Superiority and equivalence in thrombolytic drugs: an interpretation

T. WALLEY, Y. DUNDAR, R. HILL and R. DICKSON

From the Department of Pharmacology and Therapeutics, University of Liverpool, Liverpool, UK

Introduction

Streptokinase and alteplase both reduce mortality from acute myocardial infarction. Newer thrombolytic drugs (reteplase and tenecteplase) have putative advantages, such as ease of delivery and better fibrin specificity, although they are more expensive. A systematic review and meta-analysis identified the clinical efficacy and adverse effects of these drugs.1 But this can only compare drugs where there are direct trial comparisons. Inevitably, we need to make indirect comparisons also, as well as to interpret the results of the direct comparisons. We present an interpretation and consider its strengths and weaknesses.

Equivalence and superiority in clinical trials

Broadly, ‘superiority studies’, where one drug is thought likely to be better than another, assume a null hypothesis that there is no difference, which may then be disproved. Such studies are usually analysed by intention to treat (ITT, i.e. to analyse all patients according to their initial randomized allocation, and not whether they ever received the therapy or changed at some point to the alternative therapy). The null hypothesis and ITT are conservative, and limit the chances of a type 1 error, i.e. ascribing a difference where none exists.2

Equivalence trials examine whether a new therapy is as effective as an existing standard. Equivalence studies might also be used to compare adverse effects, or costs of two effective drugs. In a randomized trial, the two drugs are unlikely to produce exactly the same results (this would be the case even if we compared one drug to itself). We must therefore define ‘equivalence’ a priori, i.e. what is the range of apparent difference in efficacy between two therapies, within which the therapies may be considered clinically equivalent. This is an arbitrary decision.

For thrombolytics, the efficacy is usually recorded as 30- or 35-day mortality. The American College of Chest Physicians (ACCP)3 and others4 have suggested that the range of difference (i.e. the 95% CIs of any difference) should be <1% absolute difference in mortality at 30/35 days.3 This is based on the extent of the difference seen in the GUSTO I study5 between alteplase and streptokinase, where alteplase is generally considered superior to streptokinase, and where this difference is considered clinically valuable.

Others have used different criteria: for instance, the COBALT study,6 examining two methods of dosing with alteplase, used a level of not more than 0.4% difference, based on the lower confidence interval of the difference between drugs in GUSTO I. This was subsequently considered excessively rigorous by many, and Ware and Antman2 suggested a difference of up to 1.5%. Others still have used a difference of not more than 50% relative mortality difference compared to streptokinase (on the basis that streptokinase shows a 2% reduction compared to placebo, this interval equates to roughly 1% difference in absolute mortality).

Address correspondence to Professor T. Walley, Department of Pharmacology and Therapeutics, University of Liverpool, 70 Pembroke Place, Liverpool L69 3GF. e-mail: twalley@liv.ac.uk

© Association of Physicians 2003
The American Food and Drugs Administration (FDA) proposes a boundary based on the relative risk ratio between two drugs, where the upper 95\% CI should not exceed 14.3\% relative difference (also based on the relative difference in GUSTO II). The European Medicines Evaluation Agency (responsible for licensing thrombolytic drugs in Europe) has not yet defined equivalence in this area, but it seems likely to be similar to the ACCP guidance (F. Rotblat, Medicines Control Agency, personal communication, 2002).

This definition of equivalence relates only to 30/35-day mortality. Some argue that equivalence should be based on benefits and on adverse effects such as stroke, but there is no consensus on this, nor on what the limits of equivalence for such an endpoint might be.

**Analysing equivalence trials**

An intention-to-treat analysis might hide true differences between drugs and may therefore be less appropriate in equivalence studies. The more conservative approach in equivalence studies is a ‘per protocol’ analysis (i.e. analysing only those patients who received the treatment as specified in the protocol), as this tends to emphasize differences. An ideal would be to consider both forms of analysis. Differences between the two may arise because of loss of data (e.g. due to patient dropouts) or difficulties handling patients originally assigned to one arm who cross-over to the alternative. Fortunately, these problems do not apply in thrombolysis studies: data ascertainment for 30/35-day mortality is usually almost complete, and the acute nature of the treatment means that there is no crossover. We therefore believe it reasonable to interpret the confidence intervals from superiority studies of thrombolytic therapies as if they had been produced in true equivalence studies. This view is supported by others, and by the one study that reported both analyses, with similar results (see below INJECT).

