS Score (categorized into 5 levels)
- Edema/mass effect on baseline CT (yes/no)
- Time dependent covariates: external bleeding/oozing and pulse pressure

Scenario 4
- Baseline NIHSS Score (categorized into 5 levels)
- Edema/mass effect on baseline CT (yes/no)
- A treatment indicator variable (t-PA/Placebo)
- A treatment effect interaction with current smoking.
- No mention of time dependent covariates

It was stated in their Methods Section (p. 2111)[11] that these models would be used to define high-risk subgroups for the development of ICH within which t-PA treatment effect could be assessed. However, they are never mentioned in the results or discussion sections.

7.2 Review Committee Analyses: As specified in the committee charge, we conducted a "careful evaluation of both the risks and potential benefits" and followed the NINDS investigators' lead to see if baseline data can help define high-risk subgroups in which t-PA treatment might be contraindicated due to the level of elevated risk of ICH. Our analytical efforts involved three separate activities.

(1) Imputation of missing data,
(2) Assessment of the net effect of t-PA therapy in the face of the increased risk of ICH,
(3) An attempt to identify a group at high-risk for the development of ICH.

Among the 622 patients to whom we restricted our analyses (Section 4.1.3), we identified 22 who experienced a symptomatic ICH and 20 who experienced an asymptomatic ICH.

Consequently, all of our analyses and comments pertain to these 622 patients among whom at most 42 experienced an ICH.

7.2.1 Missing Values: We planned to use the same set of 45 variables that the NINDS investigators utilized to define groups of patients at high risk for ICH[11]. However, as described in Section 4.3.2, some variables had missing values which we imputed while others were not observed in enough patients to warrant their inclusion in the analysis. Further, we elected, as is described in Section 6, not to use admission or baseline blood pressure determinations in our analyses. Consequently, the analyses reported herein are restricted to 34 variables with observations on all of the 622 patients.

7.2.2 ICH Analyses: To state our questions precisely, to describe the data available to address the questions, and to put our analyses into perspective, we offer the following statements and observations. While these observations pertain specifically to the occurrence of Symptomatic Intracerebral Hemorrhages they also apply in essence to all ICHs, both symptomatic (SICH) and asymptomatic (ASICH).

7.2.2.1 ICH Risk Increases with t-PA: The chance of an ICH increases with the use of t-PA therapy.

a) 2 SICHs out of 312 placebo treated patients.
b) 20 SICHs out of 310 t-PA treated patients.

7.2.2.2 Favorable Outcome Chance Decreases with t-PA: The chance of a favorable outcome decreases in the presence of an SICH.

In the t-PA treated group, for the Barthel index (B), among those 290 patients not experiencing an SICH, 55% had a favorable outcome at 90 days. In contrast, among those 20 patients experiencing an SICH, only 10% had a favorable outcome. For the Rankin (R), Glasgow (G) and NIHSS (N) outcome measures, the corresponding percents are: 45% & 10%, 48% & 10%, and 36% & 15%, respectively.

In the placebo group, the favorability rates among those 310 patients not experiencing an SICH were 38%, 27%, 31% and 21% respectively for the B, R, G, & N outcome measures respectively. There were only 2 patients in the placebo group who experienced an SICH and they both had an unfavorable outcome. Thus, there are no data regarding the rate of a favorable outcome among ischemic stroke patients experiencing an SICH in the absence of t-PA therapy.

7.2.2.3 Favorable Outcome Chance Increases with t-PA and no ICH: The chance of a favorable outcome increases with t-PA therapy in patients without SICH.

In the comments above, we note that the percent of patients without an SICH who had a favorable outcome was higher in the t-PA treated group than in the placebo group. This can be summarized, for the 4 outcome measures, in the table.

t-PA	*Placebo*	*rate diff*	*rate ratio*	*odds ratio*
B	*55%*	*38%*	*17%*	*1.44*	*1.98*
R	*45%*	*27%*	*18%*	*1.69*	*2.25*
G	*48%*	*31%*	*17%*	*1.53*	*2.02*
N	*36%*	*21%*	*15%*	*1.75*	*2.18*

Thus, among those not experiencing an SICH, the favorable outcome rate is greater in the t-PA treated patients than in those on placebo by between 15 & 18 percentage points with the corresponding odds of a favorable outcome among t-PA treated patients essentially twice that among those on the placebo. Again, the lack of data on the favorable outcome of SICH patients in the placebo group presents an analytic problem.

The fundamental question that must be addressed is how to balance the evidence of the efficacy of t-PA therapy with the equally clear evidence that such therapy carries an associated increased risk of ICH, substantially decreasing the chances of a favorable outcome.

There are two components to this question. One pertains to the net effect of t-PA therapy while the other pertains to the issue of whether there are subgroups of patients who are particularly susceptible to ICH and, therefore, should not be treated with t-PA. While we intend to offer comments on both issues, we must express caution, as did the NINDS investigators[1, 9, 11], that the clinical trials whose data we are examining, were designed and powered to address the question of a net effect, not the question of an interaction or subgroup effect, and that we are performing exploratory subgroup analyses.

7.2.2.4 ICH Related Morbidity and Mortality: Among the 42 patients with either a symptomatic or an asymptomatic ICH, 10 were in the Placebo group and 32 in the t-PA group. Very few had a favorable outcome (7, 6, 6, 5 on the B, R, G, and N scales respectively with only one from among the 10 placebo treated patients). Fewer of the 22 SICH patients had a favorable outcome (2 each by B, R, and G and only 1 by N) with none of them from the 2 SICHs in the Placebo group. At 90 days, 22 (52%) of the 42 ICH patients were dead which included 16 (73%) of the 22 SICH patients

7.2.3 Net t-PA Effect: The following concerns the net t-PA effect, addressing the issue of whether a patient should be administered t-PA.

7.2.3.1 Favorable Outcome Chance Increases with t-PA Among All Patients: The chance of a favorable outcome **increases** with t-PA therapy even when those patients experiencing an SICH are included in the analysis.

*In the table below, the observed data **including all patients** randomized to the studies whether experiencing an SICH or not, are summarized in terms of the rates of favorable outcomes.*

	t-PA	Placebo	rate diff	rate ratio	odds ratio
B	*52%*	*38%*	*14%*	*1.37*	*1.78*
R	*43%*	*27%*	*16%*	*1.61*	*2.07*
G	*45%*	*31%*	*14%*	*1.46*	*1.85*
N	*34%*	*21%*	*13%*	*1.67*	*2.01*

These observed rates and comparisons, with the effects of t-PA, clearly diluted by the inclusion of the SICH patients, nearly all from the t-PA treatment group with their associated reduced chance for a favorable outcome, are still highly suggestive of a net positive effect associated with t-PA therapy.

7.2.3.2 Modeled Likelihood ORs Significantly > 1 Among All Patients: The chance of a favorable outcome increases with t-PA therapy, even when those patients experiencing an SICH are included in the analysis and formal models are created to adjust for stratification factors and other covariates associated with the chances of a favorable outcome.

The models developed are discussed in Section 5.5.1. They do not contain any interaction terms involving the t-PA indicator, as none was found to be significant. The adjusted OR estimates for Barthel, Rankin, Glasgow and NIHSS are, 2.19, 2.43, 2.13 and 2.19 respectively. The adjusted results, in terms of estimates of differences in favorable outcome rates, (Section 5.6.1) are 19.3%, 20.2%, 17.9% and 15.6% respectively, all somewhat larger than seen in the above table. The fundamental message in these last two analyses is that the net effect of t-PA therapy remains positive even though some patients are put at higher risk of an unfavorable result due to their increased risk of an SICH.

7.2.3.3 Conclusion Regarding Net Effect: In Section 5.6 of this report, the public health implication of these analyses is discussed. In summary, if 1000 acute ischemic stroke patients receive t-PA therapy according to the NINDS protocol, somewhere between 120 and 160 more of them will experience a favorable outcome at three months than if t-PA was not available. This even though about 65 of these 1000 patients would experience an SICH as a result of the t-PA with the resultant reduced chance of a favorable outcome.

7.2.4 Identification of Variables Predicting ICH: This analysis is based on the study of 622 patients, 310 randomized to t-PA and 312 to placebo. Of the 310, 20 experienced a symptomatic ICH within 36 hours of randomization (odds = 20/290 = 0.069) and an additional 12 were diagnosed as having an asymptomatic ICH within the same time period (odds of any ICH = 32/278 = 0.115). Among the 312 patients randomized to placebo, 2 experienced a symptomatic ICH and 8 an asymptomatic ICH with respective odds of 2/310 = 0.0065 and 10/302 = 0.033. Each of these placebo odds differs significantly from its counterpart in the t-PA group ($p = .0004$ & $p < .0001$ for all ICH & SICH respectively, Table 7.1).

To facilitate the remainder of this discussion we define some notation used in the tables. We used the same 4 "scenarios" defined by the NINDS investigators terms of types of ICH (symptomatic or all ICH) and treatment groups (t-PA treated or all patients)

I. Symptomatic ICH; t-PA treated only (n=310, SICH=20)

II. Symptomatic ICH; t-PA & placebo patients (n=622, SICH=22)

III. Symptomatic & asymptomatic ICH; t-PA treated only (n=310, ICH=32)

IV. Symptomatic & asymptomatic ICH; t-PA & placebo (n=622, ICH=42)

The odds of an ICH are, respectively: 0.069, 0.037, 0.115 and 0.072.

Within each scenario we used 34 prerandomization variables, including the treatment indicator variable, t-PA if appropriate, to examine the question of which of them, individually and collectively, might predict those at a higher risk for ICH. In Table 7.1, we have summarized the results of univariate logistic model analyses of the influence of each of the 34 variables within each of the four scenarios. The variable names are defined in Section 4.1.2. The bolded variable names indicate variables for which some imputation was necessary. The "DF" column indicates the number of degrees of freedom a variable requires in a model (i.e., one for continuous and dichotomous variables and greater than one for variables dividing the patients into more than two classes). The remaining columns are divided into four sets of two with the four sets corresponding to the four scenarios (indicated by Roman Numerals) and the 2

columns within each set providing the univariate p-values and odds ratio estimates (no ORs provided if DF>1). The variables are in ascending order according to their p-values in the scenario I analyses and those with p-values <0.20 are in bold. This facilitates observation that while there is a great deal of commonality among the scenarios as to which variables are "significant", there is also some diversity. Note that in scenarios II & IV the t-PA variable is included and its p-value indicates the significant difference between the t-PA treated patients and the placebo treated patients with regard to the risk of an ICH.

In Table 7.2**,** the next stage of the investigation, as carried out by the NINDS investigators, is summarized. In that stage they took all of the variables within each scenario whose univariate p-values were <0.20 and put them into a multivariate logistic model. The same structure is used so it can easily be seen which variables are not in the models. Note here that some of the 34 variables are literally constructs of others and they cannot be in a multivariate model together. In such situations where both were significant, we either made a considered judgment as to which variable to include or we used the variable selected by the NINDS investigators. We note that in these models there are frequently variables that have p-values >0.20 and they are no longer presented in boldface. This type of thing happens when variables are correlated.

In Table 7.3, the results of simple stepwise modeling processes for each scenario are summarized. As did the NINDS investigators, we required a variable to have a p-value < 0.20 to enter and remain in each model. Here there are columns labeled STP to indicate the order (step) in which the variables entered the models – an indication of "importance". There is much more diversity in these models although some variables appear nearly always, indicating some consistency, if not validity, in the process.

7.2.4.1 Methodological Issues Our primary analyses are based on BsNIHSS, the "continuous" version of the NIHSS variable, but we did investigate use of BsNIHSS(5), a partition of the NIHSS score into 5 categories used by the NIHSS investigators. Such categorization allows for non-linear relations, but our analyses did not indicate a sufficient degree of lack of linearity to warrant 5 categories. However, not surprisingly, BsNIHSS(5) did suggest that the major ICH risk was at the upper end of the NIHSS score. However, BsNIHSS was the most statistically significant and we based our analyses on it.

Glucose (GLU) presented a similar issue with a continuous version and a dichotomous version, dividing the glucose scale at 300 mg/dl, competing with each other. Because the continuous version was so statistically significant when the t-PA & placebo groups were combined we included it in further analyses.

There were two important and related concepts, edema and mass effect as assessed at the prerandomization CT scan. We identified three modeling options for these dichotomous variables: let each be a candidate for the model, with only one allowed in; combine them into a 2 degree of freedom variable; or create an "either/neither" indicator variable. The NINDS investigators used this latter option as did we, creating the "ED/ME" variable which equals 1 if a patient has either Edema or Mass Effect and 0 otherwise, which is highly significant in all our models.

In Table 7.4, the results of stepwise modeling with the constraint that certain variables must be in the models are summarized. For scenarios I, and III the variables constrained to be in the

models were AGE & BsNIHSS while in scenarios II & IV, t-PA was also constrained to be in the models. The variables constrained to be in the models regardless of their p-values are designated as having entered the models at step 0.

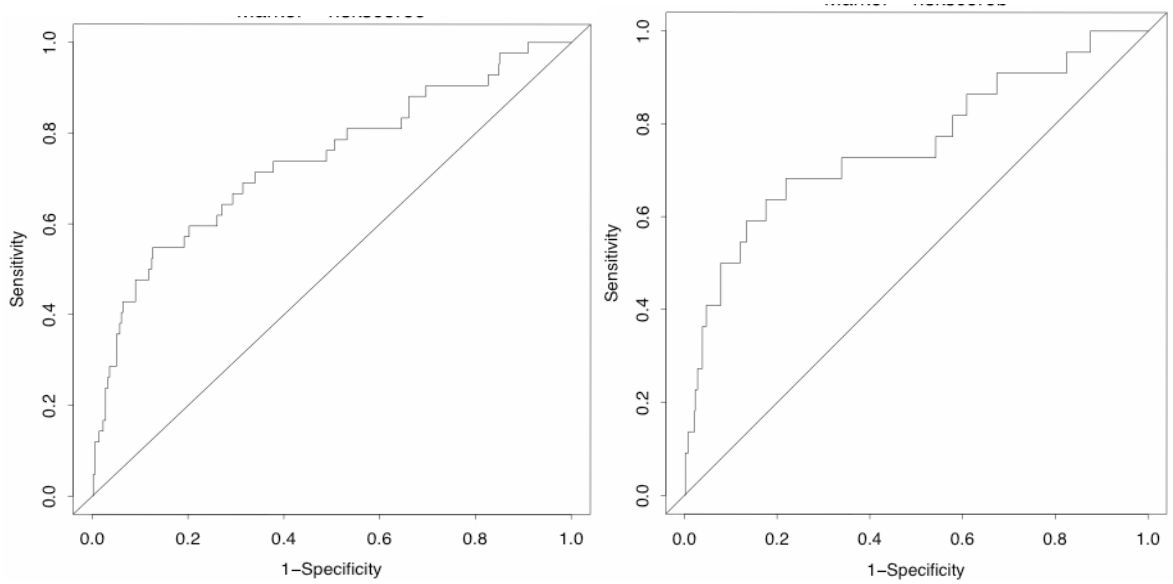
7.2.4.2 Results of ICH Risk Factor Identification: In all scenarios the variables that are important ($p< 0.20$)(in addition to t-PA) are associated with increased risk of ICH. In addition, in scenarios II & IV we can investigate whether any variables modify the t-PA effect in increasing the risk of ICH. To investigate this question we selected the two models in Tables 7.3 and 7.4 that were the most effective and tested whether any t-PA interactions with other variables were statistically significant. While the t-PA/current smoking (CSMK) interaction was "suggestive" ($p =.15$ in Scenario II and $p = .07$ in Scenario IV) in light of all the "data mining" taking place we elected not to consider it further.

