Randomized Clinical Trials in Stroke Research

Chul Ahn, PhD* and Daniel Ahn, DO†

Abstract: A randomized clinical trial is widely regarded as the most rigorous study design to determine the efficacy of intervention because spurious causality and bias associated with other experimental designs can be avoided. The purpose of this article is to provide clinicians and clinical researchers the types of randomized clinical trials used in stroke studies and to discuss the advantages and the limitations for each type of randomized stroke clinical trials.

Key Words: RCT, advantages, disadvantages

(*J Investig Med* 2010;58: 277–281)

Randomized clinical trials (RCTs) are prospective clinical trials in which the participants are randomly (ie, by chance) assigned to either a treatment or a control to measure and compare the effect and value of a treatment against a control. Randomization avoids systematic bias by producing groups that differ only by chance in both known and unknown prognostic factors. Thus, randomization makes the study groups similar in all relevant aspects at baseline so that outcome differences may be attributed to the intervention.

Randomized clinical trials are widely regarded as the most rigorous study design to determine the efficacy of interventions because spurious causality and bias associated with other experimental designs can be avoided.[1] Randomized clinical trials are considered the best design to establish cause and effect, and the only effective design known to eliminate biases that may lead to systematic differences between treatment groups.[2] Randomized clinical trials are statistically attractive because many statistical methods assume random assignment, which permits the use of probability theory to express the likelihood that any difference in outcome between intervention groups merely reflects chance. The validity of appropriate statistical tests is assured by the process of randomization.

Randomized clinical trials are costly and time consuming because of randomization, blinding, placebo, and large sample size for comparison, especially with low incidence outcomes. Randomized clinical trials are not feasible for outcomes that are rare or have long lag times. However, even for rare disorders, RCTs remain as the best method to obtain unbiased estimates of treatment effects.[3,4] The disadvantage of RCTs includes lack of representativeness because study subjects may not represent the general study population. Randomized clinical trials often involve many inclusion and exclusion criteria for study subjects, thus limiting generalizability to a more general population. The disadvantage of RCTs also includes dealing with the resistance of patients and clinicians. Clinicians should do what they think is best for their patients. That is, if a clinician knows the best treatment, a clinician should not participate in the RCT.

''Physicians who are convinced that one treatment is better than another for a particular patient cannot ethically choose at random which treatment to give, they must do what they think best for the patient. For this reason, physicians who feel they already know the answer cannot enter their patients into a trial. If they think, whether for a wise or silly reason, that they know the answer before the trial starts, they should not enter any patients… [Weijer et al[5]]''

Proper design of clinical trials is critical because analysis cannot rescue improper design. That is, a poor design cannot be salvaged by good statistics. There are many aspects to be considered for the design of RCTs such as randomization, blinding, stratification, blocking, sample size estimation, multiple comparisons, and multiplicity issues. However, in this paper, we mainly discuss the types of RCT designs used in stroke clinical trials and their advantages and disadvantages.

TYPES OF RCT DESIGNS

There are a number of available RCT designs with each developed for specific situations in stroke research. Most RCTs use a parallel design. In a parallel clinical trial, each group of study subjects is exposed to only one of the study interventions in a random fashion. For instance, in a parallel clinical trial to evaluate the effects of cilostazol in preventing the progression of the symptomatic intracranial arterial stenosis compared with those of a placebo in patients with acute symptomatic stenosis in the M1 segment of middle cerebral artery or the basilar artery, investigators randomly gave cilostazol to one group of patients and placebo to the other group of patients.[6] Figure 1 shows the simplest group comparison parallel design that is widely used in stroke studies. If a subject meets the study eligibility criteria and signs the informed consent form, the subject is randomly allocated to receive either treatment A or B.

The advantages of a parallel design are simplicity and universal acceptance. The disadvantages are time and effort involved in their effective implementation, dealing with the resistance of patients and clinicians, and large sample size for comparison, especially with low incidence outcomes. Randomized clinical
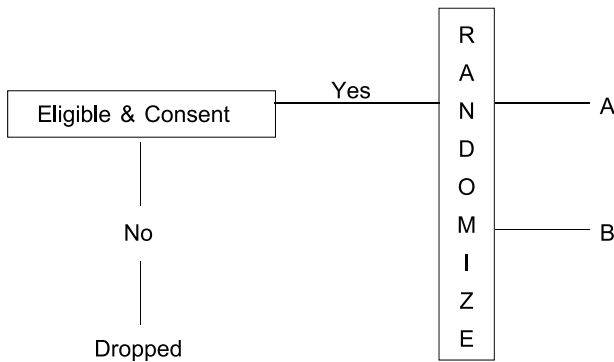
FIGURE 1. Parallel design.

trials may not be feasible for outcomes that are rare or have long lag times.

