

Systematic review and metaanalysis of the efficacy of FK506 in experimental stroke

Malcolm R Macleod^{1,2}, Tori O'Collins¹, Laura L Horkey¹, David W Howells^{1,3} and Geoffrey A Donnan^{1,3}

¹National Stroke Research Institute, Melbourne, Australia; ²School of Molecular and Clinical Medicine, University of Edinburgh, Edinburgh, UK; ³Department of Medicine, University of Melbourne, Melbourne, Australia

FK506 is a candidate drug for acute stroke. For such drugs, any decision to proceed to clinical trial should be based on a full and unbiased assessment of the animal data, and consideration should be given to the limitations of those data. Such an assessment should include not only the efficacy of a drug but also the *in vivo* characteristics and limits to that efficacy. Here we use systematic review and meta-analysis to assess the evidence for a protective effect of FK506 in animal models of stroke. In all, 29 studies were identified describing procedures involving 1759 animals. The point estimate for the effect of FK506 was a 31.3% (95% confidence interval 27.2% to 35.4%) improvement in outcome. Efficacy was higher with ketamine anaesthesia and temporary ischaemia and was lower in rats, in animals with comorbidities, and where outcome was measured as infarct size alone. Reported study quality was modest by clinical trial standards, and efficacy was lower in high-quality studies. These findings show a substantial efficacy for FK506 in experimental stroke, but raise concerns that our estimate of effect size might be too high because of factors such as study quality and possible publication bias.

Journal of Cerebral Blood Flow & Metabolism advance online publication, 9 February 2005; doi:10.1038/sj.jcbfm.9600064

Keywords: FK506; meta-analysis; neuroprotection; stroke; systematic review

Introduction

Only tissue plasminogen activator has proven efficacy in human studies (The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group, 1995) despite an experimental literature describing the efficacy of more than 700 drugs in experimental stroke. While this might be due, at least in part, to factors such as the inappropriately long time windows used in most stroke trials, there remain concerns with the interpretation of the available animal evidence supporting efficacy, and Pound *et al* (2004) have argued for a more systematic

evaluation of animal data before proceeding to clinical trials. Important information to be gained from such an approach may include not only the demonstration that a drug can have neuroprotective activity under ideal conditions, but also some idea of any limits to that efficacy, which may impact on clinical usefulness.

The ideal candidate neuroprotective drug is both effective and is sufficiently safe to allow it to be administered in a prehospital setting without extensive patient workup. We have been developing a clinical trial protocol for the prehospital administration of combination neuroprotectants in patients with suspected stroke (the Time Window in Neuroprotection (TWiN) study). To identify candidate drugs, we performed a brief semisystematic literature review, and then scored drugs according to criteria including the evidence for their efficacy, their safety profile and their potential ease of use (our unpublished observations). We are now exploring in more detail the limits to and determinants of efficacy for drugs thus identified, including FK506.

Systematic review and meta-analysis have contributed greatly to the interpretation and aggregation of data in clinical sciences. Systematic review uses a

Correspondence: Dr MR Macleod, Department of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK.

E-mail: malcolm@apoptosis.freeserve.co.uk

This work was supported by an intramural grant from the National Health and Medical Research Council of Australia. MRM acknowledges the support of the Brain and Spine Foundation and the Wellcome Trust.

Possible conflicts of interest: None declared.

Received 28 August 2004; revised 27 September 2004; accepted 30 September 2004

methodical approach to minimise the risk of bias in the selection of studies for inclusion, whereas metaanalysis combines results from individual studies to produce a better estimate of treatment effect (Egger *et al*, 2001). Stratified metaanalysis can then be used to explore the impact of particular study characteristics (Macleod *et al*, 2004, in press).

FK506 was identified in our semisystematic review as a candidate neuroprotective drug for the TWiN study; it has been used in transplant medicine for many years (Murphy *et al*, 2003), and clinical trials in stroke are planned (Labiche and Grotta, 2004). Biological activities of FK506 include inhibition of calcineurin activity (Liu *et al*, 1991) and of neuronal apoptosis (Macleod *et al*, 2001), and efficacy in animal models of renal (Nalesnik *et al*, 1990), hepatic (Suzuki *et al*, 1993) and myocardial (Weinbrenner *et al*, 1998) ischaemia. Here we have investigated the neuroprotective properties of FK506 in experimental stroke using systematic review, metaanalysis and stratified metaanalysis. Specifically, we have calculated a global estimate of the efficacy of FK506, and examined the impact of reported study quality and various study characteristics on the estimate of effect size.

Materials and methods

Studies of FK506 in animal models of stroke were identified from Pubmed (1974 to January 2004), Embase (1980 to January 2004) and BIOSIS (1969 to January 2004): search strategy [*FK506*] AND [*stroke*] OR [*ischaemia*]; hand-searching abstracts of scientific meetings including the Society for Neuroscience and the International Society for Cerebral Blood Flow and Metabolism; reference lists of identified publications; requests to senior authors of identified publications for references to other studies. We included all controlled studies of the effect of FK506 in animal models of focal cerebral ischaemia, where the outcome was measured as infarct size or a neurologic score.

We defined a 'comparison' as the assessment of outcome in treatment and control groups after treatment with an administered dose of drug or with vehicle, with treatment starting a given time before or after the induction of cerebral ischaemia. For each comparison, we extracted data for mean outcome, standard deviation (s.d.) and number of animals per group. Values for data expressed graphically were requested from authors. Where FK506 was administered in multiple doses, the comparison was grouped according to the first dose at the first time it was administered and the dose administered was recorded as the total dose in the first 24 h after ischaemia.

Where neurological tests were performed at different times, only the final test was included. Where one group of animals was scored in more than one neurological domain (for instance, motor and sensory scores), or where both neurological score and infarct size were measured, data were combined using metaanalysis (see below) to give an overall estimate of effect size and its standard

error. We defined effect size as the proportional improvement in outcome (infarct size, neurologic score or combined score) in treated animals relative to untreated ischaemic controls.

