

Evidence for the Efficacy of NXY-059 in Experimental Focal Cerebral Ischaemia Is Confounded by Study Quality

Malcolm R. Macleod, PhD, FRCP; H. Bart van der Worp, MD, PhD; Emily S. Sena, BSc;
David W. Howells, PhD; Ulrich Dirnagl, MD, PhD; Geoffrey A. Donnan, MD, FRACP

Background and Purpose—The neutral results of the SAINT II trial have again highlighted difficulties translating neuroprotective efficacy from bench to bedside. Animal studies are susceptible to study quality biases, which may lead to overstatement of efficacy. We report the impact of study quality on published estimates of the efficacy of NXY-059 in animal models of stroke.

Methods—We conducted a systematic review and stratified meta-analysis of published studies describing the efficacy of NXY-059 in experimental focal cerebral ischemia.

Results—Overall, NXY-059 improved infarct volume by 43.3% (95% CI, 34.7 to 52.8). Only 2 of 9 publications reported randomization, concealment of treatment allocation, and blinded outcome assessment. Studies not reporting these quality items gave substantially higher estimates of efficacy than did higher-quality studies.

Conclusions—The reported efficacy of NXY-059 in animal models of stroke is confounded by low study quality. The failure of SAINT II highlights the need for substantial improvements in the design, conduct, and reporting of animal studies; journals can play an important role in this by adopting standards for animal studies similar to those agreed over 10 years ago for clinical trials. (*Stroke*. 2008;39:000-000.)

Key Words: neuroprotection ■ NXY-059 ■ study quality

NXY-059 is a free radical scavenger that was considered to have substantial neuroprotective properties in animal models of stroke. The evidence supporting this efficacy has been reported^{1,2} to meet all of the criteria established by the Stroke Academic Industry Roundtable³ (STAIR) for the further development of neuroprotective agents. The first clinical trial of NXY-059 (SAINT I) showed a small but significant benefit on a coprimary end point,¹ but this benefit was not seen in the subsequent, larger SAINT II trial designed to confirm efficacy.² The failure of the SAINT trials and the apparent quality of the animal data has led many to question whether the STAIR criteria are helpful and indeed whether current paradigms for the development of neuroprotective drugs are likely ever to prove successful.

Different hypotheses have been advanced to explain translational failure in stroke. Broadly, this might be due either to false-positive animal studies or to false-negative clinical trials. There is substantial evidence both for low methodological quality in animal studies (increasing the risk of false-positive results) and for clinical trials in which important factors such as the delay to treatment or the tissue concentration of drug achieved do not match those under which efficacy was seen in animals.^{4,5} A further potential explana-

tion is that animal models might not recapitulate human stroke with sufficient fidelity to be useful.

Since the publication of SAINT II, a number of plausible explanations for the failure of the substantial efficacy observed in animal studies of NXY-059 to translate to clinical trial have been advanced,⁶ including a failure to demonstrate pharmacodynamic markers of free radical scavenging activity in animals⁷; the possibility that the marginal efficacy reported in the first SAINT study was in fact due to an active metabolite of NXY-059⁸; and a lack of testing in multiple academic institutions using many different animal models compounded by publication bias.⁹

We set out to test the hypothesis that the published animal data for NXY-059 are confounded by low study quality.

Methods

We conducted a systematic review and meta-analysis as previously described.¹⁰ Studies of NXY-059 in animal models of focal cerebral ischemia were identified from electronic searches of PubMed, EMBASE, and BIOSIS conducted on July 4, 2007, using search terms (NXY059 OR NXY-059 OR Cerovive) AND (stroke OR ischemia OR cerebrovascular OR middle cerebral artery OR MCA OR ACA OR anterior cerebral artery OR MCAO) AND Animals NOT (coronary OR myocardia*) and from checking the reference lists of

Received January 27, 2008; final revision received February 6, 2008; accepted March 3, 2008.