Run-In Design

A run-in period is a period before randomization during which potential participants who meet all the eligibility criteria for an RCT are assigned to take the study medication. That is, during a run-in period, subjects who are entering a long-term trial are asked to take the study medication before randomization. Figure 2 shows the diagram of a run-in design.

Either placebo or active therapy can be used during a run-in period with usually a single-blind run-in phase. A placebo run-in period allows trial staff to be sure that reported adverse effects are not caused by treatment, whereas an active run-in period can exclude subjects who may not be able to tolerate the medication in a long-term trial.

A major advantage of run-in phases is the increase in trial efficiency gained by screening out potentially noncompliant patients, which has a direct effect on the power of the study. The increase in sample size required to achieve the efficiency of a totally compliant study is approximately $1 / (c_1 + c_2 - 1)_2$, where c_1 and c_2 are the compliance rates in each group. That is, a run-in period can reduce the required sample size.[7] In addition, a placebo run-in period can serve as the washout period to remove the effects of previous treatment and also serve as the training period for investigators, staffs, and patients. However, a run-in period increases the length to complete a clinical trial, which results in increased cost and potential reduction of the enthusiasm of the patients and investigators.

Clinical applicability of the clinical trial results can be diluted or enhanced with the use of run-in periods, depending on the p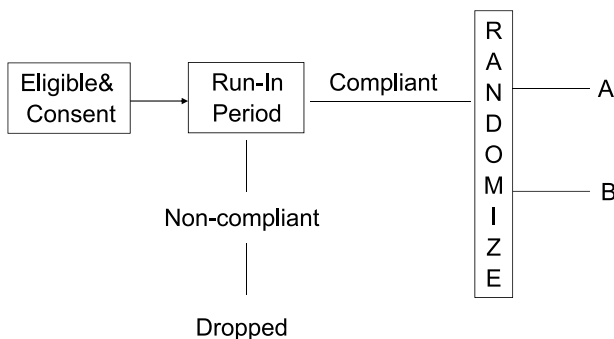atient group to whom the results will be applied. The run-in trial needs to carefully report whether the exclusion of potential subjects will reduce the generalizability of the trial results, and how this aspect of run-in design affects the application of the results to clinical practice.

A run-in design was used to assess the effects of long-term angiotensin-converting enzyme inhibitor therapy on cerebral circulation in patients with previous minor stroke.[8] In a low-dose–aspirin randomized trial for the primary prevention of cardiovascular disease in women,[9] eligible women were enrolled in a 3-month placebo run-in period to identify a group likely to be compliant with the study protocol. Those who complied throughout the run-in period (n = 39,876) were randomized to receive either 100-mg aspirin every other day (n = 19,934) or matching placebo (n = 19,942). The run-in design yielded highly impressive long-term complete follow-up rates of 97.2% and 99.4% for complete morbidity and mortality data, respectively. The randomized participants would generally be included in an intention-to-treat analysis.[10]

Randomized Consent Design

Randomized consent design is proposed by Zelen[11,12] as a method of randomizing participants before obtaining consent to enhance recruitment to clinical trials. Zelen's design, also known as the postrandomization consent design, has 2 variants. One is the single-consent design,[11] and the other is the double-consent design.[12] Figure 3 shows the diagram for the single-consent design. In this design, the participants receiving standard care need not be consented for participation in the study. On the other hand, consent is only sought for participants randomized to the experimental group. If participants decline consent, they will receive the standard care instead but analyzed with the experimental group. The design provides unbiased response to patient preference if analysis is done by intent-to-treat analysis.

The single-consent design has been criticized because of lack of consent because participants are randomized before consent and subjects receiving standard care are included without informed consent of their participation in the trial. Clinicians are comfortable with the design because clinicians only seek consent for a treatment without the uncertainty of randomization. Patients will not have the uncomfortable feeling about the uncertainty of the treatment they will receive. The disadvantages of the design are contamination due to crossing over between treatment arms and lack of allocation concealment. Because the treatment is known to participants in this design, contamination is more likely due to crossing over between treatment arms. The design is likely to yield bias because the treatment is known to participants.