The reported methodological quality of individual studies was scored against the following criteria: publication after peer review; statement of control of temperature; random allocation to treatment or control; masked induction of ischaemia; masked assessment of outcome; use of anaesthetic without significant intrinsic neuroprotective activity; appropriate animal model (aged, diabetic or hypertensive); sample size calculation; compliance with animal welfare regulations; statement of potential conflict of interests (Horn *et al*, 2001; Macleod *et al*, 2004). Each study was given a quality score out of a possible total of 10 points, and the group median was calculated.

Data were processed as described previously (Macleod *et al*, 2004). Briefly, for each comparison, the mean outcome for the treatment group and the s.d.'s in treatment and control groups were expressed as a proportion of the outcome in the control group, and the effect size (the difference between the treatment and control groups) and its standard error were calculated. Data were aggregated using a weighted mean difference method with the random effects model of DerSimonian and Laird (1986), a more conservative technique than fixed-effects metaanalysis.

To explore the impact of study characteristics on estimates of effect size, we then performed a stratified metaanalysis with experiments grouped according to reported quality score; use of aged, diabetic or hypertensive experimental animals; anaesthetic used; whether the data had been published in full or in abstract; permanent or temporary ischaemia; method of occlusion; outcome measure; time to outcome measurement; route of drug delivery; single or multiple dosing regime; and species and gender of animal used. The significance of differences between groups was assessed by partitioning heterogeneity and using the χ^2 distribution with $n-1$ degrees of freedom, where n equals the number of groups. To allow for multiple comparisons, we set our significance level at $P < 0.001$.

Results

Electronic searching identified 581 publications meeting the search criteria, of which 39 (23 papers, 15 abstracts) described the effect of FK506 in focal cerebral ischaemia, where the outcome was expressed as a volume of infarction or a neurological score. Hand searching identified a further 10 abstracts. In all, 20 abstracts described work also described in full papers, and since the search was performed one further abstract has been published in full. In response to our request for further relevant information one author provided unpublished data. This meta-analysis is therefore based on data from 24 full papers published between 1994 and 2004 (Sharkey and Butcher, 1994; Sharkey *et al*, 1996; Kuroda and Siesjo, 1996; Butcher *et al*, 1997; Aoyama *et al*, 1997; Takamatsu *et al*, 1998, 2001; Yoshimoto and Siesjo, 1999; Kuroda *et al*, 1999;

Toung *et al*, 1999; Bochelen *et al*, 1999; Miyazawa *et al*, 2000; Aronowski *et al*, 2000; Arii *et al*, 2001; McCarter *et al*, 2001; Ebisu *et al*, 2001; Fredduzzi *et al*, 2001; Maeda *et al*, 2002; Brecht *et al*, 2003; Furuichi *et al*, 2003*a,b*; Nito *et al*, 2004; Chung *et al*, 2004; Shichinohe *et al*, 2004), four abstracts (Sharkey *et al*, 1997; Toung *et al*, 1997; Janelidze *et al*, 1999; McGregor *et al*, 2001) and one personal communication (J Aronowski); it represents the work of 14 independent groups, and nine of 29 publications came from groups funded directly by the Fujisawa Chemical Company, the manufacturers of FK506. Within the 29 studies, 109 comparisons (see Materials and methods for definition) were identified.

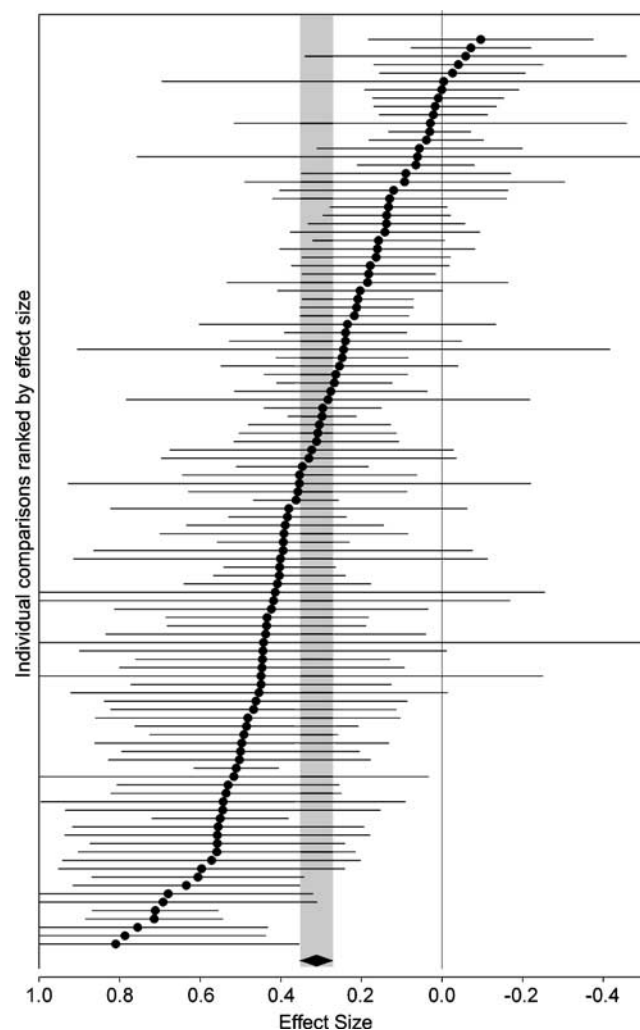


Figure 1 Point estimate and 95% CIs for global estimate and each of 109 comparisons ranked by effect size. Effect size is the improvement in treated animals expressed as a proportion of the outcome in control animals. The diamond indicates the global estimate and its 95% CI, also shown as the grey band. The solid vertical line marks where treatment and control are equal.

