

Comparison of treatment effects between animal experiments and clinical trials: systematic review

Pablo Perel, Ian Roberts, Emily Sena, Philipa Wheble, Catherine Briscoe, Peter Sandercock, Malcolm Macleod, Luciano E Mignini, Pradeep Jayaram, Khalid S Khan

Abstract

Objective To examine concordance between treatment effects in animal experiments and clinical trials.

Study design Systematic review.

Data sources Medline, Embase, SIGLE, NTIS, Science Citation Index, CAB, BIOSIS.

Study selection Animal studies for interventions with unambiguous evidence of a treatment effect (benefit or harm) in clinical trials: head injury, antifibrinolytics in haemorrhage, thrombolysis in acute ischaemic stroke, tirilazad in acute ischaemic stroke, antenatal corticosteroids to prevent neonatal respiratory distress syndrome, and bisphosphonates to treat osteoporosis.

Review methods Data were extracted on study design, allocation concealment, number of randomised animals, type of model, intervention, and outcome.

Results Corticosteroids did not show any benefit in clinical trials of treatment for head injury but did show a benefit in animal models (pooled odds ratio for adverse functional outcome 0.58, 95% confidence interval 0.41 to 0.83).

Antifibrinolytics reduced bleeding in clinical trials but the data were inconclusive in animal models. Thrombolysis improved outcome in patients with ischaemic stroke. In animal models, tissue plasminogen activator reduced infarct volume by 24% (95% confidence interval 20% to 28%) and improved neurobehavioural scores by 23% (17% to 29%). Tirilazad was associated with a worse outcome in patients with ischaemic stroke. In animal models, tirilazad reduced infarct volume by 29% (21% to 37%) and improved neurobehavioural scores by 48% (29% to 67%). Antenatal corticosteroids reduced respiratory distress and mortality in neonates whereas in animal models respiratory distress was reduced but the effect on mortality was inconclusive (odds ratio 4.2, 95% confidence interval 0.85 to 20.9). Bisphosphonates increased bone mineral density in patients with osteoporosis. In animal models the bisphosphonate alendronate increased bone mineral density compared with placebo by 11.0% (95% confidence interval 9.2% to 12.9%) in the combined results for the hip region. The corresponding treatment effect in the lumbar spine was 8.5% (5.8% to 11.2%) and in the combined results for the forearms (baboons only) was 1.7% (-1.4% to 4.7%).

Conclusions Discordance between animal and human studies may be due to bias or to the failure of animal models to mimic clinical disease adequately.

Introduction

Before clinical trials are carried out, the safety and effectiveness of new drugs are usually tested in animal models.¹ Although the use of animals in medical research is controversial, a poll by the Medical Research Council found that most people support their use provided that there are benefits to human health care, no alternative exists, and no unnecessary suffering occurs.²

The usefulness of animal testing has, however, been questioned.³⁻⁵ Some believe that the results from animal experiments cannot be applied to humans because of the biological differences between the species and because the results of animal experiments often depend on the type of animal model.³ To date the methods used to assess the value of animal trials include historical analyses, critiques of animal models, surveys of clinicians, and citation analyses. In this paper we compared treatment effects from systematic reviews of clinical trials with those of our own systematic review of the corresponding animal experiments.⁶⁻⁸

Methods

We identified six interventions for which there was evidence of a treatment effect (benefit or harm) in systematic reviews of clinical trials and we carried out a systematic review of the corresponding animal experiments. We searched for all published and unpublished controlled studies in animal models for the following interventions: corticosteroids in traumatic head injury,⁹⁻¹¹ antifibrinolytics in haemorrhage,¹² thrombolysis in acute ischaemic stroke,¹³⁻¹⁴ tirilazad in acute ischaemic stroke,¹⁵ antenatal corticosteroids to prevent neonatal respiratory distress syndrome,¹⁶ and bisphosphonates to prevent and treat osteoporosis.¹⁷ We were unaware of the results of the animal studies when selecting the six interventions. We carried out our systematic review in accordance with the recommended methods for health technology assessment, described elsewhere.¹⁸⁻²¹ Briefly, we systematically searched for randomised and non-randomised controlled studies of the six interventions in animal models. To be eligible for inclusion the studies had to report outcomes corresponding to those for which a treatment effect had been shown in clinical trials.

