Special Report

The Case Against A Randomized Trial of Unruptured Brain Arteriovenous Malformations

Misinterpretation of a Flawed Study

Nicholas C. Bambakidis, MD; Kevin M. Cockroft, MD; Joshua A. Hirsch, MD; E. Sander Connolly, MD; Sepideh Amin-Hanjani, MD; Philip M. Meyers, MD; Robert M. Friedlander, MD

In November 2013, the primary results of the prospective, multicenter National Institutes of Health–funded study, A Randomized Trial of Unruptured Brain Arteriovenous Malformations (ARUBA), were published in *Lancet.* This randomized, nonblinded trial compared medical management alone with medical management plus interventional therapy in patients with unruptured brain arteriovenous malformations (BAVM). Patients were enrolled from 39 clinical sites worldwide. Those patients with previous BAVM treatment or hemorrhage were excluded as were those patients with BAVMs deemed untreatable. The primary end points of the study were symptomatic stroke or death, with a secondary outcome of death and disability as measured by the modified Rankin Scale. The initial study design called for an enrollment of 800 patients but was reduced to 400 after slow trial recruitment prompted a reassessment of the design. When after ≈6 years of enrollment, the trial was halted on the recommendation of the trial’s independent Data and Safety Monitoring Board; outcome data were available for only 223 subjects. Among the randomized subjects, the most common presenting symptoms were seizure and headache. Ninety-four patients (42%) were asymptomatic at the time of diagnosis. BAVMs in the study cohorts were relatively well matched for size, location, and venous drainage pattern, as well as for Spetzler–Martin grade. The majority of the patients were identified with Spetzler–Martin grade 1, 2, or 3 (approximately equal) with a few with grade 4 (10%), and no patients with grade 5 were enrolled. The average follow-up period was 33 months. Of the 109 patients in the medical group, 11 patients (10.1%) experienced the primary end point of death or symptomatic stroke, whereas 35 of the 114 patients in the interventional arm (30.7%) met the same end point, making the risk of stroke or death significantly lower in the medical group compared with the interventional group (hazard ratio, 0.27; 95% confidence interval, 0.14–0.54). These primary results were also associated with a pronounced difference in functional outcome, with the risk of death or serious disability (modified Rankin Scale ≥2) being significantly lower in the medical management group (8 of 53, 15%) when compared with the interventional group (24 of 52, 46%; relative risk, 0.33; 95% confidence interval, 0.16–0.66). Interestingly, the difference in the number of deaths between the 2 groups was not statistically significant, with 3 in the interventional arm and 2 in the medical management arm. When analyzed according to Spetzler–Martin grade, patients with grade 1 BAVMs fared better in the interventional group, although this trend did not reach statistical significance. The rates of adverse events were statistically much higher in patients with Spetzler–Martin grade 2 and 3; too few patients with grade 4 were enrolled to allow for analysis of differences in outcomes.

To many observers, these results were not particularly surprising. Many issues concerning the study’s design have been raised previously, and the published results only serve to heighten concerns over factors, such as selection bias, participating site characteristics, clinical practice patterns, and the exceedingly short length of follow-up. Unfortunately, now the difficulty comes in deciding how to interpret these results and apply them to everyday clinical practice. In a study of a complex heterogeneous disease using broad inclusion criteria and multiple interventions, it becomes hard to generalize these findings to the unruptured BAVM population as a whole.

One of the most significant limitations of studies such as ARUBA is selection bias. The external validity of ARUBA and its applicability to patients with BAVM in general depends on the representativeness of the study’s sample population. In the case of ARUBA, the lack of detailed screening logs makes the

Received June 20, 2014; final revision received June 20, 2014; accepted July 9, 2014.

From the Department of Neurological Surgery, University Hospitals Case Medical Center, Cleveland, OH (N.C.B.); Department of Neurosurgery, Penn State, Milton S. Hershey Medical Center, Hershey, PA (K.M.C.); Massachusetts General Hospital, Institute for Heart, Vascular and Stroke Care, Boston (J.A.H.); Department of Neurosurgery, Columbia University Medical Center, New York, NY (E.S.C., P.M.M.); Department of Neurosurgery, University of Illinois at Chicago (S.A.-H.); and Department of Neurological Surgery, University of Pittsburgh, PA (R.M.F.).

