

# Emulating multicentre clinical stroke trials: a new paradigm for studying novel interventions in experimental models of stroke

P. M. W. Bath<sup>1,2\*</sup>, M. R. Macleod<sup>3</sup>, and A. R. Green<sup>2</sup>

The recent meta-analysis of NXY-059 in experimental stroke models using individual animal data found the drug to be an effective neuroprotective agent. However, the failure of translation of both this compound and many others from preclinical studies to the clinic indicates that new approaches must be used in drug discovery so that animal models become more reflective of the clinical situation, and studies using animal models of stroke mimic the design of studies performed in humans, as far as possible. In this review, we suggest that a fundamental paradigm shift is needed away from performing preclinical studies in individual laboratories to performing them in an organised group of independent laboratories. Studies should be run by a steering committee and should be supported by a coordinating centre, external data monitoring committee and outcome adjudication committee. This structure will mimic the practice of multicentre clinical trials. By doing so, future studies will minimise potential sources of bias including randomisation, concealment of allocation, blinding of surgery and outcome assessment and ensure publication of all data. It is likely that individual studies will involve increased heterogeneity and therefore will need to be larger. However, regular independent monitoring of data will allow development of interventions to be ceased immediately if neutral or negative data are obtained. The additional costs involved should be seen as reasonable when compared with the resources that would have been

expended in running a clinical trial that subsequently proved negative.

Key words: experimental stroke models, meta-analysis, multicentre preclinical trials, neuroprotection, NXY-059

## Introduction

The failure of translation from preclinical stroke studies to clinical trials has exercised many minds over the last 10 years. Although numerous examples exist, they can be summarised as interventions that were positive in preclinical studies but were either neutral [(e.g. calcium channel blockers, NXY-059 (1, 2)] or actually harmful [DCLHb, selfotel, tirilazad, enlimomab (3–6)] in clinical trials. Only alteplase was positive in both animal and human studies (7–9). Consequently, the relevance of preclinical models of stroke has been questioned and discussed by many (10–15). In an attempt to improve the compound selection process, the STAIR criteria were developed (16) and extended (11) in order to better define the properties that a compound should display in preclinical stroke models before it progressed to clinical trial.

The recent failure of NXY-059 in phase III clinical trials (2) resulted in several publications speculating on the cause of failure. Some assessed possible weaknesses in the experimental study design (17), while others criticised more generally the quality of the methodology (18). Improving the quality of drug trials in experimental stroke models will undoubtedly result in less or non-effective compounds being eliminated before they progress to clinical investigation, a worthy ethical goal in itself because such compounds should not be administered to patients. However, it will not increase the chances of selecting a clinically effective compound because even if, as suggested (19), the efficacy of NXY-059 in preclinical stroke models was overestimated, it is not clear that it would have worked clinically even if the efficacy in animals had been substantially higher.

Correspondence: Philip Bath\*, Division of Stroke Medicine, City Hospital Campus, University of Nottingham, Hucknall Road, Nottingham NG5 1PB, UK. Tel: +44 115 823 1768; Fax: +44 115 823 1767; e-mail: philip.bath@nottingham.ac.uk

<sup>1</sup>Stroke Trials Unit, Division of Stroke Medicine, City Hospital Campus, University of Nottingham, Nottingham, UK

<sup>2</sup>Institute of Neuroscience, School of Biomedical Sciences, Queen's Medical Centre, University of Nottingham, Nottingham, UK

<sup>3</sup>Department of Clinical Neurosciences, University of Edinburgh, Western General Hospital, Edinburgh, UK

Conflict of interest: None.

We recently completed two meta-analyses of the data generated on NXY-059 in experimental stroke models. One examined the study quality using published investigations (19) while the other examined efficacy using individual animal data and included positive and neutral data, both published and unpublished (20). Both analyses concluded that NXY-059 was an effective neuroprotectant in experimental stroke. However, they also highlighted weaknesses in the preclinical development process that have led us to consider new approaches to the problem of translation. Crucially, these results confirm that a paradigm shift in the practice of preclinical studies is required and we now suggest a novel approach for the conduct of preclinical stroke studies, as presented in this communication.

## The current situation

### Haphazard development

The current preclinical development of interventions for stroke is largely haphazard. Novel treatments are usually assessed in small studies within the company or organisation having the original idea (and often owning the patent) in order to generate data on dose, toxicity and potential efficacy. Because these studies are required to justify further investigation, they usually focus on paradigms likely to show efficacy, e.g. starting treatment soon after the onset of transient ischaemia. A second generation of studies follows to extend the 'envelope' and normally includes increasing the time interval between the onset of ischaemia and starting treatment (the window of opportunity), and examining efficacy in several models of ischaemia. Academic laboratories are sometimes brought in to increase the speed with which such information can be gathered. This sharing of the work of development addresses a key STAIR I criterion, namely that information should be gathered from more than one laboratory (16).