The global estimate of the effect of FK506 was 0.313 (95% confidence interval (CI) 0.272–0.354, $P < 10^{-50}$), an improvement in outcome of around 30% (Figure 1). There was significant statistical heterogeneity ($\chi^2 = 361$, $df = 108$, $P < 10^{-28}$) between comparisons. Study characteristics are shown in Table 1. No study described a sample size calculation or contained a statement of potential conflict of interest. Random allocation to treatment group was described in only six studies; induction of ischaemia by an investigator masked to treatment allocation in one study and masked assessment of outcome in two studies. The median reported quality score (see Materials and methods) was 4 (range 0 to 7), and classifying studies by quality score accounted for a significant part of the between-group heterogeneity ($\chi^2 = 113.3$, $df = 7$, $P < 10^{-20}$), with studies of higher quality giving a lower estimate of effect size (Figure 2).

Significant protection was seen for all doses of FK506 above 0.03 mg/kg, with doses above 3 mg/kg giving less protection than intermediate doses (Figure 3A). Significant protection was also seen for all time points, with no apparent interaction

Table 1 Quality characteristics of included studies

Publication	Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Score
Sharkey	1994	✓	✓				✓					3
Kuroda	1996	✓	✓				✓					3
Sharkey	1996	✓	✓	✓			✓			✓		5
Aoyama	1997	✓	✓				✓					3
Butcher	1997	✓	✓				✓					3
Sharkey	1997						✓					1
Toung	1997		✓				✓					2
Takamatsu	1998	✓	✓				✓			✓		4
Bochelen	1999	✓	✓				✓	✓		✓		5
Janelidze	1999						✓					0
Kuroda	1999	✓	✓				✓	✓				4
Toung	1999	✓	✓	✓			✓			✓		5
Yoshimoto	1999	✓	✓				✓					3
Aronowski	2000	✓	✓				✓			✓		4
Miyazawa	2000	✓	✓				✓					2
Arii	2001	✓	✓				✓			✓		4
Ebisu	2001	✓	✓	✓			✓			✓		5
Fredduzzi	2001	✓	✓	✓	✓		✓	✓		✓		7
McCarter	2001	✓	✓				✓					3
McGregor	2001											0
Takamatsu	2001	✓								✓		2
Maeda	2002	✓	✓				✓			✓		4
Brecht	2003	✓	✓				✓			✓		4
Furuichi	2003 <i>a</i>	✓	✓				✓			✓		4
Furuichi	2003 <i>b</i>	✓	✓				✓			✓		4
Aronowski	2004											0
Chung	2004	✓		✓			✓					3
Nito	2004	✓	✓	✓			✓	✓		✓		6
Shichinohe	2004	✓	✓				✓			✓		4

Studies fulfilling the criteria of (1) peer reviewed publication; (2) control of temperature; (3) random allocation to treatment or control; (4) blinded induction of ischaemia; (5) blinded assessment of outcome; (6) use of anaesthetic without significant intrinsic neuroprotective activity; (7) animal model (aged, diabetic or hypertensive); (8) sample size calculation; (9) compliance with animal welfare regulations; and (10) statement of potential conflict of interests.

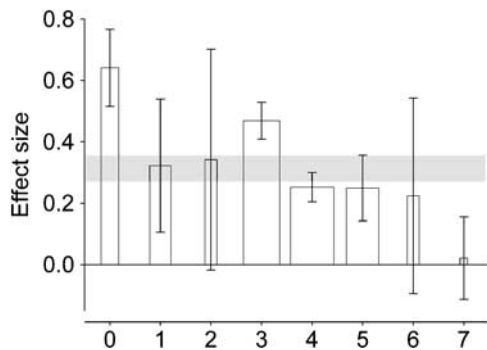


Figure 2 Point estimates and 95% CIs of effect size by reported study quality score. The thickness of each bar reflects the number of comparisons contributing to that data point. The 95% CI for the global estimate is again shown as a grey band. There is a significant relationship between reported quality score and effect size.

between time of administration (up to 5 h post occlusion) and effect size (Figure 3B).

Study design characteristics are shown in Table 2. Effect size was significantly higher in temporary ischaemia models (Figure 4A: $\chi^2 = 52.3$, $df = 3$, $P < 10^{-10}$) and in studies using ketamine anaesthesia (Figure 4B: $\chi^2 = 102.8$, $df = 5$, $P < 10^{-19}$). Experiments using MRI measures of infarct size gave more conservative estimates of effect size than those using histological measures of infarct size, and these in turn were more conservative than those using a combined outcome score or mortality (Figure 4C; $\chi^2 = 86.0$, $df = 3$, $P < 10^{-18}$). Furthermore, the time at which outcome was measured was important; the longer the interval between ischaemia and measurement of outcome, the greater the effect size (Figure 4D; $\chi^2 = 45.6$, $df = 4$, $P < 10^{-8}$).

Experiments on monkeys gave a higher estimate of effect size than those using rodents (Figure 4E; $\chi^2 = 51.4$, $df = 3$, $P < 10^{-10}$), and effect size was higher with healthy animals (Figure 4F; 0.333, 0.299 to 0.368) than where hypertensive or hypoglycaemic animals were used (0.170, 0.079 to 0.261; $\chi^2 = 44.8$, $df = 1$, $P < 10^{-10}$). Studies published in abstract only gave a higher estimate of effect size (Figure 4G; 0.436, 0.309 to 0.563) than those published in full (0.297, 0.263 to 0.332; $\chi^2 = 21.6$, $df = 1$, $P < 10^{-5}$).

There was no increase in effect size in studies using multiple doses of FK506 (0.518, 0.334 to 0.703 versus 0.309, 0.274 to 0.343; $\chi^2 = 4.6$, $df = 1$, $P = 0.03$) or in studies using FK506 alone rather than in combination with other interventions such as tPA, magnesium or hypothermia (0.322, 0.285 to 0.358 versus 0.201, 0.135 to 0.266; $\chi^2 = 6.2$, $df = 1$, $P = 0.01$); while there was a trend for effect size to be higher with endothelin and filament models and lower with photothrombotic and surgical (occlusion by dividing artery or the external application of thread or aneurysm clip) models (Figure 3H; $\chi^2 = 12.8$, $df = 3$, $P = 0.005$), this did not reach our prespecified significance level.