In addition to checking for interactions, we reran the stepwise models for scenarios II & IV with the t-PA variable eliminated. The resulting models contained the same variables as when t-PA was available, indicating that the presence of the t-PA variable did not influence which other variables were associated with an increased risk for an ICH in this patient population.

Therefore, we created two risk scores (RS_{ICH} and RS_{SICH}) using in both the variables with $P < 0.1$ in the model developed for scenario IV, forcing AGE & BsNIHSS into the models. Each of these risk scores is based on a multivariate logistic/linear model using dependent variables AGE, BsNIHSS, ED/ME, GLU, CSMK and RACE to discriminate between patients with and patents without an Intracerebral Hemorrhage. For RS_{ICH} we discriminated between patients with any ICHs (either symptomatic or asymptomatic) and those with no ICH. For RS_{SICH} we discriminated between SICH patients and all other patients. Each patient then obtained a value for each of these risk scores based on the patient's values for the 6 variables and the estimated intercept and coefficients of the variables in the two separate models.

7.2.4.3 Risk Score Sensitivity and Specificity for any ICH: The Receiver Operating Characteristic (ROC) curves (sensitivity plotted against (1-specificity)) associated with each risk score are illustrated below. The curves show that neither RS is very effective in predicting ICH and that they are almost equally effective. Indeed, one indicator of the value of a risk score is the area under the ROC curve. These two ROC curves have almost identical areas (.74 for RS_{ICH} and .75 for RS_{SICH}).

The ROC Curve on the left corresponds to RS_{ICH} and the one on the right to $RS_{SICH.}$



When each risk score is inserted in the separate logistic models for the four outcome measures (Barthel, Rankin, Glasgow and NIHSS) and in the Global model predicting a favorable outcome (see Sections 5.4.3 and 5.3.4), it brings some new information into some of the models as is illustrated by the p-values in the table below.

	RS_{ICH}	RS_{SICH}
Barthel	0.02	0.02
Rankin	0.005	0.003
Glasgow	0.14	0.13
NIHSS	0.41	0.53
Global	0.04	0.05

Thus, in all but the Glasgow and NIHSS models the risk scores each bring some new information into the models. Considering that the risk scores are constructed using Age and BsNIHSS, both of which are included in the adjusting covariates, it might have been expected that they would add nothing to these models.

However, that is not the primary question at hand. What is really critical is whether the risk scores help in the identification of a subset of patients who, because of their risk for an intracerebral hemorrhage, might be at especially high risk of an unfavorable outcome should they be exposed to t-PA. Thus, in models with the adjusting covariates and an indicator for t-PA treatment (see Section 5.5) we also insert each risk score and the associated t-PA by Risk Score interaction term. The results of these interaction tests are summarized in terms of p-values, in the table below.

	RS_{ICH}	RS_{SICH}
Barthel	0.35	0.40
Rankin	0.32	0.25
Glasgow	0.22	0.27

NIHSS	0.77	0.64
Global	0.73	0.81

In each of these models, for each of the risk scores, the interaction between the t-PA group indicator and risk score is not statistically significant. Thus, while the concept of a risk score based on a careful statistical analysis comparing those with and without an intracerebral hemorrhage is appealing, this formal process led to risk scores which were not particularly sensitive of specific and did not identify a group of patients who would be placed at special risk if treated with t-PA.

7.2.4.4 A Simplified Risk Function: To simplify real-time implementation of the RS_{ICH} approach, we dichotomized the 4 most important variables used in computing the RS_{ICH} (AGE, NIHSS, ED/ME, & GLU) as indicated and subdivided the 622 patients into 3 categories according to those factors

- Age >70 years
- Glucose >300 mg/dl
- Baseline NIHSS >20
- Edema and/or Mass Effect on the CT scan

As seen in the table below, the risk of symptomatic ICH and of any ICH increases noticeably with the number of risk factors ($p < .0001$ in both instances). Clearly, this grouping based on these four factors does predict the occurrence of ICH.

Table: Simplified ICH Risk Function

No. of risk factors	No. of patients (%)	% with ICH		
		Symptomatic	Asymptomatic	Total
None	238 (38%)	1.3	2.1	3.4
1	278 (45%)	2.9	2.9	5.8
≥2	106 (17%)	10.4	6.6	17.0

In Table 7.5 we summarize a basic analysis of the observed data yielding rates of favorable outcomes by all four outcome measures, within each of the three groups among all patients. The bottom part of the table, pertaining to the 106 patients with one or more of the ICH risk factors, provides some interesting results. Of the 106 patients, the overall percents of favorable outcome for the B, R, G, & N outcome measures respectively, were 15%, 8%, 12% & 8.5%, much less than reported for the study overall. Most importantly, the rates of a favorable outcome for the placebo treated patients are slightly though not significantly larger than for the t-PA treated patients for three of the 4 outcome measures. When the three groups are compared in models containing only the stratification factors of Center, Part and OTT we found them to have significantly different odds of a favorable outcome for each of the 4 outcome measures and for the Global analysis ($p <0.0001$ for all models, models not shown). In the same context when we searched for t-PA by ICH group interaction none were significant ($p = 0,21$, $p = 0.16$, $p = 0.09$, $p = 0.15$, and $p = 0.41$ for the Barthel, Rankin, Glasgow, NIHSS and Global models respectively).

In Table 7.6, we summarize the results of inserting indicator variables separating these three groups into the individual outcome models and the global model with all of the adjustment

factors (Section 5.5) and the treatment indicator t-PA included. In such models, we find no evidence that the three groups have different rates of a favorable outcome because the variables BsNIHSS and AGE, which are key to forming the groups, are among the adjusting variables. Furthermore, and most importantly from the net effect standpoint, we find no evidence of a significant interaction between t-PA and these groups in any of the models ($p = 0.57, 0.28, 0.24, 0.18$ and 0.41 for the Barthel, Rankin, Glasgow, NIHSS and Global models respectively). Recall all of the caveats regarding the detection of significant subgroup effects.

7.3 Summary and Conclusions:

In the NINDS trial, the overall risk of symptomatic ICH was 6.5% in t-PA treated patients vs. 0.6% in patients receiving placebo. When a symptomatic ICH occurred after treatment with t-PA, there were significant clinical consequences. Only a small minority had a favorable outcome (e.g., for the Barthel index, the favorable outcome in patients with symptomatic ICH was 10% vs. 55% in patients without ICH) and the three month mortality rate was very high (75%).

A number of putative risk factors for ICH were identified, with many of them being interrelated. Our exploratory analysis found four risk factors, age >70 years, baseline NIHSS >20 points, plasma/serum glucose >300 mg/L and edema and/or mass effect on the initial CT scan, that were associated with both an increased risk of having an SICH and a lower likelihood of having a favorable outcome. For patients with either no risk factors or only one risk factor, the likelihood of having a favorable outcome favored the t-PA treatment group, while for the group at highest risk (> 1 risk factor), there was essentially no difference between the t-PA and placebo groups with regards to the likelihood of having a favorable outcome. However, the analysis also found that the adjusted t-PA to placebo odds ratios for favorable outcome in the three subgroups with different numbers of risk factors were not significantly different, and were consistently in favor of the t-PA treatment group.

We conclude that there was no statistically significant evidence of the existence of any subgroup of acute ischemic stroke patients in whom the risk, and consequences, of having a symptomatic ICH clearly outweighed the beneficial effects of t-PA. However, it is important to keep in mind that because of the study design and the small number of patients who had an SICH, this trial was not powered to identify risk factors related to having either an SICH or a decreased likelihood of a favorable outcome. Risk factors for ICH acute ischemic stroke patients treated with t-PA should be evaluated in future studies that are designed, and powered, to evaluate this question.

How the findings of this exploratory analysis are used in the management of the individual patient with acute ischemic stroke, balancing risks and benefits based on very limited scientific information, is for the patient and the attending physician to decide.

8. SPECIAL TOPICS

8.1 Age, Baseline Stroke Severity, and Baseline Stroke Severity Imbalance

8.1.1 Introduction: It is well documented that stroke severity at onset and age are major predictors of favorable outcome. For each of the four outcome measures, and for the global statistic (Section 5.4.3), baseline NIHSS (BsNIHSS) and Age were the two most significant indicators of outcome following stroke among all available covariates in this analysis. It was further demonstrated that they interacted with each other in a synergistic way so that the joint effect of increasing age and increasing stroke severity was greater than the "sum" of their individual effects. This is not an unexpected result since at an advanced age even a modest increase in stroke severity can have significant clinical consequences. In the development of the covariate model we used BsNIHSS and Age as continuous variables and included in all models the product of BsNIHSS and Age to account for the interaction (Section 5.4.3). The NINDS investigators also initially analyzed the baseline NIHSS score as a continuous variable to adjust their analyses[1]. In that format the two treatment groups are in balance, with the NINDS investigators reporting nearly equal median values as being not significantly different by a rank sum test ($p = 0.10$). We corroborated that result (Table 4.2), and also acknowledged and discussed an imbalance noted later by the NINDS investigators (4). This imbalance became obvious when patients were grouped into five classes (approximately quintiles) according to baseline NIHSS (Q_1: 0-5, Q_2: 6-10, Q_3: 11-15, Q_4: 16-20, Q_5: >20). As seen in Table 4.2, this categorical distribution of BsNIHSS differed quite significantly ($p = 0.005$) in the t-PA and Placebo groups. To facilitate discussion, we refer to the categorical variable as BsNIHSS(5) to distinguish it from BsNIHSS. Age demonstrated a smaller imbalance ($p = 0.02$) as is seen in Table 4.2, in the opposite direction as there were more younger patients randomized to the Placebo arm of the trial than to the t-PA arm.

The primary goal of this section is to investigate the impact of the BsNIHSS imbalance since it seems the most likely factor to have impacted results and it has received widespread attention as potentially invalidating the study results. However, because of the high synergy between Age and BsNIHSS, any analysis of one must involve and impact the other, seriously complicating this process.

8.1.2 Baseline NIHSS Imbalance: The table illustrating this imbalance, shown below, demonstrates that, in the first quintile (NIHSS 0-5), 72% of the 58 patients were randomized to t-PA therapy. This was an unexpected observation since, in a randomized trial, it would be expected that within each quintile, there would be approximately equal numbers of patients randomized into each treatment group. The imbalance in the first quintile is countered in the second and fifth quintiles where the corresponding percents are 45%. The third and fourth quintiles are balanced.

ALL PATIENTS*

Treatment Group	Baseline NIHSS Quintiles					TOTAL
	0 - 5	6 - 10	11 - 15	16 - 20	> 20	
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

This imbalance led critics of the NINDS study to suggest that it could have affected the overall study results. In this section we pursue this matter. As is illustrated in the two tables immediately following, the majority of the imbalance occurred among patients randomized in the stratum defined by the time from onset to treatment (OTT) being between 91 & 180 minutes. We could not establish that this fact contributed in any substantial way to our analysis and elected to proceed with our description of what is a rather complex analysis with no further reference to the relationship of the BsNIHSS(5) categorical variable to the OTT variable.

OTT ≤ 90 PATIENTS*

Treatment Group	Baseline NIHSS Quintiles					TOTAL
	0 - 5	6 - 10	11 - 15	16 - 20	> 20	
Placebo	9 (41%)	37 (55%)	31 (44%)	37 (48%)	31 (47%)	145 (48.0%)
t-PA	13 (59%)	30 (45%)	39 (56%)	40 (52%)	35 (53%)	157 (52.0%)
Total	22	67	70	77	66	302

* p-value for test for imbalance = 0.001

OTT > 90 PATIENTS*

Treatment Group	Baseline NIHSS Quintiles					TOTAL
	0 - 5	6 - 10	11 - 15	16 - 20	> 20	
Placebo	7 (19%)	46 (55%)	35 (57%)	33 (50%)	46 (62%)	167 (52.2%)
t-PA	29 (81%)	37 (45%)	26 (43%)	33 (50%)	28 (38%)	153 (47.8%)
Total	36	83	61	66	74	320

* p-value for test for imbalance = 0.7

8.1.3 Outcomes in NIHSS Quintiles: In Table 8.1.1, we illustrate for each of the 4 outcome variables (B:Barthel, R:Rankin, G:Glasgow and N:NIHSS) the numbers of unfavorable (UF) and favorable (F) responses within each of the five classes (hereinafter referred to as quintiles) as defined by BsNIHSS(5). Also, in this table we present three measures of comparison of t-PA to Placebo within each quintile for each outcome scale. These three measures are:

D% = difference in % favorable outcome (t-PA minus Placebo)
RR = ratio of these favorable outcome percents
OR = ratio (t-PA/Placebo) of the odds of a favorable outcome.

For each outcome variable, in the 1st quintile (Q_1), all patients, whether placebo or t-PA treated, had an excellent chance for a favorable outcome. It is also interesting to note that, for three of the four outcome variables, the placebo group does modestly better than the t-PA group in Q_1, although this is not statistically significant. For each outcome variable, the proportion of favorable outcomes for both treatment groups decreases with increasing NIHSS category. However, for categories Q_2 through Q_5, all indicators of treatment effectiveness favor t-PA therapy for all the outcome variables. At the upper end of the NIHSS score, indicating more severe strokes, the likelihood of a favorable outcome is quite poor and the absolute difference in favorable outcome (D%) is much smaller in Q_5 (3.3-5.6% for the four outcomes) than in Q_2 - Q_4. However, even in Q5, the RR and OR indicators of treatment effect show values in favor of t-PA not much different than those seen in Q_2, Q_3, and Q_4.

8.1.4 Outcomes in Age Quintiles: As we have noted, age was the second most significant variable related to favorable outcomes (Tables 5.10 through 5.13). In Table 8.1.2, which is similar to Table 8.1.1 but applies to Age Quintiles we saw that effect. Overall the favorable outcome percentages (regardless of treatment) decrease with increasing age although in the age decade 65 – 74, which we divided into two groups because of the number of patients in that age decade, we saw no evidence of a decrease. However, patients in the final group (age 75+) clearly have the smallest chance of a favorable outcome.

From the standpoint of treatment effect, the differences in the favorable outcome percentages are positive, in favor of t-PA, for all outcome measures for all quintiles. The odds ratios show a

decreasing trend with increasing age group for some outcome variables. However, in the light of the major interaction of Age with BsNIHSS we did not investigate this further.

8.1.5 Age by Baseline NIHSS Interaction: We stated in the introduction to this section that the relationship between Age and NIHSS was synergistic. From the models shown in Tables 5.17 through 5.21 we see that the coefficient of the interaction term is always negative. In a setting such as this where the influence of the two variables is also negative, this means that each gains more influence as the other increases in value. For example, consider the Barthel covariate model (Table 5.17). The estimated coefficients suggest that, for a patient aged 50, a one unit increase on the NIH Stroke Scale results in a 12% decrease in the odds of a favorable outcome while, for a patient age 80 a one unit increase in the NIH Severity Scale results in a 22% decrease in the odds of a favorable outcome.

8.1.6 Model-Based Assessment of Baseline NIHSS Imbalance: We will now review formal attempts to evaluate the role of BsNIHSS(5) on the assessment of treatment effect. In Section 5.4 we describe our process of identifying the variables significantly related to the occurrence of a favorable outcome that were included as adjusting covariates in the treatment effect models. The continuous version, BsNIHSS, with a single degree of freedom in the models, was always highly significant ($p<0.0001$) as a predictor of favorable outcome and was more significant than BsNIHSS(5), with four degrees of freedom in the models, for all but one of the analyses. For this reason, and to avoid using a grouping that was identified from data exploration, we used the continuous BsNIHSS, in our principal analyses. Thus, all analyses other than those to be discussed herein are based on the use of BsNIHSS, a single degree-of-freedom variable, in our logistic and global models to adjust for the impact of baseline stroke severity, as measured by NIHSS, on 90 day outcome.