We searched Medline, Embase, SIGLE (System for Information on Grey Literature), NTIS (National Technical Information Service), Science Citation Index, CAB, and BIOSIS. Details of the search strategies are presented elsewhere.¹⁸ Reference lists were



References w1-w228 are on bmj.com

Quality of animal studies. Values are numbers (percentages)

Intervention (No of studies)	Random allocation to group	Adequate allocation concealment	Blinded assessment of outcome
Corticosteroids for traumatic head injury (n=17)	2 (12)	3 (18)	3 (18)
Antifibrinolytic agents (n=8)	3 (38)	0	4 (50)
Thrombolysis for acute ischaemic stroke (n=113)	43 (38)	23 (20)	24 (21)
Tirilazad for acute ischaemic stroke (n=18)	12 (67)	1 (6)	13 (72)
Antenatal corticosteroids (n=56)	14 (25)	0	3 (5)
*Bisphosphonates (n=16)	5 (31)	0	0

*In accordance with definition by Schulz et al.

checked and we contacted the authors of included studies, relevant drug companies, and the authorities that regulate animal testing—the Home Office in the United Kingdom. No language restrictions were applied. To reduce the number of missed studies two reviewers examined the results for potentially relevant interventions.²² These reports were retrieved in full. Two reviewers independently applied the selection criteria.

Eligible reports were assessed for methodological quality.²³ Two reviewers extracted data on allocation concealment and blinding. If the method of allocation concealment was not clearly reported, we tried to contact the authors for clarification. We used Schulz et al's definition for adequate concealment.²⁴ Two reviewers extracted data on study design, allocation concealment, number of randomised animals, type of model, intervention, outcome, and funding source. We contacted authors if relevant outcome measures were not reported but we believed the data to be available. The reviewers were not blinded to the authors or journal.²⁵

For dichotomous measures (for example, mortality) we estimated odds ratios and confidence intervals and for continuous measures (for example, infarct volume) we estimated the effect size:

$$\text{Effect size} = ((\text{outcome}[\text{control}] - \text{outcome}[\text{treated}]) / \text{outcome}[\text{control}]) \times 100\%$$

We calculated pooled odds ratios, effect sizes, and 95% confidence intervals using a random effects model. Heterogeneity was examined using the I^2 statistic.²⁶ We investigated the possibility of small study bias by checking for funnel plot asymmetry and by using graphical and statistical methods.²⁷

Results

Corticosteroids for traumatic head injury

Clinical trials of corticosteroids for traumatic head injury did not show any benefit and showed an increased risk of mortality.⁹ Seventeen reports were found in animal models of traumatic head injury (19 comparisons).^{w1-w17} The quality of the experiments was poor (table 1). Only three reports (four comparisons) reported adequate allocation concealment. Two experiments reported the effect of corticosteroids on mortality. An effect estimate could not be calculated for one of these experiments because the number of animals in each group was not stated and in the other all the animals died.

Seven experiments reported neurological outcomes. Neurological status was assessed by the grip test and neurological severity score. The grip test measures how long (up to a maximum of 30 seconds) a mouse remains on a taut string suspended between two metal bars. Animals were considered "severely disabled" if they held on for less than five seconds. Four experiments reported the grip test: pooled odds ratio 0.58 (95%

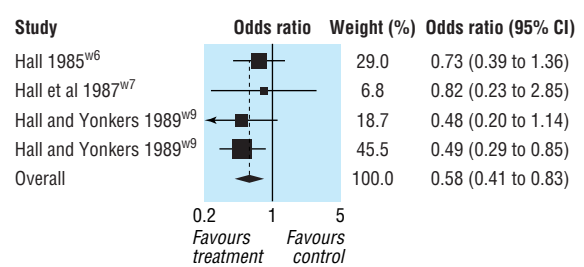


Fig 1 Meta-analysis showing effects of corticosteroids on ability of mice to remain on a taut string (grip test)

confidence interval 0.41 to 0.83; fig 1). No heterogeneity was found ($I^2=0\%$). Neurological severity score was used to assess the clinical condition of rats after trauma and was based on a series of tests. High scores indicated a worse outcome. Three experiments reported the neurological severity score but none reported the scores in each group, stating instead that there was "no significant difference."