Endorsements: The American Association of Neurological Surgeons/Congress of Neurological Surgeons Joint Cerebrovascular Section, the Society of Neurointerventional Surgery, the American Society of Neuroradiology, and the European Association of Neurosurgical Societies.

Correspondence to Nicholas C. Bambakidis, MD, Department of Neurological Surgery, University Hospitals Case Medical Center, 11100 Euclid Ave, Cleveland, OH 44106. E-mail nicholas.bambakidis2@uhhospitals.org

*Stroke. 2014;45:00-00. DOI: 10.1161/STROKEAHA.114.006519*

*Stroke is available at http://stroke.ahajournals.org*
comes and expertise is critically important; in ARUBA, shown, rigorous physician accreditation in treatment outcomes of comparatively homogeneous cerebrovascular disorders have far fewer than expected. As previous studies investigating practices did not enroll any patients, whereas others recruited the 13 enrolling US centers, 3 only enrolled 1 patient. Many patients, and of these, 14 only enrolled 1 or 2 subjects. Of these 177 patients were managed outside of the randomization process (74 underwent interventional management, whereas 61 were managed medically). Unfortunately, there is no explanation as to why these eligible patients who did not refuse enrollment were not randomized. In addition, the reasons why another 323 patients refused enrollment altogether are not known. Because most patients do not simply refuse treatment without some physician interaction or counseling, more detailed information about the characteristics of these eligible patients and their lesions would be exceedingly useful.

Treatment of BAVMs has evolved along several well-established patterns of therapy. Although it may be argued that the Spetzler–Martin grade distribution in ARUBA shows a preponderance of grade 1 to 3 arteriovenous malformations when compared with other cited population-based series, making this population seems to be optimum for interventional therapy, this distribution does not necessarily reflect a group with a high natural history risk. The Spetzler–Martin grading system as originally conceived is not a tool for stratifying the risk of natural history morbidity but rather the risk of morbidity with surgical intervention, a distinction confirmed by the ARUBA investigators secondary analyses. The grading system does not assess potential risk factors for hemorrhage, such as the presence of an intranidal aneurysm or venous outflow obstruction. In addition, the grade 3 category is not particularly specific, with varying combinations of components capable of achieving the same grade. These factors make it possible that those patients thought to be at high risk for BAVM rupture by study investigators were excluded from enrollment and treated outside of the confines of the trial. The International Study of Unruptured Intracranial Aneurysms (ISUIA) demonstrated that similar limitations as high-risk aneurysms were, by their nature, preselected out of the study, leaving a group of patients enrolled in ISUIA with a lower risk for aneurysmal rupture.

Preconceived biases leading to a lack of equipoise and discomfort with natural history risks among clinicians who treat large numbers of BAVMs in tertiary care centers in the United States may have contributed to the low number of participating US centers, as well as the low study accrual rate from States may have contributed to the low number of participants uncertain. Evidence for it can be seen in the interventional choices made by various participating centers. Overall, only 5 patients were treated with surgery alone, and only 13 patients were treated with surgical resection in combination with either endovascular procedures or radiotherapy. In general, embolization is not considered a curative procedure for BAVM treatment in the United States, whereas radiosurgery requires a lag time of several years to demonstrate a significant therapeutic effect. In fact, although surgical resection offers an obliteration of ≤96%, studies have shown much lower rates of obliteration for radiotherapy (38%) or embolization alone (13%). The fact that 30 patients underwent embolization alone as intervention speaks to serious questions about the treatment decisions of the study investigators and whether experienced surgical expertise was even available at all of the treatment sites.