8.1.6.1 Baseline NIHSS Analysis: In what follows we first show that the choice of which of these versions of NIHSS to use is irrelevant. The choice affects the estimated coefficients of other variables in the model by only a small amount and has very little effect on the estimated t-PA vs. Placebo favorable outcome odds ratios.

For each outcome variable, Table 8.1.3 shows parameter estimates with standard errors and/or p-values for four different models. For each outcome variable, the first two models include BsNIHSS and differ only as to inclusion of the t-PA indicator variable. The second two models include BsNIHSS(5) with four degrees of freedom. For each outcome measure, the t-PA vs. Placebo odds ratios estimated by models with the different versions of NIHSS are – for all practical purposes – identical. However, of greater importance is the fact that the coefficients or p-values of all of the other covariates are essentially the same whether BsNIHSS or BsNIHSS(5) is used.

In Section 5.5.2.1 we report that t-PA did not interact with the three degrees-of-freedom variable associated with the variables BsNIHSS, Age, and Age*BsNIHSS for any of the outcome variables individually or in the Global analysis. Because of the strong interaction between Age and BsNIHSS, it is necessary to treat these three variables collectively as a triumvirate. The presence of an interaction of t-PA with the BsNIHSS, Age, Age*BsNIHSS triumvirate, would mean that patients in some group(s), defined by the combination of stroke severity as measured by BsNIHSS and Age, responded differently to t-PA treatment than patients in other groups. That neither we, nor the NINDS investigators, found statistically

significant evidence of an interaction, does not imply its absence. Allowing for this caveat, the absence of a statistically significant interaction indicated that there was no evidence of a differential t-PA treatment effect related to baseline stroke severity. This finding indicates that the baseline stroke severity imbalance did not affect the study outcome.

8.1.6.2 BsNIHSS Quintile Specific Odds Ratios: The quintile-specific OR estimates for each of the outcome measures are documented in Table 8.1.1. The unadjusted global OR estimates for the five quintiles, Q_1 through Q_5, were: 0.9, 2.4, 1.9, 1.6, and 1.7 respectively. For each outcome measure, the ORs favor t-PA except in Q_1 where the values are all close to 1. Even in Q_5, the OR is in favor of t-PA therapy. Tests of the hypotheses that the odds ratios are equal across the quintiles, adjusting for the stratification factors, were not statistically significant. However, when adjusting for all covariates, these tests are complicated by the presence of a highly significant interaction between age and baseline NIHSS. In this context, such tests involve a complex interaction with 9 degrees of freedom. The table below documents the results of the chi-square tests for models including both the four and nine degrees of freedom tests. These analyses demonstrate that for each of the four outcome measures and the global analysis there was insufficient evidence to declare a difference in treatment effects (ORs) across the five quintiles.

Treatment Group	Test for Equal ORs			
	Adjusted for stratification factors*		Adjusted for all covariates#	
	Chi-square (4 DF)	p-value	Chi-square (9 DF)	p-value
Barthel index	4.27	0.37	5.41	0.80
Modified Rankin scale	2.69	0.61	5.54	0.78
Glasgow outcome scale	2.89	0.58	6.09	0.73
NIHSS	0.66	0.96	2.96	0.97
Global analysis	2.30	0.68	3.65	0.93

* Stratification factors: study part, center, OTT
# All covariates: stratification factors + history of diabetes, preexisting disability, age, baseline NIHSS, and age*baseline NIHSS

After a detailed examination of all of these models the two most important messages are; (i) with the exception of Q_1, the t-PA to placebo odds ratio estimates are uniformly greater than 1, indicating a superiority of t-PA over placebo in patients with a baseline NIHSS score of > 5, and (ii) If we focus on the age of 70 – essentially the median age of the study group – the t-PA to placebo OR estimates from the model containing the interactions are not much different from the estimates from the no interaction models. Thus, we conclude that there is no evidence that the baseline stroke severity grouping defined by BsNIHSS(5) has identified a group of patients who respond differently to t-PA therapy than the study cohort in general. All earlier caveats about the proper interpretation of non-significant tests of no interactions continue to apply.

8.1.7 An Alternative Variable: The analyses described above are complicated by the need to include among the "adjusting" variables term(s) defining an interaction between Age and some version of baseline NIHSS. In Section 5.4, it was noted that in some models, the inclusion of the interaction term literally made the terms corresponding to Age and BsNIHSS appear insignificant. In one final effort to examine this complex question of the impact of the baseline NIHSS imbalance within models that of necessity include this interaction among the adjusting variables, we defined an alternative variable as the simple product of age times BsNIHSS and classified the patients into quintiles on the basis of that variable. Table 8.1.4 illustrates, for each outcome measure, the rates of favorable outcomes for Placebo and t-PA treated patients within each of these quintiles.

The role of this variable defined by multiplying age by BsNIHSS is not easy to understand and a few examples may help. A 75 year old patient with an NIHSS of 4 would have a value of 300 for this Age*BsNIHSS product, placing him/her in the middle of the first (lowest) quintile. Similarly, a 65-year-old patient with an NIHSS of 10 would have a score of 650, placing him/her in the middle of the 2nd quintile. A 75 year old whose NIHSS is 12 or 13 would be in the 3rd quintile and one with an NIHSS of 17 would be in the 4th quintile. The 5th quintile will contain mostly very elderly patients with a very high NIHSS (e.g. an 85 year old with an NIHSS of 24).

The table illustrates that the odds of a favorable outcome decrease dramatically as we look from the 1st to the 5th of these quintiles. For example, for the Barthel index, the odds of a favorable outcome for patients whose combination of age and NIHSS at baseline place them in the first of these quintiles is 4.4, indicating that such a patient has a very good chance of a favorable outcome. While the odds are not as high for the other scales, the estimates of odds are all greater than one, indicating that by any of the 4 scales, patients in the first quintile are more likely to have a favorable outcome than not. In contrast, in the fifth quintile the prospects are grim with the odds of a favorable outcome ranging from 0.14 for the Barthel scale to 0.02 for the NIHS scale.

The above assessment of the likelihood of a favorable outcome notwithstanding, the comparison of t-PA therapy to Placebo is remarkably constant over the 5 quintiles. We will briefly discuss the consequences of using this classification of patients into the 5 groups according to the product of age by NIHSS in our formal statistical models as an alternative to the more complicated use of Age, some version of BsNIHSS and the interaction between the two as adjusting variables in the models.

The sequence of Tables 8.1.5, 8.1.6, 8.1.7, and 8.1.8, one for each outcome measure, displays the results of this analysis. For each outcome measure, the table actually consists of summaries of 7 different models. The first 4 of these models are the same as seen in earlier table in this section and involve the standard models first with BsNIHSS and then with BsNIHSS(5) in the models. The last three models use the 4 degrees-of-freedom variable necessary to account for the 5 quintiles of the age*NIHSS product variable. The final model is testing the "no-interaction" hypothesis regarding the interaction between t-PA and the 5 quintiles. As we had observed in Table 8.1.4, there is no evidence of an interaction. Indeed, the 4 p-values for the no-interaction hypotheses are: 0.99, 0.98, 0.94, and 0.91, for Barthel, Rankin, Glasgow and NIHSS, respectively. The other point to make is that the use of this variable has had minimal effect on the other coefficients in the model or on the t-PA to Placebo Odds ratio estimate.

8.1.8 Influence of the Age by Baseline NIHSS Interaction on the t-PA Treatment Effect: As discussed previously (Section 5.4.1), our analyses determined that both stroke severity and age were negatively and significantly related to a favorable outcome and that these two variables interacted in a highly significant way in predicting a favorable outcome, with the combination of advanced age and a severe stroke reducing the chances of a favorable outcome to an extremely low level. While many who care for acute stroke patients recognized this, it is a phenomenon that, heretofore, does not seem to have been quantified. Consequently, all models aimed at comparisons of the t-PA and control groups must include variables for age, NIHSS and their interaction to adjust for these effects.

The fundamental concern in this discussion is whether the data in the t-PA trials provided any evidence that this relationship between age, stroke severity, and the chances of a favorable outcome, had an impact on the effect of t-PA. The problem of estimating the interaction between t-PA and stroke severity at randomization (as estimated by the baseline NIHSS), within models in which NIHSS and age interact, is complex, requiring multiple degrees-of-freedom. We summarized three methods of analysis and the tests for interaction were not statistically significant in any of them.

However, as summarized by Brookes et al. in a recent publication[28], interaction tests are typically very underpowered in studies where the sample size was determined on the basis of main effect tests. To quantify the power of interaction tests and sub-group analyses in randomized trials these authors performed simulations. They reported that a clinical trial with 80% power to detect a specified main effect would have only a 29% chance (power) of detecting an interaction of the same magnitude as the main effect. They also reported that the overall sample size would have to be quadrupled to increase the power of the interaction test in the above situation to 80% and that, if the desired detectable interaction was only 20% of the main effect, the sample size would have to be increased "dramatically."

In our study, the less complex, one degree-of-freedom, tests of no interaction have poor power (Section 5.5.2.2). Therefore, we concluded that the non-significant results of the complex tests were likely due to low power and included warnings in our conclusions that these results should not be taken as evidence of lack of an interaction. However, because of the importance of the baseline imbalance issue, we undertook further analyses to determine if more specific information regarding these interactions could be obtained. The results of this further examination in the three methods of analysis are summarized below.

1. In order to focus directly on the simultaneous impact of age and severity on outcome and t-PA effect, we created a new regressor by multiplying age by baseline NIHSS value (Section 8.1.7). We then subdivided the patients into quintiles of this predictor. As expected, those in the lowest quintile (relatively younger with less severe stroke) fared much better than those in the highest quintile (relatively older with more severe stroke) regardless of their randomization group. However, when we formally tested whether the effect of t-PA was the same for all five quintiles, the corresponding p-value, based on a 4-degree-of-freedom chi-square of 0.37, was $p = 0.99$ (within the global model while adjusting for all the covariates other than age and NIHSS which were included in this artificial variable). Thus, we conclude that there is no evidence of a difference in the t-PA to placebo comparison over these quintiles. The fact that the chi-square value is so small indicates that the estimate of any difference in effect based on an analysis of these data is so small as to be clinically irrelevant. The analyses for the 4

individual outcome measures (Barthel, Rankin, Glasgow, NIHSS) were just as dramatically null.

Barthel: Chi-square = 0.3415, p = .99
Rankin: Chi square = 0.4499, p = .98
Glasgow: Chi square = 0.7663, p = .94
NIHSS: Chi square = 1.0022, p = .91

2. With age and baseline NIHSS each being treated as continuous variables (Section 5.5.2.1), the presence of the highly significant interaction between them is manifested by the inclusion of a variable identical to the one discussed above (i.e., the product of age and NIHSS). In such models, the hypothesis of no interaction between NIHSS and t-PA requires the addition of three new terms in the model; hence the hypothesis is tested based on a chi-square value with 3 degrees-of-freedom. Again referring only to the global model, the chi-square value (df = 3) was 2.72, corresponding to p = 0.44. Here, again, the lack of statistical significance is not nearly as important as the fact that, among the 3 degrees of freedom there is no evidence of a meaningful indication of effect, since even a one degree-of-freedom test requires a chi-square of 3.84 or higher to provide such evidence. The analyses of the 4 individual outcome measures led to similar conclusions.

Barthel: Chi-square = 3.5645, p = .31
Rankin: Chi square = 5.4728, p = .14
Glasgow: Chi square = 4.3866, p = .22
NIHSS: Chi square = 2.8117, p = .42

3. Because tables had been published with a subdivision of the baseline NIHSS scores into quintiles, which demonstrated an imbalance of assignment of patients to t-PA and placebo treatments, we also performed analyses with patients in these groups (Section 8.1.6.2). Such analyses require that the models contain 1 degree-of-freedom for age, 4 degrees-of-freedom to identify the 5 groups, and 4 degrees-of-freedom to describe the interaction of age and NIHSS. Thus, testing the interaction of t-PA with NIHSS (and, of necessity, age) requires that 9 degrees-of-freedom be added to the models. For the global analysis of this no interaction hypothesis, the chi-square value (df = 9) was 3.65, with an associated p-value of .93. Again, we conclude that these analyses provide no evidence of an interaction between t-PA and stroke severity. The analyses of the 4 individual outcome measures led to similar conclusions.

Barthel: Chi-square = 5.4077, p = .80
Rankin: Chi square = 5.5400, p = .78
Glasgow: Chi square = 6.0910, p = .73
NIHSS: Chi square = 2.9600, p = .97

Finally, in a further assessment of the impact of stroke severity on the t-PA effect for fixed ages, we used an argument described mathematically in the appendix below to estimate the t-PA effect associated with a 5 unit increase in NIHSS for individuals at 60, 70, and 80 years of age. The results of these investigations are shown in the table below. For each of these ages, as NIHSS increases by 5 units, the odds ratio estimates (in favor of t-PA) increases. In the Global analysis, for age = 60 years, the increase is 12% (95% CI: -15% to 49%); for age = 70 years, the increase is 32% (95%CI: -6% to 86%); for age = 80 years, the increase is 56% (95%

CI: -6% to 156%). These estimates, based on the variance estimates, suggest that a 5-unit higher level of stroke severity is associated with an increase in t-PA benefit with the magnitude of the increase rising with age. While none of these percent increases was significant at the 5% level (very generous in the context of the many tests being performed), some were close.

	AGE	theta	Std. Err of theta	ORR = OR(N+5) / OR(N)	95 % CI of ORR lower	upper	Minimally Detectable ORR
Global							
	60	0.114	0.1439	1.121	0.85	1.49	1.63
	70	0.279	0.1749	1.322	0.94	1.86	1.82
	80	0.444	0.2529	1.559	0.95	2.56	2.37
Barthel							
	60	0.004	0.1629	1.004	0.73	1.38	1.75
	70	0.242	0.1727	1.274	0.91	1.79	1.80
	80	0.480	0.2655	1.616	0.96	2.72	2.48
Rankin							
	60	0.205	0.2251	1.227	0.79	1.91	2.16
	70	0.277	0.2641	1.319	0.79	2.21	2.47
	80	0.350	0.3661	1.418	0.69	2.91	3.49
Glasgow							
	60	0.052	0.1675	1.053	0.76	1.46	1.77
	70	0.178	0.1949	1.195	0.82	1.75	1.95
	80	0.304	0.2977	1.355	0.76	2.43	2.77
NIHSS							
	60	0.216	0.1819	1.241	0.87	1.77	1.86
	70	0.394	0.2354	1.482	0.93	2.35	2.24
	80	0.571	0.3525	1.770	0.89	3.53	3.34

Using a more rigorous, two-sided 0.01 level of significance, we estimate that the minimally detectable (80% power) changes in the probability of a favorable outcome would be 63%, 82% and 137% respectively. These suggest that if a 5-unit increase in stroke severity did produce increases in t-PA effect of the magnitude indicated the analyses would have had an 80% chance of being significant. But, there was no statistically significant evidence of an interaction despite this unexpectedly robust power.

In summary, this study was not powered to detect subgroup interaction differences in the t-PA treatment effect. Nonetheless, our analyses provide no evidence that the effect of t-PA is clinically different for acute stroke patients with different levels of stroke severity. A post hoc power analysis allows us to conclude that there was no clinically important interaction between

baseline NIHSS and t-PA. Therefore, we conclude the baseline imbalance in NIHSS played a very minor role in the estimated benefit of t-PA.