Antifibrinolytics in haemorrhage

Clinical trials show that antifibrinolytics reduce blood loss during surgery.¹² Eight reports were found on the effects of antifibrinolytics in animal models (eight comparisons).^{w18-w25} The quality of the experiments was poor (table 1). One did not give the number of animals. One reported mortality, five reported bleeding time, five reported blood loss, and one reported haemoglobin loss. None reported the method of allocation concealment. Four reported blinded outcome assessment but failed to describe the method. One assessed the effect of antifibrinolytics on mortality, but no deaths occurred. Five reported the effects on blood loss but only one had sufficient data to enable the calculation of the effect estimate and confidence intervals. One reported a decrease in gastric bleeding of 0.09 ml (95% confidence interval 0.08 to 0.10). One reported a decrease in blood loss by an average of 523 ml but failed to give the standard deviation. One reported a decrease in blood loss by 0.83 ml but did not give the number of animals in the control group, so the confidence interval could not be calculated.

Thrombolysis in acute ischaemic stroke

Thrombolysis with recombinant tissue plasminogen activator reduces death or dependency after ischaemic stroke, despite an increase in intracranial haemorrhage.^{13 14} Overall, 113 reports were found on the effects of using tissue plasminogen activator or related agents for thrombolysis in animal models of acute ischaemic stroke (369 comparisons).^{w26-w138} The quality of the experiments was poor (table 1). Infarct volume was reported in 212 comparisons (3301 animals), neurobehavioural scores in 84 (1438 animals), and haemorrhage in 146 (2791 animals). The funnel plot suggested an excess of imprecise studies overstating efficacy (fig 2) and this was supported by an Egger regression analysis ($P<0.001$). Tissue plasminogen activator reduced infarct volume by 24% (95% confidence interval 20% to 28%), improved neurobehavioural scores by 23% (17% to 29%), and increased the probability of haemorrhage (odds ratio 1.96, 95% confidence interval 1.63 to 2.35; fig 3). Substantial heterogeneity was found for infarct volume ($I^2=78.2\%$, $P<0.0001$) and for neurobehavioural scores ($I^2=75.2\%$, $P<0.0001$) but not for haemorrhage ($I^2=0\%$).

Tirilazad in acute ischaemic stroke

Tirilazad increases the risk of death and dependency in patients with acute ischaemic stroke.¹⁵ Eighteen reports were found of tirilazad in animal models of acute ischaemic stroke (34

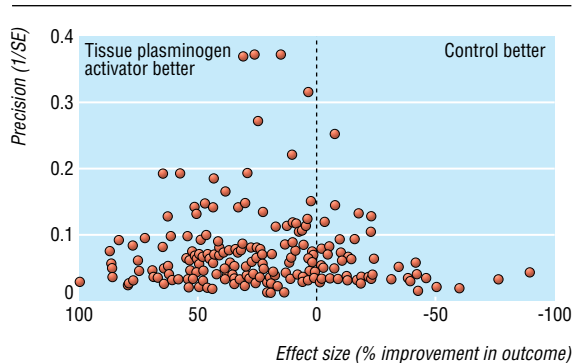


Fig 2 Funnel plot showing effects of thrombolysis using recombinant tissue plasminogen activator or related agents in animal models of acute ischaemic stroke