In the double-consent design, participants in both treatment groups are informed of the trial. Consent is sought from all participants. If participants refuse the consent for the treatment to which they were randomized, they are allowed to receive the opposite treatment.

Zelen's design was used to determine the effectiveness of a multidisciplinary stroke education program for patients and their informal caregivers[13] and to evaluate the effect of contact with a

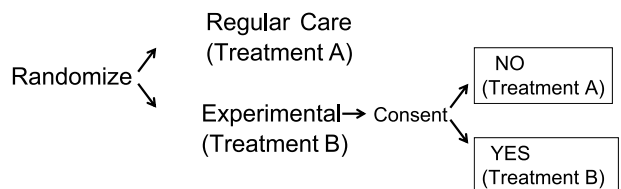**FIGURE 2.** Run-in design.

FIGURE 3. Randomized consent design.

stroke family care worker on the physical, social, and psychological status of stroke patients and their caregivers.[14]

Cluster Randomization Design

Cluster randomization trials are experiments in which clusters of individuals (hospitals and communities) rather than independent individuals are randomly allocated to intervention groups.[15] In cluster randomization, the unit of randomization is clusters, whereas the unit of analysis is observations within a cluster. The lack of independence among individuals in the same cluster, that is, between-cluster variation, creates special methodological challenges in both design and analysis. In some studies such as imaging studies, patients are clusters and lesions within each patient are observations. Therefore, patients rather than lesions are randomly allocated to intervention groups.

Suppose that n is the number of subjects per group and m is the number of observations in each subject. Then, the effective sample size (ie, the number of observations) in each group is $nm / (1 + [m - 1]\rho)$, where ρ is the intracluster correlation. The effective sample size is nm or n for $\rho = 0$ and 1, respectively. Because the effective sample size depends on the intracluster correlation, the sample size for the cluster randomization design should be larger than that of the individual randomization. The required sample size depends on the degree of intracluster correlation (ρ) and on the mean cluster size (m). Application of standard sample size approaches leads to an underpowered study. Application of standard statistical methods generally tends to bias P values downward; that is, it could lead to spurious statistical significance.

In a clinical trial, suppose that a total of 136 stroke patients are randomly allocated to either cilostazol or placebo groups. The study enrolled patients who had symptomatic stenosis in the M1 segment of 3 middle cerebral arteries. The extent of stenosis of 3 arteries in each patient was classified into 5 grades. The total number of arteries evaluated is 204 (= 68 × 3) for each group. Suppose that the end point of the study is the change in the extent of stenosis of 3 arteries, whereas the unit of randomization is a stroke patient. If 3 arteries are completely dependent, then the intracluster correlation (ρ) is equal to 1 and the effective number of arteries is 68. When 3 arteries are independent, then the effective number of arteries is 204 (= 68 × 3) with $\rho = 0$. For $0 < \rho < 1$, the effective number of arteries is $204 / \{1 + (3 - 1)\rho\}$, where 3 is the number of arteries examined for each patient.

Cluster randomization design has the advantages of administrative convenience, ease of obtaining cooperation of investigators, enhancement of subject compliance, and avoidance of treatment contamination. The disadvantage of the design includes the loss of statistical efficiency and the need to recruit more study participants due to intracluster correlation within clusters.

Cluster randomization design has been used to determine whether clinical pathways could improve the quality of the care provided to the stroke patients in a hospital and through the continuum of the care.[16] The Promoting Acute Thrombolysis for Ischemic Stroke trial used a cluster randomization to compare the effects of regular and high-intensity implementation strategies for intravenous thrombolysis in acute ischemic stroke.[17]

Superiority, Noninferiority, and Equivalence Design

Most RCTs are superiority trials that aim to determine whether a new treatment is superior to the standard treatment. By contrast, equivalence trials[18] seek to determine whether a new treatment is therapeutically similar to an existing treatment with the treatment effect being between $-\Delta$ and Δ, where Δ is the preset margin of the treatment effect. Noninferiority trials aim to determine whether a new treatment is not worse than a standard treatment by more than a preset margin (Δ). In a noninferiority trial, superiority of the new treatment would be a bonus. An equivalence trial asks, "Can I say that the response rate lies within 5% of each other for these 2 therapies with 95% certainty?" A noninferiority trial asks, "Can I say that the new therapy has response rate no worse than 5% than the standard therapy with 95% certainty?"