**Direct comparisons between drugs**

Our meta-analysis has been reported in an accompanying paper. Absolute differences in mortality and the associated confidence intervals have been calculated from the original data for single trials. However, this approach is not valid for the results of the meta-analyses, as it treats all the patient data as if it were derived from a single trial, ignoring differences in entry criteria and patient populations. For these, absolute mortality differences were calculated from the odds ratios and based on the event rate in the ‘control’ population.

**Alteplase and streptokinase**

Three major studies influence this comparison most. The first two, GISSI-2/ISG and ISIS-3, compared the standard alteplase regimen (100 mg over 3 h) and show no clear benefit of alteplase over streptokinase. The third is GUSTO-I, which used a frontloaded or accelerated alteplase infusion (15 mg bolus dose, then 85 mg infusion over 90 min). In one small angiographic study, this regimen improved artery patency at 60 min (76\% accelerated vs. 63\% standard, \( p = 0.03 \)) but not at 90 min. This result was later confirmed in an angiographic substudy of GUSTO-I. On this basis, the accelerated regimen might produce better mortality results than standard alteplase, although this has not been demonstrated in a direct comparison. GUSTO-I had four arms with approximately 10,000 patients in each: (a) streptokinase with subcutaneous heparin; (b) streptokinase with intravenous heparin; (c) accelerated alteplase with intravenous heparin; and (d) standard alteplase with streptokinase.

GUSTO-I showed an absolute decrease in mortality of 1\% (95\% CI 0.37–1.6\%) at 30 days, in favour of accelerated alteplase with heparin over streptokinase (both arms merged). There was a non-significant increase in stroke rate of 0.25\% (95\% CI –0.04\% to 0.59\%). It has been argued that:

(i) The benefit in alteplase over streptokinase was largely seen in patients treated in North America, where there was a 1.2\% absolute reduction in mortality vs. 0.7\% reduction in non-US patients. This may reflect American familiarity with alteplase-based regimens and relative unfamiliarity with streptokinase, particularly since the trial was not blinded.

(ii) There were differences in management of patients between centres, for instance, the larger proportion of alteplase patients who received coronary artery bypass grafting and differences in post-infarct management in US compared to non-US study sites.

(iii) Merging of the two streptokinase regimes, while not merging the two alteplase arms, was invalid.

The GUSTO investigators responded that although not blinded, allocation to treatment was randomized, that there was an intention to treat analysis, and that the recording of the primary endpoint (death) was unlikely to be biased. Among
patients who did not have coronary artery bypass surgery during the hospitalization, those treated with alteplase had 30-day mortality of 6.5% compared to 7.6% in those treated with streptokinase;\(^\text{18}\) still a clinically and statistically significant difference.

Collins and colleagues\(^\text{19}\) argue that too much emphasis has been put on this one trial instead of on all of the evidence comparing alteplase and streptokinase. They argue that accelerated alteplase might achieve earlier patency, but by no more than 30–60 min; given the mean time to treatment in GUSTO-I (approx 165 min) and the time/mortality relationship between thrombolysis and mortality,\(^\text{20,21}\) alteplase could save only 1 to 2 lives per 1000 (0.2% absolute decrease in 30 days mortality) more than streptokinase. They therefore consider the extent of benefit in GUSTO-I implausible and a statistical outlier. This is supported by a Bayesian re-analysis of GUSTO-I.\(^\text{22}\) Collins et al. also argue that the accelerated alteplase regimen was not crucial, since the total dose of alteplase in the first hour of treatment was almost identical in both alteplase arms (82 mg in the accelerated arm vs. 78 mg in the alteplase/streptokinase arm).

They therefore conducted a meta-analysis of GUSTO-I (merging both alteplase arms, and both streptokinase-only arms), GISSI-2 and ISIS-3.\(^\text{19}\) This showed a statistically significant difference of 0.49% (4.9/1000 patients treated) in 30-day non-stroke mortality in favour of alteplase. There was a statistically significant excess risk of total stroke from alteplase over streptokinase:18 still a clinically and statistically significant difference.

Our meta-analysis\(^\text{1}\) used GISSI-2/ISG rather than GISSI-2 alone, as well as other smaller studies, and used the more common approach of comparing the alteplase-only arm of GUSTO-I with the merged streptokinase arms. We showed that including or excluding GUSTO-I made little real difference in efficacy: the first meta-analysis including GUSTO-I indicated no clear mortality benefit for alteplase over streptokinase (difference 0.52% in favour of alteplase, 95%CI –0.34 to 1.28) while a second excluding GUSTO-I showed similar results (difference 0.00%, 95%CI –0.57 to 0.57). There was a significant absolute excess of strokes on alteplase in both (with GUSTO 0.33%, 95%CI 0.15–0.52%; without GUSTO 0.36%, 95%CI 0.16–0.61%).

