



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Medical management with interventional therapy versus medical management alone for unruptured brain arteriovenous malformations (ARUBA)

Citation for published version:

Aruba co-investigators, Mohr, JP, Overbey, JR, Hartmann, A, Kummer, RV, Al-shahi Salman, R, Kim, H, Van Der Worp, HB, Parides, MK, Stefani, MA, Houdart, E, Libman, R, Pile-spellman, J, Harkness, K, Cordonnier, C, Moquete, E, Biondi, A, Klijn, CJM, Stapf, C & Moskowitz, AJ 2020, 'Medical management with interventional therapy versus medical management alone for unruptured brain arteriovenous malformations (ARUBA): final follow-up of a multicentre, non-blinded, randomised controlled trial', *The Lancet Neurology*, vol. 19, no. 7, pp. 573-581. [https://doi.org/10.1016/S1474-4422\(20\)30181-2](https://doi.org/10.1016/S1474-4422(20)30181-2)

Digital Object Identifier (DOI):

[10.1016/S1474-4422\(20\)30181-2](https://doi.org/10.1016/S1474-4422(20)30181-2)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

The Lancet Neurology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The Lancet Neurology

Medical management with interventional therapy versus medical management alone for unruptured brain arteriovenous malformations (ARUBA): final follow-up of a multicentre, open, parallel group, randomised controlled trial --Manuscript Draft--

Manuscript Number:	THELANCETNEUROLOGY-D-19-01064R2
Article Type:	Article (Original Research)
Keywords:	arteriovenous malformations
Corresponding Author:	JP Mohr Columbia University Medical Center New York, New York UNITED STATES
First Author:	J. P Mohr
Order of Authors:	J. P Mohr Jessica R. Overbey, DrPH Andreas Hartmann, MD Jay P. Mohr, MD Rüdiger von Kummer, MD Rustam Al-Shahi Salman, FRCP Helen Kim, PhD H. Bart van der Worp, MD Michael K. Parides, PhD Marco A. Stefani, MD Emmanuel Houdart Richard Libman, MD John Pile-Spellman, MD Kirsty Harkness, MD Charlotte Cordonnier, MD Elen Moquete, RN Alessandra Biondi, MD C.J.M. (Karin) Klijn, MD Alan J. Moskowitz, MD
Manuscript Region of Origin:	UNITED STATES
Abstract:	<p>Background : In A Randomized trial of Unruptured Brain Arteriovenous malformations (ARUBA), randomisation was halted by recommendation of the National Institute of Neurological Disorders and Stroke appointed data and safety monitoring board at a mean follow-up of 33.3 months after a pre-specified interim analysis demonstrated that medical management alone was superior to the combination of medical management and interventional therapy in preventing symptomatic stroke or death. We aimed to study whether these differences persist in the longer-term of 5 years follow-up.</p> <p>Methods : ARUBA was an open, randomised (1:1), parallel group trial of adult patients diagnosed with an unruptured brain arteriovenous malformation, who had never undergone interventional therapy, and were considered by participating clinical centres suitable for intervention to eradicate the lesion. The trial compared medical management alone with medical management and interventional therapy (neurosurgery, embolisation, or stereotactic radiotherapy, alone or in any combination,</p>

sequence or number). Patients were randomised at 39 international clinical centres with randomisation stratified by clinical centre using a random permuted block design generated by a trial statistician and implemented via a central web based data collection system. The primary outcome was time to death or symptomatic stroke confirmed by imaging, assessed by a neurologist at each centre not involved in the management of participants' care, and monitored independently using an adaptive approach with interim analyses. Enrolment began on April 4th, 2007 and was halted on April 15th, 2013, after which follow-up continued until July 15th, 2015. All analyses were by intention-to-treat. The trial is registered with ClinicalTrials.gov, number NCT00389181.

Findings : 226 patients were randomly allocated, 110 to medical management alone and 116 to medical management plus interventional therapy. During a mean duration of follow-up of 50.4 ±22.9 months, the incidence of death or symptomatic stroke was lower with medical management alone (15/110, 3.39 per 100 patient-years) compared to medical management with interventional therapy (41/116, 12.32 per 100 patient-years; HR 0.31, 95% CI 0.17-0.56). Two patients in the medical management group and four patients in the interventional therapy group died during follow-up (two attributed to intervention). Adverse events were observed less often (283 vs 369; 58.97 vs 78.73 per 100 patient-years; risk difference -19.76 ((95% confidence interval, -30.33 to -9.1)) in patients allocated to medical management compared with interventional therapy.

Interpretation: After extended follow-up, medical management alone remained superior to interventional therapy for the prevention of death or symptomatic stroke in patients randomised to ARUBA with unruptured brain arteriovenous malformations. These data should affect standard specialist practice and the information presented to patients. The longer-term risks and difference between the two therapeutic approaches are uncertain.

Funding :

National Institute of Neurological Disorders and Stroke (NINDS) U01NS051483 & U01NS051566 (2007-2013). Vital Projects Fund, private gift to Columbia University (2014-2015)

1 **Medical management with interventional therapy versus medical management alone for**
2 **unruptured brain arteriovenous malformations (ARUBA): final follow-up of a multicentre,**
3 **open, parallel group, randomised controlled trial**

4
5 Prof Jay P. Mohr MD,¹ Jessica R. Overbey DrPH,² Prof Andreas Hartmann MD,³ Prof Rüdiger
6 von Kummer Dr.med.,⁴ Prof Rustam Al-Shahi Salman FRCP Edin,⁵ Prof Helen Kim PhD,⁶ H.
7 Bart van der Worp MD,⁷ Prof Michael K. Parides PhD,⁸ Prof Marco A. Stefani MD,⁹ Prof
8 Emmanuel Houdart MD,¹⁰ Richard Libman MD,¹¹ John Pile-Spellman MD,¹² Kirsty Harkness
9 MD,¹³ Prof Charlotte Cordonnier MD,¹⁴ Ellen Moquete RN,² Prof Alessandra Biondi MD,¹⁵ Prof
10 Catharina J.M. Klijn MD,^{7,16} Prof Alan J. Moskowitz MD.²

- 11
- 12 1. Doris & Stanley Tananbaum Stroke Center, The Neurological Institute, Columbia University
13 Irving Medical Center, New York, NY, USA.
 - 14 2. International Center for Health Outcomes and Innovation Research, Department of
15 Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York,
16 NY, USA.
 - 17 3. Charité-Universitätsmedizin Berlin; Department of Neurology, Klinikum Frankfurt/Oder,
18 Frankfurt/Oder, Germany.
 - 19 4. Department of Neuroradiology, University Hospital Dresden, Dresden, Germany.
 - 20 5. Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK.
 - 21 6. Department of Anesthesia and Perioperative Care, Center for Cerebrovascular Research,
22 University of California - San Francisco, San Francisco, CA, USA.
 - 23 7. Department of Neurology and Neurosurgery, University Medical Center Utrecht, Brain
24 Center, Utrecht University, Utrecht, The Netherlands.
 - 25 8. Department of Cardiovascular & Thoracic Surgery, Albert Einstein Medical Center, Bronx,
26 New York, USA.
 - 27 9. Department of Neurology and Neurosurgery, Hospital de Clinicas de Porto Alegre, Porto
28 Alegre, Brazil.
 - 29 10. Department of Neuroradiology, APHP-Hôpital Lariboisière, Univ Paris Diderot-Sorbonne
30 Paris Cité, Paris, France.

- 31 11. Department of Neurology, North Shore Long Island Jewish Medical Center, New Hyde Park,
32 NY, USA.
- 33 12. Department of Neuroradiology, NYU-Winthrop Medical Center, Mineola, NY, USA.
- 34 13. Department of Neurology, Royal Hallamshire Hospital, Sheffield, UK.
- 35 14. Université Lille, Inserm U1171, Degenerative and Vascular Cognitive Disorders, CHU Lille,
36 Department of Neurology, Lille, France.
- 37 15. Department of Interventional Neuroradiology, Jean Minjoz Hospital, University of Franche
38 Comté, Besançon, France.
- 39 16. Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud
40 University Medical Center, Nijmegen, The Netherlands.

41

42 Corresponding author:

43 JP Mohr, MD

44 Neurological Institute

45 Columbia University Irving Medical Center

46 710 West 168th Street

47 New York, NY, 10032, USA

48 Tel: (1) 212 305 8033

49 Fax: (1) 212 305 5796

50 E-mail: jpm10@cumc.columbia.edu

51

52 **Abstract word count: 422**

53 **Text Word Count: 3613**

54

55

56 **ABSTRACT**

57

58 **Background:** In A Randomized trial of Unruptured Brain Arteriovenous malformations
59 (ARUBA), randomisation was halted by recommendation of the National Institute of
60 Neurological Disorders and Stroke appointed data and safety monitoring board at a mean follow-
61 up of 33.3 months after a pre-specified interim analysis demonstrated that medical management
62 alone was superior to the combination of medical management and interventional therapy in
63 preventing symptomatic stroke or death. We aimed to study whether these differences persist in
64 the longer-term of 5 years follow-up.

65

66 **Methods:** ARUBA was an open, randomised (1:1), parallel group trial of adult patients
67 diagnosed with an unruptured brain arteriovenous malformation, who had never undergone
68 interventional therapy, and were considered by participating clinical centres suitable for
69 intervention to eradicate the lesion. The trial compared medical management alone with medical
70 management and interventional therapy (neurosurgery, embolisation, or stereotactic
71 radiotherapy, alone or in any combination, sequence or number). Patients were randomised at 39
72 international clinical centres with randomisation stratified by clinical centre using a random
73 permuted block design generated by a trial statistician and implemented via a central web based
74 data collection system. The primary outcome was time to death or symptomatic stroke confirmed
75 by imaging, assessed by a neurologist at each centre not involved in the management of
76 participants' care, and monitored independently using an adaptive approach with interim
77 analyses. Enrolment began on April 4th, 2007 and was halted on April 15th, 2013, after which
78 follow-up continued until July 15th, 2015. All analyses were by intention-to-treat. The trial is
79 registered with ClinicalTrials.gov, number NCT00389181.

80

81 **Findings:** 226 patients were randomly allocated, 110 to medical management alone and 116 to
82 medical management plus interventional therapy. During a mean duration of follow-up of 50.4
83 ±22.9 months, the incidence of death or symptomatic stroke was lower with medical
84 management alone (15/110, 3.39 per 100 patient-years) compared to medical management with
85 interventional therapy (41/116, 12.32 per 100 patient-years; HR 0.31, 95% CI 0.17-0.56). Two
86 patients in the medical management group and four patients in the interventional therapy group

87 died during follow-up (two attributed to intervention). Adverse events were observed less often
88 (283 vs 369; 58.97 vs 78.73 per 100 patient-years; risk difference -19.76 ((95% confidence
89 interval, -30.33 to -9.1)) in patients allocated to medical management compared with
90 interventional therapy.

91

92 **Interpretation:**

93 After extended follow-up, medical management alone remained superior to interventional
94 therapy for the prevention of death or symptomatic stroke in patients randomised to ARUBA.
95 These data should affect standard specialist practice and the information presented to patients.
96 The longer-term risks and difference between the two therapeutic approaches are uncertain.

97

98 **Funding:**

99 National Institute of Neurological Disorders and Stroke (NINDS) U01NS051483 &
100 U01NS051566 (2007-2013). Vital Projects Fund, private gift to Columbia University (2014-
101 2015).

102

103 **INTRODUCTION**

104

105 A Randomised trial of Unruptured Brain Arteriovenous malformations (ARUBA) addressed the
106 longstanding uncertainty whether medical management alone or with interventional therapy is
107 superior for those with a brain arteriovenous malformation that has never bled.¹⁻⁷ The two-fold
108 primary aim of the study was, first, to determine whether medical management was *superior* to
109 interventional therapy for preventing the composite outcome of death from any cause or stroke
110 (symptomatic and confirmed by imaging). Failure to declare superiority would not necessarily
111 imply that the two treatments were equivalent; therefore, if medical management was not
112 superior to interventional therapy, the second and subsequent aim of the study was to determine
113 whether medical management was *not inferior* to interventional therapy.

114

115 The first patient was randomised on April 4th 2007. On April 15th 2013, randomisation in the trial
116 was terminated prematurely following the recommendation of its independent data and safety
117 monitoring board. Their decision was based on the results of a planned interim analysis
118 demonstrating a log-rank Z value (4.10) exceeding the pre-specified stopping boundary value
119 (2.87), early and consistent separation of survival curves in two arms, and strong magnitude of
120 effect. After a mean follow-up of 33 months, medical management alone was found to be
121 superior to medical management with interventional therapy for the prevention of death or stroke
122 (hazard ratio (HR) 0.27, 95% CI 0.14-0.54).⁸ Further analyses showed that medical management
123 was also superior to medical management with interventional therapy for the prevention of fatal
124 or disabling stroke (defined as a mRS score ≥ 2).⁹

125

126 In response to the recommendation by the data and safety monitoring board, but without
127 continued funding by the National Institute of Neurological Disorders and Stroke, follow-up
128 continued in order to evaluate treatment effects at five years. Centre participation continued until
129 July 15th 2015. The present report describes outcomes observed over the initial randomisation
130 phase and during continued follow-up to assess whether medical management alone remained
131 superior to medical management plus interventional therapy.

132

133

134 **METHODS**

135

136 *Study design*

137

138 ARUBA was an open, parallel group, trial in which participants were randomly allocated 1:1 to
139 medical management alone (i.e., pharmacological therapy for neurological symptoms as needed)
140 or medical management with interventional therapy (i.e., neurosurgery, embolisation, or
141 stereotactic radiotherapy, alone or in any combination or sequence).¹ Patients were randomised
142 in nine countries at 39 clinical centres (Supplemental Table 1). Protocol and consent forms were
143 approved by the relevant institutional review boards or equivalent ethics committees at all
144 institutions.

145

146 *Participants*

147

148 Patients eligible for randomisation were adults (age ≥ 18 years) whose brain arteriovenous
149 malformation had never bled, and who were considered suitable for attempted AVM eradication
150 by the local centres. Each patient provided written informed consent. No control in case selection
151 was exerted by the clinical coordinating centres (ARUBA-WEST: Columbia University Medical
152 Center; ARUBA-EAST: Hôpital Lariboisière). A full list of eligibility criteria can be found in
153 the study protocol available in the Supplement).

154

155 *Randomisation and masking*

156

157 Randomisation was stratified by clinical centre using a random permuted block design with
158 block sizes of 2, 4, and 6. The randomisation sequence was generated by a trial statistician at the
159 data coordinating centre (the International Center for Health Outcomes and Innovation Research
160 at the Icahn School of Medicine at Mount Sinai) and assignment was controlled via a central
161 web-based data collection system, which did not reveal treatment allocation until all baseline
162 data had been submitted. Site coordinators randomised participants after verifying eligibility and
163 obtaining patient consent. All individual clinical centres were aware of the treatment assignment
164 for their own patients but were not informed of the outcomes from other clinical centres in the

165 trial. Outcomes at each clinical centre were assessed by a senior study neurologist who did not
166 perform interventional procedures.

167

168 Participants and study leaders, including those at the clinical coordinating center, remained
169 blinded as to the overall randomisation assignments and outcomes until April 15th 2013 when the
170 results of a planned interim analysis were provided to the trial executive committee at a meeting
171 with the data and safety monitoring board.

172

173 ***Procedures***

174

175 Interventional therapy options comprised endovascular embolisation, neurosurgical resection, or
176 stereotactic radiotherapy, as single or multiple therapies, in any order, sequence or number. No
177 published guidelines for selection or sequence of choice(s) of intervention existed during the
178 trial.¹⁰ Centres implemented the approaches considered standard practice in their specialist centre
179 and country.

180

181 During the blinded phase of the trial, patient data were collected at six-month intervals for the
182 first two years of follow-up, thereafter annually, with the goal of final status report at year 5, the
183 planned end of the trial by the original protocol. Follow-up was expected to continue also for
184 those experiencing a primary outcome event. Although clinical centres were free to undertake
185 whatever management they deemed appropriate *after* the primary outcome event, reports of
186 outcomes depended on the willingness of those affected to continue in follow-up. After the end
187 of the randomisation phase, supported by Vital Projects Fund, New York, NY and without
188 funding by the National Institute of Neurological Disorders and Stroke, clinical centres
189 continued efforts at follow-up and reported the status of their patients until July 15th 2015 when
190 the database was closed, eight years and two months from the start of accrual of randomisations.
191 Due to lack of continued funding, further follow-up was not feasible. Supplemental Figure 1
192 provides the number of last follow-up reports by year.

193

194 ***Outcomes***

195

196 The primary outcome was time to the composite event of death from any cause or symptomatic
197 stroke (haemorrhage or infarction), documented by imaging (CT or MR scan). Until the end of
198 the randomisation phase in 2013, all primary outcome events were adjudicated by a four-member
199 committee comprised of internationally-renowned experts in stroke neurology, endovascular
200 therapy, radiosurgery, or vascular neurosurgery. All primary outcomes reported after the end of
201 the randomisation phase were adjudicated by a single member of the original adjudication
202 committee (Prof Marie-Germaine Bousser, Hôpital Lariboisière, Paris, France).

203

204 The secondary outcome was death or neurological disability at five years after randomisation.
205 Neurological disability was defined as a score ≥ 2 on the modified Rankin scale (mRS), which
206 ranges from 0 to 6, with higher scores indicating more severe disability with 6 indicating death.¹¹
207 Additional secondary outcomes included the incidence of adverse events which were collected
208 systematically and adjudicated by the event adjudication committee. Definitions of adverse
209 events are available in the protocol included in the supplement. More information on study
210 conduct, data collection, and outcome assessments are also available in the primary publication.¹

211

212 *Statistical analysis*

213

214 The initial protocol, submitted and approved by two National Institute of Neurological Disorders
215 after two separate Stroke Study Section reviews, had a sample size of 800 patients which would
216 have an estimated 87.5% power to detect a 40% reduction in the hazard for death or
217 symptomatic stroke over 5 years based on an assumed 5-year event rate of 22% in the medical
218 management and interventional therapy arm. The study was overseen by a National Institute of
219 Neurological Disorders and Stroke-appointed data and safety monitoring board.