From Clinical Neurosciences (M.R.M., E.S.S.), University of Edinburgh, Scotland, UK; the Department of Neurology (H.B.v.d.W.), Rudolf Magnus Institute of Neuroscience, University Medical Centre, Utrecht, The Netherlands; the National Stroke Research Institute (E.S.S., D.W.H., G.A.D.), Melbourne, Australia; and the Department Neurology (U.D.), Charite Hospital, Berlin, Germany.

Correspondence to Malcolm R. Macleod, PhD, FRCP, Clinical Neurosciences, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, Scotland, UK. E-mail malcolm.macleod@ed.ac.uk

© 2008 American Heart Association, Inc.

Stroke is available at <http://stroke.ahajournals.org>

DOI: 10.1161/STROKEAHA.108.515957

Table 1. Study Characteristic Report

Author	Year	Species	No. of Animals	Dose Range	Time of Administration, minutes	Anesthetic	Type of Ischemia	Route of Delivery	Outcome Measure(s)
Balogh ¹⁴	2005	Mouse	14	10 mg/kg	30	Unknown	Permanent	IP	Infarct volume
Kuroda ²³	1999	Rat	74	7.5–750 mg/kg	180–480	Halothane	Temporary	IV	Infarct volume
Lapchak ¹²	2004	Rabbit	189	0.1–100 mg/kg	60–360	Halothane	Thrombotic	IV	Neurological score
Lapchak ²⁴	2002	Rabbit	72	100 mg/kg	5	Unknown	Thrombotic	IV	Infarct volume
Lapchak ¹³	2002	Rabbit	58	100 mg/kg	5–180	Unknown	Thrombotic	IV	Neurological score
Marshall ²⁵	2003	Marmoset	44	0.9 mg/kg	240	Alphaxalone	Permanent	IV	Infarct volume Neurological score
Marshall ²⁶	2001	Marmoset	11	0.27 mg/kg	5	Alphaxalone	Permanent	IV	Infarct volume
Sydserrf ²⁷	2002	Rat	175	65–1750 mg/kg	5–240	Halothane	Temporary Permanent	IV	Infarct volume Neurological score
Yoshimoto ²⁸	2002	Rat	38	120–750 mg/kg	180	Halothane	Temporary	IV	Infarct volume
Yoshimoto ²⁹	2002	Rat	20	720 mg/kg	125–180	Halothane	Temporary	IV	Infarct volume
Zhao ³⁰	2001	Rat	30	750 mg/kg	5	Halothane	Permanent	IV	Infarct volume

included articles. Publications reporting outcome as infarct size or neurobehavioral score were selected for further analysis. Briefly, publication details, including study quality, experiments performed, and their reported outcomes, were entered to the CAMARADES database and stratified normalized mean difference DerSimonian and Laird random effects meta-analysis carried out.

The range of evidence supporting the efficacy of NXY-059 was considered against the STAIR criteria,³ which we considered to be 1) recovery of sensorimotor function; 2) recovery of cognitive function; 3) efficacy replicated in 2 laboratories; 4) tested in models of permanent and temporary occlusion; 5) tested in males and females; 6) behavioral outcome measured for at least 1 month; 7) efficacy tested in primates; 8) clinically appropriate route of administration used; and 9) experiments carried out in a “blinded, randomized” fashion.

We were specifically interested in the effects of reported study quality, and therefore we analyzed differences in reported efficacy among studies reporting compliance with each component of our previously published 10-item study quality checklist¹¹ comprising (1) publication in a peer-reviewed journal; (2) statement of control of temperature; (3) randomization to treatment or control; (4) blinded induction of ischemia (ie, concealment of treatment group allocation at time of induction of ischemia); (5) blinded assessment of outcome; (6) avoidance of anesthetics with marked intrinsic neuroprotective properties; (7) use of animals with hypertension or diabetes; (8) sample size calculation; (9) statement of compliance with regulatory requirements; and (10) statement regarding possible conflicts of interest; and the relationship between reported efficacy and the number of checklist items scored. The analyses to be performed were specified in advance of any data collection, and all analysis performed are reported here.

