Subarachnoid hemorrhage trials

Ali Zomorodi, M.D., James E. Herndon II, Ph.D., and John H. Sampson, M.D., Ph.D., M.H.Sc., M.B.A.

1Division of Neurosurgery, Department of Surgery; Departments of 2Pathology, 3Immunology, 4Radiation Oncology, and 5Biostatistics and Bioinformatics; and 6Preston Robert Tisch Brain Tumor Center, Duke University Medical Center, Durham, North Carolina

In this issue of the Journal of Neurosurgery, Ilodigwe et al.1 use advanced statistical techniques to understand the implications of dichotomizing outcomes in clinical studies whose outcomes are ordinal. Ordinal variables rank or order outcomes but do not measure the relative differences between outcomes. A classic example of an outcome that is an ordinal variable is the Glasgow Coma Scale (GCS) score. In many clinical studies, these ordinal variables are collapsed into 2 outcomes. Again, using the GCS as an example, we could dichotomize head injuries into severe or not severe based on a GCS score ≤ 8 or > 8, respectively. When such outcomes are dichotomized, information is lost and statistical comparisons lose power. The authors of the following paper attempt to make this point by analyzing several recent randomized clinical trials that were conducted in patients with subarachnoid hemorrhage (SAH).

Clinical outcomes in cerebrovascular surgery are measured using well-established ordered scales such as the Glasgow Outcome Scale (GOS) or the modified Rankin Scale (mRS). For the purposes of statistical analysis, most trials collapse these scales into the 2 categories of favorable and unfavorable outcome. While such dichotomization facilitates data analysis, it has several problems, which were well described by Murray et al.3 in 2005. First, this approach creates an artificial threshold for efficacy that may not be clinically relevant. For example, a treatment that enables patients with Hunt and Hess Grade V SAH to consistently achieve an mRS score of 3 would not cross the boundary between favorable and unfavorable, but it would nonetheless be very worthwhile. Second, and more importantly, the dichotomous approach is based on the assumption that every patient within the control group has the same chance of attaining a favorable outcome. This assumption ignores the prognostic heterogeneity that is an essential feature of patients with aneurysmal SAH. The considerable work done to establish admission injury severity and prognostic scales in aneurysmal SAH has demonstrated that very few patients at either extreme of the admission injury severity scale would be expected to cross the outcome boundary established for clinical trials. These patients would contribute very little to the overall power of a study and might even make the study insensitive to some meaningful outcomes. For example, as demonstrated by Machado et al.,2 if prognostic heterogeneity is ignored, a treatment would have to have an OR of 1.75 for a favorable outcome to be observed in a typical study powered to detect a 10% absolute increase in favorable outcomes. This may explain why so many interventions with promising preclinical performance have failed to demonstrate benefit in clinical trials.

In their paper, Ilodigwe and colleagues review the individual patient data from 4 tirilazad trials, the CONSCIOUS-1 trial, and the ISAT. All of these trials used the fixed dichotomy approach for analyzing outcomes. Importantly, the tirilazad and clazosentan trials demonstrated significant improvement in vasospasm rates in the treatment arm without significant increases in the number of favorable clinical outcomes. As such, these trials raise the same questions described above and are well suited for reevaluation using alternative methods.

The authors evaluated individual patient data from each trial using the traditional fixed dichotomy approach as well as the proportional odds model and the sliding dichotomy approach. The proportional odds model is an extension of logistic regression that allows the ordered categorical outcome to take on more than 2 possible outcomes. Within this model, all fixed dichotomizations are considered within individual logistic models under the assumption that treatment has the same effect within each dichotomization. The latter is a strong assumption that was violated for many of the analyses considered in their paper.

The authors also used the sliding dichotomy approach proposed by Murray et al.,3 in which patients are grouped into strata (or tiers) according to baseline prognosis. Within each stratum, stratum-specific definitions of good and bad are used to dichotomize outcome. Treatment groups are then compared with respect to the number of good outcomes pooled across strata. A diagnostic score is used to define strata, and the score is estimated from a logistic model that predicts the fixed dichotomized outcome traditionally used in the analysis of these studies as a function of baseline risk factors. Although the goal of using all data without discarding information is commendable, there are challenging issues associated with this approach that must be considered. First, given that the stratification definition is data driven and not obtained from external sources, there is concern about the ability to compare study results with other trials. Murray et al.3 also acknowledged concern from the regulatory perspective that the analysis plan is not completely prespecified and is driven by the trial data. Second, given that the use...
of fixed dichotomy results in a loss of information, it is not clear how the use of that outcome in the development of the prognostic model for stratification will affect the sensitivity of subsequent treatment comparisons using the sliding dichotomy. Murray and associates provide a more comprehensive discussion of the open issues associated with this analytical approach.

In conclusion, Ildigwe and colleagues should be congratulated on their reevaluation of how we statistically analyze randomized controlled trials that are important to our field. In this particular instance, only minor differences were found among fixed dichotomy, sliding dichotomy, and proportional odds models with regard to the magnitude and precision of the odds ratios within these clinical trials. And while these approaches did not help for these particular studies, the principles the authors espoused should be duly considered in analyzing outcomes in future clinical trials important to our field.

(http://thejns.org/doi/abs/10.3171/2012.5.JNS12730)

Disclosure

The authors report no conflict of interest.

References


Response

R. Loch Macdonald, M.D., Ph.D.

Division of Neurosurgery, St. Michael’s Hospital, Labatt Family Centre of Excellence in Brain Injury and Trauma Research, Keenan Research Centre of the Li Ka Shing Knowledge Institute of St. Michael’s Hospital, and Department of Surgery, University of Toronto, Ontario, Canada

We thank Zomorodi and colleagues for their comments on our study published in the Journal of Neurosurgery. They mention the value of prognostic models for estimating prognostic heterogeneity and using prognostic models to select patients for clinical trials as well as to modify the method of measuring outcome. Excellent prognostic models have been developed in some neurosurgical fields, such as traumatic brain injury. The IMPACT (International Mission for Prognosis and Analysis of Clinical Trials in TBI) collaboration has published numerous papers on prognostic modeling for traumatic brain injury, and an online prognostic prediction tool is available. We would point out that the development of prognostic models for SAH has been relatively limited, especially compared to what has been done in head injury. In one of our early attempts, we found numerous factors associated with outcome, but the identified factors explained only about a third of the variation in outcome. Another limitation is that none of the SAH prognostic models is externally validated.

We believe that our paper raises several interesting questions and points for further study. First, we were disappointed in the performance of some of the alternative statistical approaches. In the ISAT there was a significant benefit in coiling over clipping as measured on the dichotomous mRS. We hypothesized that alternative methods of analyzing the mRS would increase the effect size of coiling or the difference in outcome between the groups; however, we were disappointed to see a minimal effect. There are multiple possible reasons for this but a key one may be that the outcome scale itself is not optimal for this disease. Outcome scales that incorporate measures of the deficits present in patients with SAH, such as cognitive and executive dysfunction, must be developed. Second, as we conclude in our study, additional studies can be done to determine the conditions under which these alternative methods of analyzing outcome may perform better.

References


Disclosure

The authors report no conflict of interest.

References


Please include this information when citing this paper: published online October 5, 2012. DOI: 10.3171/2012.5.JNS12730.