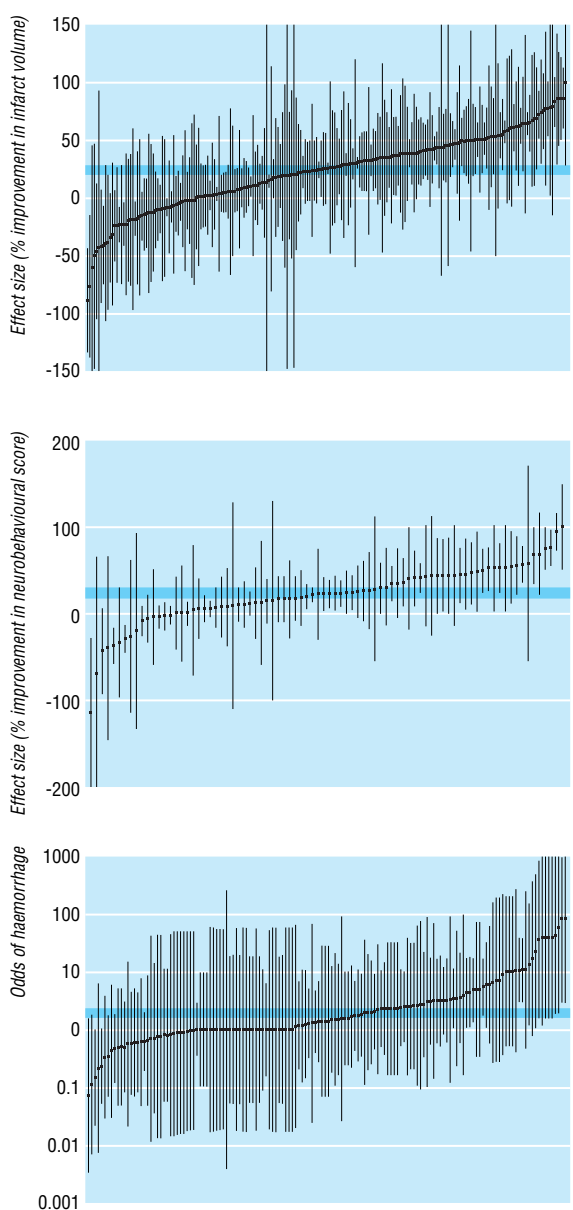


Fig 3 Effects of thrombolysis using recombinant tissue plasminogen activator or related agents in animal models of acute ischaemic stroke. Comparisons are ranked according to effect on infarct volume, neurobehavioural scores, and odds of haemorrhage. Grey bars are 95% confidence limits of global estimate of efficacy. Vertical error bars are 95% confidence limits for individual estimates

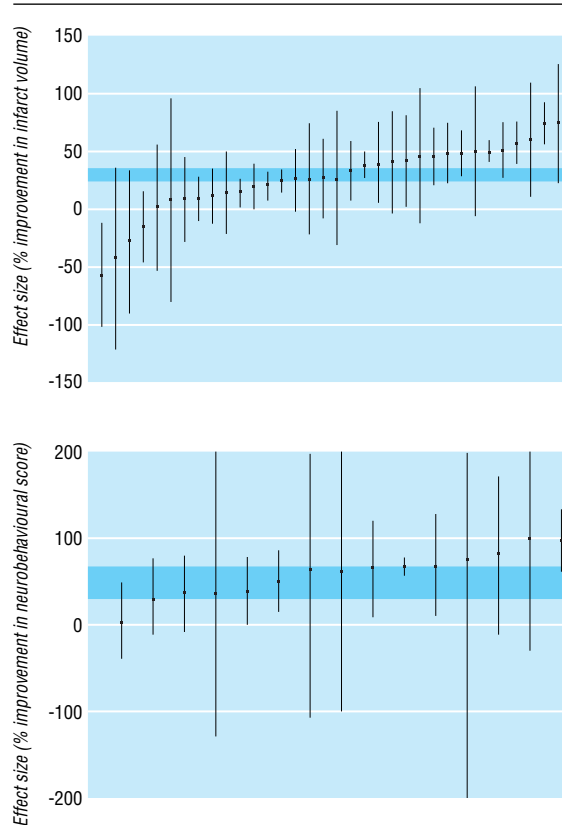


Fig 4 Effect of tirilazad in animal models of acute ischaemic stroke. Comparisons ranked according to effect on infarct volume or neurobehavioural scores. Grey bars are 95% confidence limits of global estimates of efficacy. Vertical error bars are 95% confidence limits for individual estimates

comparisons).^{w139-w156} All 18 reports presented infarct volumes and eight reported neurobehavioural outcomes (fig 4). The quality of the experiments was poor (table 1). The funnel plot suggested a preponderance of small experiments overstating efficacy but this was not supported by an Egger regression analysis. Tirilazad reduced infarct volume by 29% (95% confidence interval 21% to 37%) and improved neurobehavioural scores by 48% (29% to 67%). Substantial heterogeneity was found for both outcome measures (infarct volume $I^2 = 73.3\%$, $P < 0.0001$; neurobehavioural scores $I^2 = 58.1\%$, $P < 0.002$).