The choice of external laboratories is driven by three considerations: those with experience of working with companies in conducting preclinical studies, those the company selects because of their expertise with a particular technique and, occasionally, those that are 'self-chosen' because they request access to the drug. Studies performed in external laboratories are usually performed by junior researchers who need publications and, not surprisingly, experimental protocols will sometimes comprise those that are most likely to lead to a positive result.

The result of this relatively haphazard development schedule is that studies will tend to use treatment paradigms that are most likely to produce positive results. As a result, the circumstances under which efficacy was demonstrated in early company studies (treatment within minutes of ischaemia, transient models of ischaemia, young male animals) are often replicated in the place of circumstances that might more closely mirror human stroke (several hours between stroke onset and treatment, permanent ischaemia, older people of

both sexes). This bias towards 'easier' positive studies will also influence the decision to move the drug from the laboratory to the clinic.

### Quality of studies

Systematic reviews of compounds tested in preclinical stroke have revealed that the reported quality of studies is suboptimal (21). Studies appear to fail in several key methodological areas, especially in lack of randomisation, and failure to blind surgeons and outcome assessors to treatment. The presence of such sources of bias means that studies are more likely to be positive (21). However, assessment of quality often depends on what was reported in the publication rather than what was actually done; some studies of apparent low quality may indeed have used randomisation and blinding but not reported them because the investigators feel these 'basic' requirements are inherent in good investigative techniques and not worthy of comment (22). This problem highlights the need to report preclinical studies adequately, an issue already identified in the reporting of clinical stroke trials (23). For some time, there has been a pressing need for guidelines on the reporting of preclinical studies, much as CONSORT defines how clinical trials should be reported (24), and a consensus statement of good laboratory practice in the modelling of focal cerebral ischaemia has recently been published (25). If several preclinical studies do not meet these standards, are biased and of poor quality but nevertheless positive, it remains likely that their results will influence the overall impression of efficacy in the experimental model. The decision to move from the laboratory to the clinic will then be flawed and the likelihood of subsequent clinical failure would be increased.

### The preclinical compound selection criteria using animal models

While NXY-059 was developed in accordance with many of the criteria underlying good preclinical development (26), some have questioned whether they were fully met (17) while others praised the development programme (27). The meta-analyses of NXY-059 (19, 20) nevertheless revealed weaknesses in the design of studies and highlighted criteria that we suggest must be included in future before a compound is considered for clinical trial.

### Age and sex

The STAIR I criteria (16) do not address animal age requirements and all animals used in the NXY-059 studies were young. However, given the evidence that the efficacy of neuroprotective drugs is reduced in older animals (28, 29) and the fact that most patients are also older leads us to suggest that old animals should be included in at least some preclinical studies. The original STAIR criteria do include a statement that animals of both sexes should be examined and this is emphasised in the

recent updating of the criteria (30). Although marmosets of both sexes were studied with NXY-059, all rats were male (20), and most of the earlier investigations of neuroprotective compounds were similarly flawed. Both sexes must therefore be examined in sufficient numbers in all future investigations to allow a firm statistical evaluation to be undertaken.

### Comorbid conditions

Although many stroke patients have a history of hypertension, NXY-059 was examined only in a small number of hypertensive rats and the overall results in this subgroup were neutral (20); although this finding might indicate a type 2 error, it could equally be indicative of a lack of effect of drug in this condition. No other conditions (such as diabetes) were modelled, even though many patients have such pre-existing conditions. As suggested recently (30), we support the proposal that an adequate number of hypertensive animals must be used in future studies; other comorbid conditions should also probably be included.

### Measurement of drug exposure

While the majority of the pharmacodynamic studies on NXY-059 also examined drug exposure, such information is lacking in many other publications on putative neuroprotective agents. Pharmacokinetic measures must be performed early in order to obtain information on plasma drug concentration and plasma protein-binding as this can vary markedly in different species. Drug exposure in animal studies must be relevant to that expected to be well tolerated in humans. In turn, this requires that early phase I/II clinical studies be conducted on tolerability.