The significance of reported study quality in explaining differences in observed efficacy was determined by partitioning of heterogeneity. Individual study quality items were considered to have a significant effect if the 95% confidence limits of the estimates did not overlap.

Results

Systematic search identified 216 “hits” (PubMed, 77; EMBASE, 123; BIOSIS, 16) of which 56 studies (112 hits) were identified in 2 databases. After screening of electronic abstracts, 21 full publications and 6 abstracts were considered potentially relevant and the full publications retrieved. Of the full publications, 8 were not relevant, one reported information also published elsewhere, and one was a re-evaluation of

previously published data. All 6 abstracts had subsequently been published in full. Taken together, the 11 included studies (Table 1) met all the criteria proposed by the Stroke Academic Industry Roundtable. Nine of 11 studies reported outcome as infarct volume and 4 reported neurobehavioral score.

One neurobehavioral outcome reported by one group in 2 publications (a rabbit small clot embolic stroke model reporting the weight of clot causing a neurological dysfunction in 50% of animals)^{12,13} could not (for methodological reasons relating to the form in which data are required) be incorporated in this meta-analysis; because these represent a substantial portion of the data for neurobehavioral outcome, this was not analyzed further.

The 9 publications reporting infarct volume described 29 separate experiments and outcome in 408 experimental animals. The median quality score was 5 (interquartile range, 4 to 6); 3 of 9 studies reported random allocation to group, 5 of 9 reported that ischemia was induced by an investigator blinded to treatment allocation, and 4 of 9 studies reported that outcome was assessed without knowledge of treatment group allocation (Table 2).

Overall, NXY-059 improved infarct volume by 43.3% (95% CI, 34.7 to 52.8; Figure 1). Stratifying studies by reported quality score explained a significant proportion of the observed heterogeneity ($\chi^2=55.6$, $df=6$, $P<0.001$) with reported efficacy being highest in low-quality studies (Figure 2). The effect of reported study quality was also seen with individual quality items. Reported efficacy was significantly lower in studies that reported randomization (20.3% versus 52.8%; $\chi^2=36.9$, $df=1$, $P<0.001$; Figure 3A); that reported measures taken to conceal treatment allocation from the time of cerebral ischemia up to the time of outcome assessment (25.1% versus 54.0%; $\chi^2=34.1$, $df=1$, $P<0.001$; Figure 3B); and that reported the use of spontaneously hypertensive rats rather than healthy animals (17.6% versus 47.8%; $\chi^2=29.1$, $df=1$, $P<0.001$; Figure 3C). Efficacy was significantly higher in studies that reported the use of an anesthetic not

Table 2. Study Quality Report

Author	Year	1	2	3	4	5	6	7	8	9	10	Quality Score
Balogh ¹⁴	2005	+										1
Kuroda ²³	1999	+	+				+					3
Lapchak ²⁴	2002	+		+	+	+			+	+		6
Marshall ²⁶	2001	+			+	+	+			+		5
Marshall ²⁵	2003	+			+	+	+			+		5
Sydserrf ²⁷	2002	+	+				+			+		4
Yoshimoto ²⁸	2002	+	+	+	+	+	+		+	+		8
Yoshimoto ²⁹	2002	+	+				+			+		4
Zhao ³⁰	2001	+	+	+	+		+	+		+		7

Study quality items are (1) publication in a peer-reviewed journal; (2) statement of control of temperature; (3) randomization to treatment or control; (4) blinded induction of ischemia (ie, concealment of treatment group allocation at time of induction of ischaemia); (5) blinded assessment of outcome; (6) avoidance of anaesthetics with marked intrinsic neuroprotective properties; (7) use of animals with hypertension or diabetes; (8) sample size calculation; (9) statement of compliance with regulatory requirements; and (10) statement regarding possible conflicts of interest.

known to have intrinsic neuroprotective activity (47.4% versus 4.3%, $\chi^2=29.1$, $df=1$, $P<0.001$; Figure 3D) than in 2 studies in which the anesthetic was not stated. There was no significant effect of statements reporting control of temperature, the blinding of outcome assessment, sample size calculation, or compliance with animal welfare regulations. All experiments were published after peer review and no publication contained a statement of potential conflicts of interest, so the impact of reporting of these items cannot be assessed.