Antenatal corticosteroids to prevent neonatal respiratory distress syndrome

Antenatal corticosteroids reduce respiratory distress syndrome and mortality in neonates.¹⁶ In total, 56 reports were found of corticosteroids in animal models of preterm delivery (56 comparisons).^{w157-w212} Thirty two assessed intramuscular or subcutaneous injections in mothers, six assessed intraperitoneal instillation in mothers, and 18 assessed direct injections into the fetus. Only three assessed the effect of maternal corticosteroid injections on respiratory distress syndrome in the neonates (one each in monkeys, cattle, and rabbits). Respiratory distress syndrome was measured differently in the three studies. The quality of the experiments was poor (table 1).

Respiratory distress syndrome was reduced in the corticosteroid groups in all three experiments. In one experiment two of 15 calves in the corticosteroid group compared with nine in the control group developed respiratory distress syndrome ($P = 0.01$). In another experiment the total (SD) lung capacity in newborn rabbits in the corticosteroid group was 1.8 (0.4) ml/g compared with 1.4 (0.4) ml/g in the control group. In a third

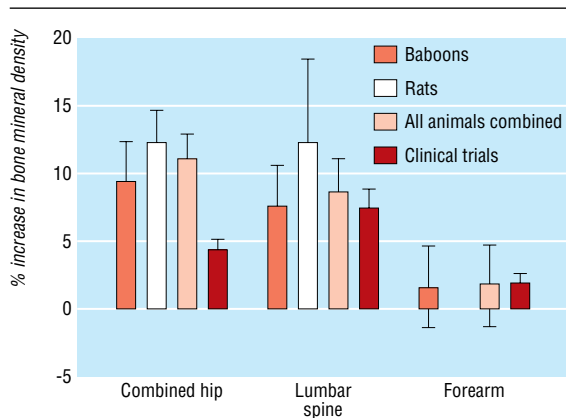


Fig 5 Point estimates and 95% confidence intervals for change in bone mineral density after alendronate administration in baboons, rats, and all animals combined compared with results from clinical trials

experiment, six of 12 monkeys in the corticosteroid treated group compared with 11 in the control group developed severe respiratory distress syndrome ($P=0.03$). Seven experiments reported the effects of corticosteroids on neonatal mortality: pooled odds ratio for mortality 4.2 (95% confidence interval 0.85 to 20.9). Significant heterogeneity was found ($I^2=72.7\%$, $P=0.003$). The pooled odds ratio for mortality in ewe models was 12.5 (1.9 to 79.2) with no evidence of significant heterogeneity ($I^2=33.1\%$, $P=0.22$).

Bisphosphonates to treat osteoporosis

Bisphosphonates increase bone mineral density in postmenopausal women with osteoporosis.¹⁷ Sixteen reports were found of bisphosphonates in animal models (two experiments in baboons and 14 in rats).^{w213-w228} The quality of the experiments was poor (table 1). All experiments were carried out in ovariectomised animals. The effect of bisphosphonates on bone mineral density was reported in 11 experiments (fig 5). When outcome data were available, 11 of 11 (100%) studies showed an increase in bone mineral density and six of six (100%) studies showed improvements in bone mass.

Meta-analysis showed that compared with placebo alendronate increased bone mineral density by 11.0% (95% confidence interval 9.2% to 12.9%) in the combined results for the hip region. The corresponding treatment effect in the lumbar spine was 8.5% (5.8% to 11.2%) and in the combined results for the forearms (baboons only) was 1.7% (-1.4% to 4.7%).

Discussion

Concordance between animal studies and clinical studies varied for six interventions: three had similar outcomes and three did not. Thrombolysis with tissue plasminogen activator was effective in animal models of acute ischaemic stroke and the results agreed with the clinical trials. The animal studies were of poor quality, however, with evidence of publication bias. Our evidence for concordance may therefore be biased. We found over 100 experiments, totalling more than 3000 animals. The pooled result was therefore precise although not necessarily valid. The concordance may be explained by the large volume of evidence and the replication of similar designs in different animals and different laboratories. Furthermore, tissue plasminogen activator was tested in older animals, in those with comorbidities, and at a range of intervals after stroke onset, ensuring a reasonable match with patients in clinical trials. The results for bisphosphonates to treat osteoporosis agreed between animal studies and clinical tri-

als. We also found that antenatal corticosteroids reduced neonatal respiratory distress syndrome in animal studies and in clinical trials, although the data were sparse and we found no evidence of agreement for mortality.