### Other trial design weaknesses

We highlighted other weaknesses in the publications included in the meta-analyses, (19, 20) and so need not repeat them here. However, the paucity of studies performed with NXY-059 in thromboembolic models should be emphasised. Three studies were performed; two using a rabbit model (31, 32) and one using a rat model (33). The rat study measured, in common with the MCA occlusion models, infarct volume (33), making the results amenable to inclusion in the individual animal meta-analysis (20). However, the rabbit studies determined the dose of blood clot required to cause a particular lesion in the presence or absence of NXY-059 (31, 32), and this very different methodological approach meant that data were not in a form suitable for inclusion in this meta-analysis. Consequently, while all three studies suggested the efficacy of NXY-059 in a thromboembolic model, only the rodent data could be included and analysed. It is clear therefore that limited information was obtained on the efficacy of NXY-059 in thromboembolic models. As recently emphasised by Lo (34), stroke is a vascular disorder with a neurological outcome, and

so mimicking this and producing robust data (including dose and time window information) should be mandatory even though long-term occlusion is not possible in this model. It is also possible that both age and comorbid conditions alter the structure and function of the blood–brain barrier; this again emphasises the need to examine appropriate animals in the thromboembolic model.

## A future paradigm for preclinical stroke research

### The conduct of the preclinical trials

The disconnection in acute stroke between the preclinical and the clinical findings for every novel intervention other than thrombolysis means that a new paradigm for laboratory development is urgently needed. The key difference between preclinical and clinical studies is the degree in coordination by which research subjects (patients and animals) are recruited, and interventions are tested. Animal studies may be loosely coordinated by the developing organisation, but never to the degree seen in clinical trials (Table 1). The central proposal of this paper is that preclinical studies should be run in the same manner as randomised clinical trials (see Table 2). This does not preclude individual investigators examining compounds in any manner they wish. However, the results of their studies should be given little weight in the portfolio of information required when a novel compound is being selected for progression to a clinical trial.

To achieve results as robust as those from well-conducted clinical trials, preclinical studies need to be coordinated so that assessment of the intervention is performed in a number of laboratories in a range of animals (species, age, sex, comorbidity) with a range of experimental conditions (model, dose, time to administration) and outcomes (lesion volume, motor impairment, cognitive function). Studies should be run by a steering committee and supported by a coordinating centre, external data monitoring committee (DMC) and outcome adjudication committee (Fig. 1).

## Study management

### Steering committee

A key feature of previous and current developments has been the semi-random development of new interventions. To oversee future developments, a steering committee is needed, this representing the funding institution, patent holder, collaborating laboratories and the coordinating centre. For commercial organisations, the steering committee would comprise members of the company as well as selected experts from academia (as done in clinical trials), the latter to provide some external advice and balance, and represent potential collaborating external laboratories. The steering committee would coordinate the design and execution of studies during preclinical

**Table 1** Comparison of preclinical experimental stroke studies and clinical stroke trials

	Animal/preclinical studies	Clinical (phase IIb/III) trials
Centres	Single or few centres	Multicentre
Staff	Industry and academic	Mostly clinical, some academic
Dose response studies	Variable	Common
Time response studies	Variable – most studies assess early administration	Uncommon
Sample size calculation*	Uncommon	Common (37)
Randomisation*	Variable	Almost all
Allocation concealment*	Variable	Common
Placebo control*	Variable	Common
Blinded surgery*	Variable	Not relevant
Blinded outcome assessment*	Uncommon (outcome often measured by surgeon)	Common
Functional outcome assessment	Uncommon	Common
Death (postrandomisation)	Animals discarded so death not counted	Counted within functional outcome scores (mRS = 6) and as a serious adverse event
Plasma drug analysis	Uncommon	Common
Publication bias	Common (18)	Variable
Publication quality	Poor	Moderate (23)
Study registration	Nil	Common (e.g. through controlled trials ISRCTN)
Data sharing	Limited [e.g. (20)]	Moderately common [e.g. VISTA, OAST (48, 49)]
Systematic reviews based on summary (group) data	Uncommon (except by the applicants)	Very common (see Cochrane Library)
Systematic reviews based on individual subject data	Only 1 to date (20)	Moderately common (e.g. (50, 51))

\*These procedures may be done but not reported. mRS, modified Rankin scale.