220

221 During the study start-up period, the anticipated number of participating clinical centres was not
222 realized, resulting in a lower than expected recruitment rate. Eighteen months after the first
223 randomisation, given the opportunity for longer follow-up to achieve the outcome event rates
224 within the period of National Institute of Neurological Disorders and Stroke funding, the data
225 and safety monitoring board reviewed the emerging data by treatment group in private, keeping
226 clinical investigators blinded. The board accepted a revised sample size of 400 patients presented

227 by the study statisticians, which would have an estimated 80% power to detect a 46% reduction
228 in the hazard of death or symptomatic stroke, equivalent to a hazard ratio of 0.54. This hazard
229 ratio corresponded to an absolute decrease in 5-year event rates of 9.5% for medical management
230 alone, from an assumed 5-year event rate of 22% for medical management with interventional
231 therapy. ARUBA had an adaptive design involving pre-specified interim analyses. Two interim
232 analyses were pre-specified in the protocol with early stopping boundaries defined by an O'Brien
233 Fleming-type spending function using a Lan-DeMets approach. There were no interim
234 assessments for futility since the study was set up to assess superiority and non-inferiority.

235
236 The statistical analysis was carried out by the data coordinating center. For the primary outcome,
237 cumulative event-free survival curves for each group were estimated using the Kaplan-Meier
238 method. The incidence rate of primary outcome events is the number of patients who had an
239 event divided by the number of event-free patient years observed. A Cox proportional-hazards
240 model was used to estimate the hazard ratio and corresponding 95% confidence interval. To
241 account for potential clustering effects by clinical centre, a Cox model including a frailty term
242 for clinical centre was also explored.

243
244 Risk of death or clinical impairment at five years after randomisation was summarised using the
245 proportion of patients with mRS ≥ 2 at five years and compared between groups by computing
246 the relative risk and corresponding 95% confidence interval. Five-year mRS scores were based
247 on mRS assessments documented between 54 and 66 months from randomisation for participants
248 whose date of randomisation made them eligible for assessment before closure of the trial
249 database. For patients who had multiple assessments in this window, the assessment closest to
250 the expected 60-month (five years) follow-up visit date was selected as the patient's five-year
251 mRS. Patients who died before month 66 were assigned a mRS score of 6, irrespective of the
252 cause of death.

253
254 Frequencies of adverse events including all strokes, focal deficits, seizures, and headaches were
255 computed by allocated treatment group, as well as the rate difference between the groups and the
256 corresponding 95% confidence interval. Duration of follow-up in months was calculated using
257 the last date of contact with each patient. Exploratory subgroup analyses of the primary outcome

258 were conducted using Cox proportional hazards models and tested for heterogeneity of treatment
259 effect using interaction terms.

260
261 All analyses were conducted by the intention-to-treat principle using SAS version 9.4 (Cary,
262 NC). Due to the descriptive nature of the study there is no bias adjustment due to the adaptive
263 design. Additional information on the timing and results of the two pre-specified interim
264 analyses are given in supplemental figure 2. The trial is registered with ClinicalTrials.gov,
265 number NCT00389181.

266
267 ***Role of the Funding Source***
268 National Institute of Neurological Disorders and Stroke officers participated in study design, data
269 interpretation, and writing of the report, but had no role in data collection or data analysis. The
270 Vital Projects Fund, New York, New York, had no role in study design, data collection, data
271 analysis, data interpretation, or writing of the report. The corresponding author had full access to
272 all the data in the study after the end of the randomisation phase and had final responsibility for
273 the decision to submit for publication.

274
275 **RESULTS**

276
277 Of 1740 patients screened, 1014 (58.3%) were ineligible and 726 were eligible, of whom 323
278 (44.5%) refused participation and 177 (24.4%) decided their management outside the trial
279 (figure 1); outcomes were not collected for eligible patients who were not randomised. Thirty-
280 nine international centres randomised a total of 226 participants at a rate of 3.2 patients per
281 month from April 4th 2007 to April 15th 2013 (Supplemental Figure 3). Of the 226 patients
282 randomised, 110 were allocated to medical management alone and 116 to medical management
283 plus interventional therapy. Three patients randomised in the interval between data lock for the
284 final report presented to the Data and Safety Monitoring Board on April 15th 2013 and the end of
285 enrolment were not included in the primary publication¹, but are included in the current report.
286 All randomised patients were included in the time-to-event analysis of the primary outcome.

287

288 At final data lock on July 15th 2015, patients randomised had a mean length of follow-up of 50.4
289 months (SD±22.9; median 48.0, IQR 35.9-71.1). The average follow-up for patients allocated to
290 medical management patients was 52.4 months (SD±23.7; median 49.1, IQR 36.1-71.8), and for
291 patients allocated to interventional therapy it was 48.5 months (SD±22.0; median 45.5, IQR
292 34.7-62.1). The distribution of dates of the last patient contact is shown in Supplemental figure
293 1.

294
295 Baseline characteristics and mRS scores were similar between groups (Table 1), apart from focal
296 neurological deficits at presentation, small AVM nidus size and Spetzler Martin grade.

297
298 In total, there were 15 primary outcome events in patients randomised to medical management
299 (incidence rate: 3.39 per 100 patient-years) versus 41 in patients randomised to interventional
300 therapy (incidence rate: 12.32 per 100 patient-years; Table 2), resulting in a hazard ratio of 0.31
301 (95% CI 0.17-0.56; Figure 2). These results remained consistent after accounting for the
302 potential clustering effect of clinical centre (adjusted HR 0.31, 95% CI 0.16-0.61). Of the 56
303 primary outcomes, 10 were new events reported between April 15th 2013 and July 15th 2015.
304 Four of these 10 occurred in the medical management arm and six in those randomised to
305 medical management plus interventional therapy. In total, two patients allocated to medical
306 management alone and four patients allocated to medical management plus interventional
307 therapy died during follow-up. In the latter group two of the four deaths were attributed to the
308 intervention (Table 2).

309
310 The risk of the secondary outcome of death or neurological disability at five years after
311 randomisation, available for 96 patients, was lower for those allocated to medical management
312 alone (Table 2). Supplemental Figure 4 illustrates the distribution of mRS scores by study arm
313 for those with data available at 5years.

314
315 Patients allocated to interventional therapy experienced more adverse events compared to those
316 allocated to medical management (Table 3)

317

318 Subgroup analyses of the primary outcome were consistently in favour of medical management
319 over medical management plus interventional therapy, except for venous drainage (superficial
320 only versus any deep), arteriovenous malformation maximum nidus size (<3 cm versus \geq 3 cm),
321 and Spetzler-Martin grade, where there was heterogeneity of the treatment effects (Figure 3).
322 The effect of medical management in patients with Spetzler-Martin Grade I AVMs appeared to
323 differ from Spetzler-Martin Grade II-V AVMs, although the effect in the Spetzler-Martin Grade
324 1 group was not significant (HR 1.82, 95% CI 0.46-7.28).

325
326 Eight patients who were randomised to medical management alone received medical
327 management plus interventional therapy. Of the 116 randomised to interventional therapy, three
328 experienced an outcome before interventional therapy began and 15 never received
329 interventional therapy. Ultimately 106 patients received interventional therapy. The median time
330 from randomisation to first intervention in this group was 76 days (IQR 42-136). For the 43
331 patients who reached a primary outcome following the initiation of interventional therapy, the
332 median time since the last intervention was one day (IQR 0-43). Sixty-eight of the 106 patients
333 who received interventional therapy (64.2%) were treated by a single modality, while for 38/106
334 (35.8%) the therapy was multimodal. At the time of final data lock, 47/106 (44.3%) of those
335 receiving medical management plus interventional therapy had angiographic evidence of brain
336 arteriovenous malformation eradication, 43/106 (40.6%) had evidence of a brain arteriovenous
337 malformation remnant on last follow-up imaging, and in 16/106 (15.1%, all after radiotherapy)
338 the brain arteriovenous malformation status was unknown due to missing follow-up imaging
339 (Supplemental Table 2).

340
341 Twenty-two of 106 (20.8%) patients who were treated with medical management and
342 interventional therapy underwent neurosurgery, either alone or as part of a multimodal
343 interventional therapy strategy following embolisation. In 21 of 22 (95.5%) the brain
344 arteriovenous malformation had been eradicated based on a post-operative angiogram. Nine of
345 22 (40.9%) patients who were operated on experienced a primary outcome event (Supplemental
346 Table 2). Stereotactic radiotherapy was used in the treatment of 57 of 106 (53.8%) patients
347 treated with medical management and interventional therapy, either alone or as part of a
348 multimodal strategy with either embolisation alone or with both endovascular and surgical

349 therapy (n=1). At the time of the final analysis, 12 of 57 patients (21.1%) had reached
350 angiographically-documented eradication of the brain arteriovenous malformation, and 21/57
351 (36.8%) had had a primary outcome event (Supplemental Table 2). Sixty-six of 106 (62.3%)
352 patients were treated by endovascular embolisation, either alone or as part of a multimodal
353 treatment strategy with neurosurgery and/or radiotherapy; 34 of 66 (51.5%) demonstrated
354 eradication of the brain arteriovenous malformation on catheter angiography, and 33 of 66
355 (50.0%) experienced a primary outcome event (Supplemental Table 2).

356

357 **DISCUSSION**

358

359 With extended follow-up of 226 participants with unruptured brain arteriovenous malformations
360 in ARUBA, the risk of death or stroke remained significantly lower after medical management
361 alone than after medical management with interventional therapy after a mean follow-up of 50
362 months. Ninety-six participants followed up for 60 months without intervention also had a
363 significantly lower risk of death or neurological disability, and fewer adverse events. Although
364 the persisting difference between the two management options persisted, ideally longer follow-
365 up would be desirable but was not possible due to funding constraints.

366

367 This longer-term follow-up report of ARUBA has limitations. The number of patients included
368 in the study was much smaller than the 800 patients initially planned, but with 223 patients, whose
369 disparity in outcomes led to the early suspension of randomisation. In addition, due to limited
370 resources, the duration of follow-up was shorter than the five years for all participants that had
371 initially been planned. Although 26 of 226 patients were lost to follow-up at the time of the final
372 data lock, their numbers were similar in both arms, so that potentially missed outcomes for them
373 are unlikely to have had a large effect on the reported results. Because of the smaller number of
374 included patients than initially planned, the estimates of the treatment effect in the subgroup
375 analyses were less precise than anticipated.

376

377 Strengths of ARUBA are its randomised design and its inclusion of patients at 39 centres, which
378 enhances the generalisability of its findings. A recent systematic review in the Cochrane
379 Database (search date January 14th 2019) found ARUBA was the only published randomised trial

380 comparing medical management with interventional therapy for unruptured brain arteriovenous
381 malformations.¹² The study also has striking similarities with the literature: ARUBA was
382 consistent with the outcomes for intervention in a meta-analysis¹³ from the few cohorts of
383 untreated unruptured brain AVMs,^{14, 15} and a non-randomised cohort study with concurrent
384 controls with follow-up for up to 12 years.¹⁶ The distribution of the Spetzler-Martin grades of
385 included brain AVMs were bias towards those more safely and easily treated, indicating few
386 participants were unsuitable for interventional therapy.

387
388 Previous case series have reported differing risks of clinical outcomes and angiographic
389 obliteration after treatment for brain arteriovenous malformations.¹³ For unruptured brain
390 arteriovenous malformations, the risk of treatment has to be weighed against the risk of rupture
391 with medical management alone, 1.3% per year over a period of 10 years among the largest
392 reported series of 1389 from four major centers.¹⁷ Comparative observational studies with
393 concurrent controls have also reported worse outcomes associated with interventional therapy for
394 brain arteriovenous malformations compared to medical management over up to 12 years.¹⁶

395
396 Patients with an unruptured brain arteriovenous malformation should, therefore, be informed
397 about the absolute and relative risks of both treatment strategies in ARUBA. In addition, the
398 current report may inform the design of other randomised controlled trials seeking to investigate
399 the reproducibility of the ARUBA model. An improved design depends on a better
400 understanding of the natural history of unruptured brain arteriovenous malformations, and in the
401 context of other settings or approaches to improvements in therapeutic interventional therapies.

402
403 Three other RCTs are currently ongoing in patients with brain arteriovenous malformations; all
404 three including both patients with ruptured and patients with unruptured brain arteriovenous
405 malformations. One is testing whether two embolisation approaches are equivalent (endovascular
406 embolisation with Onyx versus with TRUFILL n-butyl cyanoacrylate n-BCA;NCT00857662);
407 one whether conservative management or intervention will reduce the risk of death or
408 debilitating stroke and whether endovascular treatment can improve the safety and efficacy of
409 surgery or radiosurgery (NCT02098252), and the third is testing whether transvenous
410 embolisation or trans-arterial embolisation is most effective and safe in achieving angiographic

411 obliteration of the arteriovenous malformation (NCT03691870). The longer-term risks of
412 interventional therapy compared to medical management than reported here will remain
413 unknown unless future randomised trials are sufficiently funded to permit an adequate duration
414 of follow-up.

415
416 How some lesions seem stable for decades or life-time is still unclear.^{18,19} The few studies
417 assessing these risks have documented that haemorrhage risk is related to high intra-nidal
418 pressure²⁰ and to single-vein drainage.²¹ Future studies may extend these findings but likely will
419 depend on further innovations in non-invasive imaging to assess arteriovenous resistivity
420 patterns. Continued interest in long-term outcomes in patients with unruptured brain
421 arteriovenous malformations is being pursued in a large, international observational cohort study
422 funded by National Institute of the Neurological Disorders and Stroke (R01 NS099268) based on
423 the methods used for the Multicenter Arteriovenous Malformations Study (MARS), with a goal
424 of identifying predictors of haemorrhage and treatment risks in >2500 patients.²² Also ongoing is
425 the Treatment Of Brain Arteriovenous malformations Study (TOBAS) comprising two open-
426 label randomised arms.²³

427
428 No formal detailed guidelines on the management of unruptured brain AVM have emerged from
429 professional associations. Two consensus reports endorsed by the American Heart Association,
430 the first from 2001 cited that treatment results vary considerably²⁴ and the most recent from 2017
431 that medical management alone and three often complementary methods of interventional
432 therapy exist.²⁵

433
434 In summary, after mean length of follow-up 50.4 months (SD±22.9; median 48.0, IQR 35.9-
435 71.1), medical management alone remained superior to medical management with interventional
436 therapy for the prevention of death or symptomatic stroke in patients with an unruptured brain
437 arteriovenous malformation in the ARUBA trial. Evidence of this hazard should have an impact
438 on standard specialist practice and should be among the materials presented to patients. The rate
439 of outcome events and degree of disparity between the two management options beyond four
440 years remain uncertain.

441

442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471

Data Sharing Statement: Trial data collected during the NINDS-funded phase are archived by NINDS and available upon request. Information on how to request the data is available here: <https://www.ninds.nih.gov/Current-Research/Research-Funded-NINDS/Clinical-Research/Archived-Clinical-Research-Datasets> (last accessed 29/03/2020).

The trial has been funded internationally by the US National Institutes of Health / National Institute of Neurological Disorders and Stroke (NIH/NINDS) via cooperative agreements U01NS051483 (J.P.Mohr, PI Clinical Coordinating Center) and U01 NS051566 (A.J. Moskowitz, PI Data Coordinating Center).

The post-NIH funding period was supported by a gift from the Vital Projects Fund, Inc. (J.P.Mohr, PI Clinical Coordinating Center).

ARUBA has been officially adopted by the UK Stroke Research Network and endorsed by the Société Française Neurovasculaire (SFNV). No commercial funding sources have been involved.

Acknowledgement: The ARUBA Investigators are indebted to the patients who participated in this trial, the members of the adjudication committee, our safety officer, members of the data and safety monitoring board. Additional thanks to Dr Fernanda Teixeira and Ms Karen Lord for assistance with the manuscript.

Declarations of Interest: JPM, JRO, AH, RvK, MKP, MAS, EH, RL, JPS, KH, CC, EM, and AB had nothing to declare during the course of the study. RAS reports reimbursement from the National Institutes of Health, during the conduct of the study. HK reports grants from Columbia University, from NIH/NINDS, during the conduct of the study. HBW reports funding from the Dutch Heart Foundation, the European Union and consultancy fees paid to his institution from Boehringer Ingelheim, Bayer, and LivaNova, CJMK cites grants from Netherlands Cardiovascular Research Initiative, which is supported by the Dutch Heart Foundation, CVON2015–01: CONTRAST, and the support of the Brain Foundation Netherlands (HA2015•01•06), from the Dutch Heart Foundation (Clinical established investigator grant 2012T077), and from The Netherlands Organization for Health Research and Development,

472 ZonMw (grant 015008048), during the conduct of the study. AJM reports grants from NIH,
473 during the conduct of the study.

474

475 **Contributions of Authors:**

476 JPM, MKP, AJM contributed to study design; JPM, RAS contributed to scientific literature
477 search;

478 JPM, AH, RvK, RAS, HBW, MKP, MAS, EH, RB, JPS, KH, CC, E, AB, CJMK contributed to
479 data collection; JPM, JRO, MKP, AJM contributed to data analysis; JPM, JRO, AH, RvK, RAS,
480 HK, HBW, MKP, MAS, EH, RB, JPS, KH, CC, E, AB, CJMK, AJM contributed to data
481 interpretation; JPM, JRO contributed to figures and tables; JPM, JRO, AH, RvK, RAS, HBW,
482 CJMK AJM contributed to writing

483 JPM, JRO, AH, RvK, RAS, HK, HBW, MKP, MAS, EH, RB, JPS, KH, CC, E, AB, CJMK,
484 AJM contributed to editing.

485

486 **NINDS Officers**

487 Scott Janis, Ph.D. Program Director, Division of Clinical Research, National Institute of
488 Neurological Disorders and Stroke

489 Claudia Scala Moy, Ph.D. Program Director, Division of Clinical Research, National Institute of
490 Neurological Disorders and Stroke

491

492 **DSMB Members**

493 Chair, Prof. Karen Lisa Furie, MD, Neurology, Brown University Warren Alpert Medical
494 School, Providence, Rhode Island, USA.

495 Peter Gilbert, MS. Clinical Research Project Manager, National Institute of Neurological
496 Disorders and Stroke, USA.

497 Prof. Barry R. Davis, PhD, University of Texas School of Public Health, Houston, TX, USA.

498 Prof. Michael D. Hill, MD, Calgary Stroke Program, Department of Clinical Neurosciences and
499 Department of Radiology, Hotchkiss Brain Institute, University of Calgary, Canada.

500 Prof. Charles A. Jungreis, MD, Department of Radiology, Temple University School of
501 Medicine, Philadelphia, PA, USA.