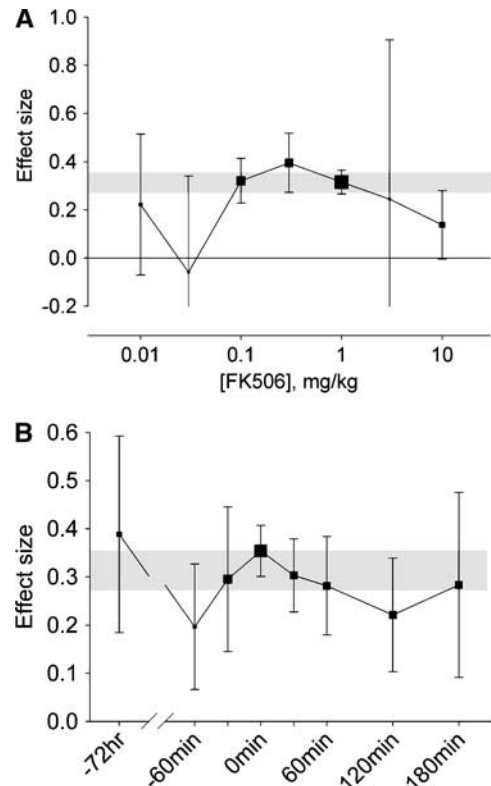


Figure 3 Point estimate of effect size and 95% CIs by (A) FK506 dose and (B) time to treatment. The size of the symbol reflects the number of comparisons contributing to that data point. For details, see Figure 2 legend.

Discussion

Treatment with FK506 lead to a substantial and highly significant improvement in outcome of more than 30%, and improved outcome was seen for all doses of FK506 above 0.03 mg/kg and at each time point studied. Maximum effect size was seen with doses of 0.1 to 0.3 mg/kg administered within the first 30 mins of the onset of ischaemia, and FK506 was effective even when administered 3 days before the onset of ischaemia. The effective dose of FK506 is clinically relevant, being equivalent to that used long term in humans for immunosuppression after renal transplantation (0.2 mg/kg/day) (Murphy *et al*, 2003).

There are a number of potential difficulties with our approach: Firstly, our analysis can only include available data, and if negative studies are less likely to be published then this metaanalysis will overstate effect size. Furthermore, any bias (due perhaps to nonrandomisation or unblinded assessment of outcome) in individual studies will be reflected in the metaanalysis. Secondly, stratified metaanalysis remains a form of subgroup analysis. The stratifications reported here were prespecified, all analyses have been reported, and our stringent significance level accounts for multiple testing; nonetheless, some of our results might have arisen

Table 2 Design characteristics of included studies

Publication	Year	Sex	n (C)	n (Rx)	Dose range (mg/kg)	Species	Time to treatment	Anesthetic	Permanent or focal ischaemia	Route of drug delivery ^a	Outcome measure
Sharkey	1994	Male	39	67	0.03 to 1	Rat	1 to 60 mins	Halothane	Permanent	i.v.	Infarct volume
Kuroda	1996	Male	16	17	1	Rat	180 to 300 mins	Halothane	Temporary	i.v.	Infarct volume
Sharkey	1996	Male	6	6	1	Rat	1	Halothane	Permanent	i.v.	Combined
Aoyama	1997	Male	10	9	0.3	Rat	0	Halothane	Temporary	i.v.	Infarct volume
Butcher	1997	Male	14	22	1 to 10	Rat	-30 to 1 mins	Halothane, Pentobarbitone	Permanent	i.v., i.p.	Infarct volume
Sharkey	1997	Male	44	54	0.1 to 1	Rat	-3 days to 180 mins	Halothane	Permanent	i.v.	Infarct volume
Toung	1997	Male	10	21	0.3 to 1	Rat	0	Halothane	Temporary	i.v.	Infarct volume
Takamatsu	1998	Male	19	17	0.3	Rat	0	Halothane	Permanent, Thrombotic	i.v.	Infarct volume
Bochelen	1999	Male	63	119	0.01 to 5	Rat	-30 to 30 mins	Isoflurane	Permanent, Temporary	i.v.	Infarct volume
Janelidze	1999	NK ^b	10	10	1	Rat	0	NK	Temporary	i.v.	Infarct volume
Kuroda	1999	Male	5	9	1	Rat	-30 mins	Halothane	Temporary	i.p.	Infarct volume
Toung	1999	Male	10	21	0.3 to 1	Rat	-30 mins	Halothane	Temporary	i.v.	Infarct volume
Yoshimoto	1999	Male	18	32	1	Rat	90 to 180 mins	Halothane	Temporary	i.p., i.car.	Infarct volume
Aronowski	2000	Mixed	22	14	1	Mouse	-60 mins	Chloral	Temporary	i.p.	Infarct volume
Miyazawa	2000	Male	11	12	0.5	Rat	-1 day	Halothane	Temporary	Intramural	Infarct area
Arii	2001	Male	22	20	0.3	Rat	30 to 120 mins	Halothane	Temporary	i.v.	Infarct volume
Ebisu	2001	Male	20	18	1	Rat	0	Halothane	Permanent	i.v.	Infarct volume
Fredduzzi	2001	Male	3	8	1	Rat	10 mins	Chloral	Permanent	i.p.	Infarct volume
McCarter	2001	Male	29	25	0.1 to 1	Rat, mouse	-30 to 5 mins	Halothane	Permanent, temporary	i.v., i.p.	Infarct volume
McGregor	2001	NK	10	10	1	Mouse	NK	NK	NK	NK	Mortality
Takamatsu	2001	Male	11	10	0.1	Monkey	5 to 175 mins	Ketamine	Permanent	i.v.	Infarct volume
Maeda	2002	Male	126	153	0.1 to 1	Rat	0 to 180 mins	Halothane	Thrombotic	i.v.	Infarct volume
Brecht	2003	Male	25	25	1	Rat	0	Chloral	Temporary	i.p.	Infarct volume/area
Furuichi	2003a	Male	13	33	0.01 to 0.2	Monkey	0	Ketamine	Thrombotic	i.v.	Combined
Furuichi	2003b	Male	106	246	0.032 to 1	Rat	0 to 180 mins	Halothane	Permanent, temporary	i.v.	Infarct volume
Aronowski	2004	NK	21	7	1	Rat	15 mins	NK	NK	i.v.	Infarct volume
Chung	2004	Male	10	16	0.5	Gerbil	-30 mins	Chloral	Permanent	i.p.	Infarct volume
Nito	2004	Male	20	20	0.3	Rat	120 mins	Halothane	Temporary	i.v.	Combined
Shichinohe	2004	Male	10	15	1 to 3	Mouse	-30 mins	Isoflurane	Permanent	i.p.	Infarct volume