Appendix

The equations below define some age specific assessments of the impact of increases in severity as measured by baseline NIHSS. In this assessment we treat NIHSS and age as continuous variables and ignore all covariates not involved in this assessment (of course they are not ignored in the actual analysis). Define an indicator variable (t-PA = 1 for those on t-PA and = 0 for those on Placebo) and specify an interaction model as follows:

$$\text{Log(OR)} = \beta_0 + \beta_1 \text{ Age} + \beta_2 \text{ NIHSS} + \beta_3 \text{ Age*NIHSS} + \beta_4 \text{ t-PA}$$

$$+ [\beta_5 \text{ Age} + \beta_6 \text{ NIHSS} + \beta_7 \text{ Age*NIHSS}]\text{*t-PA}.$$

The odds ratio (OR) is the ratio of the odds of a favorable outcome for those treated with t-PA to the odds of a favorable outcome for those treated with the placebo.

The last three terms in this model describe the interaction of t-PA with stroke severity because of the complex relationship of age and stroke severity with the likelihood of a favorable outcome. For a fixed age (A), and a fixed NIHSS (N), the log of the OR comparing t-PA to placebo is:

$$\text{Log(OR given A \& N)} = \beta_4 + \beta_5 A + \beta_6 N + \beta_7 A*N.$$

The no interaction hypothesis is that: $\beta_5 = \beta_6 = \beta_7 = 0$. If true, the t-PA to Placebo OR is $\exp(\beta_4)$. An interaction exists if any of these three betas are non-zero and for the non-null model we estimate their values from our data. Keeping age fixed at A and changing NIHSS from N to N + Δ,

$$\text{Log(OR given A \& N + } \Delta) = \beta_4 + \beta_5 A + \beta_6 (N + \Delta) + \beta_7 A*(N + \Delta)$$

If these two equations give the same answers for age A, the t-PA to Placebo ORs are the same no matter what the NIHSS level is. The difference between these two equations, which we arbitrarily call Θ, is the Log of the ratio of OR at severity level N + Δ, call it OR(N + Δ), to the log of the OR at severity level N, call it OR(N). Define ORR(Δ) = OR(N + Δ)/OR(N), then:

$$\Theta = \beta_6 \Delta + \beta_7 A*\Delta = \Delta\{\beta_6 + \beta_7 A\} = \text{Log(OR(N + } \Delta)) - \text{Log(OR(N))}$$

$$= \text{Log(ORR(} \Delta)).$$

If we test and reject the null hypothesis that the expected value of Θ is zero, we would have established that, at least for A-year olds, there is an interaction between t-PA and NIHSS.

From the output of our models, we obtain the estimates of β_6 and β_7 as well as estimates of their variances and the covariance between them. This allows us to estimate the variance and hence, the standard error of the corresponding estimate of Θ. Then we obtain confidence intervals for Θ, and, using these empirical variance estimates, estimate the minimally

detectable value of Θ and hence, exp(Θ), the minimally detectable ratio of odds ratios, ORR(Δ), corresponding to a difference of Δ on the NIHSS scale for a fixed value, A, on the age scale.

Treating Θ as its own estimate, $\text{Var}(\Theta) = \Delta^2\{\text{Var}(\beta_6) + 2A\text{Cov}(\beta_6, \beta_7) + A^2\text{Var}(\beta_7)\}$

In the associated table we have summarized what we find in this situation for three choices of A (60, 70, & 80) and $\Delta = 5$.

8.1.9 Summary and Conclusions: 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 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.

8.2 Onset to Treatment Time

As detailed in Section 4.2, the NINDS study was stratified on onset to treatment time (OTT) with plans for an equal number of patients to be randomized with an OTT < 91 minutes and an OTT >90 minutes but not greater than 180 minutes. For the sake of the following discussion we will refer to these two strata as the first (#1) and second (#2) respectively.

8.2.1 Restricted Randomization: Following the onset of symptoms there were variable delays prior to patient arrivals in the emergency departments. Subsequently, further delay resulted before a patient could be consented, randomized and treated due to the requirements of the study protocol, including the performance of a CT scan to determine patient eligibility. Consequently, therapy was initiated on few patients in less than an hour after symptom onset. As a result, the NINDS investigators found it much easier to enter patients into the second stratum than the first. In order to assure satisfaction of the treatment protocol that equal numbers of patients be randomized within the two strata at each center, it was necessary for a restriction to be placed on the entry of patients into the study within stratum #2. Specifically, each center was instructed that whenever the number of patients in stratum #2 exceeded the number in stratum #1 by three (3), they could not randomize a patient into stratum #2. This design modification worked quite well with only a minor imbalance; 302 patients were randomized into stratum #1 and 320 into stratum #2. Of course, this quota rule resulted in 267 otherwise eligible, patients not being entered into the clinical trial[29]. The review committee has no reason to believe that this recruitment restriction in any way violated the randomization process or that it was anything more than an inconvenience in the conduct of the study.

8.2.2 Distribution of OTT: The NINDS investigators examined in some detail the role of the actual value of OTT (not the dichotomized version) on the effectiveness of t-PA and concluded that their study demonstrated that earlier treatment was better[19]. In that manuscript they displayed a histogram of OTT values demonstrating that a high proportion of the patients entered in stratum #1 were entered with values of OTT between 80 and 90 minutes (see Figure 1 below).

Indeed, 150 (50%) of the patients randomized into stratum #1 had values of 89 or 90 minutes. We present the distribution of all OTT values in the form of a cumulative distribution function (see Figure 2 below) showing the sharp rise as the OTT values approach 90 minutes, and the cumulative percent approaches 50%.

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. Consequently, the Review Committee is somewhat skeptical of the analysis reported wherein the NINDS investigators used the OTT variable as a continuous variable[19] rather than as the protocol mandated dichotomized version.

Figure 1.

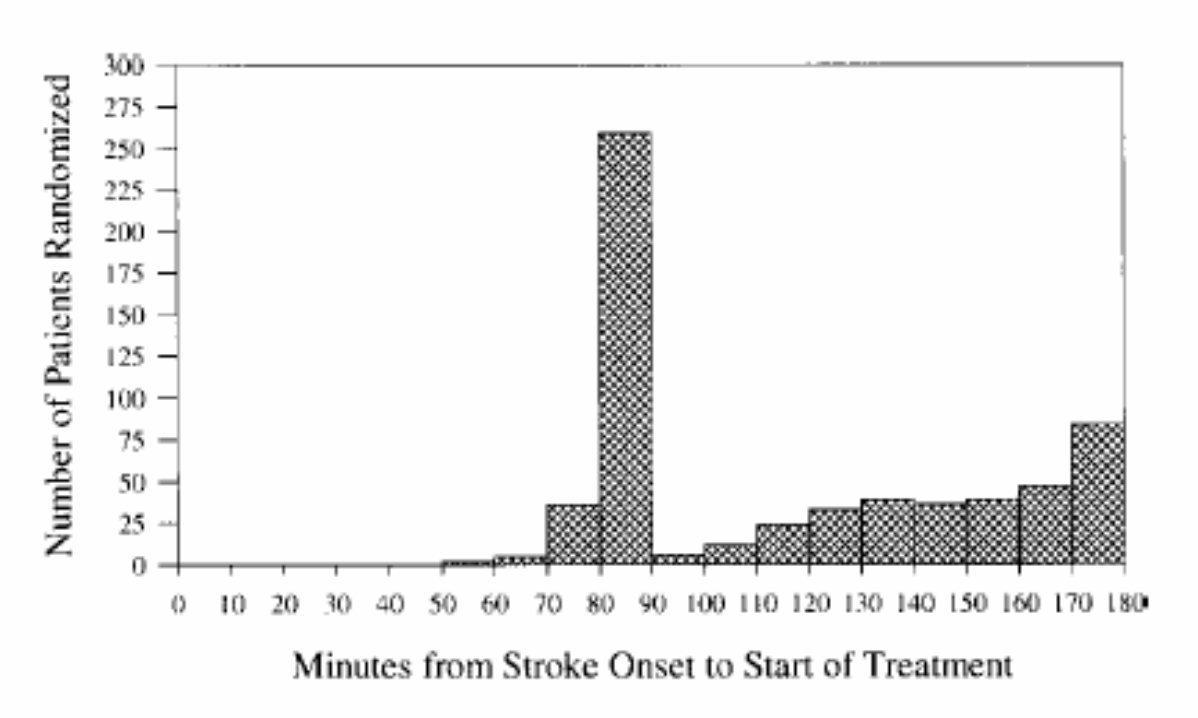
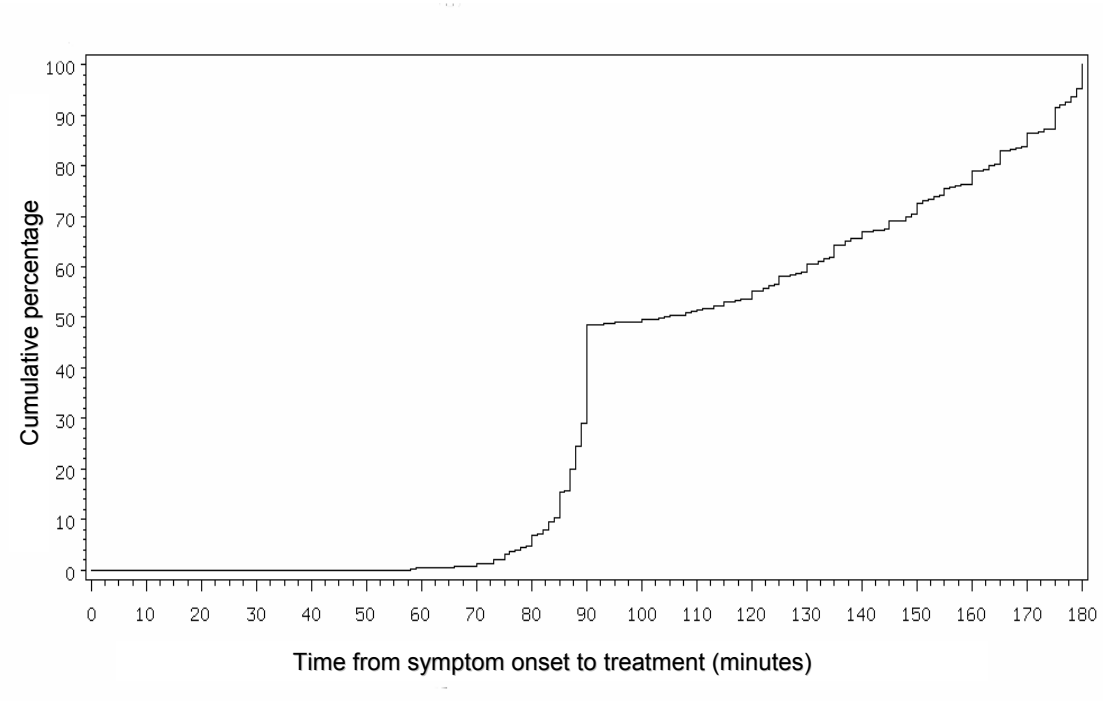


Figure 2.



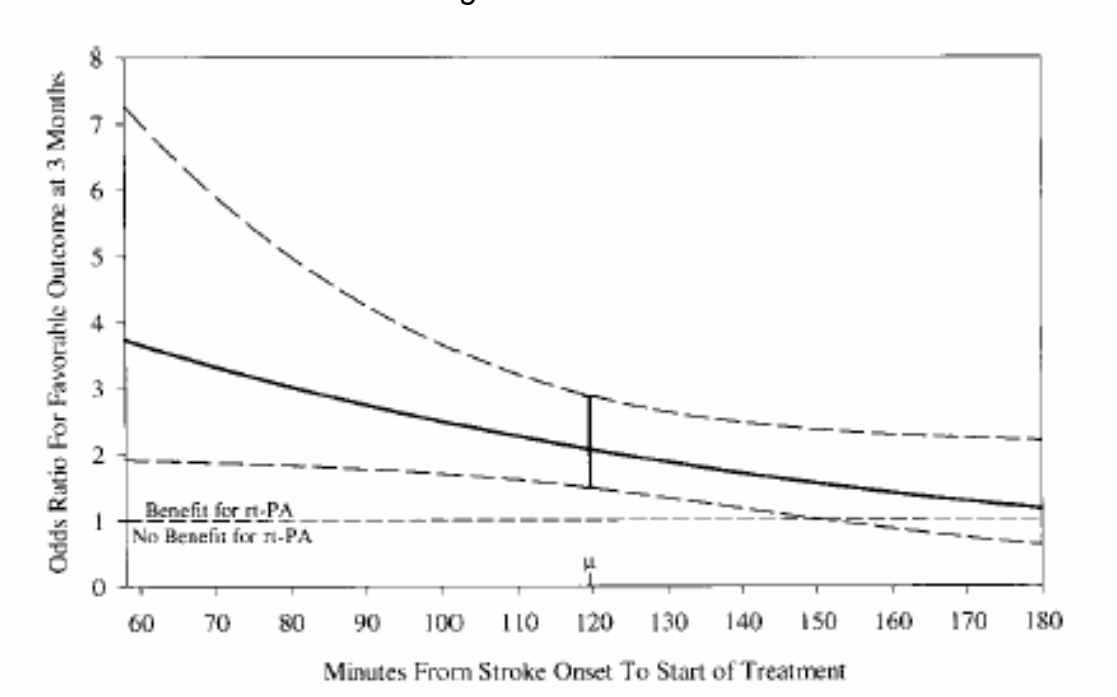
8.2.3 Does t-PA Effectiveness Decrease with Increasing OTT?: In order to investigate the issue of whether the NINDS study can lead to the conclusion that earlier t-PA therapy is better than later treatment, we performed a number of analyses.

8.2.3.1 OTT by t-PA Interactions (1): The first analysis considers whether the variable indicating OTT stratum interacted with the treatment group indicator in predicting outcomes.

We report these interaction tests in Tables 5.17 through 5.21 and discuss them in Section 5.5.2. For the Barthel, Rankin, Glasgow, NIHSS, and Global analyses, the tests of no OTT*t-PA interaction had p-values of 0.17, 0.87, 0.90, 0.22, and 0.19 respectively. These interaction tests, however, have 80 % power of detecting that two ORs differ only if their ratio is between 3.0 and 4.0, depending on outcome scale. Thus, if the true OR in stratum #2 is 1.0, the true OR in stratum #1 would have to be between 3.0 and 4.0 in order for the NINDS study to have 80% power. Put another way, these interaction tests had an 80% chance of being significant only if the two true ORs differ by a factor of 3 or more. We observed no such dramatic relative difference in ORs and the lack of statistical significance for the interaction tests is not surprising. Importantly, an interaction less than 3.0 may be clinically important, but the study has insufficient power to detect differences of such magnitude.

In the aforementioned article[19], the NINDS investigators presented a figure suggesting a range of ORs from 4.0 to 1.0 between OTT values of 60 and 180 minutes (see Figure 3). However, almost no patients had an OTT ~60 minutes. Indeed < 10% had OTT values as large as 82 mins, with a similar percent having OTT values between 176 and 180 minutes. According to the figure, the OR corresponding to 82 is <3 whereas the OR corresponding to 180 is > 1. Therefore, their own best estimate of OR differences suggests a less than 3-fold change over a reasonable OTT range, a change that the study has little power to detect.

Figure 3.



8.2.3.2 OTT by t-PA Interactions (2): Because of all the attention the NINDS publication on this topic has received, we pursued the issue further, working with indicator variables and substratifications rather than a continuous OTT. Since nearly all the OTT values in the first stratum were greater than 60 minutes, suggesting that the stratum was actually only about 30 minutes wide, we elected not to partition that stratum further. However, we did partition stratum 2 several ways to see if the "trend" reported by the investigators can be supported by an alternative analysis of the data. To facilitate the following discussion consider Table 8.2.1.

