Comments Regarding the Recent OAST Article

To the Editor:

A recent article by the Optimizing Reanalysis of Stroke Trials (OAST) Collaboration analyzed 55 stroke datasets with multiple tests and ranked them in order of the standardized test statistic; these ranks were then analyzed with a 2-way ANOVA.1 Although this article raises important issues, we have concerns regarding the methods.

Firstly, ANOVA assumes approximate normality of the response outcome, which the ordinal ranks are not. ANOVA also assumes that the responses are independent, but this is not the case here, unless we have misunderstood the method. Analyzing the same datasets with multiple tests will induce a positive correlation among the test results; ignoring this correlation generally results in smaller probability values than warranted.

We believe the Friedman test (a nonparametric analog of 2-way ANOVA) would be more appropriate because it does not assume either normality or independence.2 Although this would likely not change the overall conclusions of the OAST work, it would yield more appropriate probability values.

The authors also note with concern the problems of arbitrarily dichotomizing outcome scales (eg, modified Rankin Scale [mRS]) by noting that the ECASS II trial showed no treatment effect when dichotomized at 0 to 1 versus 2 to 6, but showed a positive effect when dichotomized at 0 to 2 versus 3 to 6. They cite 3 published reports, each of which used a different cutpoint for the mRS. We share the concerns on the arbitrariness; however, different cutpoints reflect different levels of clinical judgment of the expected risks and benefits of the intervention.

Although analyzing ordinal data as ordinal rather than dichotomizing may increase power, this does not guarantee significance. For example, using 90-day mRS score, the Stroke-Acute Ischemic-NXY Treatment I (SAINT I) trial reported a significant treatment effect of NXY-059 when using a proportional odds model.3 In a secondary dichotomized analysis some cutpoints would yield more appropriate probability values. Regarding the methods.

The authors also note with concern the problems of arbitrarily dichotomizing outcome scales (eg, modified Rankin Scale [mRS]) by noting that the ECASS II trial showed no treatment effect when dichotomized at 0 to 1 versus 2 to 6, but showed a positive effect when dichotomized at 0 to 2 versus 3 to 6. They cite 3 published reports, each of which used a different cutpoint for the mRS. We share the concerns on the arbitrariness; however, different cutpoints reflect different levels of clinical impairment, and presumably each trial makes a decision based on clinical judgment of the expected risks and benefits of the intervention.

Although analyzing ordinal data as ordinal rather than dichotomizing may increase power, this does not guarantee significance. For example, using 90-day mRS score, the Stroke-Acute Ischemic-NXY Treatment I (SAINT I) trial reported a significant treatment effect of NXY-059 when using a proportional odds model.1 In a secondary dichotomized analysis some cutpoints were significant, while others were not. Importantly, the investigators did not provide a justification for the different cutpoints used, nor did they account for multiple analyses. A subsequent follow-up trial with nearly double the sample size (n=3306 in SAINT II compared with n=1722 in SAINT I) found no evidence of a significant treatment effect.4 Furthermore, use of the proportional odds model requires the strong assumption that any choice of cutpoint would result in the same odds ratio. As the authors point out, several of the studies examined failed the test of proportionality. Although the proportional odds model is believed to be robust to violations of proportionality, little to no work has been published verifying this belief.

Finally, the authors suggest that the robustness of t tests to non-normal data negates concerns about its use when the data appear non-normal. However, there is a clear bi-modality to the mRS outcome data from SAINT I (Figure).3 Using the mean is therefore questionable, and nonparametric analogs are preferable. Indeed, the SAINT I data were not significant when analyzed nonparametrically.5 We caution against an over-reliance on “robustness” of methods such as t tests or the proportional odds model to the exclusion of other tests. The primary end point(s) of a trial should be determined based on clinical justification, and the appropriate method of analysis selected based on this end point, not the perceived strength of a given statistical test.

Disclosures

None.

Comments Regarding the Recent OAST Article
Scott W. Miller and Yuko Y. Palesch

Stroke. 2008;39:e14; originally published online November 29, 2007;
doi: 10.1161/STROKEAHA.107.500355
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 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/39/1/e14

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/