Number of animals in control group (n(C)); number of animals in experimental group (n(Rx)); dose range; number of doses given in the first 24 h; interval from onset of ischemia to start of treatment; anaesthetic used; permanent of focal ischaemia; route of drug delivery; and outcome measure used.

^ai.v., intravenous; i.p., intraperitoneal; i.car., intracarotid.

^bNK, Not known.

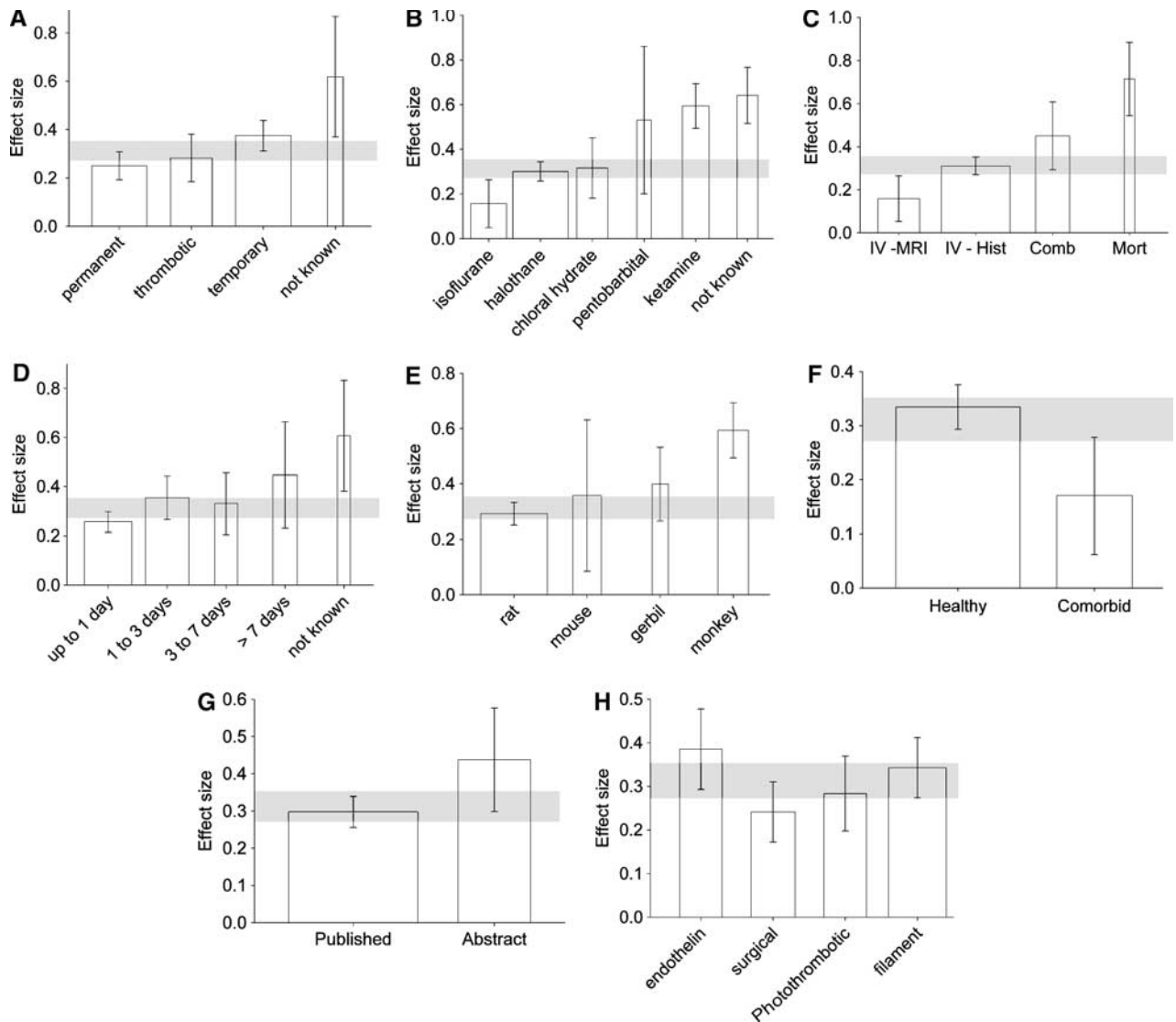


Figure 4 Point estimate of effect size and 95% CIs by (A) duration of occlusion; (B) anaesthetic used; (C) outcome measure; (D) delay to outcome measure; (E) animal species; (F) use of animals with co-morbidity; (G) published or abstract only; and (H) mechanism of occlusion. The thickness of each bar reflects the number of comparisons contributing to that data point. IV-MRI—infarct size assessed by MRI; IV-Hist—infarct size assessed histologically; Comb—combined infarct volume and neurobehavioural end point; Mort—mortality. For other details, see Figure 2 legend.

by chance and our data should be interpreted with caution.