Discussion

This systematic review and meta-analysis suggests that NXY-059 has substantial efficacy in animal models of ischemic stroke and confirms that when considered together, the animal evidence for NXY-059 meets all of the STAIR criteria. However, there were important differences between studies that did or did not report the use of important methodological safeguards against bias in that studies reporting adherence to these quality items gave significantly lower estimates of efficacy than those that did not.

Why is this literature for NXY-059, which is compliant with the STAIR criteria, the subject of careful evaluation from the industry and subjected to rigorous review so susceptible to bias? First, a nonsystematic review may not

have identified all relevant publications (for instance, the paper included here from *Bioorganic & Medicinal Chemistry Letters*¹⁴) and so may have given an incomplete appraisal of published data.

Second, efficacy may be overstated due to publication bias; it may be that there are neutral or negative unpublished data that give a more realistic appraisal of the effects of NXY-059; although we are aware that some such data do exist and remain unpublished despite having been submitted for publication, the number of publications is not large enough to test statistically for publication bias. Comparing our data with a systematic review that included such unpublished data and proprietary industry data would demonstrate whether published data have substantially overstated the efficacy of NXY-059.

Third, although the evidence supporting the efficacy of NXY-059 does indeed meet all of the STAIR criteria when considered together, individual studies do not. Those criteria relate not just to the range of evidence, but also to the quality of that evidence; only one publication reported meeting all of the

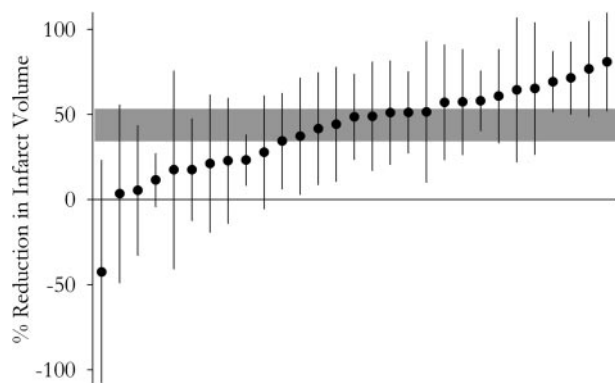


Figure 1. Individual comparisons ranked according to their effect on infarct volume. The shaded gray bar represents the 95% confidence limits of the global estimate. The vertical error bars represent the 95% CIs for the individual estimates.

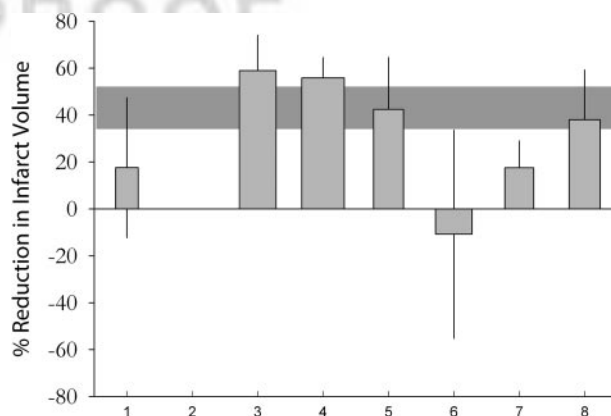


Figure 2. Effect of number of quality items scored on the estimate of efficacy. The shaded gray bar represents the 95% confidence limits of the global estimate. The vertical error bars represent the 95% CIs for the individual estimates. The size of each point reflects the log of the number of animals contributing to that comparison. Stratification by quality accounts for a significant proportion of the heterogeneity observed between studies ($P<0.001$).