In that table, for each of the 4 outcome variables, there are 8 columns in three groups. The first column (labeled OTT: 0 – 90) summarizes data from the first stratum and the last column (labeled OTT: 91 – 180) presents the same summaries for the second stratum. Columns 2 through 5 correspond to a partition of the second stratum into 4 substrata, each containing nearly the same total number of randomized patients (i.e. quartiles of the distribution of OTT in stratum #2). Columns 6 & 7 contain the sums of columns 2 & 3 and 4 & 5 respectively. Each column contains 4 tables providing, for each outcome measure, the number of favorable and unfavorable responses among the placebo and t-PA patients randomized within the stratum (substratum) defined by that column. Two summary measures are then provided for each table. They are the odds of a favorable response among the placebo patients (PL odds F) and the t-PA vs. Placebo odds ratio (OR).

The most interesting aspect of this table is found in the substratum labeled OTT: 91 – 133 which summarizes information from the 81 patients randomized within stratum #2 whose OTT values were assessed to be between 91 minutes and 133 minutes inclusive. Of these patients, 50 (62%) were randomized to placebo. This is different from the expected 50% ($p = 0.035$). There is nearly perfect t-PA to placebo balance in the other three substrata and at the conclusion of the study 52% of the patients in stratum #2 were randomized into the placebo group. This still represents an excess of 7 patients from the expected 50%, almost all of which is attributable to the unexplained imbalance among those randomized with OTT values between 91 & 133. Of further interest in this regard is that on all outcome scales those 50 placebo patients randomized with OTT values between 91 and 133 had by far the lowest odds of a favorable outcome of any of the OTT substrata. As a result, on the OR scale this substratum stands out with exceptionally high values favoring t-PA. However, if the placebo patients in that substratum had odds for a favorable outcome more in line with the rest of the study, the corresponding ORs would be between one third and one half the quoted values.

The foregoing observations are relevant to the decreasing trend in OR as reported by the NIHSS investigators[19] seen here in Figure 3 . The inexplicable and likely artificial elevation of the OR during the 91 to 133 minutes interval could tilt the OR scale up at the earlier part of stratum #2 resulting in an estimate of a negative slope with increasing OTT. Furthermore, for the Barthel & Glasgow outcome measures, the period OTT: 174 – 180 demonstrates what appear to be equally inexplicable elevations in the odds of a favorable outcome among the placebo patients. These clearly contributed to the lower estimates of the OR in that period and would have also contributed to the negative slope estimate.

8.2.3.3 OTT by t-PA Interactions (3): These observations about the nature of the relationship between the distributions of favorable response among the placebo patients and OTT notwithstanding, we carried out two sets of additional formal logistic and GEE regression analyses, including all final covariates, which we summarize briefly in this paragraph. Both analyses involved partitioning the second stratum into substrata as illustrated in Table 8.2.1. In the first analysis we divided it into two substrata (OTT: 91-154 & OTT: 155-180) each containing 160 patients. In the second analysis we used the 4 substrata defined by the quartiles discussed earlier. In both analyses, stratum #1 was used as the comparison group. In the first analysis, 2 degrees-of-freedom were required to separate the resulting three OTT classes and in the second analysis 4 degrees-of-freedom were required to separate the 5 OTT classes. In the first step of each analysis we estimated the t-PA to Placebo odds ratio to assess the impact of this change in the OTT variable on the OR estimates. The results (first analysis, second analysis) for each outcome variable are: Barthel (2.19, 2.18), Rankin (2.43,

2.39), Glasgow (2.13,2.11), NIHSS (2.21, 2.22) and Global (2.14, 2.14). These results do not differ from each other nor do they differ from the adjusted odds ratios reported in Section 5.5.1.

The second step of the analyses was to determine whether t-PA and OTT interacted with each other in separate models each containing one of these two versions of OTT. For the first analysis, with the 2 degree-of-freedom OTT variable, the results of the no interaction test may be summarized as: Barthel (p = 0.14), Rankin (p = 0.08), Glasgow (p = 0.31), NIHSS (p = 0.39) and Global (p = 0.13). For the second analysis (4 dfs) the results are: Barthel (p = 0.17), Rankin (p = 0.21), Glasgow (p = 0.15), NIHSS p = 0.17) and Global (p = 0.06). Clearly these analyses failed to identify any significant interaction between the treatment and OTT variables (as did the fundamental analysis reported earlier in this section) that is, no collection of patients randomized at any of the five OTT levels discussed can be said to have a significantly different response to t-PA therapy than any other group.

8.2.4 Summary and Conclusions: In light of these results, the substantially nonlinear nature of the distribution of OTT when considered as a continuous variable, and the idiosyncratic distribution of favorable response rates among the placebo patients, we conclude that the data provided by this study failed to support a conclusion that the effect of t-PA therapy diminishes with increasing values of OTT within the protocol specified 3 hour time limit. However, this does not mean such a relationship does not exist, and further studies are needed to address the question of a differential t-PA treatment effect related to time from symptom onset to treatment. It is also important to recognize that the results from this study provide no data on the effectiveness of thrombolytic therapy administered to acute ischemic stroke patients more than 180 minutes after symptom onset.

8.3 Clinical Centers:

Randomization took place within each of 9 centers; however, one center randomized and treated only one patient who was followed by another center. The NINDS investigators considered those two centers as a single center. We therefore consider the study as involving 8 centers, or strata, and the "Center" variable carries 7 degrees-of-freedom in all of the models.

8.3.1 Center Comparisons of Favorable Outcome Rates: Center differences regarding such issues as recruitment and outcome are illustrated in Tables 8.3.1 and 8.3.2. These two tables have identical structure and support all of the statements regarding specific numerical values unless otherwise specified.

The capacity of the centers and their access to appropriate patients differed appreciably. Two centers (#'s 4 and 5) randomized 146 and 150 patients respectively (nearly 50% of the entire study population). The remaining centers randomized 103, 71, 62, 39, 37, and 14 patients. As was observed in Section 5.3 (Table 5.1), the chances of a favorable outcome, regardless of therapy, varies by outcome variable. The overall favorability percents are 45%, 35%, 38%, and 27% for the Barthel, Rankin Glasgow and NIHS outcome measures respectively, corresponding to odds of favorability of 0.82, 0.53, 0.62 and 0.38. Across the centers the percent favorable ranges over about 25 percentage points with Center 7 always low, but not always the lowest and Center 3 always the highest.

Chi-square tests (not presented) of the hypothesis that the rates of a favorable outcome are the same over all centers were not significant. Furthermore, in the final models described in Sections 5.4 and 5.5, the 7 degrees-of-freedom Center variables, always included because it is part of the study design, was never significant even after adjusting for all the other variables in the model. Thus, observed differences in the likelihood of a favorable outcome result from statistical variation and should not be taken as evidence of important, underlying center-to-center variation.

8.3.2 Center Comparisons of t-PA Effect: In Tables 8.3.1 and 8.3.2, we have listed for the 8 centers, within each of the 4 outcome measures, their associated numbers of favorable and unfavorable outcomes by treatment group with two measures of treatment effect. These are, the difference (delta) between the percents of t-PA and placebo patients experiencing a favorable outcome and the ratio (t-PA to Placebo) of the odds of a favorable outcome. In both tables the centers are ordered by their rank according to the t-PA by placebo odds ratio for the Barthel scale. Thus, center #4 with a Barthel OR of 2.77 and a delta favorable outcome percent of 24.8% is ranked first even though it does not rank first for all 4 scales. Center #7, with only 14 patients randomized, which had consistently among the lowest overall rates of favorable outcomes did rank last on all scales when comparing t-PA to placebo. Indeed, all of the center #7 estimated odds ratios were less than one by a considerable, although not statistically significant, margin, consistently indicating more favorable outcomes among the placebo treated patients than among the t-PA treated patients at that center.

The data in the tables suggest what appears to be considerable variability among the centers as regards the odds ratios comparing t-PA therapy to placebo. For the Barthel scale, for example, the maximum odds ratio of 2.77 (Center # 4) is nearly 9-fold higher than the

minimum of 0.33 (center #7). For the Rankin and Glasgow scales this ratio of maximum and minimum odds ratios is even greater. However, the 95% confidence intervals, most of which overlap the null value 1, indicate that very few of the within center odds ratio estimates are "significantly" greater than 1 and none is significantly less than 1. Most of these confidence intervals, especially those based on the centers with smaller numbers of patients randomized, are very wide, reflecting the substantial random error present in estimates obtained from such small numbers of observations.

8.3.3 Center by t-PA Interaction: These observations raise the question of whether there is evidence that the response of patients to t-PA therapy as estimated through the odds ratios comparing the response of t-PA treated patients to the response of placebo treated patients is different among the 8 centers. This is a verbal description of an interaction between the variable defining therapy and the variables defining Center and the question is formally addressed by introducing into the models upon which our comparisons of t-PA to Placebo are based an appropriate interaction.

We have already described and briefly discussed formal testing of those interactions in Section 5.5.2. Specifically, in Tables 5.17 through 5.21 we reported the p-values for these 7 degrees-of-freedom tests for each outcome variable as well as for the global analysis. These tests were all conducted within the models containing the covariates deemed appropriate for "adjusting" the treatment comparisons. The p-values reported were 0.16, 0.24, 0.17, 0.87 & 0.47 for the Barthel, Rankin, Glasgow, NIHSS, and Global analyses respectively and are based on Chi-squared statistics of 10.49, 9.12, 10.28, 3.20, and 6.61 respectively. Based on this lack of statistical significance, we conclude that there is little evidence of an important interaction. However, the study was not powered to detect interactions, so the lack of significance does not guarantee the absence of an interaction.

8.3.4 Estimates of Differences in Favorable Outcome Percentages: Motivated by the foregoing, we pursued the question of the influence of the "interactions" a step further. Using equation 2 from Section 4.3.7, which permits the estimation of the difference in favorable outcome percentages while weighting the individual contribution of groups of patients (in this case from the different centers) we estimated these differences based on three different scenarios. The first scenario is based on the crude data seen in Tables 8.3.1 and 8.3.2. Thus, for Barthel, Rankin, Glasgow and NIHSS scales, respectively, the direct estimates of the differences in percent favorable response (with 95% confidence intervals) are: 14.1% (6.4%, 21.9%), 16.3% (8.9%, 23.7%), 14.4% (6.8%, 22.0%), and 13.7% (6.8%, 20.6%). If we model these estimates without including an interaction term, forcing the OR estimates to be the same for all centers, but otherwise adjusting for all covariates (Section 5.4.3) the estimates become: 17.3%, 16.5%, 16.7% and 14.2%, respectively. Finally, if we estimate these differences using models that contain a t-PA by Center interaction, permitting the OR estimates to be different for the 8 centers, and thus taking into account the fact that some centers have an estimated negative difference, the estimates are: 19.1%, 15.3%, 17.4% and 14.6%, respectively.

A comparison of the several estimates of the differences (t-PA minus Placebo) in rates of favorable outcomes suggests that these estimated differences do not change notably as we go from direct estimates to estimates based on complex models. Specifically, if we allow the models to estimate different ORs for each center and use those ORs in the estimates of the difference we obtain difference estimates that are essentially the same as those obtained in

the absence of an interaction. Thus, using the most flexible and completely adjusted within-center estimates of ORs that are available in the estimation of the differences between t-PA and Placebo rates of a favorable outcome, the difference estimates are essentially the same as those obtained from the raw data. Thus, our position that there is a statistically and clinically significant net positive effect of t-PA remains.

8.3.5 Summary and Conclusions: We found no significant difference between the centers in the baseline characteristics of the patients. The likelihood of having a favorable outcome differed considerably between the centers, those with fewer patients often having the worst outcome. However, the between-center variation in t-PA treatment effect for either the global outcome, or the individual outcome measures, was not statistically significant and did not invalidate the trial results. Nevertheless, it will be important in future studies to identify the factors that lead to good outcomes at institutions administering t-PA to treat acute ischemic stroke patients. This information will be very helpful to other institutions that are looking to develop the resources needed to administer t-PA safely to acute ischemic stroke patients.

8.4 Stroke Subtype

8.4.1 Introduction: The NINDS investigators examined all pre-randomization records in an attempt to determine the ischemic stroke subtype into which each patient could be classified[1]. The result was a post-randomization classification of the patients into one of four subtypes: small vessel, cardioembolic, large vessel and other. In Table 4.2 and Section 4.1.3, we noted that the randomization of the patients into t-PA and Placebo treatment groups had resulted in a marginal imbalance regarding the subtype groups (p=0.064). In Table 4.2 and in Table 8.4.1 associated with this section, it can be seen that only the 273 cardioembolic patients were divided as nearly as possible equally between the two treatment groups. In contrast, 63% of the 81 small vessel patients were randomized into the t-PA treatment arm while 46% of the remaining 268 patients, mostly classified as having a large vessel stroke, were randomized to t-PA.

8.4.2 Analyses: The four subtype groups were examined analytically in three ways. First, the three variables necessary to uniquely indicate each patient's membership in one of the groups were, collectively, examined with all other potential covariates as part of the process of arriving at a "final" collection of covariates to be included in the treatment comparison models. In Table 5.10, it can be seen that these variables were quite significant in the first stage of this process when all covariate candidates were examined within models containing only the stratification variables. The corresponding 3-degrees-of-freedom Chi-squares and associated p-values for the 4 outcome measures were: Barthel (17.78, p = 0.0005), Rankin (15.16, p = 0.0017), Glasgow (10.74, p = 0.0132) and NIHSS (9.60, p = 0.0222). These analyses clearly indicate that stroke subtype is associated with the likelihood of a favorable outcome on all four scales. The nature of this association can be seen in the Table 8.4.1. For the Barthel, Rankin and Glasgow scales, the small vessel stroke patients had better than an even chance of a favorable outcome regardless of treatment, with odds of 1.9, 1.2, and 1.25 while for the other subtypes combined the odds were 0.72, 0.47 and 0.55 respectively. For the NIHSS scale the direction of the difference was the same but less dramatic.

However, as is seen in Table 5.10, there were other variables that were much more strongly related to the likelihood of a favorable outcome and, as the stepwise process of identifying the critical covariates continued, these variables entered the models, modifying the level of significance of the stroke subtype variables such that they never entered a single model. The logical conclusion to draw here is that the combined information contained in BsNIHSS, Age, PrDisability and PrDM, the variables that did enter the models, was highly correlated with the information that separated the small vessel stroke patients from the others and that the differences among the subtypes was no longer necessary in the models.

Nevertheless, we continued with the second and third stages of our examination of the subtype variables. We first included the three indicator variables, regardless of their p-values, in the outcome models that have been discussed so extensively in Sections 5.4 and 5.5. In these models, the 3-degrees-of-freedom Chi squares and associated p-values for the subtype variables are; Barthel (1.76. p = 0.62), Rankin (5.00, p = 0.17), Glasgow (5.06, p = 0.17). NIHSS (2.77, p = 0.43) and Global (3.09, p = 0.38). Since these variables are not statistically significant in these models, they will not be included. However lack of statistical significance does not prove that there are no differences among the stroke subtypes regarding a patient's likelihood of experiencing a favorable outcome. Even in the face of this lack of significance, it is interesting to note that, in these models with all of the influential covariates, the adjusted

estimates associated with the subtype variables now suggests that the cardioembolic group has the best chance of experiencing a favorable outcome, regardless of therapy.

Finally, we examined the question of whether the influence of t-PA as measured by the t-PA to Placebo odds ratio, is different across the subtype groups. In Table 8.4.1 we see, ignoring the "other" group because it is small (18 patients) and ill defined, the only consistent pattern is that the OR is smallest in the cardioembolic group for all outcome measures. In the 5 formal analyses, four outcome measures and the Global analysis, none of the interaction tests were significant (p – values between 0.40 and 0.58), indicating that the data do not provide statistically significant evidence suggesting a differential t-PA effect by stroke subtype. All previous caveats about insignificant tests of no interaction apply in this case as well.