502 Prof. Renée H. Martin, Department of Biostatistics, Bioinformatics and Epidemiology Medical
503 University of South Carolina, Charleston, SC, USA.

504 Prof. Emeritus Duke S. Samson, MD, Department of Neurological Surgery, University of Texas
505 Southwestern Medical Center, Dallas, Texas, USA.

506

507 **Adjudication Committee**

508 Prof. Marie-Germain Bousser, MD, Department of Neurology, Hopital Lariboisière, Université
509 Paris VII Denis Diderot, Paris, France.

510 Prof. Louis R. Caplan, MD, Stroke Division, Department of Neurology, Beth Israel Deaconess
511 Medical Center, Boston, MA, USA.

512 Prof. Charles M. Strother, MD, Department of Radiology, University of Wisconsin-
513 Madison, Madison, WI, USA.

514 Prof. Emeritus Bryce Keith A. Weir, MD, Department of Neurosurgery, University of Chicago,
515 Chicago, IL, USA.

516

517 **Safety Officer**

518 Prof. Anthony J. Furlan, MD, Department of Neurology, Case Western School of Medicine,
519 Cleveland, OH, USA.

520

521 **PANEL: RESEARCH IN CONTEXT**

522

523 **Evidence before this study:**

524 Case series have reported different risks of clinical outcomes and angiographic obliteration for
525 brain arteriovenous malformations, which are unreliable for comparison with the best available
526 data indicating a 1% annual risk of haemorrhage from un-treated unruptured brain arteriovenous
527 malformations. Comparative observational studies with concurrent controls have reported worse
528 outcome associated with interventional therapy for brain arteriovenous malformations compared
529 to medical management over up to 12 years. A recent systematic review in the Cochrane
530 Database (search date January 14th 2019) found one published randomised trial comparing
531 medical management with interventional therapy for unruptured brain arteriovenous
532 malformations. The ARUBA trial terminated recruitment when its data monitoring committee
533 concluded that medical management was superior to interventional therapy for the prevention of
534 stroke or death on the basis of the first 223 recruited participants after a mean follow-up of 33
535 months (HR 0·27, 95% CI 0·14-0·54). The data monitoring committee concluded that there was,
536 “a compelling need for additional long-term data.”

537

538 **Added value of this study:**

539 The current report includes longer term outcomes than in the initial publication of the
540 randomised phase of the ARUBA trial, now including all 226 participants recruited at 39
541 international hospitals with mean follow-up extended from 33 months to 50 months. The final
542 results of ARUBA show that medical management remained superior to interventional therapy
543 (HR 0·31, 95% CI 0·17-0·56).

544

545 **Implications of all the available evidence:**

546 The final results of ARUBA demonstrate harm from interventional therapy compared to medical
547 management over an average duration of follow-up of more than four years. Patients with
548 unruptured brain arteriovenous malformation should be informed about the absolute and relative
549 risks in ARUBA, which may inform the design of other randomised controlled trials seeking to
550 investigate the reproducibility of ARUBA in the context of other settings or approaches to
551 interventional therapy. The long-term risks of interventional therapy compared to medical

552 management will remain unknown unless future randomised trials are sufficiently funded to
553 permit an adequate duration of follow-up.

554

555 **REFERENCES**

556

- 557 1. Mohr JP, Parides MK, Stapf C, et al. Medical management with or without interventional
558 therapy for unruptured brain arteriovenous malformations (ARUBA): a multicentre, non-
559 blinded, randomised trial. *Lancet* 2014;**383**:614–21.
- 560 2. Massaro AR, Young WL, Kader A, et al. Characterization of Arteriovenous
561 Malformation Feeding Vessels by Carbon-Dioxide Reactivity. *Am J Neuroradiol*
562 1994;**15**:55–61.
- 563 3. Mast H, Mohr JP, Osipov A, et al. Steal Is an Unestablished Mechanism for the Clinical
564 Presentation of Cerebral Arteriovenous-Malformations. *Stroke* 1995;**26**:1215–20.
- 565 4. Wedderburn CJ, van Beijnum J, Bhattacharya JJ, et al. Outcome after interventional or
566 conservative management of unruptured brain arteriovenous malformations: a
567 prospective, population-based cohort study. *Lancet Neurol* 2008;**7**:223–30.
- 568 5. Chang SD, Marcellus ML, Marks MP, Levy RP, Do HM, Steinberg GK. Multimodality
569 treatment of giant intracranial arteriovenous malformations (Reprinted from Archival
570 article vol 53, pg 1, 2003). *Neurosurgery* 2007;**61**:432–42.
- 571 6. Ondra SL, Troupp H, George ED, Schwab K. The Natural-History of Symptomatic
572 Arteriovenous-Malformations of the Brain - a 24-Year Follow-up Assessment. *J*
573 *Neurosurg* 1990;**73**:387–91.
- 574 7. Hartmann A, Mast H, Choi JH, Stapf C, Mohr JP. Treatment of arteriovenous
575 malformations of the brain. *Curr Neurol Neurosci Rep* 2007;**7**:28–34.
- 576 8. Rukovets O. Why the ARUBA Trial on arteriovenous malformations was stopped.
577 *Neurology Today* 2013;**13**:1,16–18.
- 578 9. Mohr JP, Overbey JR, von Kummer R, et al. Functional impairments for outcomes in a
579 randomized trial of unruptured brain arteriovenous malformations . *Neurology*
580 2017;**89**:1499–506.
- 581 10. Ogilvy CS, Stieg PE, Awad I, et al. AHA Scientific Statement: Recommendations for the
582 management of intracranial arteriovenous malformations: a statement for healthcare
583 professionals from a special writing group of the Stroke Council, American Stroke
584 Association. *Stroke* 2001;**32**:1458–71.
- 585 11. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver
586 agreement for the assessment of handicap in stroke patients. *Stroke* 1988;**19**:604–7.

- 587 12. Zuurbier SM, Al-Shahi Salman R. Interventions for treating brain arteriovenous
588 malformations in adults. *Cochrane Database Syst Rev.* 2019;**9**:CD003436.
- 589 13. van Beijnum J, van der Worp HB, Buis DR, et al. Treatment of Brain Arteriovenous
590 Malformations A Systematic Review and Meta-analysis. *J Am Med Assoc*
591 2011;**306**:2011–9.
- 592 14. Hartmann A, Mast H, Mohr JP, et al. Morbidity of intracranial hemorrhage in patients
593 with cerebral arteriovenous malformation. *Stroke* 1998;**29**:931–4.
- 594 15. Laakso A, Dashti R, Juvela S, Niemela M, Hernesniemi J. Natural history of
595 arteriovenous malformations: presentation, risk of hemorrhage and mortality. *Acta*
596 *Neurochir Suppl* 2010;**107**:65–9.
- 597 16. Al-Shahi Salman R, White PM, Counsell CE, et al. Outcome After Conservative
598 Management or Intervention for Unruptured Brain Arteriovenous Malformations. *J Am*
599 *Med Assoc* 2014;**311**:1661–9.
- 600 17. Kim H, Al-Shahi Salman R, McCulloch CE, Stapf C, Young WL, MARS
601 Coinvestigators. Untreated brain arteriovenous malformation: patient-level meta-analysis
602 of hemorrhage predictors. *Neurology* 2014;**83**:590–7.
- 603 18. Mast H, Young WL, Koennecke HC, et al. Risk of spontaneous haemorrhage after
604 diagnosis of cerebral arteriovenous malformation. *Lancet* 1997;**350**:1065–8.
- 605 19. Stapf C, Mast H, Sciacca RR, et al. Predictors of hemorrhage in patients with untreated
606 brain arteriovenous malformation. *Neurology* 2006;**66**:1350–5.
- 607 20. Young WL, Kader A, Pile-Spellman J, Ornstein E, Stein BM. Arteriovenous
608 malformation draining vein physiology and determinants of transnidial pressure gradients.
609 The Columbia University AVM Study Project. *Neurosurgery* 1994;**35**:389-95.
610 PMID:7800130
- 611 21. Sahlein DH, Mora P, Becske T, et al. Features predictive of brain arteriovenous
612 malformation hemorrhage. Extrapolation to a physiologic model. *Stroke* 2014;**45**:1964-
613 1970.
- 614 22. Kim H, Al-Shahi Salman R, Flemming KD, et al. Long-Term Outcomes in Unruptured
615 Brain Arteriovenous Malformation Patients: The Multicenter Arteriovenous
616 Malformation Research Study (MARS). *Stroke* 2019;**50**.

- 617 23. Magro E, Gentric JC, Batista AL, et al. The Treatment of Brain AVMs Study (TOBAS):
618 an all-inclusive framework to integrate clinical care and research. *J Neurosurg*
619 2018;**128**:1823–9
- 620 24. Ogilvy CS, Stieg PE, Awad I, et al. the Special Writing Group of the Stroke Council, and
621 the American Stroke Association. AHA Scientific Statement: recommendations for the
622 management of intracranial arteriovenous malformations: a statement for healthcare
623 professionals from a special writing group of the Stroke Council, American Stroke
624 Association. *Stroke*. 2001; **32**:1458–71.
- 625 25. Derdeyn CP, Zipfel GJ, Albuquerque FC, et al. Management of Brain Arteriovenous
626 Malformations A Scientific Statement for Healthcare Professionals From the American
627 Heart Association/American Stroke Association. *Stroke* 2017;48:E200–E24.

Table 1. Baseline Characteristics

	Interventional Therapy (N=116)	Medical Management (N=110)
Demographics		
Mean age (\pm SD) years	44.5 (\pm 12.5)	44.3 (\pm 12.2)
Female sex (%)	50/116 (43.1)	44/110 (40.0)
Male sex (%)	66/116 (56.9)	66/110 (60)
White (%)	100/116 (86.2)	88/110 (80.0)
Right-handed (%)	109/116 (94.0)	101/110 (91.8)
Clinical presentation		
Seizure (%)	52/116 (44.8)	45/110 (40.9)
Headaches (%)	56/116 (48.3)	60/110 (54.5)
Focal deficit (%)	21/116 (18.1)	10/110 (9.1)
Other (%)	3/116 (2.6)	8/110 (7.3)
Asymptomatic (%)	45/116 (38.8)	49/110 (44.5)
Modified Rankin Scale score at randomisation		
0 (%)	57/116 (49.1)	51/110 (46.4)
1 (%)	59/116 (50.9)	59/110 (53.6)
Spetzler-Martin grade**		
I (%)	32/114 (28.1)	33/110 (30.0)
II (%)	45/114 (39.5)	27/110 (24.5)
III (%)	29/114 (25.4)	35/110 (31.8)
IV (%)	8/114 (7.0)	15/110 (13.6)
AVM nidus morphology		
Mean maximum diameter (\pm SD), mm	24.8 (\pm 12.1)	27.6 (\pm 11.1)
Maximum diameter <3cm (%)	79/116 (68.1)	61/110 (55.5)
Left-sided (%)	50/116 (43.1)	51/110 (46.4)
Any lobar location (%)	105/116 (90.5)	100/110 (90.9)
Infratentorial location (%)	8/116 (6.9)	5/110 (4.5)
Eloquent location (%)***	55/116 (47.4)	52/110 (47.3)
Concurrent arterial intracranial aneurysms		
Associated aneurysm (%)‡	15/116 (12.9)	21/110 (19.1)
Unrelated aneurysm (%)	4/116 (3.4)	7/110 (6.4)
Venous drainage pattern*		
Superficial only (%)	79/114 (69.3)	69/110 (62.7)
Any deep (%)	35/114 (30.7)	41/110 (37.3)

Plus-minus values are means \pm SD.

Categorical variables are the number with the value / number with available data (%).

*Patients may have more than one presenting symptom.

** Baseline score unavailable for 2 patients enrolled without angiography in the interventional therapy group.

*** Eloquent (as defined by the Spetzler-Martin scale) is any AVM location involving the sensorimotor, language, or visual cortex; the hypothalamus and thalamus; the internal capsule; the brainstem; the cerebellar peduncles; or the deep cerebellar nuclei.

‡ Associated arterial aneurysms are flow-related aneurysms located on a feeding artery or within the AVM nidus ("intranidal aneurysms").

AVM, arteriovenous malformation; mRS, modified Rankin scale; SD = standard deviation, no. number; obs, observations.

Table 2 Primary and secondary outcomes by randomised assignment to medical management alone vs medical management with interventional therapy

	Interventional Therapy (N=116)		Medical Management (N=110)		Effect of Medical management alone vs Interventional Therapy
	N	Incidence Rate Per 100 Patient-Years	N	Incidence Rate Per 100 Patient- Years	Hazard Ratio (95% CI)
Primary outcome					
Symptomatic stroke or death	41	12.32	15	3.39	0.31 (0.17 - 0.56)
Symptomatic stroke					
Any incident stroke	40	11.99	13	2.94	0.27 (0.15 - 0.51)
Haemorrhagic	30		9		
Ischaemic	10		4		
Death					
Any*	4	0.85	2	0.42	0.49 (0.09 - 2.67)
AVM-related	2		0		
Not AVM-related	2		1		
Unknown cause	0		1		
Functional Outcome	N/N with follow-up available	%	N/N with follow-up available	%	Relative Risk (95% CI)
mRS 2-6 at 5 years	17/45	37.8	9/51	17.7	0.47 (0.23 - 0.94)

* Three patients in the IT arm experienced at least one stroke and eventually died during the course of the trial

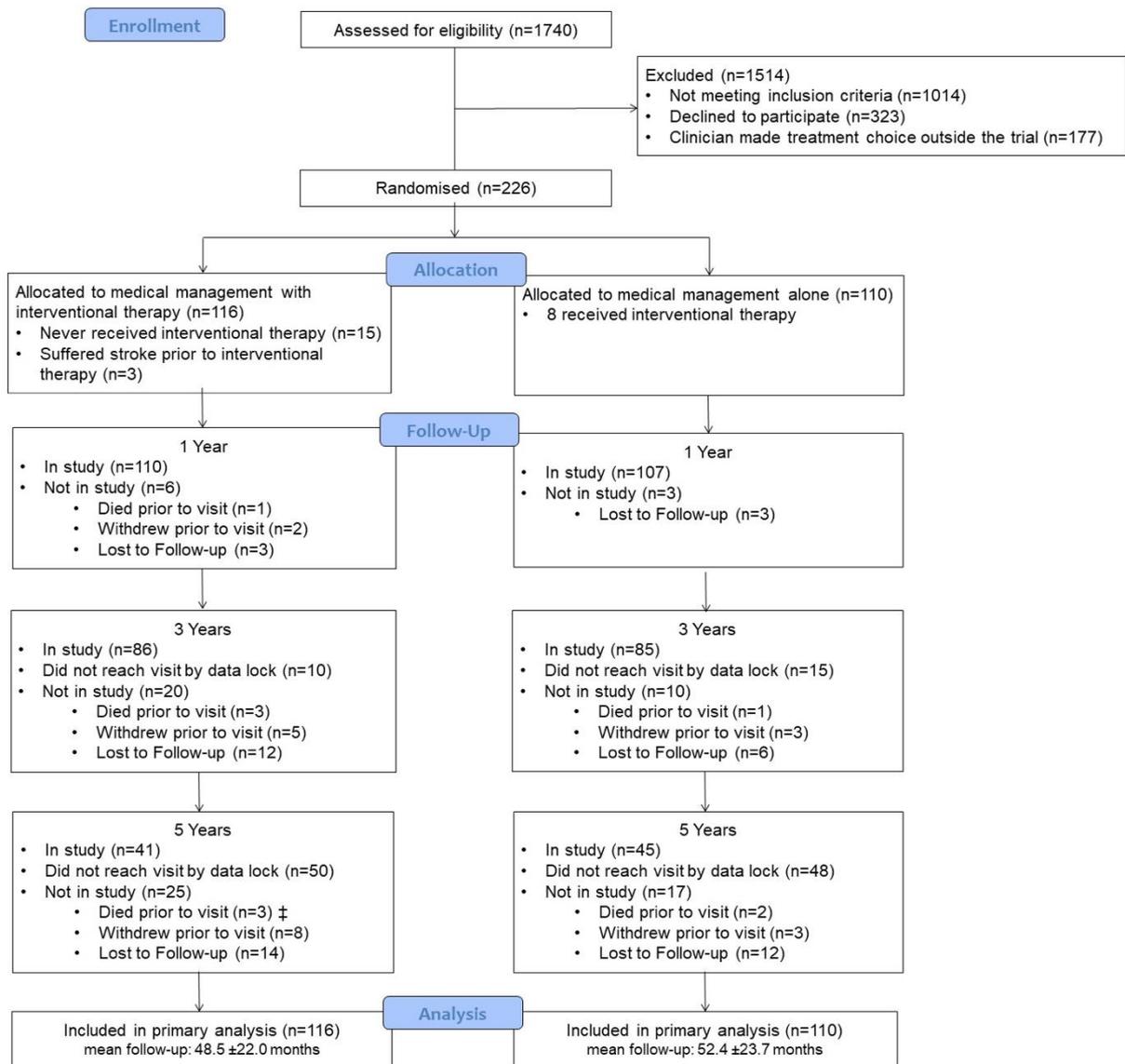
AVM, arteriovenous malformation; N, number, mRS, modified Rankin scale; CI = Confidence interval

Table 3 Adverse events by randomisation assignment to medical management alone vs medical management with interventional therapy

Event type	Interventional Therapy (n=116)		Medical Management (n=110)		Risk Difference for MM-IT (95% CI)
	Number of Events	Rate per 100 patient-years	Number of Events	Rate per 100 patient-years	
Stroke					
All	52	11.09	16	3.33	-7.76 (-11.19, -4.33)
Haemorrhagic	39	8.32	11	2.29	-6.03 (-8.97, -3.09)
Ischemic	13	2.77	5	1.04	-1.73 (-3.49, 0.03)
Focal deficit, unrelated to stroke					
All	20	4.27	3	0.63	-3.64 (-5.64, -1.64)
Persistent	7	1.49	1	0.21	-1.29 (-2.46, -0.11)
Reversible	13	2.77	2	0.42	-2.36 (-3.97, -0.74)
Seizure					
All	95	20.27	68	14.17	-6.1 (-11.39, -0.81)
Simple focal	38	8.11	17	3.54	-4.57 (-7.64, -1.49)
Partial complex	19	4.05	7	1.46	-2.6 (-4.71, -0.48)
Generalized	34	7.25	34	7.08	-0.17 (-3.58, 3.24)
Not classified	4	0.85	10	2.08	1.23 (-0.31, 2.77)
Headache, unrelated to stroke					
All	116	24.75	111	23.13	-1.62 (-7.85, 4.61)
Migraine (with or without aura)	23	4.91	57	11.88	6.97 (3.29, 10.65)
Tension-type (episodic, chronic)	55	11.73	39	8.13	-3.61 (-7.62, 0.41)
Unclassified, other	38	8.11	15	3.13	-4.98 (-8.01, -1.96)
Other AVM related:					
Contrast reaction	1	0.21	0	0.00	-0.21 (-0.63, 0.2)
Catheter adherence	1	0.21	0	0.00	-0.21 (-0.63, 0.2)
Systemic embolisation	1	0.21	0	0.00	-0.21 (-0.63, 0.2)
Infection after interventional	2	0.43	0	0.00	-0.43 (-1.02, 0.16)
Procedural vascular injury	2	0.43	0	0.00	-0.43 (-1.02, 0.16)