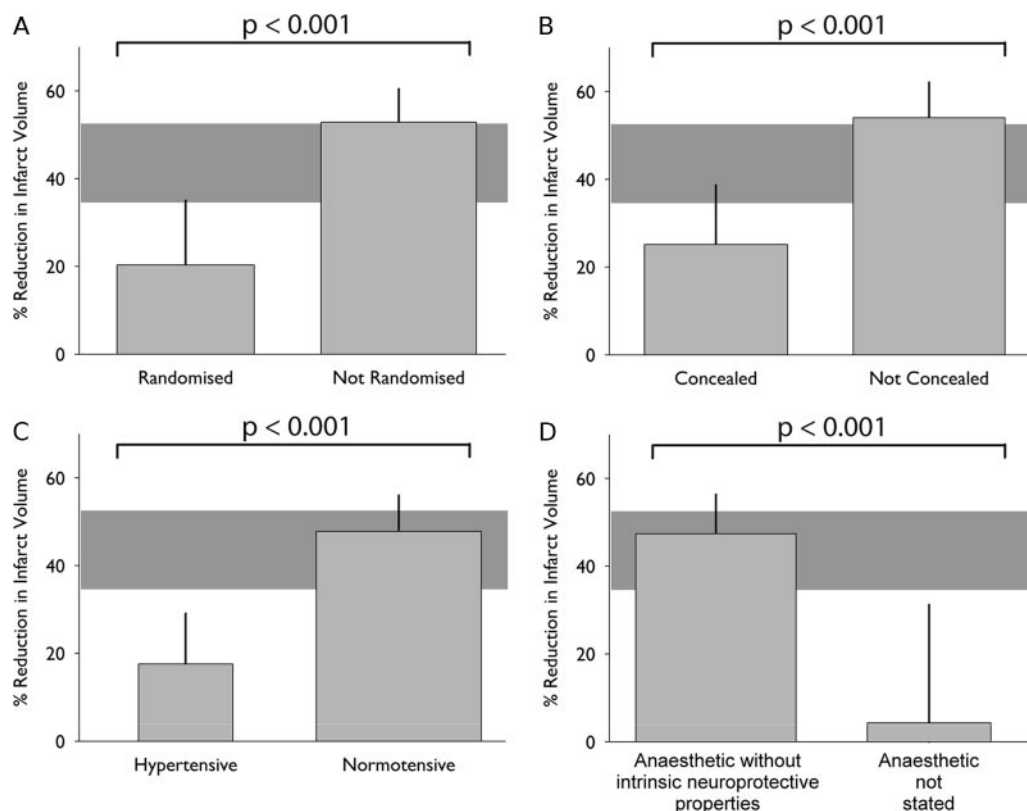


Figure 3. Influence of (A) randomization to experimental group; (B) concealment of treatment group allocation during the experiment; (C) blinding of outcome assessment; and (D) hypertension on the estimate of efficacy. The shaded gray bar represents the 95% confidence limits of the global estimate. The vertical error bars represent the 95% CIs for the individual estimates. The width of each vertical bar reflects the log of the number of animals contributing to that comparison. In A, B, and D, the stratification accounts for a significant proportion of the heterogeneity observed between studies ($P < 0.001$).

STAIR criteria relating to study quality. That is to say that the quality of this literature compares poorly with that for the animal data supporting other treatments for stroke; allocation concealment (55%) and blinded assessment of outcome (44%) were more common here than in systematic reviews of 6 other interventions (11% and 29%, respectively⁵). The 33% of studies randomizing animals to treatment group was similar to that seen for other interventions (36%).⁵

We do not consider it necessary for each publication to meet each of the STAIR criteria, and we have proposed criteria relating to the range of evidence (which might be represented across different publications) and criteria relating to the quality of evidence (which should be met by every study).⁵ Given the small numbers of animals included in individual studies, it might be reasonable to require that each item relating to the range of evidence be replicated in more than one laboratory.

Finally, efficacy may have been overstated due to methodological flaws in individual animal experiments. The data reported here suggest that in particular randomization, allocation concealment, and blinded outcome assessment are important methodological measures, which might reduce bias but which are poorly represented in this literature.

