Novel Methodologic Approaches to Phase I, II, and III Trials

Sharon D. Yeatts, PhD

Among the future directions of stroke research identified by the Stroke Progress Review Group was a call for improved trial design, conduct, and outcome assessment. This demand is underscored by recent trial findings. Of the >100 clinical trials in ischemic stroke published during the first decade of the 21st century, only 5 of 31 phase III trials demonstrated efficacy on the primary outcome.2 As the number of experimental therapies undergoing clinical investigation in stroke increases, there is a rising need for more efficient statistical designs to (1) determine the optimal dose, (2) identify interventions with therapeutic potential, and (3) evaluate clinically relevant treatment effects.

The correct selection of the optimal dose, via a thorough understanding of the dose–toxicity and dose–efficacy relationships, is critical for establishing the efficacy of a potential therapeutic agent.3 Traditional rule–based dose-escalation designs, such as the 3+3 design and its variations, have suboptimal statistical operating characteristics. The inefficiency in the 3+3 design is that it stems from the decision to escalate or de-escalate, a decision is based solely on event data from the current dose, without considering event information available from neighboring doses. An alternative design called the Continual Reassessment Method4 is an adaptive dose-finding algorithm, wherein escalation or de-escalation through the dose region is determined by continuous re-estimation of the dose–toxicity curve, and each cohort of subjects is treated at the current dose currently believed to be the maximum tolerated dose. Although computationally more intensive than the 3+3 design, the Continual Reassessment Method and its variations use all available toxicity data in the estimation of the maximum tolerated dose, and thus are statistically more efficient. Ethical considerations also favor the Continual Reassessment Method, because the Continual Reassessment Method typically treats fewer patients at subtherapeutic doses.

Once the appropriate dose has been determined through a phase I trial, the next step is to evaluate efficacy potential in a phase II trial. The traditional concurrently controlled phase II design, intended to simultaneously estimate treatment effect and assess variability, is often criticized as an underpowered phase III comparative clinical trial. Proposed alternative designs, such as the selection design and the futility design, are instead intended to weed out ineffective or mediocre therapies.

The objective of the futility design, which has been successfully implemented in stroke,5–6 is to discard treatments that do not show promise. Statistical hypotheses are stated such that the goal is to demonstrate that the intervention is futile. Failure to conclude futility would be considered evidence in favor of a need for a definitive phase III clinical trial.

In the single arm futility design,7 the experimental treatment arm is compared with a target response rate, \( \pi_t \), defined as the minimum proportion of successes in the treated group which would warrant further study. If the true success proportion \( \pi \) is less than \( \pi_t \), the intervention is declared futile. Comparison of the experimental arm with a target response rate, which is fixed and has no variability, results in a smaller sample size than would be required for direct comparison with a concurrent control arm. The target response rate can be determined based on a clinically relevant treatment effect and historical control data.

Concerns over the use of historical control data include temporal changes in outcome associated with improvements in patient management, as well as variations in eligibility criteria, protocol adherence, and primary outcome measures across clinical trials. If the historical control data are outdated and therefore no longer relevant, trial results may not reflect an accurate assessment of the futility of the experimental treatment. Inclusion of a small cohort of control subjects for calibration of the threshold value has been suggested.8 If this cohort is too small, its usefulness in terms of calibration is quite limited; if this cohort is too large, the trial begins to resemble an underpowered phase III.

The inclusion of a concurrent control arm in the futility design avoids the drawbacks of historical control data and allows for a direct comparison of treatment arms. The futility hypothesis is based on a direct comparison of the randomized treatment arms, such that futility would be declared if the absolute treatment effect is less than \( \delta \), a prespecified clinically meaningful futility margin, in favor of the experimental treatment. The sample size is increased over the single-arm design, but the concurrently controlled futility design is not an alternative to phase III efficacy testing. The objective remains to establish futility, rather than to demonstrate efficacy, of the active treatment. The National Institute of Neurologic Disorders and Stroke-funded phase II trial of deferoxamine mesylate in intracerebral hemorrhage (Hi-Def in ICH [intracerebral hemorrhage], clinicaltrials.gov NCT01662895) uses this design.

Selection designs7 can be used to prioritize candidate interventions, such that resources are allocated to the most promising of candidates. In a selection design, the objective is to...
select the best among K interventions (or K interventions and a control) for further testing. In the selection design, subjects would be randomized to one of the K interventions. The best intervention is defined as the intervention with the numerically, rather than statistically, highest response rate. The sample size is determined to ensure that, if the best treatment is superior by at least some margin D, then the best treatment will be selected with high probability.

The selection design can be combined with a futility or superiority test in a sequential 2-stage design, as described in the trial of Co-Q10 in ALS (amyotrophic lateral sclerosis). At the conclusion of stage 1, a treatment would be selected, and the statistical hypothesis tested at the end of stage 2. Inclusion of stage 1 subjects in the statistical hypothesis test has the advantage of using all available outcome data but introduces bias, which must be accounted for in the test statistic. If the stage 1 subjects are excluded from the stage 2 hypothesis test, the parameter estimate is unbiased, but the overall sample size is increased.

Adaptive designs promise increased flexibility of the trial to respond to accumulating information, a promise which has generated both enthusiasm and confusion. According to the Food and Drug Administration draft guidance, an adaptive design “includes... a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study.” It is important to emphasize that the potential issues, and in what manner the design stage to maintain trial validity.

Although certainly not novel, group sequential methods are adaptive according to this definition, in that they allow the trial to be stopped early, based on accumulating data, in the face of overwhelming efficacy or futility. Other valid mechanisms for adaptation in phase III trials include blinded sample size re-estimation and covariate adaptive randomization. Adaptations based on an unblinded assessment of interim data, including sample size re-estimation and response adaptive randomization, may be more enticing but may also be more controversial in the confirmatory setting. Early phase designs allow for more flexibility with regard to adaptation, and adaptive designs have gained greater acceptance in this exploratory setting. Whether exploratory or confirmatory, Bayesian or Frequentist, each trial must demonstrate that the statistical operating characteristics remain sound in the face of the chosen adaptation(s).

Even a well-designed trial of an effective therapy can fail based on an inappropriate primary outcome or primary analysis. The selection of valid, reliable, and efficient outcome measures and analytic strategies for stroke clinical trials is receiving much attention in the stroke literature. The modified Rankin Scale reflects a global assessment of function, but a patient may consider many other life aspects in describing outcome after a stroke. Patient-reported outcomes, including the NeuroQOL and PROMIS tools, may provide important information regarding quality of life, cognition, and social functioning. Consideration of these varying aspects of global outcome may provide a more complete understanding of the evolution of stroke and allow a more sensitive estimate of treatment effect.

These are just a sampling of the novel approaches being considered, and in some cases implemented, in current stroke trials. Continued development of innovative trial designs, outcome assessments, and analytic approaches is an essential component of stroke research.

Disclosures
Dr Yeatts is the SDMC PI (Statistics and Data Management Center Principal Investigator) for the phase II trial of deferoxamine in ICH (Hi-Def; U01 NS074425). She is the unblinded statistician for the IMS 3 (U01 NS077304) and ProTECT (U01 NS062778; Bio-PROTECT R01 NS071867) phase III clinical trials. The other authors have no conflicts to report.

References
Novel Methodologic Approaches to Phase I, II, and III Trials
Sharon D. Yeatts

*Stroke.* 2013;44:S116-S118
doi: 10.1161/STROKEAHA.111.000031
*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/44/6_suppl_1/S116

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Stroke* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Stroke* is online at:
http://stroke.ahajournals.org//subscriptions/