The four experiments in our meta-analysis of corticosteroids in animal head injury models used the weight drop model.²⁸ All had good allocation concealment and blinded outcome assessment. Taken together they showed benefit. The experiments were, however, from one laboratory, had little evidence on adverse effects, and did not examine the influence of comorbidities. We also found a difference in results for tirilazad to treat stroke. The data from the animal studies suggested a benefit but the clinical trials showed no benefit and possible harm. It should be noted that the interval between stroke onset and treatment was longer in the clinical studies (median five hours) than in the animal models (median 10 minutes). Some of the clinical trials recruited patients up to 24 hours after stroke onset. For antifibrinolytics in haemorrhage, clinical trials showed clear evidence of benefit despite the lack of any reliable data from animal models.

Methodological strengths and weaknesses

It would be inappropriate to make general statements about the utility of animal research on the basis of only six interventions. Animal studies are often carried out to learn about biological mechanisms and we cannot comment on the value of animal research in these areas nor provide precise estimates of agreement. Although we tried to contact the authors of individual animal studies, we analysed what was reported and cannot rule out that other relevant data were not published. Our systematic review does, however, provide insights into the limitations of animal models, including the extent to which they represent disease in humans. As the number of systematic reviews of animal experiments increases, a quantitative approach to determine similarities between animal models and clinical trials should be possible in the future.

Implications for research

Systematic reviews could facilitate the translation of research findings from animals to humans. The animal studies in our systematic review varied in methodological quality and sample sizes rather than providing a single definitive high quality experiment for each intervention. Randomisation and blinding were rarely reported, which can have important implications as it has been shown that animal experiments carried out without either are five times more likely to report a positive treatment effect.²³ In the systematic review of thrombolysis in acute ischaemic stroke we found strong evidence of publication bias. The number of experiments in the other systematic reviews made assessment of this source of bias difficult. In most cases we pooled the results to provide precise estimates of efficacy, although given the extent of heterogeneity the precision is open to question. These methodological issues are important given concerns about the differences between promising animal studies and negative clinical trials across a range of interventions. Because animal experiments are part of the evidence used to decide which interventions are taken forward in clinical trials, efforts to avoid bias and random error are as important when reviewing the results of animal models as when reviewing the results of clinical trials.

Prospective registration of animal experiments might reduce publication bias. Although the agencies that regulate animal research hold records of animal studies we were unable to access these. Animal research in the United Kingdom is regulated by the Home Office. We asked the Home Office for details of any animal experiments relevant to our study but they were unable to provide them. In response they stated: "It is not Home Office

policy or practice to gather or retain information derived from work licensed under the Animals (Scientific Procedures) Act 1986 in the way you envisage. Such information is, generally, held by the licensed establishments concerned and made available to Home Office inspectors for inspection on site, should the need arise. I am, therefore, unable to provide you with any of the information you request. Nor am I able to confirm from Home Office records that any relevant trials were conducted under projects licensed under the 1986 act.²⁹ We did not invoke the Freedom of Information Act. Nevertheless, the Home Office response calls into question the usefulness of its records in relation to efforts to create an accessible register of animal experiments.

Research is needed on the aspects of study design that can bias treatment effects in animal models. Empirical evidence of bias from study design characteristics helped to improve the quality of clinical trials and might do the same for animal experiments. Standards for evidence based reporting, similar to the consolidated standards of reporting trials statement for clinical trials, might ensure that relevant aspects of experiment methodology are reported.²⁹

Systematic reviews can provide insights into the limitations of animal models. For example, the animal models for stroke, where there was agreement with the results from clinical trials, seemed more representative of the condition in humans than the animal models for head injury, where there were differences in the results. In stroke, the time from the occlusive event to the start of treatment was similar in animal and human studies. In head injury, treatment was given within five minutes of injury in the animal models but up to eight hours after injury in the clinical trials. None of the animal experiments used models that mimic the complex situations that usually follow traumatic head injury. Comorbidities are clearly relevant in stroke, which occurs in older people with hypertension and diabetes but also in people with head injuries, often accompanied by other injuries and by hypotension and hypothermia. Comorbidities were examined in the stroke models but not in the head injury models.