Potential Weaknesses of This Analysis

Because this analysis is based on information extracted from publications, it may be that measures taken to avoid bias have

been incompletely described. However, it seems unlikely that such differences in reported (rather than actual) study quality would be associated with differences of the magnitude observed here. In a survey of the authors of 193 publications included in previous systematic reviews of animal stroke studies,¹⁵ questionnaire responses did indeed report somewhat higher rates of, for instance, randomization and blinding than ascertained from the published work, although the response rate was low and probably overstated these differences. In addition, respondents misunderstood the meaning of certain key terms; for instance, some authors reported randomization being achieved by the animals “being chosen at random from the cage.”

Clearly, not all items of the CAMARADES study quality checklist are of equal importance. A meta-analysis of meta-analyses has confirmed important roles for both allocation concealment and the use of animals with relevant comorbidities,¹⁶ whereas the impact of randomization and the blinded assessment of outcome is not yet clear. Metaregression of data from individual animal experiments should provide further information to guide the refinement of study quality checklists.

We have only been able to present data for infarct volume, and this is a consequence of our use of normalized mean difference meta-analysis. In this technique, the outcome (infarct volume or neurobehavioral score) in normal (unlesioned) animals is set to 0, and the outcome observed in

lesioned but untreated controls is set to 1. The outcome in the treatment group is then expressed on the scale thus defined with values less than 1 representing an improvement in outcome (eg, “x”). From this we can calculate a percentage improvement in outcome ($=100*[1-x]$) and we can express this as percentage improvement in neurobehavioral score or reduction in infarct volume. This calculation requires that we know the outcome both in unlesioned untreated controls and in lesioned untreated controls. With regard to the neurobehavioral data presented in the publications of Lapchak et al,^{12,13} there are no data for outcome in unlesioned untreated animals because the outcome measure depends on the amount of clot causing a neurobehavioral deficit in 50% of animals and therefore requires the presence of a lesion. For this reason we have been unable to analyze neurobehavioral outcomes, although we believe that such an analysis would be valuable.

An alternative approach would be to use standardized mean difference meta-analysis, in which differences are scaled according to the variance of the estimates. However, although this approach is justified when sample size is large (like with clinical trials), when sample size is small (the median sample size for these studies, including those excluded studies reporting neurobehavioral outcome, was 8), the observed SD reflects the population SD with less fidelity, which introduces further error to the analysis, and this makes the standardized mean difference approach less reliable.¹⁷

Implications for Future Animal Studies

Although other factors may of course confound the apparent association between the reporting of methodological safeguards against bias and low reported efficacy, the most plausible explanation is that this relationship is causal. These observations highlight the need for more detailed reporting of the measures taken to avoid bias.

A further important implication for the design of future studies is that unbiased estimates of efficacy are likely to be substantially lower than previously considered important; such high-quality studies should therefore be powered to detect improvements in infarct volume in the range of 10% to 20%.

Other Potential Reasons for the Failure of NXY-059 in Clinical Trials

There may be other reasons why the efficacy reported in animals was not seen in clinical trials. First, it may be that the clinical trials were false-negative because of the play of chance or because they tested efficacy in the wrong group of patients or at the wrong time. For instance, the median time to treatment in these animal studies was 145 minutes compared with a mean delay from symptom onset to the initiation of treatment in SAINT II of 228 minutes. The neurobehavioral studies of Lapchak et al reported that improvement in neurobehavioral outcome seen at earlier time points was not seen when treatment was initiated 3 or 6 hours after vessel occlusion.^{12,13}

Second, evidence from healthy animals may not generalize to humans with comorbidities; the SAINT II study included 77% patients with a history of hypertension,² yet in the animal literature, only 7% of animals were hypertensive; in

these animals, efficacy was significantly lower than in normotensive animals.