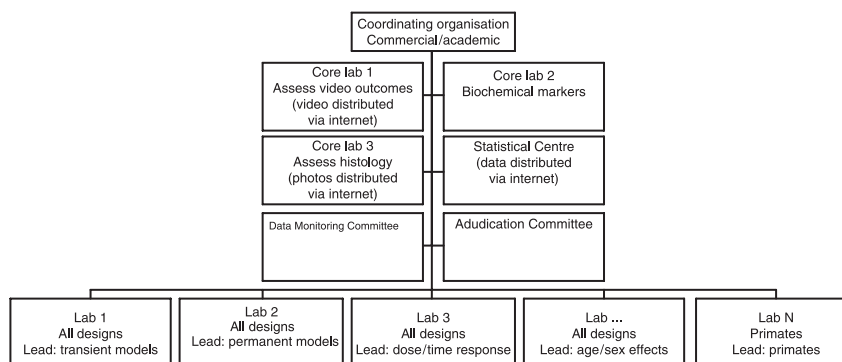
**Table 2** Proposed design for preclinical stroke studies

Logistical components	Experimental design	Data
Multiple laboratories	Male and female animals	Internal data sharing between laboratories, coordinating centre and core labs
Coordinating centre	Young and old animals	External data sharing for systematic reviews based on summary (group) data
Steering committee	Dose response studies	External data sharing for systematic reviews based on individual animals
Data monitoring committee	Time response studies	
Adjudication	Multiple species	
Study registration	Sample size calculation	
	Randomization	
	Allocation concealment	
	Placebo control	
	Blinded surgery	
	Blinded outcome assessment	
	Assessment of lesion volume, motor impairment, and functional outcome	
	Include animals who die postrandomisation in analyses	
	Publication of all studies	
	Publications must detail all key design components	

These build on the STAIR criteria (16, 30) emphasising the need for coordinated collaborative working between a group of laboratories.

development and oversee the design of the whole programme of development as well as individual studies; conduct of studies, data analysis and interpretation; and publication. It would ensure that the necessary ranges of animal character-

istics and experimental conditions are included across the development cycle, this comprising a set of individual studies running sequentially and, in some cases, in parallel. Such coordination would prevent laboratories undertaking easier



**Fig. 1** Design of future multicentre preclinical studies.

studies and thereby prevent disproportionate numbers of animals from being exposed to conditions that do not mimic those present in stroke patients.

### Coordinating centre

The coordinating centre will coordinate the studies (much as a trial coordinating centre runs clinical trials), including (i) identification and site assessment of suitable laboratories with expertise in stroke models; (ii) performing start-up, in-study monitoring and close down meetings to ensure that laboratories understand the design, and systems for randomisation, treatment, outcome assessment and data collection; (iii) collation and checking of data and its quality; (iv) coordination of interim analyses for submission to the DMC; and (v) coordination of the final analyses and distribution of data to participating laboratories, possibly working with a separate bio-statistics centre. Coordinating centres can be based at commercial or academic sites depending on the funder, and would report to the steering committee.

### DMC

It is vital to have external oversight of the experimental programme to provide a slightly less interested perspective on accumulating data and the prospects of success. Just as clinical trials have an external DMC to protect the safety of patients, and to prevent trials from continuing longer than necessary, preclinical developments should be assessed periodically by an external group, perhaps comprising members of the sponsoring institution as well as independent experts. The remit of the DMC would be periodically to review unblinded data from each study as well as integrating this with all known existing data (whether performed by the sponsoring institution or not), using meta-analytical techniques (35). The DMC would recommend to the steering committee on whether further developments are warranted and whether additional data are required, or whether to stop development if it became apparent that the intervention was unlikely to be successful preclinically (futility). Ultimately, it would be for the sponsor

(company, government, charity) to determine whether the programme should continue.

### Adjudication committee

Outcomes such as lesion size, motor impairment, cognitive function and cause of death should be assessed centrally (rather than at the laboratory that performed the experiment) by an adjudication committee who would be blinded to treatment group allocation. This would entail moving histological samples, photographs, MR images and/or video records from the laboratory to the coordinating Centre or another laboratory blinded to the experimental details. Once again, this replicates practice in many clinical trials.

### Data analysis

At present, the vast majority of laboratories analyse their own data. This process is potentially flawed, not least because some of the most powerful statistical analyses are often complex and beyond the expertise of most scientists with only moderate statistical training and experience. In future, analyses should be performed by a core facility, ideally one independent of the funder. Analyses and their reporting should be full and frank.

### Individual studies

The coordinating centre would manage the delivery of individual studies. To replicate the heterogeneity of patients in clinical trials and prevent laboratories from focusing on 'easy' experiments, the coordinating centre might determine which individual experiment is to be performed next at each participating laboratory. Explicit information would be given. For example: choose a female rat of a given weight; perform a transient model of ischaemia; administer drug/vehicle from 'vial 349' from 4 h after vessel occlusion for the next 48 h; video motor function at 48 h and behavioural testing at 2 weeks; and then kill the animal and process tissue for protein expression and lesion size. These parameters could be determined in advance by randomisation (in blocks to prevent the next experiment from being predictable) or could be determined

in the light of accruing data, perhaps based on Bayesian approaches, as used for dosing in a trial of neutrophil-inhibitory factor (36).