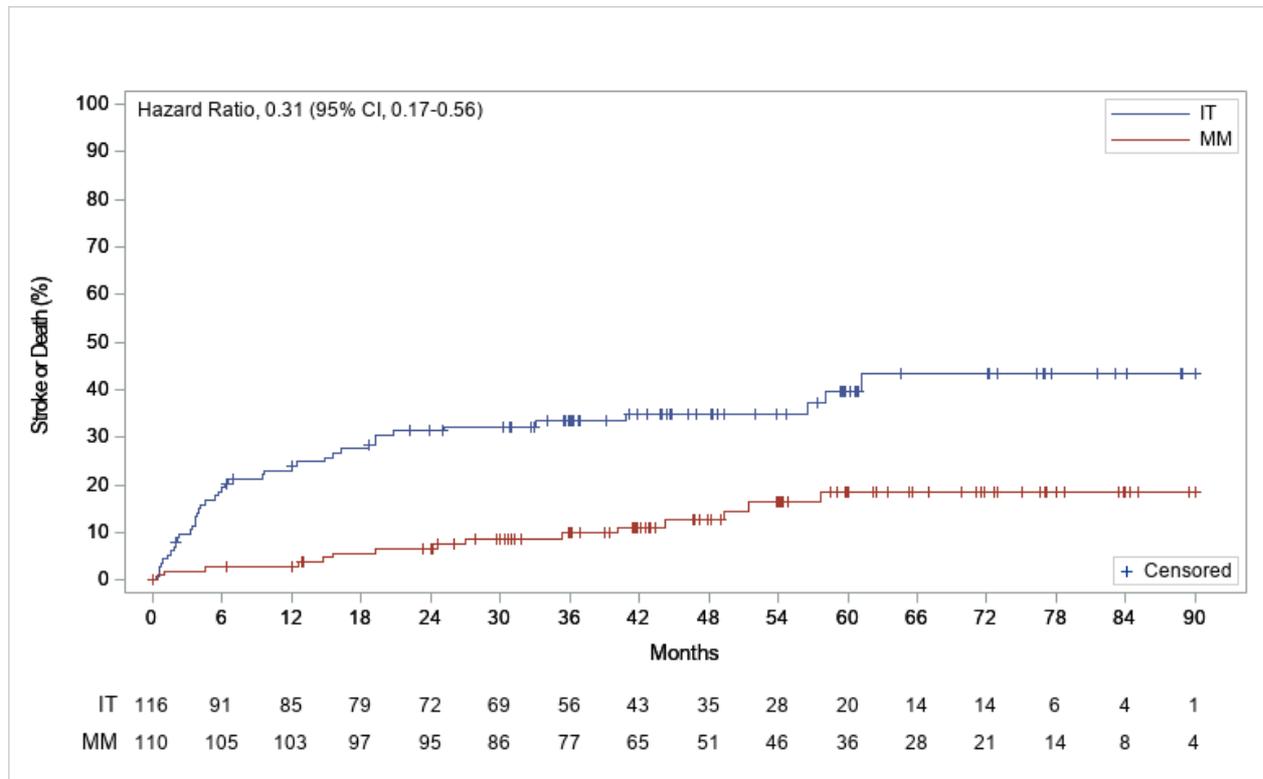
MM, medical management; IT, interventional therapy; AVM, arteriovenous malformation; CI, confidence interval

Figure 1 Trial profile



‡ One additional patient in the IT group died ~1 month after the 5 year visit

Figure 2 Risk of the primary outcome by randomised assignment to medical management alone vs medical management with interventional therapy. Crosses depict censored patients. Numbers below the x-axis indicate the numbers at risk at the start of each follow-up interval.



MM, medical management; IT, interventional therapy;

**A Randomised trial of Unruptured Brain Arteriovenous malformations (ARUBA): Final Results
Medical management with interventional therapy versus medical management
alone for unruptured brain arteriovenous malformations (ARUBA): final follow-up
of a multicentre, open, parallel group, randomised controlled trial**

J.P. Mohr, MD¹; J.R. Overbey, DrPH²; A. Hartmann, MD³; R. von Kummer, MD⁴; R. Al-Shahi Salman, MD⁵; H. Kim, PhD⁶; H.B. van der Worp, MD⁷; M.K. Parides, PhD⁸; M.A. Stefani, MD⁹; E. Houdart, MD¹⁰; R. Libman, MD¹¹; J. Pile-Spellman, MD¹²; K. Harkness, MD¹³; C. Cordonnier, MD¹⁴; E. Moquete, RN²; A. Biondi, MD¹⁵; C.J.M. Klijn, MD^{7,16}; A.J. Moskowitz, MD²

1. Doris & Stanley Tananbaum Stroke Center, The Neurological Institute, Columbia University Irving Medical Center, New York, NY, USA.
2. International Center for Health Outcomes and Innovation Research, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
3. Charité-Universitätsmedizin Berlin; Department of Neurology, Klinikum Frankfurt/Oder, Frankfurt/Oder, Germany.
4. Department of Neuroradiology, University Hospital Dresden, Dresden, Germany.
5. Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK.
6. Department of Anesthesia and Perioperative Care, Center for Cerebrovascular Research, University of California - San Francisco, San Francisco, CA, USA.
7. Department of Neurology and Neurosurgery, Brain Center, University Medical Center Utrecht, Utrecht, Netherlands.
8. Department of Cardiovascular & Thoracic Surgery, Albert Einstein Medical Center, Bronx, New York, USA.
9. Department of Neurology and Neurosurgery, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil.
10. Department of Neuroradiology, APHP-Hôpital Lariboisière, Univ Paris Diderot-Sorbonne Paris Cité, Paris, France.

11. Department of Neurology, North Shore Long Island Jewish Medical Center, New Hyde Park, NY, USA.
12. Department of Neuroradiology, NYU-Winthrop Medical Center, Mineola, NY, USA.
13. Department of Neurology, Royal Hallamshire Hospital, Sheffield, UK.
14. Université Lille, Inserm U1171, Degenerative and Vascular Cognitive Disorders, CHU Lille, Department of Neurology, Lille, France.
15. Department of Interventional Neuroradiology, Jean Minjot Hospital, University of Franche Comté, Besançon, France.
16. Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands.

Corresponding author:

JP Mohr, MD

Neurological Institute

Columbia University Medical Center

710 West 168th Street

New York, NY, 10032, USA

Tel: (1) 212 305 8033

Fax: (1) 212 305 5796

E-mail: jpm10@cumc.columbia.edu

Word Count: 3,066

ABSTRACT

Background: ~~The clinical benefit of preventive eradication of unruptured brain arteriovenous malformations (AVMs) remains controversial. A Randomised trial of Unruptured Brain Arteriovenous malformations (ARUBA) showed that medical management alone was superior to the combination of medical and interventional therapy over a mean follow up of 33 months.~~

~~However, whether these differences persist in the longer term is unknown. :~~ In A Randomized trial of Unruptured Brain Arteriovenous malformations (ARUBA), randomisation was halted by recommendation of the National Institute of Neurological Disorders and Stroke appointed data and safety monitoring board at a mean follow-up of 33.3 months after a pre-specified interim analysis demonstrated that medical management alone was superior to the combination of medical management and interventional therapy in preventing symptomatic stroke or death. We aimed to study whether these differences persist in the longer-term of 5 years follow-up.

Methods: ARUBA was an open-label, randomised (1:1), parallel group trial of adult patients diagnosed with an unruptured brain AVM, who had never undergone interventional therapy, and were considered by participating clinical centres suitable for intervention to eradicate the lesion. The trial comparing medical management alone with standard interventional therapy versus medical management ~~alone~~. (neurosurgery, embolisation, or stereotactic radiotherapy, alone or in any combination, sequence or number. Patients were randomised at 39 international clinical centres with randomisation stratified by clinical centre using a random permuted block design generated by a trial statistician and implemented via a central web based data collection system. The primary outcome was death or symptomatic stroke stroke confirmed by imaging, assessed by a neurologist at each centre not involved in the management of participants' care, and monitored independently using an adaptive approach with interim analyses.

~~After recruitment was halted in April 2013, follow-up continued until July 2015. Here we extend the analysis of the primary outcome until final follow-up 29 months after the blind was broken and assess the risk of death or disability (measured as modified Rankin Scale score ≥ 2) for those eligible for status report at five years after randomisation. Enrolment began on April 4th, 2007 and was halted on April 15th, 2013, after which follow-up continued until July 15th, 2015. All analyses were by intention-to-treat. The trial is registered with ClinicalTrials.gov, number NCT00389181.~~

Findings: ~~Among the 226 patients were randomly allocated, 110 to medical management alone and 116 to medical management plus interventional therapy. During a mean duration of follow-up of 50.4 ± 22.9 months. Medical management alone remained superior to interventional therapy, the incidence of death or symptomatic stroke was lower with medical management alone (15/110, 3.39 per 100 patient-years) compared to medical management with interventional therapy (41/116, 12.32 per 100 patient-years; (Hazard Ratio=0.31, 95% CI 0.17-0.56, p<0.0001) with 15 primary outcome events in the 110 patients randomised to medical management (incidence rate: 3.39 per 100 patient-years) compared to 41 in the 116 patients randomised to interventional therapy (incidence rate: 12.32 per 100 patient-years). Two patients in the medical management group and four patients in the interventional therapy group died during follow-up (two attributed to intervention). The risk of death or disability at 5 years was also lower after medical management (n=9/51, 17.6%) compared to interventional therapy (n=17/45, 37.8%; Relative Risk: 0.47, 95% CI 0.23-0.94). Interventional therapy led to significantly more adverse events, including epileptic seizures and neurological deficits. Adverse events were observed less often (283 vs 369; 58.97 vs 78.73 per 100 patient-years; risk difference -19.76 ((95% confidence interval, -30.33 to -9.1)) in patients allocated to medical management compared with interventional therapy.~~

Interpretation: ~~After extended follow-up, Medical management alone remained superior to interventional therapy for the prevention of death or symptomatic stroke in patients randomised to ARUBA, and death or disability, over five years in patients with an unruptured brain AVM. The rate of outcome events in the even longer term remains unknown. These data should affect standard specialist practice and the information presented to patients. The longer-term risks and difference between the two therapeutic approaches are uncertain~~

Funding:

National Institute of Neurological Disorders and Stroke (NINDS) U01NS051483 & U01NS051566 (2007-2013)

Vital Projects Fund, private gift to Columbia University (2014-2015)

Word Count: 299

INTRODUCTION

A Randomised trial of Unruptured Brain Arteriovenous malformations (ARUBA) addressed the longstanding uncertainty about whether medical management with or without interventional therapy is superior for ~~those the management of patients~~ with brain arteriovenous malformations (AVM) ~~that has who have~~ never bled.¹⁻⁷

The first patient was randomised in April 4th 2007. ~~At a planned meeting with the investigators on April 15th 2013, the~~ randomisation ~~in phase of~~ the trial was terminated prematurely ~~following~~ ~~on~~ the recommendation of its independent ~~D~~data and ~~S~~safety ~~M~~monitoring Board (~~DSMB~~). Their decision was based on the results of a planned interim analysis demonstrating a log-rank Z statistic value (4.10) exceeding the pre-specified stopping boundary value (2.87), early and consistent separation of survival curves in two arms, and strong magnitude of effect. ~~The initial results found that a~~ After a mean follow-up of 33 months, medical management alone (~~MM~~) was

superior to interventional therapy with medical management (IT) for the prevention of death or stroke (as randomised hazard ratio (HR) 0.27, 95% CI 0.14-0.54, $p=0.0001$; as treated HR 0.19, 95% CI 0.09-0.38, $p<0.0001$).⁸ Further analyses incorporating modified Rankin Scale (mRS) score ≥ 2 found that showed that medical management was also superior to medical management with interventional therapy MM was superior to IT for the prevention of fatal or disabling stroke (defined as a mRS score ≥ 2).⁹

In response to the recommendation by the data and safety monitoring board DSMB, but without continued funding by the National Institute of Neurological Disorders and Stroke post-randomisation follow-up efforts continued as originally planned in order to evaluate treatment effects at 5-years. Centre participation continued until the database was closed in July 15, 2015. The initial and combined extended results were presented at the 2016 International Stroke Conference¹⁰ and comprise the present manuscript. The present report describes outcomes observed over the initial randomisation phase and during continued follow-up to assess whether medical management alone remained superior to medical management plus interventional therapy

METHODS

Trial design and randomisation

The study design and randomisation procedures have been described previously.¹ Briefly, ARUBA was a an open, parallel group trial in which participants were randomised 1:1 to medical management with interventional therapy (IT; ie, neurosurgery, embolisation, or stereotactic radiotherapy, alone or in any combination or sequence) or medical management alone (MM; ie, pharmacological therapy for neurological symptoms as needed). A sample size of 400 guaranteed 80% power to detect a 46% reduction in the hazard of death or stroke (due to haemorrhage or infarction), equivalent to a hazard ratio of 0.54. This hazard ratio corresponds to an absolute decrease in 5-year event rates of 9.5% for medical management, from an assumed 5-year event rate of 22% for interventional therapy. The original design for ARUBA included a total sample size of 800 patients which guaranteed 87.5% power to detect a 40% reduction in the hazard for death or symptomatic stroke over 5 years based on assumed 5-year event rate of 22% in interventional therapy; however, because of lower than expected accrual rates after 18 months of randomisation, and the opportunity for longer follow up to achieve the outcome event rates,

~~the DSMB accepted the revised design of 400 patients.~~ Patients were randomised in nine countries at 39 clinical centres (Supplemental Table 1). Protocol and consent forms were approved by the relevant institutional review boards or equivalent ethics committees at all institutions.

Participants

~~These Patients~~ eligible for randomisation were adults (age ≥ 18) whose brain arteriovenous malformation had never AVM had not bled, and were considered by the local centres suitable for attempted lesion AVM eradication by the local centres. Each patient provided written informed consent. No control in case selection was exerted by the clinical coordinating centres (ARUBA-WEST: Columbia University Medical Center; ARUBA-EAST: Hôpital Lariboisière). A full list of eligibility criteria can be found in the study protocol.¹

Procedures Interventions

Interventional therapy options comprised endovascular embolisation, neurosurgical resection, or stereotactic radiosurgery, as single or multiple therapies, in any order, number, or sequence. No published guidelines for selection or sequence of choice(s) of intervention existed during the trial.¹¹ Centres implemented the approaches considered standard practice in their specialist centre and country.

Formatted: Font: Arial, English (United Kingdom)

During the blinded phase of the trial, patient data were collected at six-month intervals for the first two years of follow-up, thereafter annually, with the goal of final status report at year 5, the planned end of the trial by the original protocol. Follow-up was expected to continue also for those experiencing a primary outcome event. Although clinical centres were free to undertake whatever management they deemed appropriate *after* the primary outcome event, reports of outcomes depended on the willingness of those affected to continue in follow-up. After the end of the randomisation phase, supported by Vital Projects Fund, New York, NY and without funding by the National Institute of Neurological Disorders and Stroke, clinical centres continued efforts at follow-up and reported the status of their patients until July 15th 2015 when the database was closed, eight years and two months from the start of accrual of randomisations. Due to lack of continued funding, further follow-up was not feasible. Supplemental Figure 1 provides the number of last follow-up reports by year.

Outcomes

The primary outcome was the composite event of death from any cause or symptomatic stroke from haemorrhage or infarction, documented by imaging (CT or MR scan). The secondary endpoint was death or functional outcome measured at 5 years after randomisation. Poor functional outcome was defined as a score ≥ 2 on the modified Rankin scale (mRS), which ranges from 0 to 6, with higher scores indicating more severe disability.^{1,2} More information on study conduct, data collection, and outcome assessments are available in the primary publication.³

Randomisation and masking

Randomisation was stratified by clinical centre using a random permuted block design with block sizes of 2, 4, and 6. The randomisation sequence was generated by a trial statistician at the data coordinating centre (the International Center for Health Outcomes and Innovation Research at the Icahn School of Medicine at Mount Sinai) and assignment was controlled via a central web-based data collection system, which did not reveal treatment allocation until all baseline data had been submitted. Site

coordinators randomised participants after verifying eligibility and obtaining patient consent.

Blinding and Outcome Assessment

All individual centres were aware of the treatment assignment for their own patients but were not informed of the outcomes from other centres in the trial. ~~Assessment of e~~Outcomes at each centre was by a senior study neurologist who did not perform interventional procedures, not involved in the individual patient's care, but qualified to perform the NIH Stroke Scale (NIHSS) and the mRS, and who had agreed to the classification of case histories and images circulated by study leaders before the start of the trial. Clinical coordinating centre personnel and outcome events committees were blinded to treatment assignment.

Participants and study leaders including those at the clinical coordinating center, remained blinded as to the overall randomisation assignments and outcomes until April 15, 2013 when the results of a planned interim analysis were provided to the trial executive committee at a meeting with the data and safety monitoring board.

~~.The first public presentation of the results was during the 22nd European Stroke Conference in May 31, 2013. Acting on the DSMB recommended follow up, ARUBA centres continued to report outcomes, albeit no longer blinded to the overall study results, until the database was closed in July 15, 2015.]~~

Outcomes

~~¶~~The primary outcome was time to the composite event of death from any cause or symptomatic stroke (haemorrhage or infarction), documented by imaging (CT or MR scan). ~~through~~ Until the end of the randomisation phase in 2013 all primary outcome events were adjudicated by a four-member committee comprised of internationally-renowned experts in stroke neurology, endovascular therapy, radiosurgery, or vascular neurosurgery. All primary outcomes reported after the end of the ~~recruitment~~ randomisation phase were adjudicated by a single member of the original adjudication committee (Prof Marie-Germaine Boussier, Hôpital Lariboisière, Paris, France).