The role of meta-analysis vs. the single large trial has been discussed.\(^\text{1}\) The question of whether accelerated alteplase is superior to streptokinase is therefore more a matter of judgment in interpreting the evidence than is generally realized. This comparison is also crucial when we come to compare other drugs. (The limits for equivalence are derived from GUSTO-I and therefore cannot be applied to it.)

### Alteplase and reteplase

Two trials, RAPID\(^\text{2,23}\) and GUSTO III\(^\text{24}\) compared accelerated alteplase to reteplase. RAPID 2,\(^\text{23}\) a small angiographic study, showed better coronary arterial patency (TIMI 2/3 flow rates 82% on reteplase vs. 66% on accelerated alteplase at 60 min). This led to a postulation of a 20% reduction in 30-day mortality for reteplase over alteplase.\(^\text{25}\) This was tested in the large GUSTO III study that was planned and powered as a superiority trial. In fact, GUSTO III\(^\text{24}\) showed no significant differences in mortality between the treatments: alteplase reduced absolute 30-day mortality by only 0.23% (NS, 95%CI –0.66% to 1.1%). One year results were similar.\(^\text{26}\) There were no significant differences in stroke rates.

Reteplase’s lack of benefit in mortality despite benefits in reperfusion may mean that the reperfusion results arose by chance, or that the correlation between reperfusion and clinical outcomes is weak (with implications for the interpretation of streptokinase and alteplase in RAAMI\(^\text{13}\) and GUSTO\(^\text{I}\)), or that the main GUSTO-III\(^\text{24}\) results are a type 2 error (i.e. failure to identify a true benefit).

The trial was reported as showing no difference between the drugs, but applying the ACCP definition of equivalence, reteplase is not as effective as alteplase.\(^\text{3}\) Clinicians must decide whether the arbitrary ‘less than 1%’ limits are appropriate, and whether a limit of 1.1% as shown in GUSTO-III instead of ‘less than 1%’ is of clinical, rather than statistical, importance.

### Reteplase and streptokinase

The INJECT study\(^\text{12}\) shows no significant difference between reteplase and streptokinase: reteplase reduced absolute 35-day mortality by 0.51% (NS, 95%CI –0.96% to 1.98%). At the lower extreme therefore, this fits within the definition of equivalence and therefore it may be said that reteplase is no worse than streptokinase.

The ‘per protocol’ analysis of this study confirms this result (absolute difference in mortality of 0.53% vs. ITT analysis of 0.51%). The similarity between the two analyses is unsurprising, since 98.8% of patients actually received randomized treatment. There was no significant difference in stroke rate between drugs.
Alteplase and tenecteplase

ASSENT-2 was designed to show equivalence between tenecteplase or accelerated alteplase. This was confirmed with a mortality difference of 0.028% (95% CI -0.69 to 0.76). There was no difference in stroke rate, but there was a statistically significant difference in the rate of major bleed (5.94% on alteplase vs. 4.66% on tenecteplase, difference 1.28%, 95% CI 0.6–1.96%).

Indirect comparisons

The lack of evidence from head-to-head trials between some thrombolytics forces some indirect comparisons. Conclusions from such indirect comparisons are more tenuous.

Streptokinase vs. tenecteplase

Extrapolating from ASSENT-2 (tenecteplase versus alteplase) and GUSTO I (alteplase versus streptokinase), tenecteplase is either superior to streptokinase (by the same degree as alteplase in GUSTO I), or equally effective.

Reteplase vs. tenecteplase

Extrapolating from GUSTO III (alteplase vs. reteplase) and ASSENT-2 (tenecteplase vs. alteplase) is more fraught. If reteplase is considered equivalent to accelerated alteplase in GUSTO-III, then it is equivalent to tenecteplase. We must remember too that reteplase is equivalent to streptokinase (INJECT study), so our interpretation of GUSTO-I is also important here. If the equivalence definitions are strictly applied to GUSTO III, then reteplase is not equivalent to alteplase or tenecteplase.

Bleeding and adverse events

The major adverse events differ among the drugs and are described under each comparison and in more detail in the accompanying paper. The rates of major bleeding other than intracranial are confusing, as there are substantial differences in the definitions of bleeding and hence the rates of bleeding in different studies, but it seems that the risk of major bleed is slightly higher on streptokinase than on the other drugs in direct comparisons, possibly because of its lower specificity for fibrin.