## Funding

Commercial developments will need to be funded by the relevant company (as now) with support for laboratory studies and the activities of the various committees and adjudicators. In contrast, academic studies will need to be supported through project and/or programme grants from government or charity funders (again as now). The magnitude of funding needed may initially be higher, partly reflecting a rise in the number of animals and complexity of the programme of work, but review and early cessation of ineffective compounds will probably offset any increase. However, the huge payback will be in preventing the need and cost of clinical trials investigating neutral interventions. Thus, the total cost of developing an effective clinical treatment is likely to decline.

## Study design

### Sample size

A key challenge in clinical trial design is determining how large trials need to be. Simple statistical formulae can calculate this for binary outcomes (e.g. dead or alive) based on how much difference the drug is expected to make, the event rate in the control group, the significance level required ( $P$ , or  $\alpha$ ) and the desired study power ( $1 - \beta$ ). Related methods are available for calculating the size of studies where the primary outcome variable is continuous or ordinal in nature. Unfortunately, many acute stroke trials have been too small to reliably detect the sort of treatment effects that might be expected (37), one probable cause for past failures. The same appears to be true in experimental stroke studies, as found in systematic reviews of several agents (21). It is vital to plan sample size correctly; studies that are too small will be underpowered and are likely to miss an active treatment effect (false neutral), thereby risking losing a potentially useful intervention. Conversely, preclinical studies that are too large (albeit probably an unusual circumstance at present) are unethical and extravagant in time and financial cost. It is likely that future animal studies will each have to be even larger as the emphasis moves from 'easy' protocols to a balanced portfolio of experimental conditions, where effect sizes are likely to be smaller. Specifically, increasing the heterogeneity of animals (species, age, sex, comorbidities) and experimental conditions will reduce effect size and therefore inflate the number of animals needed. As a result, studies of 10–20 animals will disappear and be replaced by ones two to 10 times in size. This does not mean that the overall number of animals being studied in stroke will necessarily increase as stopping the development of ineffective interventions immediately data show futility and preventing duplica-

tion of 'easy' protocols will offset size increase in individual studies.

## Randomisation

It is unclear how many experimental studies involve randomisation because its lack of mention in publications does not necessarily mean that it was not performed. Randomisation reduces assignment bias (38) and should not be confused with 'picking the nearest animal in the cage', which is inadequate as a means of randomisation (18, 25). Randomisation requires generating a list of treatment codes before the study. Alternatively, randomisation can be performed in real time over a secure internet site, as used in current acute stroke trials (39). This approach allows for more sophisticated forms of randomisation to be used, e.g. based on adaptive minimisation (40), which can reduce potential imbalances (i.e. ensure that baseline impairment is similar across the treatment groups) and slightly increase the statistical power.

In a further extension, allocation of animals to differing doses can be performed adaptively over a real-time internet link using a Bayesian sequential design, as done in a trial of neutrophil-inhibitory factor (36). In this strategy, outcomes are continuously monitored and fed back to the computer randomisation system so that it can then focus most on experiments at efficacious doses. The approach could be extended to assess time response or a combination of dose and time response.

## Concealment of allocation

Randomisation is pointless if researchers are aware of what the next subject (human or animal) will receive, that is, allocation is not concealed, which leads to selection bias and can overwhelm any treatment effect in magnitude (41, 42). Similarly, allocation should be concealed once animals have been enrolled to prevent performance bias. Inadequate allocation concealment appears to be significant in preclinical studies (21). Real-time computer-based (including internet-based) randomisation can circumnavigate such prior knowledge and thereby ensure concealment of allocation.

## Placebo control

All drug studies should be placebo controlled to reduce assignment bias. The coordinating centre would ensure that adequate samples of intervention and placebo were dispensed to participating laboratories in a timely and coordinated manner, thereby mirroring the current status in most commercial (and some academic) clinical trials.

Data for drug exposure, in the form of plasma unbound drug levels, should be routine and data should be collected in such a way as to allow results to be matched to individual animals to allow correlation with outcome, as can be done in clinical investigations. This requires integration with the company

sponsoring the study because they alone are likely to have the facilities for plasma drug analysis.

### Blinded surgery

Indirect data from existing studies suggest that surgeons may not always be blinded when performing middle cerebral artery occlusion (or any other sort of stroke). The use of placebo-control and real-time randomisation (i.e. randomisation performed once MCA occlusion has been initiated) would reduce this form of bias. Ensuring that surgeons are blinded to treatment requires another staff member to prepare and administer the intervention. It is important that a different person is used because many interventions can be unblinded due to their physicochemical properties or effects on physiological variables (such as blood pressure and temperature).