The use of radiosurgery as a stand-alone treatment was even more common (31 patients) than embolization, yet the study provides no specifics as to the types of radiosurgery used, the target volumes treated, or radiation doses delivered, all factors that may significantly affect BAVM obliteration. Although radiosurgery has a well-known, low, up-front morbidity rate in terms of death or symptomatic stroke, the latency to therapeutic effect and the relatively low overall obliteration rate make meaningful conclusions from a comparison to medical management in a <3-year time frame exceedingly difficult. No conclusion can be reliably made with respect to arteriovenous malformation obliteration or eradication because there are no data about confirmation of obliteration in the published results. Despite including multiple treatment modalities and combinations, the trial was not powered to examine the differences in morbidity between interventions.

Finally, the study authors correctly point out the limitation of a mean follow-up of only 33 months. When comparing intervention to medical management in a lesion with a relatively low, but life-long risk of hemorrhage and disability, the importance of this limitation cannot be overstated. The average age of the study subjects was 45 years for the interventional group and 44 for the medical management group. According to US government actuarial statistics, at age 45, the life expectancies for a US male and female are ≈38 and ≈34 years, respectively. For a disease process that spans a patient’s lifetime, it is difficult to justify the applicability of <3 years of follow-up. Although the event rate in the interventional group may have been higher than some groups expected, it should be lost on no one that at 2.2% per year, the spontaneous arteriovenous malformation rupture rate in the medical group is not insignificant and is well in line with previous natural history
studies involving unruptured BAVM. Going forward, a similar rate of continuing events would be expected in the medical group, whereas the interventional group’s events should dramatically decrease once the initial morbidity of treatment has passed. Looked at as a specific example, if one conservatively assumes a continuing event rate of 11 every 3 years in the medical group and 3 every 3 years in the interventional group, at 15 years, there will have been 55 events in the medical group and 47 in the interventional group. Although 15 years may seem like a long duration for the National Institutes of Health–funded study, it represents less than half the life expectancy of the average study subject.

In response to criticisms that the existing trial follow-up is too short, the ARUBA investigators are requesting additional funding to continue monitoring the 223 enrolled patients, potentially for ≤10 years. Given the above evidence and the enormous cost of ARUBA to date, it is important to carefully assess the usefulness of such a proposal. Although as described, longer follow-up could potentially show equality in outcomes or even the superiority of intervention, such a reversal will not alter the external validity of the trial. A crossover of outcomes such that patients managed noninvasively suffer a higher incidence of stroke and death than those treated should not lead anyone to conclude that all patients with unruptured BAVM must be treated. Instead, perhaps the best, most appropriate next step in the investigation of unruptured BAVMs would be to create and promote a long-term, comprehensive, and adjudicated registry inclusive of all patients with BAVM.

In conclusion, the overall data presented from the ARUBA trial are not unexpected. Given the problematic nature of the trial, it nevertheless reinforces well-established information about the safety of intervention for grade 1 lesions. It confirms that there are indeed significant risks to the observation of unruptured BAVMs, with an overall event rate of 8% among as-treated subjects in the medical management arm; despite the short study follow-up period. It does not provide meaningful information about the risks of treatment for Spetzler–Martin grade 3 BAVMs, which are known to be incredibly heterogeneous and difficult to manage, and the trial provides no information about the treatment of high-grade lesions. The results as presented are primarily a commentary on embolization complication rates in a selected subset of patients with BAVM thought to be appropriate for randomization in a clinical trial. Unfortunately, any randomized trial attempting to determine the superiority of management paradigms for lesions as rare and heterogeneous as BAVM with such entrenched and variable practice patterns is going to be limited in its external validity and concomitantly in the interpretation of its results. In the future, when examining the question of management approaches for lesions such as BAVMs, it may be better to apply the limited resources available to the establishment of a national, multicenter, adjudicated registry of consecutive patients.

Disclosures

Dr. Cockcroft is a consultant for Covidien Neurovascular. The other authors report no conflicts.

References


The Case Against A Randomized Trial of Unruptured Brain Arteriovenous Malformations: Misinterpretation of a Flawed Study
Nicholas C. Bambakidis, Kevin M. Cockroft, Joshua A. Hirsch, E. Sander Connolly, Sepideh Amin-Hanjani, Philip M. Meyers and Robert M. Friedlander

Stroke. published online July 31, 2014;
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 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/early/2014/07/31/STROKEAHA.114.006519.citation

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/