The secondary outcome was death or neurological disability at five years after randomisation. Neurological disability was defined as a score ≥ 2 on the modified

Formatted: Font: (Default) Arial

Rankin scale (mRS), which ranges from 0 to 6, with higher scores indicating more severe disability with 6 indicating death.¹¹ Additional secondary outcomes included the incidence of adverse events which were collected systematically and adjudicated by the event adjudication committee. Definitions of adverse events are available in the protocol included in the supplement. More information on study conduct, data collection, and outcome assessments are also available in the primary publication.¹

Follow-up

~~During the blinded phase of the trial, patient data were collected at six-month intervals for the first two years of follow-up, thereafter annually, with the goal of final status report at year 5, which was the planned end of the trial by the original protocol. Follow-up was expected to continue even for those experiencing a primary outcome event. Although centres were free to undertake whatever management they deemed appropriate after the primary outcome event, reports of outcomes depended on the willingness of those affected to continue in follow-up.~~

~~After the end of the randomisation phase, centres continued initial efforts to follow-up and reported the status of their patients through July 2015. Supplemental figure 1 provides the number of last follow-up reports by year.~~

Statistical analysis

The initial protocol, submitted and approved by two National Institute of Neurological Disorders after two separate Stroke Study Section reviews, had a sample size of 800 patients which would have an estimated 87.5% power to detect a 40% reduction in the hazard for death or symptomatic stroke over 5 years based on an assumed 5-year event rate of 22% in the medical management and interventional therapy arm. The study was overseen by a National Institute of Neurological Disorders and Stroke-appointed data and safety monitoring board.

During the study start-up period, the anticipated number of participating clinical centres was not realized, resulting in a lower than expected recruitment rate. Eighteen months after the first randomisation, given the opportunity for longer follow-up to achieve the outcome event rates within the period of National Institute of Neurological Disorders and

Stroke funding, the data and safety monitoring board reviewed the emerging data by treatment group in private, keeping clinical investigators blinded. The board accepted a revised sample size of 400 patients presented by the study statisticians, which would have an estimated 80% power to detect a 46% reduction in the hazard of death or symptomatic stroke, equivalent to a hazard ratio of 0.54. This hazard ratio corresponded to an absolute decrease in 5-year event rates of 9.5% for medical management alone, from an assumed 5-year event rate of 22% for medical management with interventional therapy. ARUBA had an adaptive design involving pre-specified interim analyses. Two interim analyses were pre-specified in the protocol with early stopping boundaries defined by an O'Brien Fleming-type spending function using a Lan-DeMets approach. There were no interim assessments for futility since the study was set up to assess superiority and non-inferiority.

The statistical analysis was carried out by the data coordinating center. For the primary outcome, Cumulative event-free survival curves for each group were estimated by the Kaplan-Meier method. The primary null hypothesis was tested in an intent-to-treat analysis using a 0.05 level log-rank test. The incidence rate of primary outcome events is the number of patients who had an event divided by the number of event-free patient years observed. A Cox proportional-hazards regression model was used to estimate hazard ratios (HR) and corresponding 95% confidence interval. To account for potential clustering effects by clinical centre, a Cox model including a frailty term for clinical centre was also explored. Risk of death or clinical impairment at five years after randomisation was summarised using poor functional outcome was based on the proportion of patients with mRS ≥ 2 at 5 years and compared between groups using a Chi-square test. Five-year mRS scores were based on mRS assessments documented between 54 and 66 months from randomisation for participants whose date of randomisation made them eligible for such assessment before the closure of the trial dataset. For patients who had multiple assessments in this window, the assessment closest to the expected 60-month (five years) follow-up visit date was selected as the patient's 5-year mRS. Patients who died before month 66 were assigned a mRS score of 6, irrespective of the cause of death.

Formatted: Font: (Default) Arial, English (United Kingdom)

~~Frequencies Group differences in the incidence rates of adverse events including all strokes, focal deficits, seizures, and headaches were computed by allocated treatment group, as well as the rate difference between the groups and the corresponding 95% confidence interval compared using Poisson regression.~~ Duration of follow-up in months was calculated using the last date of contact with each patient. Exploratory subgroup analyses of the primary outcome were conducted using Cox proportional hazards models and tested for heterogeneity of treatment effect using interaction terms.

All analyses were conducted by the intention-to-treat principle using SAS version 9.4 (Cary, NC). Due to the descriptive nature of the study there is no bias adjustment due to the adaptive design. Additional information on the timing and results of the two pre-specified interim analyses are given in supplemental figure 2.

~~Patients who were randomised to MM alone but who changed to IT and subsequently received IT were analysed as MM in 'as randomised' analyses but were reported in IT in 'as treated' analysis if the reason for intervention was other than brain AVM rupture. Patients randomised to IT who never received it or who suffered a stroke before the initiation of interventional therapy were reported in the IT for the 'as randomised' analysis but reported in MM for the 'as treated' analysis.~~

The trial is registered with ClinicalTrials.gov, number NCT00389181.

Role of the Funding Source

NINDS participated in study design, data interpretation, and writing of the report, but had no role in data collection or data analysis. The Vital Projects Fund had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study after the end of the randomisation phase and had final responsibility for the decision to submit for publication.

RESULTS

Of 1740 patients screened, 1014 (58.3%) were ineligible and 726 were eligible, of whom 323 (44.5%) refused participation and 177 (24.4%) decided their management

Formatted: Font: (Default) Arial, English (United Kingdom), Pattern: Clear (White)

~~outside the trial (figure 1); outcomes were not collected for eligible patients who were not randomised.~~ Thirty-nine international centres randomised a total of 226 participants at a steady rate of 3.2 per month from April 4, 2007 to April 15, 2013 (supplemental figure 2). A CONSORT diagram of patient retention at key intervals is shown in figure 1. It is based on the years of participation and not calendar dates. ~~No outcomes were reported for the 177 eligible patients whose treatment choices were made outside the trial, many from centres who enrolled no patients. Almost half of the participants were randomised within 3 years (before April 15, 2010) and finished their 5 year participation before or during 2015; for the other half, recruited later, most were not eligible for a formal 5 year report when the database closed, although many had their status reported during the final two years of follow up while the database remained open.~~

Of the 226 patients randomised, 110 were allocated to medical management alone MM and 116 to medical management plus interventional therapy IT. Three patients randomised in the interval between data lock for the final DSMB report and presentation to the data and safety monitoring board DSMB in April 2013, were not included in the primary publication,¹ but are included in the current report. At final data lock on July 15, 2015, patients randomised, ~~regardless of outcome status,~~ had a mean length of follow-up of 50.4 months (SD±22.9; median 48.0, IQR 35.9-71.1). The average follow-up for patients allocated to medical management MM patients was 52.4 months (SD±23.7; median 49.1, IQR 36.1-71.8), and for patients allocated to interventional therapy IT patients was 48.5 months (SD±22.0; median 45.5, IQR 34.7-62.1). The distribution of dates of the last patient contact is shown in supplemental figure 1.

Baseline characteristics and mRS scores were ~~previously reported for the first 223 randomised patients and were similar~~ between groups.¹ apart from focal neurological deficits at presentation, small AVM nidus size and Spetzler Martin grade. ~~An updated baseline table including all 226 randomised patients is given in the supplement (Supplemental Table 1). Minor imbalances existed towards more focal neurological symptoms at presentation in patients allocated to IT.~~

Between April 15, 2013 and July 15, 2015, four new primary outcome events were reported in the MM arm and six in the IT arm. Combined with previously reported events, there were a total of 15 primary outcome events (incidence rate: 3.39 per 100 patient-years) in patients randomised to MM versus 41 (incidence rate: 12.32 per 100 patient-years) in patients randomised to IT (Table 1). Compared with the primary publication, the updated analysis of the primary outcome continued to favour MM over IT, in both the ‘as randomised’ (HR 0.31, 95% CI 0.17-0.56, $p < 0.0001$) and ‘as treated’ (HR 0.22, 95% CI 0.12-0.41, $p < 0.0001$) analyses (Figure 2). Based on the primary outcome analysis, the number needed to harm by IT at 5 years was 5 (95% CI 3-13) when analysed by ‘as randomised’ groups and 3 (95% CI 2 to 6) when analysed by ‘as-treated’ groups. In total, there were 15 primary outcome events in patients randomised to medical management (incidence rate: 3.39 per 100 patient-years) versus 41 in patients randomised to interventional therapy (incidence rate: 12.32 per 100 patient-years; Table 2), resulting in a hazard ratio of 0.31 (95% CI 0.17-0.56; Figure 2). These results remained consistent after accounting for the potential clustering effect of clinical centre (adjusted HR 0.31, 95% CI 0.16-0.61). Of the 56 primary outcomes, 10 were new events reported between April 15th 2013 and July 15th 2015. Four of these 10 occurred in the medical management arm and six in those randomised to medical management plus interventional therapy. In total, two patients allocated to medical management alone and four patients allocated to medical management plus interventional therapy died during follow-up. In the latter group two of the four deaths were attributed to the intervention (Table 2).

reated’ groups.

The risk of the secondary outcome of death or neurological disability at five years after randomisation, available for 96 patients, was lower for those allocated to medical management alone (Table 2). Supplemental Figure 4 illustrates the distribution of mRS scores by study arm for those with data available at 5 years.

Subgroup analyses of the primary outcome were consistently in favour of MM over IT, except for venous drainage, AVM maximum nidus size, and Spetzler-Martin grade, where there was statistically significant heterogeneity of the treatment effects. Figure 3 shows the primary outcomes by subgroup in an ‘as randomised’ analysis. For patients with Spetzler-Martin Grade I

the 'as randomised' analyses appeared to favour IT (HR 1.82, 95% CI: 0.46, 7.28) but the 'as treated' analyses appeared to favour MM (HR 0.66, 95% CI: 0.18, 2.44, supplemental figure 3).

Figure 4 illustrates the distribution of mRS scores by study arm for those with data available at 5 years. The risk the secondary outcome of death or neurological disability at five years after randomisation, available for 96 patients, poor functional outcome (i.e., mRS score ≥ 2) was lower for those allocated to medical management alone MM (n=9/51, 17.6%) compared to those allocated to IT (n=17/45, 37.8%; RR 0.47, 95% CI 0.23-0.94, p=0.03). A similar effect is seen in the 'as treated' analysis (RR 0.41, 95% CI 0.20-0.83, p=0.009).

Patients allocated to interventional therapy IT experienced significantly more adverse events compared to those allocated to medical management MM (78.73 versus 58.97 per 100 patient-years; p<0.001), including more epileptic seizures (20.27 versus 14.17 per 100 patient years; p=0.02), and more focal neurological deficits (4.27 versus 0.62 per 100 patient years; p<0.001). No difference was observed in the rates of documented episodes of headache (Supplemental Table 2).

Subgroup analyses of the primary outcome were consistently in favour of medical management over medical management plus interventional therapy, except for venous drainage (superficial only versus any deep), arteriovenous malformation maximum nidus size (<3 cm versus ≥ 3 cm), and Spetzler-Martin grade, where there was heterogeneity of the treatment effects (Figure 3). The effect of medical management in patients with Spetzler-Martin Grade I AVMs appeared to differ from Spetzler-Martin Grade II-V AVMs, although the effect in the Spetzler-Martin Grade 1 group was not significant (HR 1.82, 95% CI 0.46-7.28).

Eight patients who were randomised to medical management alone MM crossed over to received medical management plus interventional therapy =IT. Of the 116 randomised to interventional therapy IT, three experienced an outcome before interventional therapy began 18 are included in the MM group in the as-treated analysis (3 experienced an outcome before IT began and 15 never received interventional therapy IT). Among the 106 patients who received interventional

~~therapy IT prior to experiencing an outcome,~~ the median time from randomisation to first intervention was 76 (IQR 42-136) days. For the 43 patients who reached a primary outcome following the initiation of ~~interventional therapyIT,~~ the median delay after the last intervention was 1 day (IQR 0-43). Sixty-eight ~~(64.2%) of the 106 patients who received interventional therapy (64.2%)~~ were treated by a single modality ~~of IT,~~ while for 38 (35.8%) the therapy was multimodal. At the time of final data lock, 47 (44.3%) of those receiving ~~I medical management plus interventional therapy T treatment~~ had angiographic evidence of brain ~~arteriovenous malformation AVM~~ eradication, 43 (40.6%) had evidence of a brain ~~arteriovenous malformation AVM~~ remnant on last follow-up imaging, and in 16 (15.1%, all post radiotherapy) the brain ~~arteriovenous malformation AVM~~ status was unknown due to missing follow-up imaging (Table 2).

~~Twenty-two of 106 (20.8%) Of the 22~~ patients who ~~who were treated with medical management and interventional~~ underwent neurosurgery, either alone or as part of a multimodal treatment strategy following embolisation. ~~In 21 of 22 (95.5%) were free of evidence of brain AVM arteriovenous malformation had been eradicated based on a post-operative angiogram. Nine of the 22 (40.9%) patients who were operated on~~ experienced a primary outcome event. Stereotactic radiotherapy was used in the treatment of 57 ~~of 106~~ patients, either alone or as part of a multimodal strategy with either embolisation alone or with both endovascular and surgical therapy (n=1). ~~At the time of the final analysis, 12 of 57 patients (21.1%) had reached angiographically-documented absence of the brain arteriovenous malformationAVM, and 21 (36.8%) had had a primary outcome event. Finally, 66 Sixty-six of 106 (62.3%) patients were treated by endovascular embolisation, either alone or as part of a multimodal treatment strategy with neurosurgery and/or radiotherapy; 34 of 66 (51.5%) demonstrated absence of the brain AVM on catheter angiography, and 33 of 66 (50.0%) experienced a primary outcome event (Table 2).~~

DISCUSSION

~~This updated analysis of ARUBA, with mean follow up extended from 33 to 50 months, continues to show that for patients with an unruptured bran AVM medical management alone is~~

superior to intervention management (single or multimodality) as prophylaxis for haemorrhage, both during the randomisation period of trial and thereafter.

With extended follow-up of 226 participants with unruptured brain arteriovenous malformations in ARUBA, the risk of death or stroke remained significantly lower after medical management alone than after medical management with interventional therapy after a mean follow-up of 50 months. Ninety-six participants followed up for 60 months without intervention also had a significantly lower risk of death or neurological disability, and fewer adverse events. Although the persisting difference between the two management options persisted, ideally longer follow-up would be desirable but was not possible due to funding constraints.

ARUBA remains the only randomised trial comparing MM with or without IT for patients with an unruptured brain AVM.^{11,13} The rationale for the trial and its design have been reviewed elsewhere.¹⁴ The 37 academic and two private clinical centres who enrolled patients in ARUBA had ample prior published experience with interventional treatment of brain AVMs: 630 publications in PubMed as of July 2019. The 226 patients randomised in the ARUBA trial represent 31% of the 726 screened patients deemed eligible for enrolment. Of the remainder of the screened eligible patients, 323 refused participation in the trial and 177 were treated according to the wishes of their own clinician, outside of the trial. These screening data include information from 23 centres who contributed screening and eligibility data but enrolled no patients, all such issues a hindrance to success for neurosurgical trials.¹⁵

This longer-term follow-up report of ARUBA has limitations. The number of patients included in the study was much smaller than the 800 patients initially planned, but with 223 patients, whos disparity in outcomes led to the early suspension of randomisation. In addition, due to limited resources, the duration of follow-up was shorter than the five years for all participants that had initially been planned. Although 26 of 226 patients were lost to follow-up at the time of the final data lock, their numbers were similar in both arms, so that potentially missed outcomes for them are unlikely to have had a large effect on the reported results. Because of the smaller number of included patients than initially planned, the estimates of the treatment effect in the subgroup analyses were less precise than anticipated.

Strengths of ARUBA are its randomised design and its inclusion of patients at 39 centres, which enhances the generalisability of its findings. A recent systematic review in the Cochrane Database (search date January 14th 2019) found ARUBA was the only published randomised trial comparing medical management with interventional therapy for unruptured brain arteriovenous malformations.¹²

Although generalizability of ARUBA's results has been questioned¹⁶, baseline characteristics of the trial population, observed event rates, and the direction as well as the magnitude of the effects of interventional therapy were remarkably similar to those seen in a contemporary prospective population-based study, in which the association between interventional therapy and the risk of stroke persisted for up to 12 years.¹⁷ Patient characteristics, outcome rates, and mRS values observed in the ARUBA trial are also comparable to those seen in other populations who presented without haemorrhage and who did not receive interventional therapy, which were published in two single-centre reports.^{18,19} These measures were also similar to a meta-analysis of published reports of treated patients aggregated regardless of haemorrhage status.²⁰

Strengths of ARUBA are its randomised design and its inclusion of patients at 39 centres, which enhances the generalisability of its findings. A recent systematic review in the Cochrane Database (search date January 14th 2019) found ARUBA was the only published randomised trial comparing medical management with interventional therapy for unruptured brain arteriovenous malformations.¹² The study also has striking similarities with the literature: ARUBA was consistent with the outcomes for intervention in a meta-analysis¹³ from the few cohorts of untreated unruptured brain AVMs,^{14, 15} and a non-randomised cohort study with concurrent controls with follow-up for up to 12 years.¹⁶ The distribution of the Spetzler-Martin grades of included brain AVMs were bias towards those more safely and easily treated, indicating few participants were unsuitable for interventional therapy.

Previous case series have reported differing risks of clinical outcomes and angiographic obliteration after treatment for brain arteriovenous malformations.¹³ For unruptured brain arteriovenous malformations, the risk of treatment has to be weighed against the risk of rupture with medical management alone, 1.3% per year over a period of 10 years among the largest reported series of 1389 from four major centers.¹⁷ Comparative

observational studies with concurrent controls have also reported worse outcomes associated with interventional therapy for brain arteriovenous malformations compared to medical management over up to 12 years.¹⁶

Patients with an unruptured brain arteriovenous malformation should, therefore, be informed about the absolute and relative risks of both treatment strategies in ARUBA. In addition, the current report may inform the design of other randomised controlled trials seeking to investigate the reproducibility of the ARUBA model. An improved design depends on a better understanding of the natural history of unruptured brain arteriovenous malformations, and in the context of other settings or approaches to improvements in therapeutic interventional therapies.