### Blinded outcome assessment

One of the most potent sources of bias in studies is unblinded outcome assessment leading to measurement bias; in this situation, the observer knows what treatment was given when they are assessing outcomes. Clinical trials often use blinded observers who are not involved with the care of patients; in many cases, these assessors are based centrally with assessments performed by telephone or post. Once again, it appears that preclinical studies have not, sufficiently, utilised blinded assessment (21). A simple solution within a laboratory is for one researcher to perform surgery in one study and assess outcomes in a separate study, while a second researcher does the opposite.

However, the use of blinded observers at every site may add noise and imprecision and ongoing work is developing the use of central adjudication of videoed functional assessment in clinical trials because inter-observer variation is significant even for simple outcome scales. In preclinical studies, photographs of stained slices and videos of behavioural studies could be submitted over the internet for central or distributed adjudication (Fig. 1); in the latter case, laboratories would assess images for each other with blinding to the laboratory and experimental conditions (and even the study). Using a clinical example, digital neuroimaging data are submitted and curated on a central computer server and then sent over the internet for distributed adjudication, as done in the MRC NeuroGrid project (<http://www.neurogrid.ac.uk/>), which supports the ENOS and IST-3 acute stroke trials (43).

It is vital when analysing outcomes to include information on animals that died or that were killed because of poor health rather than discarding such information. The analysis of data for an intervention that reduces death as well as impairment and disability will be more powerful with inclusion of death. Conversely, exclusion of death from analysis of data for an intervention that reduces impairment/disability but increases death, for example thrombolysis (7), will overestimate benefit. This approach is used in clinical trials with outcome assessed using the modified Rankin scale. Additionally, excluding death

from analyses will risk attribution bias. Unfortunately, few animal studies incorporate death into analyses. Information on any animals that have been randomised but are excluded from analysis should also be reported (25).

### Publication

Systematic reviews of animal stroke studies suggest that there is substantial publication bias (18, 20, 44, 45); neutral or negative studies may remain unpublished, while the publication of positive studies of patented interventions may be suppressed or delayed during development. Publication bias tends to lead to an overestimation of effect and so it is vital that all studies are finally published, even if it is necessary to delay this for a year or so to protect patent applications. Standard statistical techniques may be used to identify the possible presence of publication bias (46, 47). Indeed, some approaches allow the size and point estimate to be estimated for missing studies. Enforcing registration of projects at their start, as is done with clinical trials (e.g. controlled trials using IRCTN numbers, <http://www.isrctn.org/>), would help identify studies that were never published.

Publications should include detailed information on the species, strain and source of animals, the sample size calculation, inclusion/exclusion criteria, method of randomisation and concealment of allocation, how outcomes were assessed and information on any conflicts of interest (25). Such guidelines on the reporting of preclinical studies share the same aims as the CONSORT criteria used for clinical trials (24) and it is worth noting that stroke trials also suffer from poor reporting (23). Editors and referees should ensure that these approaches to experimentation and publication are followed.

Laboratory-based stroke studies are usually performed by PhD students and postdoctoral researchers, and it is vital for their career progression that their work is published. Because central management and coordination of multi-site experimental studies might threaten publication, coordinated analyses and publication will be needed. For example, laboratory 1 might lead publication on work involving transient models, laboratory 2 would lead on permanent models, laboratory 3 would lead on dose/time relationships, etc. (Fig. 1). Publication under a study acronym, as is often done with clinical trials, would allow all relevant staff at all laboratories to be listed in respect of their input into the project in all publications.

### Downside to central coordination of preclinical stroke studies

There are several potential negative aspects to performing experimental studies in a coordinated multicentre manner. First, collaborative projects add complexity so that mistakes may be more likely. Second, dealing with multiple centres will also add delay to initiating studies although the time to complete them should be far quicker with multiple sites, and so the overall time should be less. Third, financial costs will be

higher in the presence of several sites. However, the ability to drop development programmes immediately if utility becomes apparent and to complete studies more quickly should assist in keeping total costs down.

## Summary

We believe that standard structures (e.g. steering committees, coordinating centres, DMCs, adjudication committees) and techniques (e.g. central randomisation, outcome blinding or central assessment, web transfer of data) used in clinical trials should also be used in preclinical drug development. Although enhancing the predictability of preclinical studies will reduce the number of interventions reaching the clinical arena, it will increase the chance that future clinical trials are positive. However, the fine details need to be firmed up, ideally at workshops involving key stakeholders. In the meantime, we are testing aspects of the overall approach to confirm its feasibility.