Thirdly, it has been argued that drug efficacy in animals should be assessed against behavioural rather than volumetric end points (Stroke Academic Industry Roundtable, 1999). However, there is little evidence to show that this approach leads to better prediction of efficacy in clinical trial, and in fact in humans lesion volume determined by MRI diffusion-weighted imaging shows correlation both with impairment at 24 h (Tong *et al*, 1998) and with clinical outcome (Wardlaw *et al*, 2002; Engelter *et al*, 2003), although the relationship for lesion volume determined by CT is less clear (NINDS, 2000).

Similarly, we have aggregated data for outcome measured as infarct size and neurobehavioural outcome. Only four of 108 comparisons reporting infarct size measured an area of infarction rather than a volume; converting these areas of infarction to volume (assuming a spherical infarct) increased the overall estimate of effect size from 31.3% (27.2% to 35.4%) to 31.7% (27.5% to 35.8%), and similarly had a negligible effect on the stratified metaanalyses. More difficult is the combination of data from histological and neurobehavioural end points. Importantly, neither of these are themselves 'pure', instead representing an amalgam of either different histologic methods or of different neurobehavioural scores, and all are intended to measure the single

entity of neuroprotective efficacy. Different outcome measures reflect different aspects of efficacy, and our aggregate estimate represents the best estimate of overall efficacy based on all the available evidence.

Finally, there is a conflict between including only studies of high technical quality and including all available studies (to minimise the risk of publication bias described above). As the metaanalysis technique weights each study by its s.d. the model does take some account of technical quality; labs which replicate their results with high fidelity will report low s.d.'s and their data will be given more weight in the metaanalysis.

Our analysis suggested that the use of ketamine anaesthesia leads to an absolute overestimation of effect size of around 30%. This is in keeping with its known actions as a noncompetitive NMDA antagonist (Martin and Lodge, 1985), with reports of synergistic activity with putative neuroprotective drugs (Chang *et al*, 2002), and with our findings for both nicotinamide (Macleod *et al*, 2004) and melatonin (Macleod *et al*, in press).

Effect size was lower in permanent rather than temporary ischaemia models; it might be that FK506 inhibits more effectively those pathophysiological pathways preferentially activated on reperfusion. Effect size was lower in animals with comorbidities, consistent with our findings for nicotinamide. However, most of these data come from the genetically distinct spontaneously hypertensive rat (SHR) strain, and it might be that the reduced effect size is a consequence of genetically determined factors other than blood pressure.

Effect size was higher in monkeys than in rodents, and it might be that this gives a better indication of potential efficacy in human stroke. It is widely held that primate ischaemia models provide a more useful guide to clinical efficacy, but because no neuroprotective agent has unequivocal efficacy in human stroke it is not possible to test this hypothesis.

The global estimate of the efficacy for FK506 is similar to that of nicotinamide (28.7% (Macleod *et al*, 2004)) and lower than that of melatonin (42.8% (Macleod *et al*, in press)). However, the melatonin literature does not include any animals with comorbidity, and because of differences such as these we believe that comparisons of these global estimates are invalid. Where there is a sufficient range of data, it might be possible to use regression modelling to approximate differences in effect size, but a more reliable estimate would require individual head-to-head studies.

True study quality is often higher than the score allocated from details available in the publication, particularly for publications available only in abstract. For instance, the quality score of one paper increased from 1 in abstract (Nito *et al*, 2002) to 6 on full publication (Nito *et al*, 2004). It is clearly difficult to convey experimental detail in the space

available in Abstracts, and even in full publication constraints of space may militate against a complete description of the steps taken to minimise bias.

Reported study quality was similar to that for nicotinamide (Macleod *et al*, 2004) and melatonin (Macleod *et al*, in press), and grouping studies by reported study quality accounted for a degree of the heterogeneity between studies, with high-quality studies reporting smaller effect sizes. Importantly, these reports include some which describe work performed over 10 years ago, and it would be invidious to judge these publications against standards which had not gained wide acceptance until many years later. However, study quality is a potentially crucial issue. Bebertha *et al* (2003) have shown that studies reporting randomisation and/or blinding are less likely to report positive findings than those which do not; the extent to which such deficiencies in study design might have led to the overestimation of the effect of FK506 is not known.

Our quality score is an empirically derived, ordinal scale, which seeks to measure study quality. It does appear to have some validity, because classifying comparisons by quality score accounts for a substantial proportion of the heterogeneity observed in each drug metaanalysis thus examined to date. The importance of individual scale items, and the possible contribution of alternative items, is the subject of ongoing research. In the meantime, we suggest that, where possible, investigators adopt these measures. Although some criteria (e.g. blinding) might be difficult to achieve in small scientific groups, developments in study design (for instance, randomising to treatment group after the induction of ischaemia) may help circumvent some of these problems.

Criteria have been published for reporting clinical trials (Moher *et al*, 2001), and we support the development of similar criteria for reporting experiments in focal cerebral ischaemia, compliance with which could be stated in full publication or abstract at minimal cost in space. We propose that these criteria should be (a) compliance with local legislation or guidelines regarding animal experimentation; (b) prespecified statement of sample size; (c) control of temperature; (d) randomisation to treatment or control; (e) induction of ischaemia and (f) determination of outcome performed masked to treatment allocation; and (g) disclosure of any potential conflict of interest and (h) the avoidance of anaesthetic agents with marked intrinsic neuroprotective activity.

While the overall results of this metaanalysis suggest that FK506 has substantial effect size, the decline in effect size seen with increasing reported study quality coupled with the possible influence of publication or other bias raise concerns that the true efficacy for FK506 might be substantially lower than reported here. Further large studies of high methodological quality (including randomisation to treatment group, and masked induction of ischaemia and

assessment of outcome) are required to give a precise, unbiased assessment of the efficacy of FK506. Based on our observations, to have 80% power to estimate the percentage improvement in outcome after treatment to the nearest 10% would require 115 animals per group. Such sample sizes may seem large, but the use of smaller cohorts represents a false economy and results are likely to be misleading; ultimately more animals would be required.

