To Phase 3 or Not to Phase 3?

Warren W. Wasiewski, MD

Decisions to progress new chemical entities, biologicals, or new treatment modalities into Phase 3 are usually based on small Phase 2 studies that are designed to assess safety and to explore efficacy. Based on these Phase 2 safety studies, both large and small companies have launched large Phase 3 programs that include 2 concurrently run studies, ie, Stroke-Acute Ischemic NXY Treatment (SAINT) I and SAINT II ASP I and Ancrod Stroke Program (ASP) II. When Phase 3 studies have failed to demonstrate efficacy for the whole population under study, subgroup analysis to identify patients who may have benefited from treatment has led to additional Phase 3, subsequently negative studies. In part as a result of this approach, the number of failed studies for treatment of acute ischemic stroke, especially for neuroprotection, has become legendary.

In this issue, Mandava and Kent present a model to predict if the results of a single study are different from a pooled control population. The model calculates expected outcome using a pooled control population and the \( \pm 95\% \) “prediction interval surfaces” using the baseline variables of the National Institutes of Health Stroke Scale and age. The outcome variables assessed in the model are mortality and the 90-day modified Rankin Scale (mRS) score of 0 to 2. Using this model, they correctly predict the recent failure of SAINT II based on the results of SAINT I because the treatment group outcome in SAINT I was no different from the outcome of the pooled control population. They also demonstrated that the multiple Abciximab Emergent Stroke Treatment Trial (AbESST) studies were not likely to be successful based on the first AbESST study. The strength of their model is demonstrated by the output related to the National Institute of Neurological Diseases and Stroke recombinant tissue plasminogen activator study. The combined study results had an outcome greater than the \( \pm 95\% \) prediction interval surface suggesting that recombinant tissue plasminogen activator is an effective treatment. When the study was evaluated in its 2 parts, the results of Part 1 exceeded the \( \pm 95\% \) prediction interval surface suggesting that Part 2 would be successful. Indeed, Part 2 also exceeded the \( \pm 95\% \) prediction interval surface.

Although the model successfully predicts outcomes of several studies using a primary outcome measure of mRS score of 0 to 2 at Day 90, this is not the primary end point of a number of the evaluated studies. The SAINT I study is a clear example of this point. The primary outcome variable for SAINT I was the mRS score at Day 90 analyzed across the whole distribution of scores. The primary end point in SAINT I was met, yet using this model with the mRS score of 0 to 2 as the end point, the study is negative. On the opposite end of the spectrum is the CLOT BUST study, which was negative on the primary end point, mRS 0 to 1 at Day 90. In the model using mRS 0 to 2, the results fall above the \( +95\% \) prediction interval surface as does the author’s primary, mRS 0 to 1. The model also predicts that the laser treatment used in the NeuroThera Effectiveness and Safety Trial-1 (NEST I) study is a potentially effective treatment. This is an interesting finding in that the NEST I study is a small study, 120 patients, with a long time to treat in both groups, \( > 16 \) hours, and with treatment randomized 2:1. As a result of the sample size and the randomization scheme used, the control group is only 41 patients. Despite the small sample size, which is likely to have some imbalance in baseline characteristics and even a smaller control group in which the imbalances may be more evident, the primary efficacy end point, National Institutes of Health Stroke Scale score 0 to 1 or a 9-point improvement from baseline, was met as were a number of secondary efficacy end points. This result suggests a second confirmatory study is justified. It will be interesting to see if the NEST II study is positive or if this is a false-positive result for the model.

How this model may be used to progress new therapies remains to be seen. Certainly an application of this model to the first of 2 Phase 3 studies seems appropriate. Whether this model is robust enough to influence the decision of “big Pharma” or “biotech” to suspend development of what is thought to be a “promising” therapy is unclear.

This model highlights other important aspects of clinical trials in stroke. The number of different primary end points used across the numerous stroke studies appears to be a search for the Holy Grail. The National Institute of Neurological Diseases and Stroke study should have taught us that the choice of end point is less important than the actual efficacy of the compound. In that study, all 4 components of the global test static were statistically significant. As more exclusive patient selection techniques are applied to clinical trials, ie, angiography or mismatch on MRI, one has to wonder how far away these populations will migrate from the pooled control group of previous studies. It is unlikely that the control group of a randomized trial that requires demonstration of M1 or M2 occlusion is going to be similar to the pool control group in this model. Therefore, the stroke community needs to consider a unified approach to end point selection that will permit more direct comparison of new treatments as they emerge, eg, an mRS score of 0 to 2 or a full-scale analysis of the mRS adjusted for baseline severity.

See related article, pages 1803–1810.
Mandava and Kent\(^5\) have developed an interesting and provocative model to predict outcome of subsequent trials based on results of a single trial in comparison to a large control population. Although the model has its limitations and may not be the primary assessment tool used to determine if a compound is to proceed further in development, it certainly gives one reason to pause and re-evaluate whether or not to continue. As more trials are completed and the model is used to prospectively predict success or failure, the usefulness of this approach will be more clearly defined.

**Disclosures**

None.

**References**


**Key Words:** clinical trials ■ statistical models ■ stroke care
To Phase 3 or Not to Phase 3?
Warren W. Wasiewski

Stroke. 2009;40:1553-1554; originally published online March 12, 2009;
doi: 10.1161/STROKEAHA.108.544429
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 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/40/5/1553

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/