## References

- Horn J, Limburg M. Calcium antagonists for ischemic stroke: a systematic review. *Stroke* 2001; **32**:570–6.
- Diener H-C, Lees KR, Lyden P *et al.*, for the SAINT I and II Investigators. NXY-059 for the treatment of acute stroke. Pooled analysis of the SAINT I and II trials. *Stroke* 2008; **39**:1751–8.
- Saxena R, Wijnhoud AD, Carton H *et al.* Controlled safety study of a hemoglobin-based oxygen carrier, DCLHb, in acute ischemic stroke. *Stroke* 1999; **30**:993–6.
- Davis SM, Lees KR, Albers GW *et al.*, for the ASSIST Investigators. Selfotel in acute ischemic stroke. Possible neurotoxic effects of an NMDA antagonist. *Stroke* 2000; **31**:347–54.
- Bath PMW, Bleic S, Bogousslavsky J *et al.* Tirilazad mesylate in acute ischemic stroke – a systematic review. *Stroke* 2000; **31**:2257–65.
- Enlimomab acute stroke trial investigators. Use of anti-ICAM-1 therapy in ischemic stroke: results of the Enlimomab acute stroke trial. *Neurology* 2001; **57**:1428–34.
- Wardlaw JM, Zoppo G, Yamaguchi T, Berge E. Thrombolysis for acute ischaemic stroke. *Cochrane Database Systematic Review*. 2003.
- 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**:768–813.
- Perel P, Roberts I, Sena E *et al.* Comparison of treatment effects between animal experiments and clinical trials: systematic review. *Br Med J* 2007; **334**:197–202.
- Grotta J. Why do all drugs work in animals but none in stroke patients? 2. Neuroprotective therapy. *J Intern Med* 1995; **237**:89–94.
- Green AR, Odegergen T, Ashwood T. Animal models of stroke: do they have value for discovering neuroprotective agents? *Trends Pharmacol Sci* 2003; **24**:402–8.
- Green AR. Pharmacological approaches to acute ischaemic stroke: reperfusion certainly, neuroprotection possibly. *Br J Pharmacol* 2008; **153**:S325–38.
- Kaste M. Use of animal models has not contributed to development of acute stroke therapies: pro. *Stroke* 2005; **36**:2323–4.
- Fisher M, Tatlisumak T. Use of animal models has not contributed to development of acute stroke therapies: con. *Stroke* 2005; **36**:2324–5.
- Donnan G, Davis SM. Stroke drug development. Usually but not always, animal models. *Stroke* 2005; **36**:2326.
- Stroke Therapy Academic Industry Roundtable (STAIR). Recommendations for standards regarding preclinical neuroprotective and restorative drug development. *Stroke* 1999; **30**:2752–8.
- Savitz S. A critical appraisal of the NXY-059 neuroprotection studies for acute stroke: a need for more rigorous testing of neuroprotective agents in animal models of stroke. *Exp Neurol* 2007; **205**:20–5.
- Sena E, van der Worp BH, Howells DW, MacLeod MR. How can we improve the pre-clinical development of drugs for stroke? *Trends Neurosci* 2007; **30**:432–9.
- MacLeod MR, van der Worp BH, Sena E, Howells DW, Dirnagl U, Donnan AD. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. *Stroke* 2008; **39**:2824–9.
- Bath PMW, Gray LJ, Bath AJG, Buchan AM, Miyata T, Green AR. Effects of NXY-059 in experimental stroke: an individual animal meta-analysis. *Br J Pharmacol* 2009; **157**:1157–1171.
- Crossley NA, Sena E, Goehler J *et al.* Empirical evidence of bias in the design of experimental stroke studies: a metaepidemiologic approach. *Stroke* 2008; **39**:929–34.
- Zivin JA. Identifying low-quality preclinical studies. *Stroke* 2008; **39**:2697–8.
- Bath FJ, Owen VE, Bath PMW. Quality of full and final publications reporting acute stroke trials. A systematic review. *Stroke* 1998; **29**:2203–10.
- Moher D, Schultz KF, Altman DG. The Consort statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Lancet* 2001; **357**:1191–4.
- MacLeod M, Fisher M, O'Colling V *et al.* Good laboratory practice: preventing introduction of bias at the bench. *Stroke* 2008; **40**:e59–2.
- Green AR, Ashwood T. Free radical trapping as a therapeutic approach to neuroprotection in stroke: experimental and clinical studies with NXY-059 and free radical scavengers. *CNS Neurol Disord Drug Targets* 2005; **4**:109–18.
- Feurstein GZ, Zaleska MM, Krams M *et al.* Missing steps in the STAIR case: a Translational Medicine perspective on the development of NXY-059 for treatment of acute ischemic stroke. *J Cereb Blood Flow Metab* 2008; **28**:217–9.
- Davis M, Mendelow AD, Perry RH, Chambers IR, James OF. Experimental stroke and neuroprotection in the aging rat brain. *Stroke* 1995; **26**:1072–8.
- Schaller BJ. Influence of age on stroke and preconditioning – induced ischemic tolerance in the brain. *Exp Neurol* 2007; **205**:9–19.
- Fisher M, Feuerstein GZ, Howells DW *et al.*, Group. Update of the stroke therapy academic industry roundtable preclinical recommendations. *Stroke* 2009; **40**:2244–2250.
- Lapchak PA, Araugo DM, Song D, Wei J, Zivin JA. Neuroprotective effects of the spin trap agent disodium-[(tert-butylimino)methyl] benzene-1,3-disulfonate N-oxide (generic NXY-059) in rabbit small clot embolic stroke model. Combination studies with the thrombolytic tissue plasminogen activator. *Stroke* 2002; **33**:1411–5.
- Lapchak PA, Song D, Wei J, Zivin JA. Coadministration of NXY-059 and tenecteplase six hours following embolic strokes in rabbits improves clinical rating scores. *Exp Neurol* 2004; **188**:279–85.
- Wang CX, Shuaib A. NXY-059: a neuroprotective agent in acute stroke. *Int J Clin Pract* 2004; **58**:964–9.
- Lo EH. Experimental models, neurovascular mechanisms and translational issues in stroke research. *Br J Pharmacol* 2008; **153**:S396–405.
- Bath PMW, Gray LJ. Systematic reviews as a tool for planning and interpreting trials. *Int J Stroke* 2009; **4**:23–7.
- Krams M, Lees KR, Hacke W, Grieve AP, Orgogozo JM, Ford GA, for the ASTIN Study Investigators. Acute stroke therapy by inhibition of neutrophils (ASTIN). An adaptive dose–response study of UK-279,276 in acute ischemic stroke. *Stroke* 2003; **34**:2543–8.
- Weaver CS, Leonardi-Bee J, Bath-Hexall FJ, Bath PMW. Sample size calculations in acute stroke trials: a systematic review of their reporting,