Three other RCTs are currently ongoing in patients with brain arteriovenous malformations; all three including both patients with ruptured and patients with unruptured brain arteriovenous malformations. One is testing whether two embolisation approaches are equivalent (endovascular embolisation with Onyx versus with TRUFILL n-butyl cyanoacrylate n-BCA;NCT00857662); one whether conservative management or intervention will reduce the risk of death or debilitating stroke and whether endovascular treatment can improve the safety and efficacy of surgery or radiosurgery (NCT02098252), and the third is testing whether transvenous embolisation or trans-arterial embolisation is most effective and safe in achieving angiographic obliteration of the arteriovenous malformation (NCT03691870). The longer-term risks of interventional therapy compared to medical management than reported here will remain unknown unless future randomised trials are sufficiently funded to permit an adequate duration of follow-up.

Efforts to delineate which patients with unruptured AVMs will bleed and the degree of syndrome severity they will experience has been a long-term priority.^{21,22} ARUBA study statisticians had estimated that 12 to 30 years would be needed for the events experienced by MM patients to meet those already experienced by IT patients.²³ This analysis had an impact on the decision by NINDS Study section reviewers who reviewed our proposed 5-year extension of follow-up. Further, although outcomes in the medical arm were infrequent, and few patients with Spetzler-

Martin Grade III to IV AVMs were randomised in ARUBA, a recent analysis of the original ARUBA data failed to show an association between haemorrhage rate and lesion size among patients in the medical management arm.²⁴ The declining slope of haemorrhage events in the Kaplan-Meier curves over the 10-year follow-up for those presenting without haemorrhage in the observational Multicentre Arteriovenous malformation Research Study (MARS) also suggests that the risk of rupture in unbled patients may subside over time.²⁵

How some lesions seem stable for decades or life-time is still unclear.^{18,19} The few studies assessing these risks have documented that haemorrhage risk is related to high intra-nidal pressure²⁰ and to single-vein drainage.²¹ Future studies may extend these findings but likely will depend on further innovations in non-invasive imaging to assess arteriovenous resistivity patterns. Continued interest in long-term outcomes in patients with unruptured brain arteriovenous malformations is being pursued in a large, international observational cohort study funded by National Institute of the Neurological Disorders and Stroke (R01 NS099268) based on the methods used for the Multicenter Arteriovenous Malformations Study (MARS), with a goal of identifying predictors of haemorrhage and treatment risks in >2500 patients.²² Also ongoing is the Treatment Of Brain Arteriovenous malformations Study (TOBAS) comprising two open-label randomised arms.²³

For the participants in ARUBA, we cannot exclude the possibility that patients not appearing at the local centre for follow-up visits in the planned 5-year period had experienced outcome events. However, overall, missing visit numbers are comparable between MM and IT and the low event rate in all reporting centres argues against a significant undetected morbidity or mortality.

No formal guidelines on the management of unruptured brain AVM have emerged from professional associations. Two consensus reports endorsed by the American Heart Association, the first from 2001 cited that treatment results vary considerably²⁴ and the most recent from 2017 that medical management alone and three often complementary methods of interventional therapy exist.²⁵

since the initial publication of the results of the ARUBA trial, although an AHA scientific statement recently endorsed the trial's findings.²⁶ The results of ARUBA demonstrate clinically and statistically significant excess hazard from interventional therapy. Evidence of this hazard should have an impact on standard specialist practice in many countries in several continents, and should be among the materials presented to patients. Continued interest in long-term outcomes in patients with unruptured brain AVMs is being pursued in a large, international observational cohort study funded by NINDS (R01 NS099268) based on the methods used for MARS, with a goal of identifying predictors of haemorrhage and treatment risks in >2500 patients.²⁷ Also ongoing is a study (TOBAS) comprising two open-label randomised arms.²⁸

In summary, after mean length of follow-up 50.4 months (SD±22.9; median 48.0, IQR 35.9-71.1), medical management alone remained superior to medical management with interventional therapy for the prevention of death or symptomatic stroke in patients with an unruptured brain arteriovenous malformation in the ARUBA trial. Evidence of this hazard should have an impact on standard specialist practice and should be among the materials presented to patients. The rate of outcome events and degree of disparity between the two management options beyond four years remain uncertain.

Data Sharing Statement: Trial data collected during the NINDS-funded phase are archived by NINDS and available upon request. Information on how to request the data is available here:

<https://www.ninds.nih.gov/Current-Research/Research-Funded-NINDS/Clinical-Research/Archived-Clinical-Research-Datasets>.

~~Acknowledgement: We would like to thank Dr. Christian Stapf for his substantial contributions to the ARUBA trial design, data collection, and analysis, as well as the initial drafting of this manuscript.~~

Declarations of Interest: Dr. Mohr has nothing to disclose. Dr. Overbey has nothing to disclose.. Dr. Hartmann has nothing to disclose. Dr. von Kummer has nothing to disclose. Prof. Al-Shahi Salman reports reimbursement from the National Institutes of Health, during the conduct of the study. Dr. Kim reports grants from Columbia University, grants from NIH/NINDS, during the conduct of the study. Dr. van der Worp reports other from ARUBA trial, during the conduct of the study. Dr. Parides has nothing to disclose. Dr. Stefani has nothing to disclose. Dr. Houdart has nothing to disclose. Dr. Libman has nothing to disclose. Dr. Pile-Spellman has nothing to disclose. Dr. Harkness has nothing to disclose. Dr. Cordonnier has nothing to disclose. Ms. Moquete has nothing to disclose. Dr. Biondi has nothing to disclose. Dr. Klijn reports grants from Netherlands Cardiovascular Research Initiative, which is supported by the Dutch Heart Foundation, CVON2015–01: CONTRAST, and the support of the Brain Foundation Netherlands (HA2015•01•06), from the Dutch Heart Foundation (Clinical established investigator grant 2012T077), and from The Netherlands Organization for Health Research and Development, ZonMw (grant 015008048), during the conduct of the study. Dr. Moskowitz reports grants from NIH, during the conduct of the study.

RESEARCH IN CONTEXT

Evidence before this study:

~~Despite a literature on management options dating back to the 1930's, no randomised clinical trial had been mounted, nor formal professional, societal, or national guidelines existed prior to ARUBA. The literature had conflicting reports concerning both the natural history of those whose brain AVM was discovered not having bled and outcomes from single and multiple therapies. These uncertainties prompted the support of an international, randomised clinical trial funded by the NINDS~~

~~. Case series have reported different risks of clinical outcomes and angiographic obliteration for brain arteriovenous malformations, which are unreliable for comparison with the best available data indicating a 1% annual risk of haemorrhage from un-treated unruptured brain arteriovenous malformations. Comparative observational studies with concurrent controls have reported worse outcome associated with interventional therapy for brain arteriovenous malformations compared to medical management over up to 12 years. A recent systematic review in the Cochrane Database (search date January 14th 2019) found one published randomised trial comparing medical management with interventional therapy for unruptured brain arteriovenous malformations. The ARUBA trial terminated recruitment when its data monitoring committee concluded that medical management was superior to interventional therapy for the prevention of stroke or death on the basis of the first 223 recruited participants after a mean follow-up of 33 months (HR 0.27, 95% CI 0.14-0.54). The data monitoring committee concluded that there was, "a compelling need for additional long-term data."~~

Added value of this study:

~~The current report includes longer term outcomes than in the initial publication of the randomised phase of the ARUBA trial, now including all 226 participants recruited at 39 international hospitals with mean follow-up extended from 33 months to 50 months. The results of ARUBA are comparable to non-randomised reports from individual and population-based sources, some extending well beyond the time frame of ARUBA. The most recent review is a Cochrane Database Systemic Review published on September 10, 2019. ARUBA was cited as the only randomised clinical trial. The quality of evidence~~

~~was deemed moderate and at low risk of bias but high risk of performance bias due to participants and treating physicians not being blinded to allocated treatment. (The adjudication committee was blinded). A wide range of recommendations were offered for the design of future randomised trials. How to achieve the desired blinding for a treatment plan that may include many hospital admissions and the participation of a large treating team remains a problem. The final results of ARUBA show that medical management remained superior to interventional therapy (HR 0.31, 95% CI 0.17-0.56).~~

Implications of all the available evidence:

The results of ARUBA demonstrate ~~clinically and statistically significant excess hazard-harm~~ from interventional therapy compared to medical management over an average duration of follow-up of more than four years~~for those 226 participants from 37 academic and 2 private international centres originally randomised to medical management alone or with interventional therapy whose outcome status was known as late as 5 years from randomisation. Patients with unruptured brain arteriovenous malformation should be informed about the absolute and relative risks in ARUBA which may inform the design of other randomised controlled trials seeking to investigate the reproducibility of ARUBA in the context of other settings or approaches to interventional therapy. The long-term risks of interventional therapy compared to medical management will remain unknown unless future randomised trials are sufficiently funded to permit an adequate duration of follow-up. The disparities in primary outcomes for the two arms of the ARUBA trial provide a background for those seeking to organize further research based on a randomised clinical trial model. An improved design depends on a better understanding of the natural history of unruptured brain AVMs, and improvements in therapeutic interventions.~~

Table 1 Primary and secondary endpoints by randomisation assignment and as treated

Outcome per Randomisation (intention to treat)	Interventional Therapy (N=116)		Medical Management (N=110)		Effect of Medical management alone	
		Incidence Rate Per 100 Pt-Years		Incidence Rate Per 100 Pt-Years	Hazard Ratio (95% CI)	P Value
Primary Outcome	N		N			
Symptomatic stroke or death (primary outcome)	41	12.32	15	3.39	0.31 (0.17 - 0.56)	<0.0001
Symptomatic stroke						
Any incident stroke	40	11.99	13	2.94	0.27 (0.15 - 0.51)	<0.0001
Haemorrhagic	30		9			
Ischemic	10		4			
Death						
Any*	4	0.85	2	0.42	0.49 (0.09 - 2.67)	0.49
AVM-related	2		0			
Not AVM-related	2		1			
Unknown cause	0		1			
Functional Outcome	No./No. Obs	%	No./No. Obs	%	Relative Risk (95% CI)	P Value
mRS 2-6 at 5 years (secondary endpoint)	17/45	37.8	9/51	17.7	0.47 (0.23 - 0.94)	0.03
Outcome on Treatment** (per protocol)	Interventional Therapy (N=106)		Medical Management (N=120)		Effect of Medical management alone	
		Incidence Rate Per 100 Pt-Years		Incidence Rate Per 100 Pt-Years	Hazard Ratio (95% CI)	P Value
Primary Outcome	N		N			
Symptomatic stroke or death (primary outcome)	43	14.08	13	2.77	0.22 (0.12, 0.41)	<0.0001
Symptomatic stroke						
Any incident stroke	42	13.72	11	2.34	0.19 (0.10, 0.37)	<0.0001
Haemorrhagic	29		10			
Ischemic	13		1			
Death						
Any*	4	0.92	2	0.39	0.43 (0.08, 2.35)	0.33

AVM-related	2		0			
Not AVM-related	2		1			
Unknown cause	0		1			
Functional Outcome						
	No./No. Obs	%	No./No. Obs	%	Relative Risk (95% CI)	P Value
mRS 2-6 at 5 years (secondary endpoint)	17/42	40.5	9/54	16.7	0.41 (0.20 - 0.83)	0.009

* Three patients in the IT arm experienced at least one stroke and eventually died during the course of the trial

**Eight patients randomised to MM crossed over to IT. Among the 116 patients randomised to IT, n=15 never received therapy and n=3 suffered a stroke prior to the initiation of IT; all were therefore considered part of the MM group in the as-treated analysis.

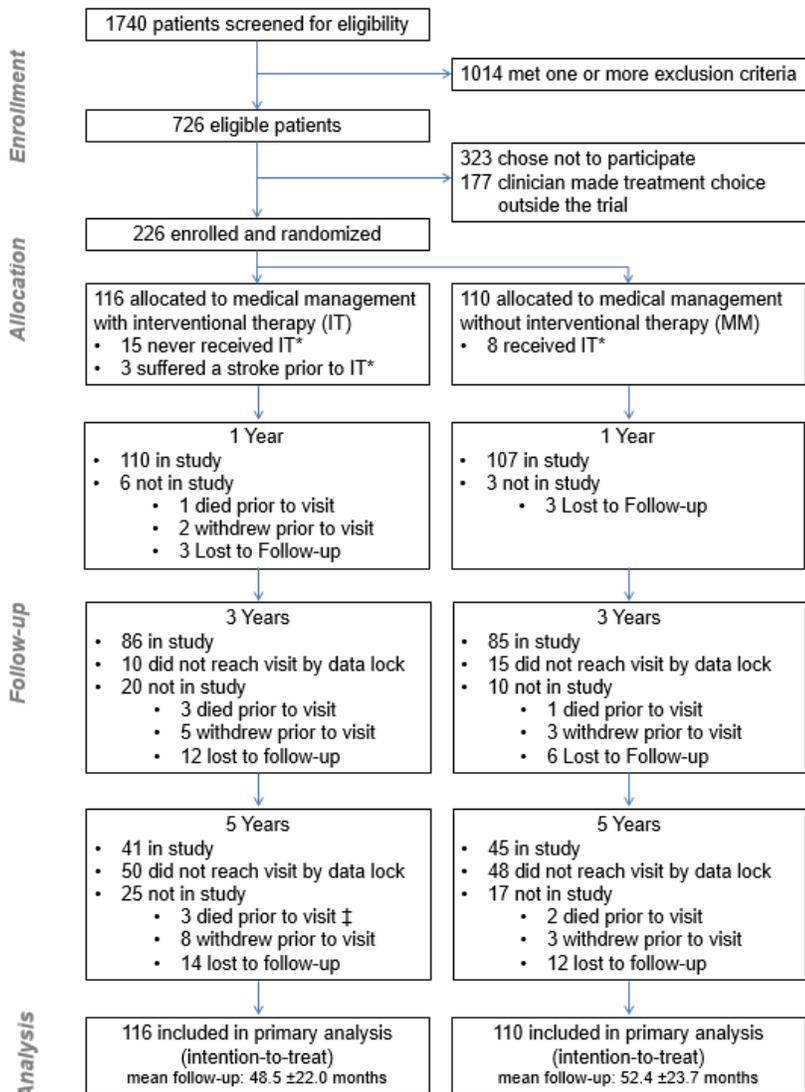
Table 2 Treatment strategy and primary outcome in n=106 patients with unruptured brain AVM undergoing interventional therapy (as-treated analysis)

Interventional treatment received	With primary outcome*	With documented AVM obliteration**
	n (row %)	n (row %)
Monomodal		
Endovascular (n=28)	14 (50.0)	14 (50.0)
Surgery (n=7)	2 (28.6)	7 (100.0)
Radiotherapy (n=33)	8 (24.2)	6 (18.2)
Multimodal		
Endovascular and Surgery (n=14)	6 (42.9)	14 (100.0)
Endovascular and Radiotherapy (n=23)	12 (52.2)	6 (26.1)
Endovascular and Surgery and Radiotherapy (n=1)	1 (100.0)	0 (0)

* Primary outcome: Symptomatic stroke or death

** Documented AVM obliteration required cerebral angiography by study protocol. For n=16 (15%) patients, the AVM obliteration status was unknown due to missing imaging information, n=43 (41%) had a documented AVM remnant on last follow-up imaging.

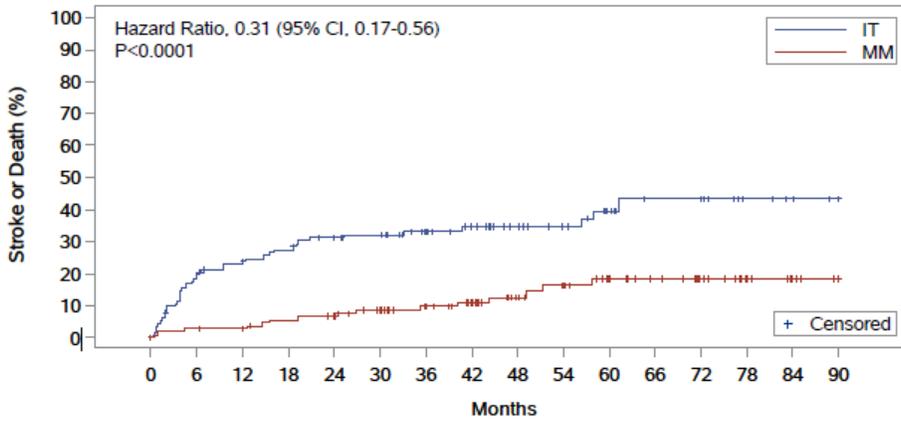
Figure 1: CONSORT Diagram of Patient Flow and Follow-up



* Considered cross-overs in the « as-treated » analysis
‡ One additional patient in the IT group died ~1 month after the 5 year visit

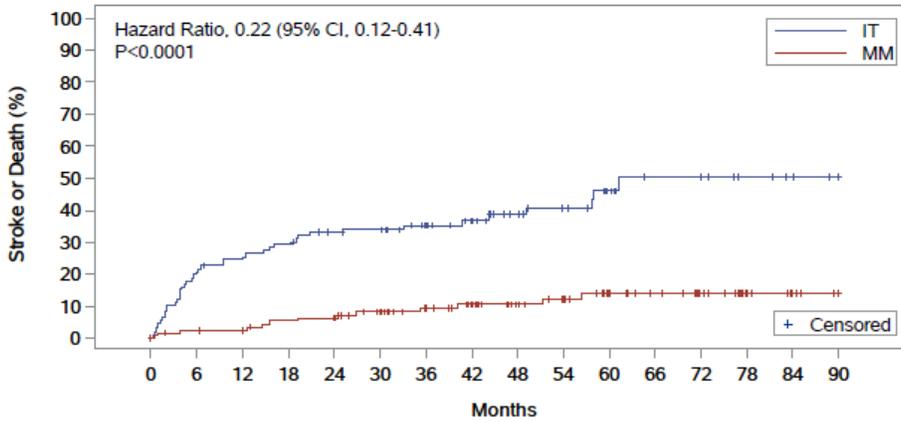
Figure 2: Kaplan-Meier estimated event rates by randomisation assignment (Panel A) and as treated (Panel B). Crosses depict censored patients.