Third, the outcome measures used in these animal studies may not be relevant to human disease. Most studies used TTC staining to measure infarct volume, but the extent to which this reflects true infarct volume has been questioned.^{18,19} Moreover, the extent to which reduction in infarct volume in small animals might predict efficacy in the much larger human brain is not clear. Behavioral outcomes might represent more appropriate indicators of efficacy, but few studies reported neurobehavioral outcomes and for methodological reasons, it was not possible to analyze those data. However, across a range of candidate neuroprotective drugs, neurobehavioral outcome does not give a substantially lower estimate of efficacy than does infarct volume; although for nicotinamide, efficacy was indeed lower for neurobehavioral outcomes,¹¹ for tirilazad, it was higher¹⁰; and for melatonin, tissue plasminogen activator and hypothermia, there was no difference.^{20–22}

The failure of the SAINT studies to show significant benefit has brought into question the usefulness of the STAIR criteria for animal stroke studies. We believe these data suggest that the criteria should indeed be amended, first to require that all studies include methodological measures to avoid bias (as outlined by Sena et al⁵) and second, that a systematic review of all data be conducted before proceeding to clinical trial. Although adopting such measures is unlikely in itself to lead to the seamless translation of efficacy from animal studies to clinical trial, we do believe that the more precise knowledge of the quality of supporting animal data and the conditions of maximum efficacy in animal models would provide a more secure basis for decisions regarding the development of stroke drugs.

Source of Funding

This work was supported in part by a Translational Medicine award from the Scottish Chief Scientist Office.

Disclosures

None.

References

1. Lees KR, Zivin JA, Ashwood T, Davalos A, Davis SM, Diener HC, Grotta J, Lyden P, Shuaib A, Hardemark HG, Wasiewski WW. NXY-059 for acute ischemic stroke. *N Engl J Med.* 2006;354:588–600.
2. Shuaib A, Lees KR, Lyden P, Grotta J, Davalos A, Davis SM, Diener HC, Ashwood T, Wasiewski WW, Emeribe U. NXY-059 for the treatment of acute ischemic stroke. *N Engl J Med.* 2007;357:562–571.
3. Recommendations for Standards Regarding Preclinical Neuroprotective and Restorative Drug Development. *Stroke.* 1999;30:2752–2758.
4. Dirnagl U. Bench to bedside: the quest for quality in experimental stroke research. *J Cereb Blood Flow Metab.* 2006;26:1465–1478.
5. Sena E, van der Worp HB, Howells D, Macleod M. How can we improve the pre-clinical development of drugs for stroke? *Trends Neurosci.* 2007;30:433–439.
6. Green AR. Pharmacological approaches to acute ischaemic stroke: reperfusion certainly, neuroprotection possibly. *Br J Pharmacol.* 2008;153(suppl 1):S325–338.
7. Feuerstein GZ, Zaleska MM, Krams M, Wang X, Day M, Rutkowski JL, Finklestein SP, Pangalos MN, Poole M, Stiles GL, Ruffolo RR, Walsh FL. 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;28:217–219.