Here, we briefly discuss noninferiority trials. However, the same principle can be also applied to 2-sided equivalence trials. A new treatment generally has some advantages, for example, greater availability, reduced cost, less invasiveness, fewer adverse effects, or greater ease of administration in noninferiority trials. The issue then comes down to the proper choice of how much worse than the standard treatment or active control is clinically acceptable. Before undertaking the trial, the preset margin (Δ) that is clinically relevant must first be established. The smallest unacceptable degree of clinical inferiority of the new treatment must be prospectively defined. In acute myocardial infarction (MI) studies, this has traditionally been a 1% difference in mortality, a difference which resulted in changes in practice patterns following the Global Utilization of Streptokinase and TPA for Occluded Coronary Arteries 1 guidelines.[19] This is an approximate 15% relative reduction in mortality.

The investigators might interpret data as showing that the 2 treatments are equivalent if the null hypothesis is not rejected in a superiority trial. This approach is potentially flawed because the study may have an insufficient number of patients to test the hypothesis. To test equivalence or noninferiority, the investigators have to estimate the sample size to test equivalence or noninferiority.

Factorial Design

Factorial design permits researchers to evaluate 2 or more interventions in a single experiment and permits the assessment of interactions among treatments. For example, in a 2 × 2 factorial design used in the Clopidogrel Optimal Loading Dose Usage to Reduce Recurrent Events and Optimal Antiplatelet Strategy for Interventions trial,[20] shown in Table 1, patients are randomly assigned to 1 of 4 groups. The Clopidogrel Optimal Loading Dose Usage to Reduce Recurrent Events and Optimal Antiplatelet Strategy for Interventions trial evaluated the efficacy and safety of a high clopidogrel-loading dose regimen compared with the standard regimen and a high-dose aspirin compared with a low-dose aspirin. The primary outcome of the trials is the composite end point of death, MI, or stroke up to day 30.

Factorial design permits the full sample to be used to estimate 2 treatment effects in the absence of interaction. That is, factorial design is ideal when the 2 treatments act independently. Factorial design tests main effects assuming no interaction and often has inadequate power to test for interaction.

The Fourth International Study of Infarct Survival[21] used a 2 × 2 × 2 factorial trial assessing the efficacy of oral captopril,

TABLE 1. 2 × 2 Factorial Design

	C: High Dose	C: Standard Dose
A: High Dose	A (high), C (high)	A (high), C (standard)
A: Low Dose	A (low), C (high)	A (low), C (standard)

A indicates acetylsalicylic acid; C, clopidogrel.

oral mononitrate, and intravenous magnesium sulfate in 58,050 patients with suspected acute MI. There were no significant interaction effects among the treatments. Each main effect was compared using approximately 29,000 treated and 29,000 control patients. Captopril was associated with a small but statistically significant reduction in 5-week mortality. Mononitrate and intravenous magnesium sulfate did not significantly reduce 5-week mortality. Factorial design has been also used for the Physician's Health Study[22] and the Canadian Cooperative Stroke Study.[23]

Adaptive Design

The Pharmaceutical Research and Manufacturers of America working group[24] provided a formal definition of adaptive design as "…a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial." The European Agency for the Evaluation of Medicinal Products[25] defined the adaptive design as follows: "A study design is called 'adaptive' if statistical methodology allows the modification of a design element (eg, sample-size, randomization ratio, number of treatment arms) at an interim analysis with full control of the type I error." Adaptive design uses accumulating data to determine how to modify aspects of the trial. Adaptive design aims to enhance the trial, not to remedy an inadequate planning. Recently, adaptive design has received much attention from clinical investigators and biostatisticians.

The major advantage of adaptive design is the flexibility that allows adaptations or modifications of the trial after the initiation without undermining the validity and the integrity of the trial. Although flexibility is a major advantage of the adaptive design, the design may cause bias and consequently has an impact on statistical inference on assessing treatment effect.