Discussion

The scientific evidence comparing the comparative effects of thrombolytic drugs on 30/35-day mortality after acute myocardial infarction is open to interpretation, in part depending on the definitions of equivalence chosen. The firm and the more tenuous comparisons that can be drawn are summarized in the box. In the absence of further evidence, the resolution of these is a matter for judgement rather than strict scientific interpretation. If we assume that there is a difference between accelerated alteplase and streptokinase, then we speculate that the resolution of the potentially contradictory results concerning reteplase is that its efficacy lies between that of streptokinase and accelerated alteplase.

The benefits of thrombolysis have to be set against the potential hazards, and it would seem reasonable that equivalence should be defined on the basis of a combined endpoint incorporating major benefit and harm, such as perhaps reduction in deaths at 30 days minus disabling stroke. The current applied limits are in effect a consensus of cardiologists driven by one study, and may be inappropriately rigidly applied.

In practice, most cardiologists in the UK accept the results of GUSTO-I, albeit with some reservations,

Box 1. Summary of comparisons

Direct comparisons lead to the following firm conclusions:

A Streptokinase is as effective as non-accelerated alteplase
B Tenecteplase is as effective as accelerated alteplase
C Reteplase is at least as effective as streptokinase

Depending on interpretation of equivalence and of some major trials, the following conclusions are also possible:

D Concerning streptokinase and alteplase
Streptokinase is as effective as all alteplase, including accelerated alteplase or Streptokinase is inferior to accelerated alteplase
E Concerning reteplase and alteplase
Reteplase is as effective as accelerated alteplase or Reteplase is not (shown to be) as effective as accelerated alteplase

The following indirect comparisons may also be drawn, and depend on the answers to D and E (obviously any conclusions drawn here are tentative):

F Concerning streptokinase and tenecteplase, depending on the interpretation of point D
Tenecteplase is as effective as streptokinase or Tenecteplase is superior to streptokinase
G Concerning reteplase and tenecteplase, depending on the interpretation of point E
Reteplase is as effective as tenecteplase or Reteplase is not as effective as tenecteplase
and would wish to use accelerated alteplase or an equivalent drug. Purchasers have been more cautious, and have often restricted funding for the more expensive alteplase. Most hospitals therefore have a policy that uses streptokinase for definite first thrombolytic treatment (only 50–70% of all myocardial infarctions presenting for thrombolysis in many centres) and alteplase for all subsequent thrombolysis, or that restricts alteplase to other subgroups. Pharmaceutical company data indicate that for the year 2000, streptokinase accounted for 55% of thrombolysis episodes nationally, alteplase 33% and reteplase 11%.

The National Institute for Clinical Excellence (NICE) has issued an appraisal of newer thrombolytic drugs based in part on the data presented here and in the accompanying paper. This does not distinguish between the drugs except for the greater lytic drugs based in part on the data presented here and in the accompanying paper. This does not distinguish between the drugs except for the greater convenience of the bolus drugs in out-of-hospital treatment and would wish to use accelerated alteplase or an equivalent drug. Purchasers have been more cautious, and have often restricted funding for the more expensive alteplase. Most hospitals therefore have a policy that uses streptokinase for definite first thrombolytic treatment (only 50–70% of all myocardial infarctions presenting for thrombolysis in many centres) and alteplase for all subsequent thrombolysis, or that restricts alteplase to other subgroups. Pharmaceutical company data indicate that for the year 2000, streptokinase accounted for 55% of thrombolysis episodes nationally, alteplase 33% and reteplase 11%.

The National Institute for Clinical Excellence (NICE) has issued an appraisal of newer thrombolytic drugs based in part on the data presented here and in the accompanying paper. This does not distinguish between the drugs except for the greater convenience of the bolus drugs in out-of-hospital thrombolysis, and leaves it to clinicians to decide the most appropriate drug balancing, benefit and risk. We present the arguments here to help in these decisions. The comparison between accelerated alteplase and streptokinase and the definitions of what constitute equivalence arising from this are the keys which determine many other comparisons. Conventional definitions of equivalence in this area do not include stroke or bleeding risks, and might usefully do so. The practice of evidence-based medicine requires the existence of good evidence. As we have shown here, even where there are large clinical trials with direct comparisons, the interpretation of the evidence can be difficult and may depend on some arbitrary judgements.

Acknowledgements

The review described in this article is based on research published as a Health Technology Assessment (HTA) Monograph commissioned by the (UK) NHS R&D HTA Programme. The views expressed in this article are those of the authors and not necessarily those of the HTA Programme, National Institute of Clinical Excellence or the Department of Health.

The authors are pleased to acknowledge the contributions of the members of the HTA review team: A. Boland, A. Bagust, A. Haycox, R. Mujica Mota and our Review Advisory Panel. We also thank the anonymous referees for valuable comments.

References