8.4.3 Summary and Conclusions: We conclude that it was appropriate that the NINDS Investigators did not include stroke subtype as a covariate in the analytic models. Further, we conclude that the data of this trial do not support any claim regarding either the presence, or absence, of a differential t-PA treatment effect within stroke subtype.

8.5 Preexisting Disability

As illustrated in Table 8.5.1, 46 of the 622 patients randomized into this study were disabled prior to their stroke. The severity of the disability was assessed using the modified Rankin scale[1], and of the 46 patients with some preexisting disability 23 had slight disability (Rankin = 2), 17 had moderate disability (Rankin = 3) and 6 had severe disability (Rankin = 4). Very few of these 46 patients were classified as having a favorable outcome at 90 days. Indeed, for the Barthel, Rankin, Glasgow and NIHSS outcome measures, the number (%) of patients experiencing a favorable outcome was 6 (13%), 5 (11%), 5 (11%) and 3 (6.5%) respectively. None of the favorable outcomes occurred among patients with a severe disability and only one among those with a moderate disability (Table 8.5.1). In contrast, for the remaining 576 patients, the corresponding percents were; 48%, 37%, 40% and 29% respectively. This 3 to 4-fold difference in favorable response rate resulted in the inclusion of a variable indicating the 46 patients with a preexisting disability in the models assessing treatment response (Section 5.4.3).

In Section 5.5.2.1, we reported that this variable, while a highly significant predictor of an unfavorable outcome in all models, does not interact significantly with the treatment variable in any of the models. This lack of an interaction is seen in Table 8.5.1 where the ORs contrasting t-PA and Placebo are found to be very similar for those with and without a preexisting disability.

8.5.1 Summary and Conclusions: Thus, despite the fact that patients with a preexisting disability had a significantly reduced chance of experiencing a favorable outcome, there was no evidence that they responded any differently to t-PA therapy than those without a preexisting disability.

8.6 Diabetes Mellitus

8.6.1 Analyses: In our preliminary investigations we identified 8 patients with baseline blood glucose in excess of 400, in violation of the study protocol. Since elimination of this small number of patients would not substantially alter our conclusions, we continued to include them.

As seen in Table 8.6.1, of the 622 patients randomized into the study, 131 (21%) had a history of diabetes mellitus (DM). Of these 131, 34%, 30%, 31%, and 18% experienced a favorable outcome at 90 days according to the Barthel, Rankin, Glasgow, and NIHSS outcome measures respectively. The corresponding figures for those without DM are 48%, 36%, 40%, and 30% (see table entitled Diabetes Workbook). The differences between the corresponding favorability percents give a measure of the impact of DM on the likelihood of a favorable outcome. In Section 5.4.3 we reported that DM was the last of the "adjusting" covariates to enter the models but it had remained a significant predictor of an unfavorable outcome even after adjusting for the stratification variables as well as for the highly significant BsNIHSS, AGE and PrDisability.

In the Table 8.6.1, the ORs comparing t-PA therapy to Placebo are quoted for the DM and nonDM groups separately. It appears that there is little evidence of a t-PA advantage over Placebo among diabetics. However, in the Section 5.5.2.1, Tables 5.17 through 5.21, we reported the results of the tests of whether DM interacted with the treatment variable for the Barthel, Rankin, Glasgow, NIHSS and Global analyses, none of which was significant ($p = 0.08, 0.25, 0.23, 0.96$, and 0.27 respectively). In that section we also reported that these tests of "no interaction" would have 80% chance (power) of being significant only if the ORs among the nonDM patients were from 4 to 6-fold higher than the ORs among the DM patients. Since we observed (see Table 8.6.1) ratios of odds ratios between 1.3 (NIHSS) & 2.6 (Barthel), the fact that these tests of no interaction were all insignificant is no surprise. The caveat that we have stated before that the lack of evidence of difference does not constitute proof of the lack of a difference needs to be kept in mind here.

8.6.2 Summary and Conclusions: Although the observed data (Table 8.6.1) and the adjusted estimated t-PA effects, indicated a strong benefit for patients without DM, but no benefit among patients with DM, this comparison must be treated cautiously because there was no statistical evidence of a t-PA*DM interaction. The trial found no statistically significant evidence that diabetic and non-diabetic acute ischemic stroke patients responded differently to t-PA therapy.

9. CONCLUSION

The committee concluded that, despite an increased incidence of symptomatic intracerebral hemorrhage in t-PA treated patients and subgroup imbalances in baseline stroke severity, when t-PA was administered to acute ischemic stroke patients according to the study protocol, there was a statistically significant, and clinically important, benefit of t-PA treatment resulting in a higher likelihood of having a favorable clinical outcome at three months.

References

1. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1995;333:1581-1587.

2. 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;24:1000-1004.

3. Haley EC, Jr., Levy DE, Brott TG et al. Urgent therapy for stroke. Part II. Pilot study of tissue plasminogen activator administered 91-180 minutes from onset. *Stroke* 1992;23:641-645.

4. 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;23:632-640.

5. Lyden P, Brott T, Tilley B et al. Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke* 1994;25:2220-2226.

6. Lyden P, Broderick J, Mascha E. Reliability of the Barthel Index Outcome Measure selected for the NINDS t-PA Stroke Trial. In: Yamaguchi T, Mori E, Minematsuk K, del Zoppo G, editors. *Thrombolytic Therapy in Acute Ischemic Stroke III*. Tokyo, Japan: Springer-Verlag; 1995. p. 327-333.

7. Levy DE, Brott TG, Haley EC, Jr. et al. Factors related to intracranial hematoma formation in patients receiving tissue-type plasminogen activator for acute ischemic stroke. *Stroke* 1994;25:291-297.

8. The National Institute of Neurological Disorders and Stroke (NINDS) rt-PA Stroke Study Group. A systems approach to immediate evaluation and managment of hyperacute stroke. Experience at eight centers and implications for community practice and patient care. *Stroke* 1997;28:1530-1540.

9. The NINDS t-PA Stroke Study Group. Generalized efficacy of t-PA for acute stroke. Subgroup analysis of the NINDS t-PA stroke trial. *Stroke* 1997;28:2119-2125.

10. Kwiatkowski TG, Libman RB, Frankel M et al. Effects of tissue plasminogen activator for acute ischemic stroke at one year. *N Engl J Med* 1999;340:1781-1787.

11. The NINDS t-PA Stroke Study Group. Intracerebral hemorrhage after intravenous t-PA therapy for ischemic stroke. *Stroke* 1997;28:2109-2118.

12. Brott T, Lu M, Kothari R et al. Hypertension and its treatment in the NINDS rt-PA stroke trial. *Stroke* 1998;29:1504-1509.

13. Lenzer J. Alteplase for stroke: money and optimistic claims buttress the "brain attack" campaign. *BMJ* 2002;324:723-729.

14. Mann H, Li J, Nathanson LA et al. Alteplase for stroke. *BMJ* 2002;324:1581.

15. Mann J. Truths about the NINDS study: setting the record straight. *West J Med* 2002;176:192-194.

16. Trotter G. Why were the benefits of tPA exaggerated? *West J Med* 2002;176:192-194.

17. Wardlaw JM, Lindely RI, Lewis S. Thrombolysis for acute ischemic stroke: still a treatment for the few by the few. *West J Med* 2002;176:198-199.

18. Warlow C. Commentary: Who pays the guideline writers? *BMJ* 2002;324:726-727.

19. Marler JR, Tilley BC, Lu M et al. Early stroke treatment associated with better outcomes. The NINDS rt-PA Stroke Study. *Neurology* 2000;55:1649-1655.

20. Tilley BC, Marler J, Geller NG et al. Use of a global test for multiple outcomes in stroke trials with application to the National Institute of Neurological Disorders and stroke t-PA stroke trial. *Stroke* 1996;27:2136-2142.

21. Clinical Review for PLA 96-0350 - Endpoints and Planned Analyses. *Food and Drug Administration* 1996;16. Available at: http://www.fda.gov/cder/biologics/review/altegen061896r2.pdf. Accessed December 12, 2003.

22. Efron B, Gong G. A leisurely look at the bootstrap, the jackknife and cross-validation. *The American Statistician* 1982;37:36-48.

23. Tilley BC, Lyden PD, Brott TG, Lu M, Levine SR, Welch KM. Total quality improvement method for reduction of delays between emergency department admission and treatment of acute ischemic stroke. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. *Arch Neurol* 1997;54:1466-1474.

24. Report of the Quality Standards Subcommittee of the American Academy of Neurology. Practice Advisory: Thrombolytic therapy for acute ischemic stroke - Summary Statement. *Neurology* 1996;47:835-839.

25. Gore JM, Granger CB, Simoons ML et al. Stroke After Thrombolysis : Mortality and Functional Outcomes in the GUSTO-I Trial. *Circulation* 1995 November 15;92:2811-2818.

26. Selker HP, Beshansky JR, Schmid CH et al. Presenting pulse pressure predicts thrombolytic therapy-related intracranial hemorrhage. Thrombolytic Predictive Instrument (TPI) Project results. *Circulation* 1994;90:1657-1661.

27. Simoons ML, Maggioni AP, Knatterud G et al. Individual risk assessment for intracranial haemorrhage during thrombolytic therapy. *Lancet* 1993;342:1523-1528.

28. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004;57:229-236.

29. Clinical Review for PLA 96-0350 - Screened Patient Characteristics. *Food and Drug Administration* 1996;23-24. Available at:

http://www.fda.gov/cder/biologics/review/altegen061896r2.pdf. Accessed December 12, 2003.

ANNEX 5: Slot *et al*

Impact of functional status at six months on long term survival in patients with ischaemic stroke: prospective cohort studies

BMJ

RESEARCH

Impact of functional status at six months on long term survival in patients with ischaemic stroke: prospective cohort studies

Karsten Bruins Slot, clinical research fellow,[1] Eivind Berge, senior consultant,[1] Paul Dorman, consultant neurologist,[2] Steff Lewis, medical statistician,[3] Martin Dennis, professor,[3] Peter Sandercock, professor,[3] on behalf of the Oxfordshire Community Stroke Project, the International Stroke Trial (UK), and the Lothian Stroke Register

[1]Department of Internal Medicine, Ullevaal University Hospital, NO-0407 Oslo, Norway
[2]Department of Neurology, Newcastle General Hospital, Newcastle upon Tyne
[3]Department of Clinical Neurosciences, Western General Hospital, Edinburgh
Correspondence to: K Bruins Slot karsten.bruins.slot@medisin.uio.no

doi:10.1136/bmj.39456.688333.BE

ABSTRACT

Objective To estimate the impact on long term survival of functional status at six months after ischaemic stroke.

Design Prospective cohort study.

Settings Three cohorts: Oxfordshire community stroke project, Lothian stroke register, and the first international stroke trial (in the United Kingdom).

Participants 7710 patients with ischaemic stroke registered between 1981 and 2000 and followed up for a maximum of 19 years.

Main outcome measures Functional status at six months after stroke assessed with modified Rankin scale or "two simple questions." Mortality during follow-up. Survival analysis with Kaplan-Meier curves, log rank test, and Cox's regression model.

Results In a combined analysis of all three cohorts, among patients who survived to assessment six months after the index stroke, the subsequent median length of survival among those independent in daily living and those dependent was 9.7 years (95% confidence interval 8.9 to 10.6) and 6.0 years (5.7 to 6.4), respectively. In a combined analysis of the Oxfordshire and Lothian cohorts, subsequent median survival fell progressively from 12.9 years (10.0 to 15.9) for patients with a Rankin score of 0-1 at six months after the stroke to 2.5 years (1.4 to 3.5) for patients with a Rankin score of 5. All previously stated differences in median survival were significant (log rank test $P<0.001$). The influence of functional outcome on survival remained significant ($P<0.05$) in each cohort after adjustment for relevant covariates (such as age, presence of atrial fibrillation, visible infarct on computed tomography, subtype of stroke) in a Cox's regression model.

Conclusion Functional status six months after an ischaemic stroke is associated with long term survival. Early interventions that reduce dependency at six months might have positive effects on long term survival.

INTRODUCTION

The global burden of stroke is large, yet there are still gaps in our knowledge.[1 2] Although there are now reliable estimates on outcome in the early months and years after an ischaemic stroke, we know much less on long term survival and what influences it.[3] This lack of information is important for many reasons. If, for example, functional status several months after a stroke has a major influence on long term survival, this will affect clinical practice (including our communication with patients), our estimates of the future global burden and costs of stroke, and the planning of health care and research.

We estimated the relative and absolute effects of the level of functional status at six months on long term survival in three large prospective cohorts of patients with ischaemic stroke.

METHODS

We sought data from three cohorts of patients with an ischaemic stroke recruited in the United Kingdom: the Oxfordshire community stroke project, the Lothian stroke register, and the UK patients enrolled in the first international stroke trial.

Initial data collection and clinical follow-up

Oxfordshire community stroke project—This project was a community based incidence study of stroke and transient ischaemic attacks.[4] Patients were registered from 1981 to 1986. Details on the study population, clinical definitions, methods of assessment, and investigations have been described in detail elsewhere.[4] A study neurologist assessed all patients as soon as possible after the onset of symptoms. Baseline characteristics were recorded in a standardised form. Trained study nurses followed up surviving patients at one, six, and 12 months from the date of stroke onset and then annually for up to five years. When possible, a study physician assessed survivors at the end of clinical follow-up.

Lothian stroke register—The register was established to collect data on patients with suspected stroke, transient ischaemic attacks, or retinal artery occlusion from those attending outpatient clinics and admitted to one

hospital in Edinburgh. The registration began in 1990 and continued to 2000. One of the study's stroke physicians examined the patient and collected baseline data as soon as possible after symptom onset. Patients were followed up at 6, 12, 24, and 36 months from the date of symptom onset. Follow-up data were obtained either by telephone interview, postal questionnaire, or home or clinic visits.

First international stroke trial—This was a randomised trial of aspirin, subcutaneous heparin, both, or neither, started within 48 hours of onset of ischaemic stroke.[5] A total of 19 435 patients were enrolled from 1991 to 1997, of whom 6257 (32%) were enrolled by hospitals in the UK. Baseline data were collected before randomisation in the trial. Final clinical follow-up at six months was by postal questionnaire or telephone interview or, in a few cases, during a clinic visit.

Collection of long term survival data

At the end of planned clinical follow-up in each of the three cohorts, notes of patients who were still alive were "flagged" at the NHS central register of the Office for National Statistics (ONS). On the death of a cohort participant, ONS forwarded notification of the death and a copy of the death certificate to the study office. Patients who were not reported to have died before the close of follow-up on 16 November 2000 were assumed to be alive.

Classification of ischaemic strokes

In all three cohorts, ischaemic stroke was diagnosed with a combination of clinical criteria and brain imaging or autopsy. As these examinations excluded intracerebral haemorrhages and conditions mimicking stroke (for example, subdural haematoma or cerebral tumour), the presence of visible infarction on imaging (or autopsy) was not necessary for the diagnosis of ischaemic stroke. According to criteria from the Oxfordshire community stroke project classification, we used the clinical features to subdivide diagnosis into total anterior circulation infarct, partial anterior circulation infarct, lacunar infarct, posterior circulation infarct, or, when no clinical subtype could be assigned, cerebral infarct of indeterminate clinical subtype.[6]

Definition of outcomes

In the Oxfordshire and Lothian cohorts the level of function at six months after stroke onset was assessed by the modified Rankin scale.[7] In the international stroke trial this was done by means of the "two simple questions" that were developed to assess functional outcome after stroke in large scale trials.[8] The patients (directly or through relatives) were asked if they had needed help from another person to perform everyday activities within the past two weeks (such as bathing, feeding, walking, dressing, or use of the toilet). The Rankin score and the two simple questions are methods that both have good validity and reliability between observers and correspond well with each other.[7 9 10] We defined an independent state as Rankin score of 0-2 and a dependent state as score of 3-5. The international stroke trial classified patients who reported not needing any help to perform everyday activities within the past two weeks as independent.