This report adds to the existing data from systematic review and metaanalysis in the assessment of putative neuroprotective drugs. Using this approach, we now have data for 3092 animals from 208 individual comparisons. With the addition of further data, multiple regression modelling should allow identification of those factors which have greatest impact on estimates of effect size. In turn, this may allow identification of a subset of variables which are sufficient to describe the properties of an administered drug, potentially reducing the number of experiments needed to characterise that drug. Where drugs or groups of drugs depart from the derived model, this may reflect distinct *in vivo* properties of those drugs, and this could provide the basis for a system of drug classification based on *in vivo* rather than *in vitro* or *ex vivo* characteristics.

Acknowledgements

We are grateful to the authors of included studies for their assistance in conducting this metaanalysis, and in particular to John Sharkey, for helpful and insightful discussions.

References

- Aoyama S, Katayama Y, Terashi A (1997) The effect of FK 506, an immunosuppressant, on cerebral infarction volume in focal cerebral ischemia in rats. *Nippon Ika Daigaku Zasshi—J Nippon Med School* 64:416–21
- Arii T, Kamiya T, Arii K, Ueda M, Nito C, Katsura KI, Katayama Y (2001) Neuroprotective effect of immunosuppressant FK506 in transient focal ischemia in rat: therapeutic time window for FK506 in transient focal ischemia. *Neurol Res* 23:755–60
- Aronowski J, Grotta JC, Strong R, Waxham MN (2000) Interplay between the gamma isoform of PKC and calcineurin in regulation of vulnerability to focal cerebral ischemia. *J Cereb Blood Flow Metab* 20:343–9
- Bebarta V, Luyten D, Heard K (2003) Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad Emerg Med* 10:684–7
- Bochelen D, Rudin M, Sauter A (1999) Calcineurin inhibitors FK506 and SDZ ASM 981 alleviate the outcome of focal cerebral ischemic/reperfusion injury. *J Pharmacol Exp Ther* 288:653–9
- Brecht S, Schwarze K, Waetzig V, Christner C, Heiland S, Fischer G, Sartor K, Herdegen T (2003) Changes in peptidyl-prolyl cis/trans isomerase activity and FK506 binding protein expression following neuroprotection by FK506 in the ischemic rat brain. *Neuroscience* 120:1037–48
- Butcher SP, Henshall DC, Teramura Y, Iwasaki K, Sharkey J (1997) Neuroprotective actions of FK506 in experimental stroke: *in vivo* evidence against an antiexcitotoxic mechanism. *J Neurosci* 17:6939–46
- Chang ML, Yang J, Kem S, Klaidman L, Sugawara T, Chan PH, Adams Jr JD (2002) Nicotinamide and ketamine reduce infarct volume and DNA fragmentation in rats after brain ischemia and reperfusion. *Neurosci Lett* 322:137–40
- Chung SY, Lin JY, Lin MC, Liu HM, Wang MF, Cheng FC (2004) Synergistic efficacy of magnesium sulfate and FK506 on cerebral ischemia-induced infarct volume in gerbil. *Med Sci Monit* 10:BR105–8
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7:177–88
- Ebisu T, Katsuta K, Fujikawa A, Aoki I, Umeda M, Naruse S, Tanaka C (2001) Early and delayed neuroprotective effects of FK506 on experimental focal ischemia quantitatively assessed by diffusion-weighted MRI. *Magn Reson Imaging* 19:153–60
- Egger M, Davey Smith G, Altman DG (2001) *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Books
- Engelter ST, Provenzale JM, Petrella JR, DeLong DM, Alberts MJ (2003) Infarct volume on apparent diffusion coefficient maps correlates with length of stay and outcome after middle cerebral artery stroke. *Cerebrovasc Dis* 15:188–91
- Fredduzzi S, Mariucci G, Tantucci M, Del Soldato P, Ambrosini MV (2001) Nitro-aspirin (NCX4016) reduces brain damage induced by focal cerebral ischemia in the rat. *Neurosci Lett* 302:121–4
- Furuichi Y, Katsuta K, Maeda M, Ueyama N, Moriguchi A, Matsuoka N, Goto T, Yanagihara T (2003a) Neuroprotective action of tacrolimus (FK506) in focal and global cerebral ischemia in rodents: dose dependency, therapeutic time window and long-term efficacy. *Brain Res* 965:137–45
- Furuichi Y, Maeda M, Moriguchi A, Sawamoto T, Kawamura A, Matsuoka N, Mutoh S, Yanagihara T (2003b) Tacrolimus, a potential neuroprotective agent, ameliorates ischemic brain damage and neurologic deficits after focal cerebral ischemia in nonhuman primates. *J Cereb Blood Flow Metab* 23:1183–94
- Horn J, de Haan RJ, Vermeulen M, Luiten PG, Limburg M (2001) Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 32:2433–8
- Janelidze S, Yoshimoto T, Kristian T, Siesjo BK (1999) The role of strain differences on the efficacy of PBN (*N*-tert-butyl- α -phenylnitron) and FK506 treatment in transient focal ischemia. *J Cereb Blood Flow Metab* 19(Suppl 1):S171
- Kuroda S, Janelidze S, Siesjo BK (1999) The immunosuppressants cyclosporin A and FK506 equally ameliorate brain damage due to 30-min middle cerebral artery occlusion in hyperglycemic rats. *Brain Res* 835:148–53
- Kuroda S, Siesjo BK (1996) Postischemic administration of FK506 reduces infarct volume following transient focal brain ischemia. *Neurosci Res Commun* 19:83–90
- Labiche LA, Grotta JC (2004) Clinical trials for cytoprotection in stroke. *NeuroRx* 1:46–70
- Liu J, Farmer Jr JD, Lane WS, Friedman J, Weissman I, Schreiber SL (1991) Calcineurin is a common target of

- cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* 66:807–15
- Macleod MR, Allsopp TE, McLuckie J, Kelly JS (2001) Serum withdrawal causes apoptosis in SHSY 5Y cells. *Brain Res* 889:308–15
- Macleod MR, O'Collins T, Horky LL, Howells DW, Donnan GA (2005) Systematic review and meta-analysis of the efficacy of melatonin in experimental stroke. *J Pineal Res* 38:35–41
- Macleod MR, O'Collins T, Howells DW, Donnan GA (2004) Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* 35:1203–8
- Maeda M, Furuichi Y, Ueyama N, Moriguchi A, Satoh N, Matsuoka N, Goto T, Yanagihara T (2002) A combined treatment with tacrolimus (FK506) and recombinant tissue plasminogen activator for thrombotic focal cerebral ischemia in rats: increased neuroprotective efficacy and extended therapeutic time window. *J Cereb Blood Flow Metab* 22:1205–11
- Martin D, Lodge D (1985) Ketamine acts as a non-competitive *N*-methyl-*D*-aspartate antagonist on frog spinal cord *in vitro*. *Neuropharmacology* 24:999–1003
- McCarter JF, McGregor AL, Jones PA, Sharkey J (2001) FK 506 protects brain tissue in animal models of stroke. *Transplant Proc* 33:2390–2
- McGregor AL, Marston HM, Sharkey J (2001) Reduced mortality and improved functional outcome associated with FK506 treatment in a mouse model of MCA occlusion. *J Cereb Blood Flow Metab* 21(Suppl 1):S244
- Miyazawa N, Saji H, Takaishi Y, Nukui H (2000) Protective effect of FK506 in the reperfusion model after short-term occlusion of middle cerebral artery in the rat: assessment by autoradiography using [¹²⁵I]PK-11195. *Neurol Res* 22:630–3
- Moher D, Schulz KF, Altman DG (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 357:1191–4
- Murphy GJ, Waller JR, Sandford RS, Furness PN, Nicholson ML (2003) Randomized clinical trial of the effect of microemulsion cyclosporin and tacrolimus on renal allograft fibrosis. *Br J Surg* 90:680–6
- Nalesnik MA, Lai HS, Murase N, Todo S, Starzl TE (1990) The effect of FK 506 and CyA on the Lewis rat renal ischemia model. *Transplant Proc* 22:87–9
- NINDS (2000) Effect of intravenous recombinant tissue plasminogen activator on ischemic stroke lesion size measured by computed tomography. *Stroke* 31:2912–9
- Nito C, Kamiya T, Ueda M, Amemiya S, Inaba T, Katayama Y (2002) Mild hypothermia enhances neuroprotective effects of immunosuppressant FK506 following transient focal ischemia in rats. *Society for Neuroscience Abstract No* 491.8
- Nito C, Kamiya T, Ueda M, Arai T, Katayama Y (2004) Mild hypothermia enhances the neuroprotective effects of FK506 and expands its therapeutic window following transient focal ischemia in rats. *Brain Res* 1008:179–85
- Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I (2004) Where is the evidence that animal research benefits humans? *BMJ* 328:514–7
- Sharkey J, Butcher SP (1994) Immunophilins mediate the neuroprotective effects of FK506 in focal cerebral ischaemia. *Nature* 371:336–9
- Sharkey J, Crawford JH, Butcher SP, Marston HM (1996) Tacrolimus (FK506) ameliorates skilled motor deficits produced by middle cerebral artery occlusion in rats. *Stroke* 27:2282–6
- Sharkey J, Teramura Y, Iwasaki K, Butcher SP (1997) Brain levels of FK506 and neuroprotection in experimental stroke. *Br J Pharmacol* 122:329
- Shichinohe H, Kuroda S, Abumiya T, Ikeda J, Kobayashi T, Yoshimoto T, Iwasaki Y (2004) FK506 reduces infarct volume due to permanent focal cerebral ischemia by maintaining BAD turnover and inhibiting cytochrome *c* release. *Brain Res* 1001:51–9
- Stroke Academic Industry Roundtable (1999) Recommendations for standards regarding preclinical neuroprotective and restorative drug development. *Stroke* 30:2752–8
- Suzuki S, Toledo-Pereyra LH, Rodriguez FJ, Cejalvo D (1993) Neutrophil infiltration as an important factor in liver ischemia and reperfusion injury. Modulating effects of FK506 and cyclosporine. *Transplantation* 55:1265–72
- Takamatsu H, Kondo K, Ikeda Y, Umemura K (1998) Neuroprotective effects depend on the model of focal ischemia following middle cerebral artery occlusion. *Eur J Pharmacol* 362:137–42
- Takamatsu H, Tsukada H, Noda A, Kakiuchi T, Nishiyama S, Nishimura S, Umemura K (2001) FK506 attenuates early ischemic neuronal death in a monkey model of stroke. *J Nucl Med* 42:1833–40
- The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group (1995) Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 333:1581–8
- Tong DC, Yenari MA, Albers GW, O'Brien M, Marks MP, Moseley ME (1998) Correlation of perfusion- and diffusion-weighted MRI with NIHSS score in acute (<6.5 hour) ischemic stroke. *Neurology* 50:864–70
- Toung TJ, Bhardwaj A, Dawson VL, Dawson TM, Traystman RJ, Hurn PD (1999) Neuroprotective FK506 does not alter *in vivo* nitric oxide production during ischemia and early reperfusion in rats. *Stroke* 30:1279–85
- Toung TJ, Hurn PD, Bhardwaj A, Dawson VL, Dawson TM, Traystman RJ (1997) Immunosuppressive agent FK506 reduces experimental stroke injury. *J Cereb Blood Flow Metab* 17:S617–55
- Wardlaw JM, Keir SL, Bastin ME, Armitage PA, Rana AK (2002) Is diffusion imaging appearance an independent predictor of outcome after ischemic stroke? *Neurology* 59:1381–7
- Weinbrenner C, Liu GS, Downey JM, Cohen MV (1998) Cyclosporine A limits myocardial infarct size even when administered after onset of ischemia. *Cardiovasc Res* 38:678–84
- Yoshimoto T, Siesjo BK (1999) Posttreatment with the immunosuppressant cyclosporin A in transient focal ischemia. *Brain Res* 839:283–91