Adaptive design was used for dose-finding studies for stroke clinical trials. The Neuroprotection with Statin Therapy for Acute Recovery Trial used an adaptive phase 1 dose-escalation design to determine the highest dose of lovastin that can be administered with a risk of mytotoxicity or hepatotoxicity lower than 10%.[26,27] The statistical design uses an adaptive design, the continual reassessment method,[28] to find the optimal dose level.

Flexible Bayesian methods were explored for phase 2 dose-finding studies in the Acute Stroke Therapy by Inhibition of Neutrophils clinical trial.[29–31] The Acute Stroke Therapy by Inhibition of Neutrophils was built on the Bayesian response-adaptive design, which was approved by the regulatory authorities based on the extensive pretrial simulations. The trial randomly allocated patients either to placebo or to 1 of 15 doses of the neutrophil inhibitory factor UK-279,276 to test the efficacy of UK-279,276 when given within 6 hours of an acute ischemic stroke. The study also estimated the ED95 target dose to be used in a confirmatory trial.

Crossover Design

In a crossover design, patients receive different sequence of treatments. The simplest crossover design is a 2 × 2 crossover design that has 2 sequences of treatments administered at 2 different periods. All the patients randomized to sequence 1 received treatment A in the first period and treatment B in the second period. All the patients randomized to sequence 2 received treatment B in the first period and then treatment A in the second period. Often, there is a washout period between the 2 periods, during which they receive no treatment. In a crossover design, the order of administering the treatment is randomized. In crossover trials, treatment effects are estimated using within-patient differences, not between-patient differences.

Crossover trials have the advantage of potentially reducing variability because each subject acts as his or her own control. A crossover trial will require less sample size than a comparable parallel design because the within-subject variability is usually smaller than the between-subject variability, and within-subject responses to treatment are usually positively correlated. The chance of carryover effects is a potential problem in crossover trials. Carryover effects can cause treatment by period interactions, which means that the treatment effect is not constant over time (ie, in the different treatment periods). Thus, a washout period with the subject off treatment is required so that the effect of the earlier treatment is not influencing the results for the next treatment. Sometimes, this constraint can make recruiting difficult, and there can be other logistic issues that can make it infeasible.

A crossover design was used to compare the effects of losartan and angiotensin-converting enzyme inhibitor quinapril on the nocturnal decrease in blood pressure and sympathetic nervous activities in hypertensive patients with a previous history of stroke.[32] Patients were randomly allocated to receive either losartan or quinapril once daily for 4 weeks, and then switched to the opposite drug for an additional 4 weeks. A crossover design was used to evaluate a day service for subjects who had a stroke.[33] Subjects were randomly allocated to attendance service or no attendance groups for 6 months and then switched to the opposite group.

DISCUSSION

Randomized clinical trials are considered as important means of advancing our knowledge of stroke interventions. Randomized clinical trials have been used for dose-finding, treatment, and educational studies. There are many aspects to be considered for the design of RCTs such as randomization, blinding, stratification, blocking, sample size estimation, multiple comparisons, and multiplicity issues. However, in this paper, we only discuss the types of RCT designs used in stroke clinical trials and discuss the advantages and the disadvantages of each RCT design. The uncertainty regarding the efficacy of many interventions stresses the need for high-quality RCTs. To promote high-quality studies, clinicians should know variations in the type of RCTs and use alternative RCT designs when conventional trials would not be feasible or suitable.

REFERENCES

1. Lachin JM, Matts JP, Wei LJ. Randomization in clinical trials: conclusions and recommendations. *Controlled Clinical Trials*. 1988; 9:365–374.

2. Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall; 1991.

3. Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. *BMJ*. 1995;311:1621–1625.

4. Schulz KF, Grimes DA. Sample size calculations in randomized trials: mandatory and mystical. *Lancet*. 2005;365:1348–1353.

5. Weijer C, Shapiro SH, Glass KC. For and against: clinical equipoise and not the uncertainty principle is the moral underpinning of the randomized controlled trial. *BMJ*. 2000;321:756–758.

6. Kwon SU, Cho YJ, Koo JS, et al. Cilostazol prevents the progression of the symptomatic intracranial arterial stenosis. *Stroke*. 2005;36:782–786.

7. Pablos-Méndez A, Barr G, Shea S. Run-in periods in randomized trials. Implications for the application of results in clinical practice. *JAMA*. 1998;279:222–225.