Statistical analysis

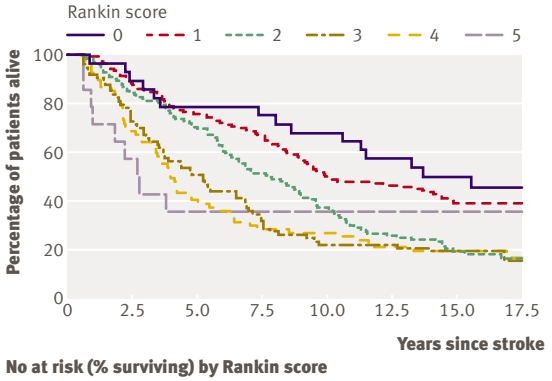
We estimated survival curves in the three cohorts with the Kaplan-Meier product limit technique. We used median rather than mean to describe and compare survival from the six month assessment of functional outcome in each cohort as means are hugely influenced by the length of follow-up (which varied in the three cohorts). We performed univariate and multivariate analyses of risk factors with Cox's proportional hazards models. Data from patients who were dead at six months after stroke onset were not entered in the models as we were interested only in the impact of functional status at six months on subsequent survival. We entered age and systolic blood pressure as continuous variables. The proportionality assumption was verified with the Schoenfeld test and did not seem to be violated.[11] We used SPSS software (version 13.0 for Mac OS X) for the statistical analysis.

RESULTS

Oxfordshire cohort

This study registered 675 patients with first ever stroke. We excluded 136 (20%). Of these, 130 did not have a diagnosis of ischaemic stroke (33 had a subarachnoid haemorrhage, 65 a primary intracerebral haemorrhage, and 32 a stroke of undefined pathological type). We excluded six other patients in whom there was an apparent error in the recording of the date of death. The 539 remaining patients had a definite (n=434) or probable (n=105) ischaemic stroke.

Table 1 shows the baseline characteristics and vital and dependency status at six months. Patients were followed up for a maximum of 19 years. Figure 1 shows



No at risk (% surviving) by Rankin score

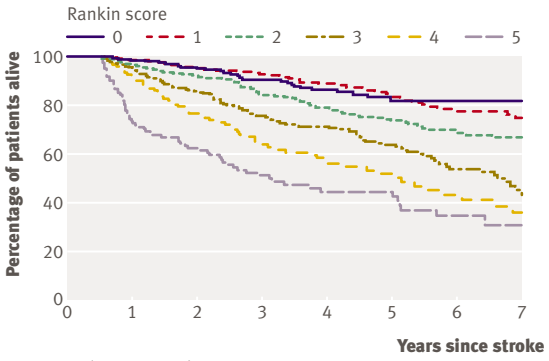
0	28 (100)	25 (89)	22 (79)	21 (75)	19 (68)	16 (57)	11 (50)	8 (45)
1	136 (100)	120 (88)	103 (76)	91 (67)	68 (50)	63 (46)	39 (39)	15 (39)
2	121 (100)	102 (84)	84 (69)	62 (51)	45 (37)	31 (26)	18 (19)	4 (16)
3	73 (100)	53 (73)	37 (51)	23 (32)	16 (22)	16 (22)	9 (19)	2 (15)
4	67 (100)	46 (69)	27 (40)	19 (28)	18 (27)	14 (21)	12 (19)	6 (17)
5	14 (100)	8 (57)	5 (36)	5 (36)	5 (36)	5 (36)	4 (36)	2 (36)

Fig 1 Oxfordshire cohort. Long term survival of patients in each category of functional status (Rankin score 0-5) from assessment at six months after index stroke

survival curves for patients stratified by Rankin score 0-5 at six months. There was a significant trend (log rank test, P<0.001) of decreasing survival with increasing Rankin score at six months. We entered all baseline variables in table 1 and functional status at six months after stroke onset into a univariate and multivariate Cox's regression model (table 2). Both the separate Rankin scores and level of dependency at six months had a significant effect (P<0.05) on subsequent survival in the multivariate analyses. The more dependent a patient was at six months, the shorter their subsequent survival. Age and the presence of atrial fibrillation on examination also had a significant negative effect (P<0.001) on survival. We used a similar model to analyse the impact of the Rankin scores at one month after stroke onset. This gave generally the same results as those of the Rankin scores at six months (data not shown).

Lothian cohort

In all, 4455 patients with a stroke, transient ischaemic attack, retinal artery occlusion, or other diagnosis were entered on the register in 1990-9. We sought patients with relevant clinical features and computed tomography or magnetic resonance imaging at baseline indicating an ischaemic infarct (n=1547) or patients with normal results on computed tomography or magnetic resonance imaging at baseline and a clinical diagnosis of a probable (n=320) or definite stroke (n=629). We excluded 442 (18%) of these patients from our final analysis; 414 patients whose first follow-up occurred (for organisational reasons) at 12 months or later (hence functional status at six months was not known), seven patients who were lost to follow-up by six months, one patient in whom there was an apparent error in the recording of the date of stroke onset, and 20 patients who had refused further participation in the study at some point after entry. In our final analyses we therefore had data on 2054 patients.



No at risk (% surviving) by Rankin score

0	283 (100)	274 (98)	214 (96)	157 (90)	102 (86)	52 (82)	28 (82)	12 (82)
1	404 (100)	392 (99)	318 (95)	253 (93)	181 (89)	117 (83)	83 (79)	46 (75)
2	455 (100)	435 (97)	373 (92)	289 (84)	224 (79)	156 (74)	97 (69)	55 (67)
3	360 (100)	342 (96)	265 (85)	202 (76)	140 (71)	91 (64)	55 (54)	22 (43)
4	122 (100)	112 (92)	83 (77)	61 (64)	47 (57)	32 (52)	20 (43)	13 (36)
5	122 (100)	89 (74)	67 (62)	43 (51)	28 (44)	24 (44)	12 (34)	6 (31)

Fig 2 Lothian cohort. Long term survival of patients in each category of functional status (Rankin score 0-5) from assessment at six months after index stroke

Table 1 shows baseline characteristics and vital and dependency status at six months. Patients were followed up for a maximum of 9.7 years. Figure 2 shows survival curves. There was a significant trend (log rank test, P<0.001) of a decrease in survival with an increase in Rankin score at six months. We entered the baseline variables in table 1 and the functional status at six months after stroke onset in a univariate and multivariate Cox's regression model (table 3). Both the separate Rankin scores and the level of dependency had a significant effect (P<0.001) on survival in multivariate analyses. Age, sex, and the presence of atrial fibrillation also had significant negative effects (P<0.05) on survival.

International stroke trial cohort

A probable or definite ischaemic stroke was diagnosed in 5139 patients recruited in the UK. We excluded 22 (0.4%) patients from the final analysis as we did not know their dependency status at six months (n=20) or there was an error in the recording of the date of death (n=2). Table 1 shows the baseline characteristics. Among the patients with probable or definite ischaemic stroke, 49% underwent computed tomography before randomisation into the trial; in the remainder the diagnosis was confirmed either by computed tomography after randomisation or by autopsy. Figure 3 shows the survival curves for patients who were

Table 1 | Baseline characteristics and status of patients at six months after stroke onset in three cohorts. Figures are numbers (percentages) of patients unless stated otherwise

	OCSP (n=539)	LSR (n=2054)	IST-1 (n=5117)	All cohorts (n=7710)
Mean (SD) age (years)	73 (12)	68 (13)	73 (11)	72 (12)
Men	269 (50)	1087 (53)	2683 (52)	4039 (52)
Mean (SD) systolic BP (mm Hg)	162 (33)	157 (30)	158 (27)	158 (28)
Atrial fibrillation on baseline ECG	84 (16)*	259 (13)†	1012 (20)‡	1355 (18)
CT performed at baseline	472 (88)	2054 (100)	2499 (49)§	5025 (65)
Visible infarct on baseline CT	263 (56)	1245 (61)	1639 (66)§	3147 (63)
Medication before stroke:				
Antiplatelet	17 (4)	641 (31)	1281 (25)¶	1939 (25)
Anticoagulant	6 (1)	84 (4)	39 (1)**	129 (2)
Stroke syndrome:				
TACI	92 (17)	246 (12)	1437 (28)	1775 (23)
PACI	182 (34)	811 (39)	2072 (40)	3065 (40)
LACI	137 (25)	546 (27)	1042 (20)	1725 (22)
POCI	128 (24)	342 (17)	551 (11)	1021 (13)
Indeterminate subtype	—	109 (5)	15 (0.3)	124 (2)
Status at six months:				
Independent	285 (53)	1142 (56)	1098 (22)	2525 (33)
Dependent	154 (29)	604 (29)	2678 (52)	3436 (45)
Dead	100 (18)	308 (15)	1341 (26)	1749 (23)

OCSP=Oxfordshire community stroke project; LSR=Lothian stroke register; IST-1=first international stroke trial; ECG=electrocardiogram; CT=computed tomography; TACI=total anterior circulation infarct; PACI=partial anterior circulation infarct; LACI=lacunar infarct; POCI=posterior circulation infarct.
*Missing data in 11 patients.
†Not recorded in 288 patients.
‡Not recorded in 348 patients during pilot phase of trial.
§Diagnosis confirmed in remainder by CT after randomisation or by autopsy. "Visible infarct on CT" refers only to those scans performed before randomisation.
¶Not recorded in 348 patients during pilot phase and subsequently recorded only if aspirin was used.
**Not recorded in 153 patients during pilot phase and use subsequently recorded only if heparin was used.

Table 2 | Univariate and multivariate Cox's regression analyses of baseline variables for patients alive at six months after stroke onset in Oxfordshire community stroke project. Figures are hazard ratios (95% confidence intervals)

Variable	Univariate analysis	Multivariate analysis
Age	1.05 (1.04 to 1.07)***	1.04 (1.03 to 1.06)***
Male sex	0.99 (0.80 to 1.24)	1.22 (0.98 to 1.54)
Mean systolic BP	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)
Atrial fibrillation	1.94 (1.43 to 2.65)***	1.85 (1.33 to 2.58)***
Visible infarct on CT	1.19 (0.96 to 1.49)	1.13 (0.89 to 1.43)
Antiplatelet use before stroke	1.08 (0.60 to 1.92)	0.95 (0.52 to 1.73)
Anticoagulant use before stroke	0.62 (0.20 to 1.93)	0.64 (0.19 to 2.14)
Stroke syndrome:		
LACI	1	1
PACI	1.05 (0.80 to 1.38)	0.93 (0.70 to 1.24)
POCI	1.12 (0.84 to 1.51)	1.14 (0.84 to 1.55)
TACI	1.29 (0.84 to 1.97)	0.90 (0.56 to 1.45)
Rankin score (at six months)†		
0	1	1
1	1.25 (0.72 to 2.17)	1.20 (0.69 to 2.08)
2	2.06 (1.19 to 3.55)*	1.54 (0.90 to 2.67)
3	2.69 (1.53 to 4.75)***	2.04 (1.15 to 3.63)*
4	2.78 (1.57 to 4.93)***	1.82 (1.02 to 3.26)*
5	2.12 (0.93 to 4.83)	1.25 (0.54 to 2.90)
Functionally dependent	1.76 (1.40 to 2.20)***	1.38 (1.09 to 1.75)***‡

CT=computed tomography; LACI=lacunar infarct; PACI=partial anterior circulation infarct; POCI=posterior circulation infarct; TACI=total anterior circulation infarct.
*P<0.05 ***P<0.001
†Overall P values of Rankin score (at six months): univariate analysis P<0.001 and multivariate analysis P=0.009
‡Separate multivariate analysis without entering the variable "Rankin score (at six months)."

independent and dependent at six months after randomisation. There was a significant effect (log rank test, P<0.001) of the level of dependency on survival. We entered all baseline variables in table 1 and the functional status at six months after stroke onset in a Cox's regression model (table 4). The level of dependency at six months had a significant effect (P<0.001) on survival in the multivariate analysis. Age, sex, presence of atrial fibrillation on baseline examination, use of aspirin before the stroke, and stroke subtype were also significant (P<0.05).

Pooled estimate of median survival

Table 5 shows estimates of the median survival time, subdivided by Rankin score, based on the combined dataset of the Lothian and Oxfordshire cohorts. There was a significant trend (log rank test P<0.001) of decreasing median survival with increasing Rankin score. Table 5 also gives estimates of median survival for independent and dependent patients based on data from all three cohorts combined. This difference was highly significant (log rank test, P<0.001).

Survival among cohorts recruited in different time periods

We compared survival in all three cohorts among independent and dependent patients who were enrolled during three different time periods (1981-6, 1990-4, and 1995-2000). Estimated median survival for patients who were dependent at six months after stroke onset was 4.2 years among those recruited during 1981-6 and 6.5 years among those recruited during 1990-4. No accurate estimations can be given for the period 1995-2000, as over half of both dependent and independent patients were alive at the end of follow-up.

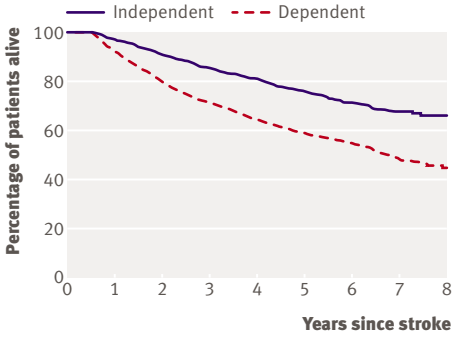
We also analysed the influence of year of recruitment on two year survival in the Lothian and international stroke trial cohorts. We compared the proportions of patients who were alive at two years. Among patients recruited in 1990-4 and 1995-2000 who were independent at the six month assessment the proportions alive at two years were 90% and 93%, respectively. Among those recruited in the same years who were dependent at the six month assessment the proportions alive at two years were 80% and 81%. These differences were not significant.

We also entered the date of stroke onset (or date of randomisation in the international stroke trial cohort) as a variable in the multivariate Cox's regression analyses of each cohort. Date of stroke onset was not a significant variable in the Oxfordshire (P=0.45) and Lothian (P=0.083) cohorts. The date of randomisation was a significant variable (P<0.001) in the international stroke trial cohort. A multivariate Cox's regression analysis in the international stroke trial cohort showed that, among patients recruited in 1995-7, survival was significantly greater than among those recruited in 1991-4 (P<0.001; hazard ratio 0.82, 95% confidence interval 0.73 to 0.91).

DISCUSSION

This study provides robust estimates of the relative and absolute effects that the level of dependency six months after an ischaemic stroke has on subsequent long term survival. The impact of functional status on median survival was substantial and remained significant after adjustment for baseline variables known to influence prognosis. The findings were consistent in size and direction across these three, somewhat different, cohorts of ischaemic stroke patients.