A) As Randomized



IT	116	91	85	79	72	69	56	43	35	28	20	14	14	6	4	1
MM	110	105	103	97	95	86	77	65	51	46	36	28	21	14	8	4

B) As Treated



IT	106	84	79	74	67	65	53	40	30	24	17	11	11	6	4	1
MM	120	112	109	102	100	90	80	68	56	50	39	31	24	14	8	4

Figure 3 Primary outcome sub-group analyses ‘as randomised’. Hazard ratios and 95% confidence intervals for each sub-group are plotted in the centre. MM=Medical management only and IT=Medical management with intervention. ‘As treated’ analyses are shown in the supplemental figure 3.

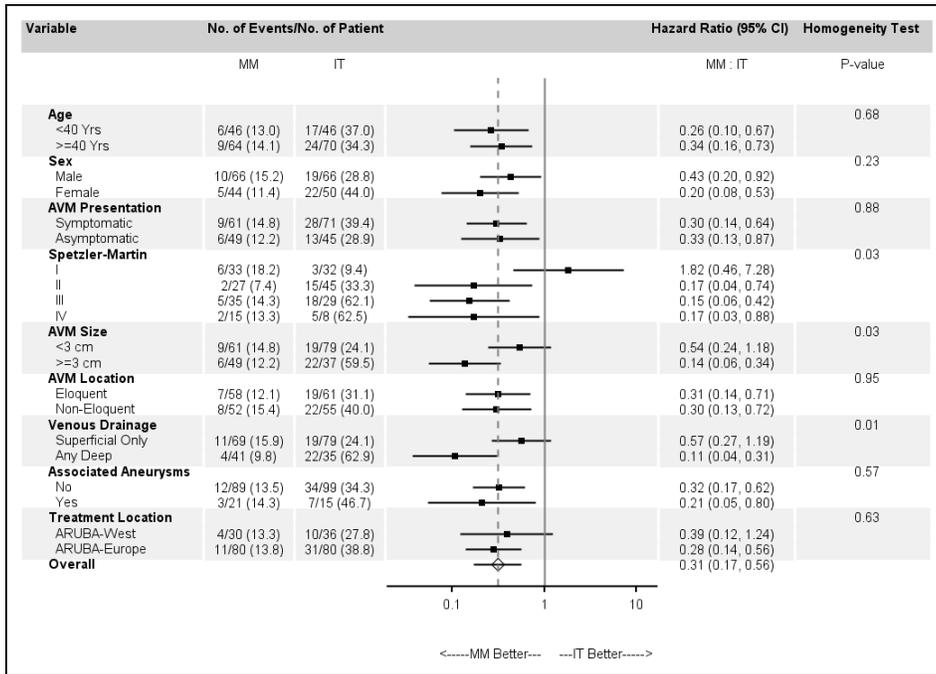
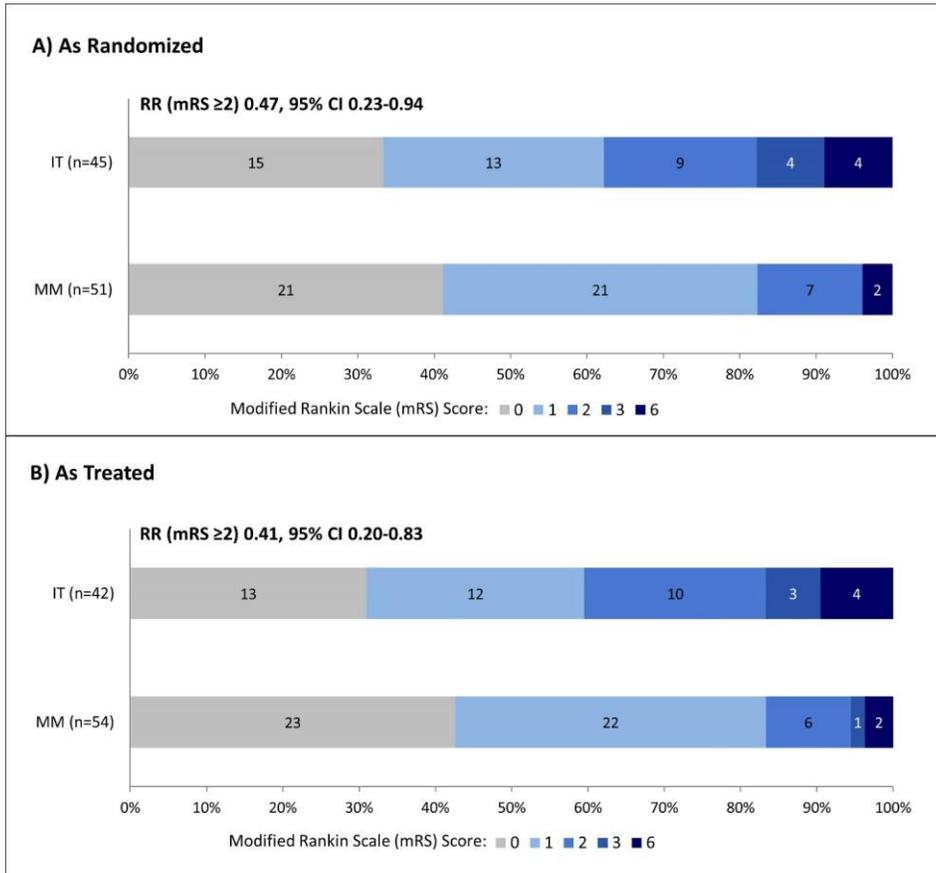


Figure 4: Secondary endpoint analysis: modified Rankin scale scores at 5 years



REFERENCES

1. Mohr JP, Parides MK, Stapf C, et al. Medical management with or without interventional therapy for unruptured brain arteriovenous malformations (ARUBA): a multicentre, non-blinded, randomised trial. *Lancet* 2014;**383**:614–21.
2. Massaro AR, Young WL, Kader A, et al. Characterization of Arteriovenous Malformation Feeding Vessels by Carbon-Dioxide Reactivity. *Am J Neuroradiol* 1994;**15**:55–61.
3. Mast H, Mohr JP, Osipov A, et al. Steal Is an Unestablished Mechanism for the Clinical Presentation of Cerebral Arteriovenous-Malformations. *Stroke* 1995;**26**:1215–20.
4. Wedderburn CJ, van Beijnum J, Bhattacharya JJ, et al. Outcome after interventional or conservative management of unruptured brain arteriovenous malformations: a prospective, population-based cohort study. *Lancet Neurol* 2008;**7**:223–30.
5. Chang SD, Marcellus ML, Marks MP, Levy RP, Do HM, Steinberg GK. Multimodality treatment of giant intracranial arteriovenous malformations (Reprinted from Archival article vol 53, pg 1, 2003). *Neurosurgery* 2007;**61**:432–42.
6. Ondra SL, Troupp H, George ED, Schwab K. The Natural-History of Symptomatic Arteriovenous-Malformations of the Brain - a 24-Year Follow-up Assessment. *J Neurosurg* 1990;**73**:387–91.
7. Hartmann A, Mast H, Choi JH, Stapf C, Mohr JP. Treatment of arteriovenous malformations of the brain. *Curr Neurol Neurosci Rep* 2007;**7**:28–34.
8. Rukovets O. Why the ARUBA Trial on AVMs Was Stopped. *Neurology Today* 2013;**13**:1,16–18.
9. Mohr JP, Overbey JR, von Kummer R, et al. Functional impairments for outcomes in a randomized trial of unruptured brain AVMs. *Neurology* 2017;**89**:1499–506.
10. Stapf C, Overbey J, Mohr JP, Moskowitz AJ, Vicaut E, Parides MA. A Randomized Trial of Unruptured Brain Arteriovenous Malformations (ARUBA): 5 year results. Presented at the American Heart Association International Stroke Conference, February 16–19, 2016. abstract (<http://www.abstractsonline.com/pp8/#!/3859/session/341>).

11. Ogilvy CS, Stieg PE, Awad I, et al. AHA Scientific Statement: Recommendations for the management of intracranial arteriovenous malformations: a statement for healthcare professionals from a special writing group of the Stroke Council, American Stroke Association. *Stroke* 2001;**32**:1458–71.
12. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;**19**:604–7.
13. Zuurbier SM, Al-Shahi Salman R. Interventions for treating brain arteriovenous malformations in adults. *Cochrane Database Syst Rev*. 2019;**9**:CD003436.
14. Mohr JP, Hartmann A, Kim H, Pile-Spellman J, Stapf C. Viewpoints on the ARUBA Trial. *Am J Neuroradiol* 2015;**36**:615–7.
15. Raymond J, Darsaut TE, Molyneux AJ; TEAM collaborative Group. A trial on unruptured intracranial aneurysms (the TEAM trial): results, lessons from a failure and the necessity for clinical care trials. *Trials* 2011;**12**:64.
16. Magro E, Gentric JC, Darsaut TE, Ziegler D, Bojanowski MW, Raymond J. Responses to ARUBA: a systematic review and critical analysis for the design of future arteriovenous malformation trials. *J Neurosurg* 2017;**126**:486–94.
17. Al-Shahi Salman R, White PM, Counsell CE, et al. Outcome After Conservative Management or Intervention for Unruptured Brain Arteriovenous Malformations. *Jama-J Am Med Assoc* 2014;**311**:1661–9.
18. Hartmann A, Mast H, Mohr JP, et al. Morbidity of intracranial hemorrhage in patients with cerebral arteriovenous malformation. *Stroke* 1998;**29**:931–4.
19. Laakso A, Dashti R, Juvela S, Niemela M, Hernesniemi J. Natural history of arteriovenous malformations: presentation, risk of hemorrhage and mortality. *Acta Neurochir Suppl* 2010;**107**:65–9.
20. van Beijnum J, van der Worp HB, Buis DR, et al. Treatment of Brain Arteriovenous Malformations A Systematic Review and Meta-analysis. *Jama-J Am Med Assoc* 2011;**306**:2011–9.
21. Mast H, Young WL, Koennecke HC, et al. Risk of spontaneous haemorrhage after diagnosis of cerebral arteriovenous malformation. *Lancet* 1997;**350**:1065–8.

22. Stapf C, Mast H, Sciacca RR, et al. Predictors of hemorrhage in patients with untreated brain arteriovenous malformation. *Neurology* 2006;**66**:1350–5.
23. Parides MK, Overbey JR, Stapf C, et al. Projecting longer term results in the ARUBA trial. *Cerebrovasc Dis* 2014;**37**:190–1.
24. Stefani MA, Sgarabotto Ribeiro D, Mohr JP. Grades of brain arteriovenous malformations and risk of hemorrhage and death. *Ann Clin Transl Neurol* 2019;**6**:508–14.
25. Kim H, Al-Shahi Salman R, McCulloch CE, Stapf C, Young WL, MARS Coinvestigators. Untreated brain arteriovenous malformation: patient-level meta-analysis of hemorrhage predictors. *Neurology* 2014;**83**:590–7.
26. Derdeyn CP, Zipfel GJ, Albuquerque FC, et al. Management of Brain Arteriovenous Malformations A Scientific Statement for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke* 2017;**48**:E200–E24.
27. Kim H, Al-Shahi Salman R, Flemming KD, et al. Long-Term Outcomes in Unruptured Brain Arteriovenous Malformation Patients: The Multicenter Arteriovenous Malformation Research Study (MARS). *Stroke* 2019;**50**.
28. Magro E, Gentric JC, Batista AL, et al. The Treatment of Brain AVMs Study (TOBAS): an all-inclusive framework to integrate clinical care and research. *J Neurosurg* 2018;**128**:1823–9.

Editorial points to be addressed:

1. Study title: doesn't comply with your Lancet style; How about "Long-term outcomes of medical management with standard interventional therapy versus medical management alone in patients with unruptured brain arteriovenous malformations (ARUBA): a multicentre, open-label, parallel, randomised controlled trial"?

Thank you for this suggestion – We've changed the title to "Medical management with interventional therapy versus medical management alone for unruptured brain arteriovenous malformations (ARUBA): final follow-up of a multicentre, open, parallel group, randomised controlled trial" which reflects the contents and avoids "long-term" as in the manuscript we say that the long-term effects remain uncertain.

2. Please check spelling, punctuation, spacing etc. and check for cohesion and clarity of each sentence. Please avoid repetitions between the individual section, this is currently the case in the method section. Please keep in mind that not all readers might be familiar with the original study and/or topic (eg, readers might not know why there was uncertainty whether medical management with or without interventional therapy is superior or not).

Thank you for this suggestion. We have done this including the use of the spelling 'center' in the co-author list for those not in the UK, centre for the others, and 'centre' throughout the manuscript.

3. Please follow CONSORT guidelines including these for the abstract.

We've modified the abstract (see 5 below) and completed the CONSORT checklist.

4. Please complete and submit CONSORT checklist.

We have submitted this alongside the manuscript.

5. Abstract (please format according to CONSORT, we have some leeway with the wordcount to comply with CONSORT):

- a. Background:

- i. not clarify why the "clinical benefit of preventive eradication of unruptured brain arteriovenous malformations (AVMs) remains controversial" given the positive results of the ARUBA trial. Here might be the place to explain ARUBA.

We have re-stated ARUBA's statistically significant finding, removed mention of it being controversial, and indicated why it is not clinically significant for some because the long-term risk/benefit balance is unknown.

- ii. Please add how long the FU of the original trial was as well as the FU of the current report to help with the flow.

Thank you for these suggestions – We have included the mean follow-up of the original report (pg 3, line 61) as suggested and added the mean follow-up of the current report to the results section (pg 3, lines 82-83) and think the flow has improved.

b. Methods:

i. Indicate the setting (community, hospital) where participants were recruited (which countries, how many centres or hospitals), and the key participant eligibility criteria.

We've added the number and international nature of centres who randomised patients (pg 3, lines 71-72). The key inclusion criteria are given in the first sentence: "adult patients diagnosed with an unruptured brain arteriovenous malformation, who had never undergone interventional therapy" (pg 3, lines 66-68)

ii. Explain the groups participants were randomly assigned to, and provide information about the methods of randomisation, masking, and stratification (eg, block size). How were participants allocated to groups and by whom? Were participants, investigators, and those assessing outcomes masked to group assignment?

We added more details to the existing sentence that detailed the random permuted block design to include how patients were randomised and who generated the sequence (pg 3 lines 72-74). We also added that primary outcome events were assessed by a neurologist at each centre not involved in participants' treatment (pg 3, lines 75-76).

iii. Give details of interventions (type, method of delivery, duration). For drugs please provide the generic name (rINN), doses, route, and schedule of administration.

We added the definition of interventional therapy to the second sentence (pg 3, lines 70-71) in parentheses, "The trial compared medical management alone with medical management and interventional therapy (neurosurgery, embolisation, or stereotactic radiotherapy, alone or in any combination, sequence or number)."

iv. What was the main outcome of this report and when was it assessed? We do not as standard include additional outcomes in the Summary.

We specified, "The primary outcome was time to death or symptomatic stroke confirmed by imaging, assessed by a neurologist at each centre not involved in the management of participants' care, and monitored independently using an adaptive approach with interim analyses" (pg 3, lines 74-77).

v. State who was included in primary and safety analyses (eg, intention to treat, per protocol, all participants who received one dose of study drug).

Methods includes sentence that states: "All analyses were by intention-to-treat" (pg 3, line 78).

c. Findings:

- i. Provide exact dates (day, month, year) between which participants were recruited and the number of participants assigned and analysed in each group, accounting for dropouts.

This is included in the methods – we added the day to the previously given month and year.

- ii. For the primary outcome give a result for each group (provide actual numbers of participants or events and their percentages), and estimated effect size (eg, odds ratio) and its precision (eg, 95% CI, p value). Report SDs for mean values and IQRs for medians, and give exact p values unless $p < 0.0001$. Use SI units. For risk changes or effect sizes, give absolute values rather than relative changes.

Done.

- iii. Please report ONLY the primary endpoint data. Secondary outcomes cannot be selectively reported in the abstract, and space restrictions typically prevent all secondary outcomes from being included in the abstract.

Done.

- iv. Summarise adverse events (actual numbers and percentages in both groups; include treatment-related deaths).

A summary of adverse events is included in the abstract and in Table 3 – rather than providing percentages we give the number of events per patient year to account for differential follow-up times. We have also added the overall number of deaths and the number attributed to treatment to the abstract (pg 3-4, lines 85-87).

- v. Results stated should agree with what is in the main paper, and all data here should also appear in.

Done.

- d. Interpretation: For clarity, please rephrase to something like this “At 5 years, medical management alone was superior to medical management with interventional therapy for the prevention of death or symptomatic stroke in patients with unruptured brain arteriovenous malformations”. This should be followed by the clinical implications/future directions, and or the key limitations and strengths of the study.

We reworded the interpretation section of the abstract accordingly, “After extended follow-up, medical management alone remained superior to interventional therapy for the prevention of death or symptomatic stroke in patients randomised to ARUBA with unruptured brain arteriovenous malformations. These data should affect standard specialist practice and the information presented to patients. The longer-term risks and difference between the two therapeutic approaches are uncertain.

6. Where dates are given, please give the exact dates (if known)—ie, day, month, year throughout the manuscript.

Done.

7. It is Lancet style to give actual numbers (numerator and denominator) together with all percentages, throughout the text and in tables etc. Please check.

We revised table 1 to include the numerators and denominators for each entry. We also reviewed the results text and added denominators where they were missing.

8. Please check and confirm that you have provided p values to two significant figures, unless $p < 0.0001$ (note number of decimal places).

We have just included p-values in Figure 3 to aid in the interpretation of the exploratory analyses of heterogeneity of treatment effects between subgroups. In the update figure that includes these p-values, they are provided to two significant figures.

9. As mentioned previously, the Methods section should be structured in this order please: Study design, Participants, Randomisation and masking, Procedures, Outcomes, Statistical analysis, Role of the funding source (see below).

We moved the sections to match the order defined above.

10. Introduction: Remove the last sentence and instead end with the aim of your study.

Done.

11. Method (needs to be completely revised to conform with CONSORT):

Done.

12. Study design:

a. Start with the study descriptor (randomised, parallel, cluster, non-inferiority, open-label, doubleblind, etc).

Methods state, "ARUBA was an open, randomised (1:1), parallel group trial..." (pg 6, line 138).

b. Move the aim of ARUBA from here to the introduction

We have moved the aim up to the first paragraph of the introduction (pg 5, lines 107-113).

c. Indicate where the study was done (community, hospital), in which countries, and in how many centres or hospitals. If too many, please add them to the appendix.