- characteristics, and relationship with outcome. *Stroke* 2004; **35**:1216–24.
- 38 Schultz KF, Grimes DA. Generation of allocation sequences in randomized trials: chance, not choice. *Lancet* 2002; **359**:515–9.
- 39 The ENOS Trial Investigators. Glyceryl trinitrate vs. control, and continuing vs. stopping temporarily prior antihypertensive therapy, in acute stroke: rationale and design of the Efficacy of Nitric Oxide in Stroke (ENOS) trial (ISRCTN99414122). *Int J Stroke* 2006; **1**:245–9.
- 40 Weir CJ, Lees KR. Comparison of stratification and adaptive methods for treatment allocation in an acute stroke clinical trial. *Stat Med* 2003; **22**:705–26.
- 41 Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *Br Med J* 2001; **323**:42–6.
- 42 Schulz KF, Grimes DA. Allocation concealment in randomized trials: defending against deciphering. *Lancet* 2002; **359**:614–8.
- 43 Wardlaw J, Bath P, Sandercock P *et al.* The NeuroGrid stroke exemplar clinical trial protocol. *Int J Stroke* 2007; **2**:63–9.
- 44 Willmot M, Gray L, Gibson C, Murphy S, Bath PMW. A systematic review of nitric oxide donors and l-arginine in experimental stroke; effects on infarct size and cerebral blood flow. *Nitric Oxide* 2005; **12**:141–9.
- 45 Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publications bias. *Stroke* 2004; **35**:1203–8.
- 46 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; **315**:629–34.
- 47 Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med* 2001; **20**:641–54.
- 48 Ali M, Bath PM, Curram J *et al.* The Virtual International Stroke Trials Archive. *Stroke* 2007; **38**:1905–10.
- 49 The Optimising Analysis of Stroke Trials (OAST) Collaboration. Can we improve the statistical analysis of stroke trials? Statistical re-analysis of functional outcomes in stroke trials. *Stroke* 2007; **38**:1911–5.
- 50 Walker MF, Leonardi-Bee J, Bath P *et al.* An individual patient data meta-analysis of randomized controlled trials of community occupational therapy for stroke patients. *Stroke* 2004; **35**:2226–32.
- 51 Leonardi-Bee J, Bath PM, Bousser MG *et al.*, Dipyridamole in Stroke Collaboration (DISC). Dipyridamole for preventing recurrent ischemic stroke and other vascular events: a meta-analysis of individual patient data from randomized controlled trials. *Stroke* 2005; **36**:162–8.