8. Proctor PH, Tamborello LP. SAINT-I worked, but the neuroprotectant is not NXY-059. *Stroke*. 2007;38:e109.
9. Savitz SI. 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;205:20–25.
10. Sena E, Wheble P, Sandercock P, Macleod M. Systematic review and meta-analysis of the efficacy of tirilazad in experimental stroke. *Stroke*. 2007;38:388–394.
11. Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke*. 2004;35:1203–1208.
12. Lapchak PA, Song DH, Wei JD, Zivin JA. Coadministration of NXY-059 and tenecteplase six hours following embolic strokes in rabbits improves clinical rating scores. *Exp Neurol*. 2004;188:279–285.
13. Lapchak PA, Araujo DM, Song DH, Wei JD, Zivin JA. Neuroprotective effects of the spin trap agent disodium-[(tert-butylimino)methyl]benzene-1,3-disulfonate N-oxide (generic NXY-059) in a rabbit small clot embolic stroke model—combination studies with the thrombolytic tissue plasminogen activator. *Stroke*. 2002;33:1411–1415.
14. Balogh GT, Vukics K, Konczol A, Kis-Varga A, Gere A, Fischer J. Nitrone derivatives of Trolox as neuroprotective agents. *Bioorganic & Medicinal Chemistry Letters*. 2005;15:3012–3015.
15. Samaranyake S. Study quality in experimental stroke. 2006. CAMARADES Monograph No 2. Available at: www.camarades.info/index_files/CM2.pdf. Accessed January 24, 2008.
16. Crossley NA, Sena E, Goehler J, Horn J, van der Worp B, Bath PMW, Macleod M, Dirnagl U. Empirical evidence of bias in the design of experimental stroke studies. A metaepidemiologic approach. *Stroke*. 2008;39:929–934.
17. Wu Q. Combining the evidence from different drug trials. 2006. CAMARADES Monograph No 1. Available at: www.camarades.info/index_files/CAMARADES%20Monograph%201.pdf. Accessed February 6, 2008.
18. Tureyen K, Vemuganti R, Sailor KA, Dempsey RJ. Infarct volume quantification in mouse focal cerebral ischemia: a comparison of triphenyltetrazolium chloride and cresyl violet staining techniques. *J Neurosci Methods*. 2004;139:203–207.
19. Benedek A, Moricz K, Juranyi Z, Gigler G, Levay G, Harsing LG Jr, Matyus P, Szenasi G, Albert M. Use of TTC staining for the evaluation of tissue injury in the early phases of reperfusion after focal cerebral ischemia in rats. *Brain Res*. 2006;1116:159–165.
20. Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P, Khan KS. Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ*. 2007;334:197–199.
21. van der Worp HB, Sena ES, Donnan GA, Howells DW, Macleod MR. Hypothermia in animal models of acute ischaemic stroke: a systematic review and meta-analysis. *Brain*. 2007;130:3063–3074.
22. Macleod MR, O'Collins T, Horky LL, Howells DW, Donnan GA. Systematic review and meta-analysis of the efficacy of melatonin in experimental stroke. *J Pineal Res*. 2005;38:35–41.
23. Kuroda S, Tsuchida R, Smith ML, Maples KR, Siesjo BK. Neuroprotective effects of a novel nitrone, NXY-059, after transient focal cerebral ischemia in the rat. *J Cereb Blood Flow Metab*. 1999;19:778–787.
24. Lapchak PA, Araujo DM, Song DH, Wei JD, Purdy R, Zivin JA. Effects of the spin trap agent disodium (tert-butylimino)methylbenzene-1,3-disulfonate N-oxide (generic NXY-059) on intracerebral hemorrhage in a rabbit large clot embolic stroke model combination studies with tissue plasminogen activator. *Stroke*. 2002;33:1665–1670.
25. Marshall JWB, Cummings RM, Bowes LJ, Ridley RM, Green AR. Functional and histological evidence for the protective effect of NXY-059 in a primate model of stroke when given 4 hours after occlusion. *Stroke*. 2003;34:2228–2233.
26. Marshall JWB, Duffin KJ, Green AR, Ridley RR. NXY-059, a free radical-trapping agent, substantially lessens the functional disability resulting from cerebral ischemia in a primate species. *Stroke*. 2001;32:190–197.
27. Sydeserff SG, Borelli AR, Green AR, Cross AJ. Effect of NXY-059 on infarct volume after transient or permanent middle cerebral artery occlusion in the rat; studies on dose, plasma concentration and therapeutic time window. *Br J Pharmacol*. 2002;135:103–112.
28. Yoshimoto T, Kanakaraj P, Ying MJ, Cheng M, Kerr I, Malaiyandi L, Watson JA, Siesjo BK, Maples KR. NXY-059 maintains Akt activation and inhibits release of cytochrome C after focal cerebral ischemia. *Brain Res*. 2002;947:191–198.
29. Yoshimoto T, Kristian T, Hu B, Ouyang YB, Siesjo BK. Effect of NXY-059 on secondary mitochondrial dysfunction after transient focal ischemia; comparison with cyclosporin A. *Brain Res*. 2002;932:99–109.
30. Zhao Z, Cheng M, Maples KR, Ma JY, Buchan AM. NXY-059, a novel free radical trapping compound, reduces cortical infarction after permanent focal cerebral ischemia in the rat. *Brain Res*. 2001;909:46–50.

Stroke

JOURNAL OF THE AMERICAN HEART ASSOCIATION

FINAL PROOF