We were surprised to see the poor survival of patients with a Rankin score of 4-5. The five year survival for



No at risk (% surviving)

Independent	1098 (100)	1001 (91)	894 (81)	410 (71)	183 (68)
Dependent	2678 (100)	2148 (80)	1726 (64)	734 (55)	48 (45)

Fig 3 International stroke trial cohort. Long term survival of patients who were alive and dependent or independent from assessment at six months after randomisation

Table 3 | Univariate and multivariate Cox's regression analyses of baseline variables for patients alive at six months after stroke onset in Lothian stroke register. Figures are hazard ratios (95% confidence intervals)

Variable	Univariate analysis	Multivariate analysis
Age	1.06 (1.05 to 1.07)***	1.05 (1.04 to 1.06)***
Male sex	1.05 (0.87 to 1.26)	1.33 (1.08 to 1.64)*
Mean systolic BP	1.00 (0.99 to 1.01)	1.00 (1.00 to 1.00)
Atrial fibrillation	3.00 (2.37 to 3.80)***	1.65 (1.25 to 2.17)*
Visible infarct on CT	1.22 (1.01 to 1.47)*	1.09 (0.89 to 1.35)
Antiplatelet use before stroke	1.48 (1.22 to 1.80)***	1.31 (1.07 to 1.62)*
Anticoagulant use before stroke	1.07 (0.63 to 1.82)	0.91 (0.51 to 1.60)
Stroke syndrome†:		
LACI	1	1
PACI	1.57 (1.24 to 2.00)***	1.32 (1.02 to 1.72)*
POCI	1.01 (0.74 to 1.39)	1.16 (0.83 to 1.62)
TACI	2.25 (1.63 to 3.11)***	1.22 (0.83 to 1.79)
Indeterminate subtype	1.25 (0.77 to 2.04)	1.14 (0.68 to 1.92)
Rankin score (at six months)‡:		
0	1	1
1	1.01 (0.66 to 1.56)	0.98 (0.63 to 1.54)
2	1.66 (1.12 to 2.46)*	1.74 (1.16 to 2.61)*
3	2.86 (1.95 to 4.20)***	2.58 (1.73 to 3.87)***
4	4.11 (2.69 to 6.30)***	3.89 (2.48 to 6.12)***
5	6.41 (4.23 to 9.73)***	4.98 (3.15 to 7.88)***
Functionally dependent	2.87 (2.38 to 3.46)***	2.43 (1.96 to 3.01)***§

CT=computed tomography; LACI=lacunar infarct; PACI=partial anterior circulation infarct; POCI=posterior circulation infarct; TACI=total anterior circulation infarct.
*P<0.05, ***P<0.001
†Overall P values of stroke syndrome: univariate analysis P<0.001 and multivariate analysis P=0.32.
‡Overall P values of Rankin score (at six months): univariate and multivariate analysis P<0.001.
§Separate multivariate analysis without variable "Rankin score (at six months)."

Table 4 | Univariate and multivariate Cox regression analyses of baseline variables for patients alive at six months after randomisation into the first international stroke trial. Figures are hazard ratios (95% confidence intervals)

Variable	Univariate analysis	Multivariate analysis
Age	1.07 (1.06 to 1.07)***	1.07 (1.06 to 1.07)***
Male sex	0.97 (0.88 to 1.07)	1.42 (1.28 to 1.58)***
Mean systolic BP	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)
Atrial fibrillation	1.64 (1.45 to 1.86)***	1.16 (1.02 to 1.32)*
Visible infarct on CT	0.95 (0.85 to 1.06)	1.05 (0.94 to 1.17)
Antiplatelet use before stroke	1.21 (1.08 to 1.35)*	1.17 (1.04 to 1.31)*
Anticoagulant use before stroke	0.71 (0.35 to 1.42)	0.62 (0.31 to 1.24)
Stroke syndrome†:		
LACI	1	1
PACI	1.25 (1.09 to 1.42)*	1.15 (1.01 to 1.31)*
POCI	0.98 (0.81 to 1.19)	1.04 (0.86 to 1.26)
TACI	1.44 (1.24 to 1.66)***	1.23 (1.06 to 1.43)*
Indeterminate subtype	0.28 (0.07 to 1.14)	0.36 (0.09 to 1.43)
Functionally dependent (at six months)	1.91 (1.68 to 2.16)***	1.63 (1.43 to 1.85)***

CT=computed tomography; LACI=lacunar infarct; PACI=partial anterior circulation infarct; POCI=posterior circulation infarct; TACI=total anterior circulation infarct.
*P<0.05, ***P<0.001.
†Overall P values of stroke syndrome: univariate analysis P<0.01 and multivariate analysis P=0.03.

Table 5 | Combined analysis estimating effect of functional status at six months on subsequent median survival

Functional status	Median survival (years) (95% CI)
Rankin score in Oxfordshire and Lothian cohorts (No of patients):	
0 (311)	>15*
1 (540)	11.7 (8.4 to 14.9)
2 (576)	8.4 (7.6 to 9.3)
3 (433)	6.0 (5.2 to 6.8)
4 (189)	3.7 (2.9 to 4.6)
5 (136)	2.5 (1.4 to 3.5)
All three cohorts (No of patients):	
Independent (2525)	9.7 (8.9 to 10.6)
Dependent (3436)	6.0 (5.7 to 6.4)

*Exact median not given as less than half of patients died during follow-up. Median survival 12.9 years (95% CI 10.0 to 15.9) for Rankin 0 and 1 combined.

this group was about 45%, which is worse than for many malignancies. Not surprisingly, median survival was negatively influenced by age at onset of stroke in all cohorts. The presence of atrial fibrillation on the first examination also significantly influenced long term survival in the three cohorts, as shown in previous studies.[12 13]

Strengths and weaknesses

The strength of these data rests on the fact that the cohorts were large and well characterised, the baseline data were generally complete, and follow-up was prospective and prolonged, with minimal loss to six month and prolonged follow-up. The scope for selection bias in the assembly of these cohorts was least for the community based Oxfordshire cohort and greatest for the randomised international stroke trial. We did not include 414 patients with ischaemic stroke from the Lothian study because data were collected at one year (instead of six months), but an analysis that included these 414 patients showed no significant differences in overall survival. Hence, the exclusion of these patients did not have a substantial influence on our findings in this cohort. Our analyses are based on the assumption that patients not reported as dead were alive and that official statistics are accurate. Patients who moved abroad after inclusion in one of the cohorts and died while overseas might not have been recorded if the death certificate was not sent to the UK. We also cannot exclude the further possibility that independent survivors might have been more likely, and able, to emigrate than dependent ones. These effects might have led to an overestimation of median survival in all cohorts, though we think the effect would be small because emigration, especially among elderly people, is relatively uncommon.[14 15]

Relevance of findings

The consistency across the three cohorts of the effect of the patient's level of dependency on subsequent survival suggests that the relative effects are

WHAT IS ALREADY KNOWN ON THIS TOPIC

Several factors influence the outcome of patients with ischaemic stroke and their survival in the early months and years after stroke onset

Little is known on the impact of functional outcome shortly after ischaemic stroke on long term survival

WHAT THIS STUDY ADDS

Functional status of patients six months after onset of an ischaemic stroke has a significant and substantial effect on their long term survival

Less than half those alive with severe disability at six months will survive five years; a survival statistic comparable with that of several malignancies

Our findings have implications for the estimation of the global burden and costs of stroke, for the planning of health care and research, and in clinical practice

generalisable. A graded effect was evident in the three cohorts, even though there were variations in case mix, time period, and location. Also, the demographics of the three studies suggest that the results are generalisable. The mean age of patients in the three cohorts (ranging from 68 to 73 years) was similar to that in large community and hospital based studies of ischaemic stroke patients.[16-22] The proportions of stroke subtypes according to the Oxfordshire community stroke project classification in our cohorts were similar to those found in other studies (though there were fewer total anterior circulation infarcts in the Lothian and Oxfordshire cohorts), as were the outcomes in terms of early case fatality and the proportion of patients who were dead or dependent at six months.[16-25] These cohorts, however, were assembled at a time when secondary prevention in stroke survivors was much less intensive than now. Our analyses of survival during different time periods showed, as one might expect, that survival did indeed slightly improve over time. Hence, when we apply these estimates to current patients, it may be reasonable to assume that on average, at a given level of dependency, median survival would be somewhat better than portrayed here.[26]

We believe that these data have several implications for clinical practice. They can be used to inform patients and their relatives about the prognosis after an ischaemic stroke. They have implications for the estimation of the impact and costs of stroke and for the planning of health care and research. Estimates of global disease burden and costs have so far relied mainly on modelling techniques. Our data could be used to assess the cost effectiveness of treatments for the acute phase of stroke. Previous studies have shown that the costs of long term care account for about half of the total costs of stroke care.[27-29] A health economics model has suggested that treatments that reduce dependency in survivors by only a modest amount might, none the less, have a substantial effect on long term survival free of dependency and hence prove highly cost effective.[30 31] Our data strongly support this hypothesis. Future studies should assess whether early interventions that reduce functional dependency at six months after onset of ischaemic stroke have positive effects on subsequent long term survival, as our study suggests.

We thank the patients who participated in all the studies. We are grateful to Mike McDowall, who has looked after the Lothian and Oxfordshire datasets, and to Sheila Grant, who collated and entered the data on deaths from the three cohorts.

Contributors: PS was the principal investigator for the international stroke trial and played an important role in design, conduct, and analysis of the Oxfordshire community stroke project. PD conceived the idea for assessing the impact of disability on long term survival and established the follow-up mechanism. MD was principal investigator for the Lothian stroke register and coinvestigator for the Oxfordshire project. KBS did the data analysis, drafted the manuscript, and is guarantor. SL supervised all the statistical aspects of the work. EB contributed to the planning of this work and participated in the interpretations of the results. All the authors contributed to design, analysis and commented on drafts of the manuscript.

Funding: The Oxfordshire community stroke project and the international stroke trial were funded by the UK Medical Research Council. Some of the costs of collection of the long term survival data were supported by an unrestricted educational grant from Boehringer Ingelheim.

Competing interests: None declared.

Ethical approval: All studies were approved by relevant local ethical committees.

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

1 Bonita R. Epidemiology of stroke. *Lancet* 1992;339:342-4.
2 Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. *Lancet Neurol* 2007;6:182-7.
3 Feigin VL, Lawes CM, Bennett DA, Anderson CS. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol* 2003;2:43-53.
4 Bamford J, Sandercock P, Dennis M, Warlow C, Jones L, McPherson K, et al. A prospective study of acute cerebrovascular disease in the community: the Oxfordshire community stroke project 1981-86. 1. Methodology, demography and incident cases of first-ever stroke. *J Neurol Neurosurg Psychiatry* 1988;51:1373-80.
5 International Stroke Trial Collaborative Group. International stroke trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569-81.
6 Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet* 1991;337:1521-6.
7 Van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;19:604-7.
8 Lindley RI, Waddel IF, Livingstone M. Can simple questions assess outcome after stroke? *Cerebrovasc Dis* 1994;4:314-24.
9 Wilson JT, Hareendran A, Hendry A, Potter J, Bone I, Muir KW. Reliability of the modified Rankin scale across multiple raters: benefits of a structured interview. *Stroke* 2005;36:777-81.
10 Dennis M, Wellwood I, Warlow C. Are simple questions a valid measure of outcome after stroke? *Cerebrovasc Dis* 1997;7:22-7.
11 Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrica* 1982;69:239-41.
12 Jorgensen HS, Nakayama H, Reith J, Raaschou HO, Olsen TS. Acute stroke with atrial fibrillation. The Copenhagen stroke study. *Stroke* 1996;27:1765-9.
13 Lin HJ, Wolf PA, Kelly-Hayes M, Beiser AS, Kase CS, Benjamin EJ, et al. Stroke severity in atrial fibrillation. The Framingham study. *Stroke* 1996;27:1760-4.
14 Hatton TJ. *Emigration from the UK, 1870-1913 and 1950-1998* . 2003. www.ecocomm.anu.edu.au/people/info/hatton/EmDubTF.pdf.
15 Office for National Statistics. *Implications of the 2001 census report* . www.statistics.gov.uk/census2001/implications.asp.
16 Carolei A, Marini C, Di Napoli M, Di Gianfilippo G, Santalucia P, Baldassare M, et al. High stroke incidence in the prospective community-based L'Aquila registry (1994-1998). First year's results. *Stroke* 1997;28:2500-6.
17 Ellekjaer H, Holmen J, Indredavik B, Terent A. Epidemiology of stroke in Innherred, Norway, 1994 to 1996. Incidence and 30-day case-fatality rate. *Stroke* 1997;28:2180-4.

18 Jamrozik K, Broadhurst RJ, Lai N, Hankey GJ, Burvill PW, Anderson CS. Trends in the incidence, severity, and short-term outcome of stroke in Perth, Western Australia. *Stroke* 1999;30:2105-11.
19 Kolominsky-Rabas PL, Weber M, Gefeller O, Neundoerfer B, Heuschmann PU. Epidemiology of ischemic stroke subtypes according to TOAST criteria: incidence, recurrence, and long-term survival in ischemic stroke subtypes: a population-based study. *Stroke* 2001;32:2735-40.
20 Lauria G, Gentile M, Fassetta G, Casetta I, Agnoli F, Andreotta G, et al. Incidence and prognosis of stroke in the Belluno province, Italy. First-year results of a community-based study. *Stroke* 1995;26:1787-93.
21 Smadja D, Cabre P, May F, Fanon JL, René-Corail P, Riocreux C, et al. ERMANCIA: epidemiology of stroke in Martinique, French West Indies: Part I: methodology, incidence, and 30-day case fatality rate. *Stroke* 2001;32:2741-7.
22 Thrift AG, Dewey HM, Macdonell RA, McNeil JJ, Donnan GA. Incidence of the major stroke subtypes: initial findings from the North East Melbourne stroke incidence study (NEMESIS). *Stroke* 2001;32:1732-8.
23 Johansson B, Norrving B, Lindgren A. Increased stroke incidence in Lund-Orup, Sweden, between 1983 to 1985 and 1993 to 1995. *Stroke* 2000;31:481-6.
24 Dewey HM, Sturm J, Donnan GA, Macdonell RA, McNeil JJ, Thrift AG. Incidence and outcome of subtypes of ischaemic stroke: initial results from the North East Melbourne stroke incidence study (NEMESIS). *Cerebrovasc Dis* 2003;15:133-9.
25 Di Carlo A, Lamassa M, Baldereschi M, Pracucci G, Consoli D, Wolfe CDA, et al. Risk factors and outcome of subtypes of ischemic stroke. Data from a multicenter multinational hospital-based registry. The European community stroke project. *J Neurol Sci* 2006;244:143-50.
26 Carter KN, Anderson CS, Hackett ML, Barber PA, Bonita R. Improved survival after stroke: is admission to hospital the major explanation? Trend analyses of the Auckland regional community stroke studies. *Cerebrovasc Dis* 2007;23:162-8.
27 Spieler JF, Lanoe JL, Amarenco P. Costs of stroke care according to handicap levels and stroke subtypes. *Cerebrovasc Dis* 2004;17:134-42.
28 Gerzeli S, Tarricone R, Zolo P, Colangelo I, Busca MR, Gandolfo C. The economic burden of stroke in Italy. The EcLIPSE study: economic longitudinal incidence-based project for stroke evaluation. *Neurol Sci* 2005;26:72-80.
29 Kolominsky-Rabas PL, Heuschmann PU, Marschall D, Emmert M, Baltzer N, Neundörfer B, et al. Lifetime cost of ischemic stroke in Germany: results and national projections from a population-based stroke registry: the Erlangen stroke project. *Stroke* 2006;37:1179-83.
30 Samsa GP, Matchar DB. Have randomized controlled trials of neuroprotective drugs been underpowered? An illustration of three statistical principles. *Stroke* 2001;32:669-74.
31 Samsa GP, Reutter RA, Parmigiani G, Ancukiewicz M, Abrahamse P, Lipscomb J, et al. Performing cost-effectiveness analysis by integrating randomized trial data with a comprehensive decision model: application to treatment of acute ischemic stroke. *J Clin Epidemiol* 1999;52:259-71.

Accepted: 21 November 2007

[redacted]