8. Hatazawa J, Shimosegawa E, Osaki Y, et al. Long-term angiotensin-converting enzyme inhibitor perindopril therapy

improves cerebral perfusion reserve in patients with previous minor stroke. *Stroke*. 2004;35:2117–2122.

9. Ridker PM, Cook NR, Lee IM, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *N Engl J Med*. 2005;352:1293–1304.

10. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the on-going Physicians' Health Study. *N Engl J Med*. 1989;321:129–135.

11. Zelen M. A new design for randomized clinical trials. *N Engl J Med*. 1979;300:1242–1245.

12. Zelen M. Randomized consent designs for clinical trials: an update. *Stat Med*. 1990;9:645–656.

13. Rogers H, Atkinson C, Bond S, et al. Randomized controlled trial of comprehensive stroke education program for patients and caregivers. *Stroke*. 1999;30:2585–2591.

14. Dennis M, O'Rourke S, Slattery J, et al. Evaluation of a stroke family care worker: results of a randomized controlled trial. *BMJ*. 1997;314:1071–1076.

15. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. London: Arnold; 2000.

16. Panella M, Marchisio S, Barbieri A, et al. A cluster randomized trial to assess the impact of clinical pathways for patients with stroke: rationale and design of the clinical pathways for effective and appropriate care study [NCT00673491]. *BMC Health Serv Res*. 2008;8:223.

17. Dirks M, Niessen LW, Huijsman R, et al. Promoting acute thrombolysis for ischemic stroke (PRACTISE). *Int J Stroke*. 2007;2:151–159.

18. Wellek S. *Testing Statistical Hypotheses of Equivalence*. Boca Raton, FL: Chapman & Hall/CRC; 2003.

19. The GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *N Engl J Med*. 1993;329:673–682.

20. Mehta SR, Bassand JP, Chrolavicius S, et al. On behalf of the CURRENT OASIS 7 Steering Committee. Design and rationale of CURRENT-OASIS 7: a randomized, 2 × 2 factorial trial evaluating optimal dosing strategies for clopidogrel and aspirin in patients with ST and non–ST-elevation acute coronary syndromes managed with an early invasive strategy. *Am Heart J*. 2008;156:1080–1088.

21. ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. ISIS-4: a randomized factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet*. 1995;345:669–685.

22. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med*. 1989;321:129–135.

23. The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. *N Engl J Med*. 1978;299:53–59.

24. Gallo P, Chuang-Stein C, Dragalin V, et al. Adaptive design in clinical drug development—an executive summary of the PhRMA Working Group (with discussions). *J Biopharm Stat*. 2006;16:275–283.

25. EMEA. Reflection paper on methodological issues in confirmatory clinical trials with flexible design and analysis plan. Doc. Ref. CHMP/EWP/2459/02. Available at: http://www.emea.europa.eu/pdfs/human/ewp/245902en.pdf. Accessed December 1, 2009.

26. Elkind MS, Sacco RL, MacArthur RB, et al. The Neuroprotection with Statin Therapy for Acute Recovery Trial (NeuSTART): an adaptive design phase I dose-escalation study of high-dose lovastatin in acute ischemic stroke. *Int J Stroke*. 2008;3:210–218.

27. Elkind MS, Sacco RL, Macarthur RB, et al. High-dose lovastatin for acute ischemic stroke: results of the phase I dose escalation neuroprotection with statin therapy for acute recovery trial (NeuSTART). *Cerebrovasc Dis*. 2009;28:266–275.

28. O'Quigley J, Pepe M, Fisher M. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics*. 1990;46:33–48.

29. Krams M, Lees KR, Hacke W, et al. 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;34:2543–2548.

30. Krams M, Lees KR, Berry DA. The past is the future: innovative designs in acute stroke therapy trials. *Stroke*. 2005;36:1341–1347.

31. Grieve AP, Krams M. ASTIN: a Bayesian adaptive dose-response trial in acute stroke. *Clin Trials*. 2005;2:340–351. Discussion 352–358, 364–378.

32. Okuguchi T, Osanai T, Fujiwara N, et al. Effect of losartan on nocturnal blood pressure in patients with stroke: comparison with angiotensin converting enzyme inhibitor. *Am J Hypertens*. 2002;15:998–1002.

33. Corr S, Phillips CJ, Walker M. Evaluation of a pilot service designed to provide support following stroke: a randomized cross-over design study. *Clin Rehabil*. 2004;18:69–75.