That there is a gap between clinical research and clinical practice is well established.³⁰ Our work highlights another gap—specifically the lack of communication between those involved in animal research and clinical trialists. Systematic reviews of animal experiments could promote closer collaboration between the research communities and encourage an iterative approach to improving the relevance of animal models to clinical trial design. When models do not represent the clinical context they could be adapted accordingly. Furthermore, as is the case for human research, systematic reviews could help identify and improve deficiencies in the conduct and reporting of animal research.

What is already known on this topic

The relevance of animal models to human health is questioned because of differences between the species

What this study adds

Many studies in animal models are of poor methodological quality

Lack of concordance between animal experiments and clinical trials may be due to bias, random error, or the failure of animal models to adequately represent human disease

We thank Sir Iain Chalmers for his constructive comments on the manuscript.

Contributors: IR, PS, MM, LEM, PJ developed the study protocol. All authors carried out the systematic reviews. PP and IR drafted the manuscript, which was revised on the basis of comments from all authors.

Funding: This work was funded by the National Health Service research and development health technology assessment programme. The views expressed in this publication are those of the authors and not necessarily those of the methodology programme, health technology assessment programme, or the Department of Health. Funding of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Competing interests: IR was an investigator in the corticosteroid randomisation after significant head injury trial. The trial was funded by the UK Medical Research Council. Pharmacia and Upjohn (Pfizer from 2003) provided the Medical Research Council with the methylprednisolone (free of charge) needed for the trial, a grant in aid for preparation of the placebo, and support for collaborators' meetings. PS is co-chief investigator of the third international stroke trial, testing intravenous recombinant tissue plasminogen activator in acute ischaemic stroke; the start-up phase (completed in 2005) of this trial was supported by Boehringer Ingelheim, the manufacturers of tissue plasminogen activator, a donation of drug and placebo for the first 300 patients. The current phase of the trial is supported by the Medical Research Council and the Health Foundation. None of the authors have any relevant competing financial interests.

Ethical approval: Not required. A table showing the quality of animal experiments included in the systematic reviews is available at www.crash2.lshtm.ac.uk.

- Sibbald WJ. An alternative pathway for preclinical research in fluid management. *Crit Care* 2000;4(suppl 2):S8-15.
- MORI. Research study conducted for the Coalition for Medical Progress. The use of animals in medical research. Mar-May 2002. www.mori.com/polls/2002/pdf/cmp.pdf
- Croce P. *Vivisection or science? An investigation into testing drugs and safeguarding health*. London: Zed Books, 1999.
- Sandercock P, Roberts I. Systematic reviews of animal experiments. *Lancet* 2002;360:586.
- Macleod M, Sandercock P. Can systematic reviews help animal experimental work? *RDS News* winter 2005.
- Pound P, Ebrahim S. Supportive evidence is lacking for report on animal studies. *BMJ* 2002;325:1038.
- Roberts I, Kwan I, Evans P, Haig S. Does animal experimentation inform human health care? Observations from a systematic review of international animal experiments on fluid resuscitation. *BMJ* 2002;324:474-6.
- Pound P, Ebrahim S, Sandercock P, Bracken M, Roberts I. Where is the evidence that animal research benefits humans? *BMJ* 2004;328:514-7.
- Alderson P, Roberts I. Corticosteroids for acute traumatic brain injury. *Cochrane Database Syst Rev* 2005;(1):CD000196. DOI:10.1002/14651858.CD000196.pub2.
- The CRASH Trial Collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury—outcomes at 6 months. *Lancet* 2005;365:1957-9.
- The CRASH Trial Collaborators. Effect of intravenous corticosteroids on death within 14 days in 10,008 adults with clinically significant head injury (MRC CRASH Trial): a randomised placebo-controlled trial. *Lancet* 2004;364:1321-8.
- Henry DA, Moxey AJ, Carless PA, O'Connell D, McClelland B, Henderson KM, et al. Anti-fibrinolytic use for minimising perioperative allogeneic blood transfusion. *Cochrane Database Syst Rev* 1999;(4):CD001886. DOI:10.1002/14651858.CD001886.
- Mielke O, Wardlaw J, Liu M. Thrombolysis (different doses, routes of administration and agents) for acute ischaemic stroke. *Cochrane Database Syst Rev* 2004;(4):CD000514. DOI:10.1002/14651858.CD000514.pub2.
- The ATLANTIS, ECASS, and NINDS rt-PA Study Group Investigators. Association of outcome with early stroke treatment: pooled analysis of ATLANTIS, ECASS, and NINDS rt-PA stroke trials. *Lancet* 2004;363:768-74.
- The Tirilazad International Steering Committee. Tirilazad for acute ischaemic stroke. *Cochrane Database Syst Rev* 2001;(4):CD002087. DOI:10.1002/14651858.CD002087.
- Roberts D, Dalziel S. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2006;(3):CD004454. DOI:10.1002/14651858.CD004454.pub2.
- Cranney A, Wells G, Willan A, Griffith L, Zytaruk N, Robinson V, et al. Meta-analyses of therapies for postmenopausal osteoporosis. II. Meta-analysis of alendronate for the treatment of postmenopausal women. *Endocr Rev* 2002;23:508-16.
- Perel P, Roberts I, Sena E, Whibley P, Sandercock P, Mcleod M, et al. Testing treatment on animals: relevance to humans. www.pcpoh.bham.ac.uk/publichealth/nccrm/PDFs%20and%20documents/Publications/JH18_Final_Report_May_2006.pdf.
- Mignini LE, Khan KS. Methodological quality of systematic reviews of animal studies: a survey of reviews of basic methodological research. *BMC Med Res Methodol* 2006 Mar 13;6:10.
- The Cochrane Collaboration. *Cochrane handbook for systematic reviews of interventions*. www.cochrane.org/resources/handbook/
- Piper RD, Cook DJ, Bone RC, Sibbald WJ. Introducing critical appraisal to studies of animal models investigating novel therapies in sepsis. *Crit Care Med* 1996;24:2059-70.
- Edwards P, Clarke M, DiGiuseppi C, Prata P, Roberts I, Wentz R. Identification of randomised controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat Med* 2002;21:1635-40.
- Bebarta B, Luyten D, Heard K. Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad Emerg Med* 2003;10:684-7.

- 24 Schulz K, Chalmers I, Hayes R, Altman D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
- 25 Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D, PEER Investigators. Does masking author identity improve peer review quality? A randomised controlled trial. *JAMA* 1998;280:240-2.
- 26 Higgins J, Thompson S, Deeks J, Altman D. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
- 27 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analyses detected by a simple graphical test. *BMJ* 1997;315:629-34.
- 28 Cernak I. Animal models of head trauma. *NeuroRx* 2005;2:410-22.
- 29 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
- 30 Haines A, Kuruvilla S, Borchert M. Bridging the implementation gap between knowledge and action for health. *Bull World Health Organ* 2004;82:724-31.
(Accepted 7 November 2006)

doi 10.1136/bmj.39048.407928.BE

Crash Trials Coordinating Centre, London School of Hygiene and Tropical

Medicine, London WC1E 7HT

Pablo Perel *research fellow*

Ian Roberts *clinical coordinator CRASH 2 trial*

Clinical Neurosciences, University of Edinburgh

Emily Sena *PhD student*

Philipa Wheble *medical student*

Peter Sandercock *professor of medical neurology*

Malcolm Macleod *senior lecturer*

Catherine Briscoe *medical student*

Centro Rosarino de Estudios Perinatales, WHO Collaborative Centre in Maternal and Child Health, Rosario 2000, Argentina

Luciano E Mignini *researcher*

Division of Reproductive and Child Health, Birmingham Women's Hospital, University of Birmingham

Pradeep Jayaram *senior house officer*

Khalid S Khan *professor of obstetrics-gynaecology*

Correspondence to: P Perel pablo.perel@lsh.tn.ac.uk