We added a sentence about the number of countries and centres that randomised patients (pg 6, line 142). We also added a table to the supplement that includes the country, centre name, and number of patients randomised at each centre (Supplemental Table 1).

- d. State the centre where ethics approval was obtained.

We moved the sentence about IRB and ethics approval that was previously under participants to the study design section (page 6, lines 142-144).

- e. Provide a link to the study protocol if available online. Please note the weblink must be permanent. Alternatively, the protocol can be included in your appendix if you wish – please indicate this in your responses.

We will provide the protocol for the appendix.

13. Participants:

- a. Describe the planned population, with inclusion and exclusion criteria and how participants were recruited. If too many criteria, please add the key one here and refer to appendix. These key criteria need to be also added to the abstract

Done.

- b. Add which centre was the Clinical Coordinating Centre

Added in parentheses to the following sentence, “No control in case selection was exerted by the Clinical Coordinating Center (ARUBA-WEST: Columbia University Medical Center; ARUBA-EAST: Hôpital Lariboisière).” (pg 6, lines 151-152).

14. Randomisation and masking.

- a. Please start with a description of the actual method of randomisation (ie, the method used to generate the sequence with which participants are allocated to comparison groups (eg, computer, random-number tables, coin-toss), including details of the methods used to restrict the randomisation—eg, block, stratification), and any stratification or minimisation factors. Words and phrases such as "randomised", and "randomly assigned" without qualification are not acceptable. Please clarify what you mean by “Statistical Coordinating Center by clinical center”. Please add which centre this was.

Done (pg 6-7, lines 157-166)

- a. Followed by a description of the method used to conceal assignment of a participant without knowledge of the next assignment in the sequence.

We believe stating they were assigned by a web-based system clarifies this – but we have added, “assignment was controlled via a central web-based data collection system, which did not reveal treatment allocation until all baseline data had been submitted.” (pg 6, lines 161-162)

- b. Please described who generated the allocation sequence, who enrolled participants, and who assigned them to the trial groups, and whether they had any involvement in the rest of the trial.

Done (pg 6, line 158-159).

- c. Describe how masking (blinding) was achieved (eg, tablets with identical appearance, syringe taped up to conceal colour of liquid inside). Please add a description of whether participants, those giving the interventions, those assessing outcomes, and those analysing the data were masked to group assignment; and how was the success of masking evaluated

Done – investigators and statisticians were not blinded to individuals’ treatment assignment. Details are given in the Randomisation and masking section.

- d. Delete “The first public presentation of the results was during the 22nd European Stroke Conference on 31 May 2013.”

Done.

15. Procedures.

1. Give details of interventions (type, method of delivery, duration). For drugs please provide the recommended international non-proprietary name, dose, route, and schedule of administration. For all commercial tests or devices, state the name of the manufacturer and place of manufacture.

Done (pg 7, lines 175-179).

2. State the follow-up intervals and assessments done at each visit.

Done – we’ve combined the previous “intervention” and “follow-up” sections into a single procedure section as recommended.

16. Outcomes. If applicable, please ensure the following items are included:

- a. State the primary outcome (for multicentre trials, whether this was centrally assessed).

Done (pg 8, lines 196-197).

- b. List secondary outcomes (a complete list).

We stipulated that modified Rankin Scale score was the lead secondary outcome, and added adverse events as the other secondary outcomes.

- c. Describe assessment of safety and adverse events.

Done (pg 8, lines 207-208).

- d. Please ensure any post-hoc or exploratory endpoints are clearly described as such. (this is not the case yet)

Done – we clarified that the subgroup analyses of the primary endpoint were exploratory in the statistical analysis section in the methods (page 9, line 257)

- e. If your paper is a primary trial report, all prespecified primary and secondary outcomes specified in the protocol should be listed in the Methods and reported in the Results. If any outcomes prespecified in the protocol are not reported in the present paper, this should be stated in the Outcomes section with a full justification.

N/A

17. Statistical analysis.

- a. Correct name: “Statistical Coordinating Center International Center for Health Outcomes and Innovation Research”?

Done – we removed the full name of the data coordinating center here as it is already referenced in the randomisation and masking section.

- b. Indicate how the target sample size was calculated and what power the study had to detect a significant difference between treatment groups.

Done – we moved the section on power, previously in the study design section, to the top of the statistical analysis section (pgs 8-9, lines 214- 231).

- c. Definitions of population assessed for primary and secondary outcomes, and for safety (eg, ITT, per protocol, etc).

We added that all analyses were by intention-to-treat (pg 10, line 261).

- d. Give details of main comparative analyses, followed by assessment of safety and adverse events, and then description of any post-hoc or exploratory endpoints (this is not the case yet)

Done – comparison of the primary endpoint is described first, followed by risk of death or clinical impairment at 5 years, adverse events, and exploratory subgroup analyses of the primary endpoint.

- e. State whether a data monitoring committee oversaw the study.

Added to the first paragraph of the section (pg 8, lines 218-219).

18. RESULTS (the first two paragraphs need some reordering to help with the flow)

- a. Paragraphs in this section should follow THE order: a description of number of participants recruited and included in analysis; baseline characteristics; findings for the primary

outcome, secondary outcomes, adverse events, and finally any post-hoc or sensitivity analyses. No subheadings should be used in the Results or the Discussion sections.

Results follow this order and no subheadings are used. We have also reduced the word count of this section by removing text redundant to data available in the tables.

b. The first paragraph should state the exact dates (eg, Jan 1, 2013, to Dec 31, 2014) between which participants were recruited, and include with a trial profile the number of participants assessed for eligibility, the number ineligible, the number eligible, the number randomised to each group, the number of exclusions or dropouts at each stage, and the number assessed for the primary endpoint. Please provide also percentages along actual numbers

These dates are included. We've also added percentages for screening numbers (i.e. % ineligible, pg 10, lines 277-278). We also moved the number randomised to each group from the second paragraph to the first (pg 10, line 282) and added a statement that all randomised were included in the analysis of the primary endpoint (pg 10, line 286).

c. Please add results of safety analysis (eg, data regarding the number of treatment-related deaths). Adverse events should be reported in a table, stratified by grade if appropriate (eg, 1-2, 3, 4 and 5). For graded, adverse events, those of grade 1 or 2 occurring in $\geq 10\%$ of patients should be reported; all grade 3, 4, and 5 events should be reported.

We have added the number of treatment-related deaths to the text. In the ARUBA trial, adverse events were not graded and therefore we are unable to provide a breakdown of events by grade. The protocol provides definitions of pre-specified adverse events which are reported in table 3.

19. Discussion (needs to be thoroughly revised for flow and cohesion, it jumps between arguments, some statements are vague, and):

a. Not sure why this has been added here – can we move to the appendix? “To address criticisms directed at the classical ‘as randomized’ result, the prior publication¹ and post-publication presentations had included ‘as treated’ analysis and one comparing outcomes based on analysis of the hazard ratios for outcome events for those ‘as treated’ and among them those whose outcome had a value of mRS ≥ 2 . For data including the follow-up, for those ‘as treated’ the disparity for outcome events had HR 0.22 (95% CI 0.12,0.41), for those mRS ≥ 2 HR 0.13 (95% CI 0.05,0.30). “

Deleted

b. The third paragraph seems random here and would be better suited (at least parts of it) in the Research In Context panel

c. We try to avoid direct quotes – possible to rephrase?

d. Please make very explicit what the study limitations are

e. Discuss limitations and strengths of your study, noting sources of bias or imprecision.

f. Discuss any controversies raised by this study.

g. Consider possible underlying mechanisms for your findings.

h. Suggest future research directions.

- i. End with a general interpretation of data in light of all evidence available, noting the clinical significance and effects on patient care and policy, expanding on the summary provided in your Research in context panel.

We have revised the discussion to meet the points above.

20. Figure 1: Please use the CONSORT trial profile as a template

We have revised Figure 1 to match the CONSORT template

21. Figure 2: provide abbreviation legend; add figure title

Done

22. Figure 3: provide abbreviation legend

Done

23. Figure 4: in text and appendix: please delete from the main text

Done

24. Tables: should be supplied in a separate Word file (not Excel or fdf/pdf). Each row of data should be in a separate line. Please ensure that rows and columns are not tabbed; data should be entered in cell form.

Done.

25. Table 1: please remove “ total cohort” column as this is not needed; add row for “male”; make clear for which group these baseline data were not available * Baseline score unavailable for 2 patients enrolled without angiography”; provide abbreviation legend

Done.

26. Table 2: provide abbreviation legend

Done.

27. APPENDIX: As mentioned previously, please submit as a separate pdf file with page numbers and refer to the appendix in the text as “appendix, p XX”.

Done

28. Author signature form: We are missing the signatures of all authors

We will include these with our resubmission.

29. ICMJE forms: we are still missing the ICMJE forms for all authors

We will include these with our resubmission.

30. Any update on Dr Stapf?

Dr Stapf has not responded to any emails inviting him to be a co-author, or acknowledged, in this manuscript, so we have removed him from our acknowledgement section per TLN requirement that we have his consent to be acknowledged in writing.

31. Author contribution: please streamline eg, XX, XZ, and RC wrote xxx.

Done.

32. Declarations on interest: please streamline, eg, XX, XZ, and RC had nothing to declare. NOTE: The statement MUST match the information on the supplied ICMJE forms. If not, please revise.

Done

33. Please add affiliation for NINDS Officers, DSMB Members, Adjudication Committee, an the Safety Officer

Done

34. Research In Context panel (needs to be rewritten to comply with the required style):

a. Evidence before this study: This section should include a description of all the evidence that the authors considered before undertaking this study. Authors should state: the sources (databases, journal or book reference lists, etc) searched; the criteria used to include or exclude studies (including the exact start and end dates of the search), which should not be limited to English language publications; the search terms used; the quality (risk of bias) of that evidence; and the pooled estimate derived from meta-analysis of the evidence, if appropriate. A summary of what the existing evidence shows should also be included.

b. Added value of this study: Please describe here how your findings add value to the existing evidence (including an updated meta-analysis, if appropriate). IMPORTANT: Please do NOT reiterate the results or describe your study approach (this is already covered by the abstract), but rather explain how the findings extend knowledge in the field and/or address unanswered questions or controversies.

c. Implications of all the available evidence: Authors should state the implications for practice or policy and future research of their study combined with existing evidence.

We have revised the research in context section accordingly.

Reviewers' comments:

Reviewer #1: The authors have improved the manuscript.

Our thanks for this comment.

"As noted in the revision, we added text (Lines 238-244) to explain the paucity of information." Please clarify correct line numbers containing the added text. It is not clear how the authors addressed the reviewer's comment.

We apologize for the confusion. The correct line number in the updated clean version of the manuscript is pg 10, line 279 which reads "outcomes were not collected for eligible patients who were not randomized."

In lines 256-258, specific mention of which subgroups of venous drainage and AVM maximum nidus size are being described would assist in informing the reader.

We thank the reviewer for pointing this out. We've added the subgroups in parentheses after each (pg 12, lines 319-320).

In lines 281-290, it would be helpful to the reader to include percentages of 116 patient corresponding to the number of patients treated with neurosurgery, radiosurgery or embolization.

We agree that it will be helpful for readers to know the distribution of interventions received. We have added the percentage of patients who received each modality (either alone of as part of a combination) to the last paragraph of the results section.

Reviewer #2: Overall Comments

I get my head down in the minutiae of a review making point by point comments and I sometimes forget the big picture. This study was a major achievement undertaking a novel adaptive design.

The authors have addressed my comments and my comments below are advisory only

Advisory comments for consideration

[Original Comment]. Things have moved on from the original analysis. There will soon be a CONSORT Guideline for adaptive designs.

a. No mention is made of bias adjustment to allow for the interim analysis within the study either in this paper or the original paper.

[Additional Comment]. Due to the stopping rules (and no futility assessment) the bias would be minimal. I personally would add a sentence, consistent with you feedback, that due to the descriptive nature of the study there is no bias adjustment due to the adaptive design

Thank you for the feedback – we have added the sentence to the end of the statistical analysis section in the methods (pg 10, lines 262-263).

[Original Comment]. I not mandate this but suggest this only. There is a lot of literature on the bias in studies which stop early (different to the bias mentioned above). This is a reporting bias not an actual bias. Your first interim analysis is not biased (the second and planned for third are) but there is a reporting bias as studies such as yours do not report the first interim analysis. Thus only those which do stop early at the first interim get reported. It is difficult for you 6 years on and it is not a CONSORT

requirement now (but it will be) but if in supplemental you give the analysis of the first interim analysis.

[Additional Comment]. Thank you for feeding back on the results. This is only a nice to have and it would be good to include in the supplemental if you are able. As I said papers erroneous report that trials which stop early are biased - see https://urldefense.proofpoint.com/v2/url?u=https-3A_doi.org_10.1177_0962280211432211&d=DwIGaQ&c=G2MiLlal7SXE3PeSnG8W6_JBU6FcdVjSsBSbw6gcR0U&r=ocMgXGLjsdweFgVd_pgarCirw2NyJtwGioPgIKR1sDI&m=fMcSM2HIOKCDsfJBcqLMagPD3fQukDGoWVg8aKBZqts&s=DTVZIDCFSEKdh-SsZmNCTW6dQej0UVW2J2-APT0aWVk&e= - when they are not. It is just that adaptive designs due not report all the steps

We have included the graph in the supplement (supplemental figure 2), and reference it in the statistical analysis section of the manuscript (pg 10, lines 263-264).

[Original Comment]. For the statistical analysis

b. Can you please clarify how the clustering effect of surgeon was accounted for in the IT arm. Although the study is individually randomised they are then clustered with surgeons the skill of which will impact on their outcomes. Either surgeon or centre for the IT arm should be entered as a cluster effect and the errors adjusted accordingly

c. Can the ICC for the clustering please be quoted

[Additional Comment]. Thank you for the clarification. Given you feedback I would have used centre myself as it was team medicine. Personally I would in text only quote the result you provided as it had no effect. A minimal ICC would confirm this.

Thank you for your additional comments - we have added the site adjusted results to the results section (pg 11, lines 301-302) and describe the analytical approach in the methods (pg 9, lines 241-242).

[Original Comment]. The paper itself should be standalone. It makes for a frustrating read to see "have been described previously" especially when this is in a paper behind a paywall. There should be these details in the paper. Can be complemented with supplemental material [Additional Comment]. Thank you for your feedback. I did not realise paywall was a UK term!

Evidence of our ignorance in the Colonies

[Original Comment]. Can all web references have date last accessed [Additional Comment]. This is just the data you last accessed. Personally the last thing I do prior to a submission is click through all web links and put that date down

We have added the last accessed date to the web address provided under the data sharing statement.

Supplemental data - Medical management with interventional therapy versus medical management alone for unruptured brain arteriovenous malformations (ARUBA): final follow-up of a multicentre, open, parallel group, randomised controlled trial

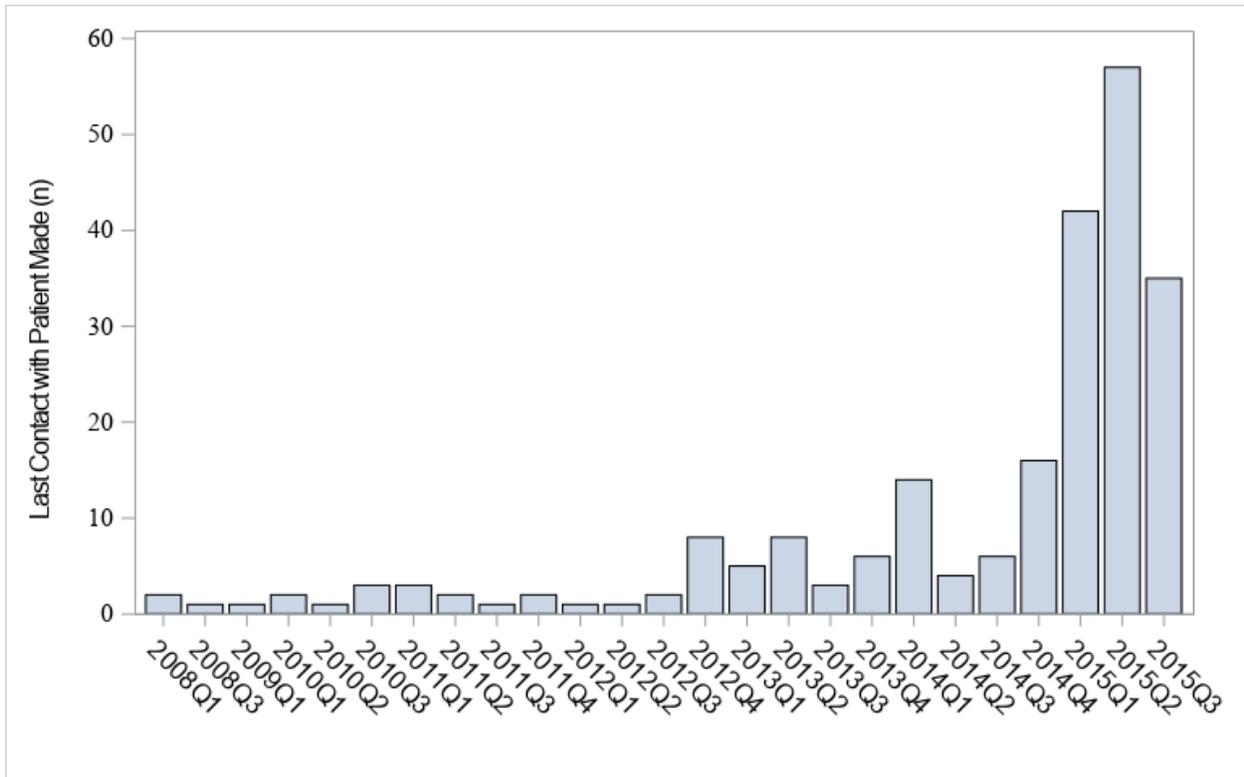
Contents

Supplemental Table 1. Randomizations by Region, Country and Clinical Centre.....	2
Supplemental Figure 1. Histogram of Patients Last Dates of Contact by Quarter.....	3
Supplemental Figure 2. Group Sequential Boundaries and Observed Z-Statistics at each Analysis.....	4
Supplemental Figure 3. Cumulative Enrolment Overtime	5
Supplemental Figure 4. Modified Rankin scale scores at 5 years post-randomisation*.....	6
Supplemental Table 2. Treatment strategy and primary outcome in the 106 patients that initiated interventional therapy with an unruptured brain AVM *.....	7

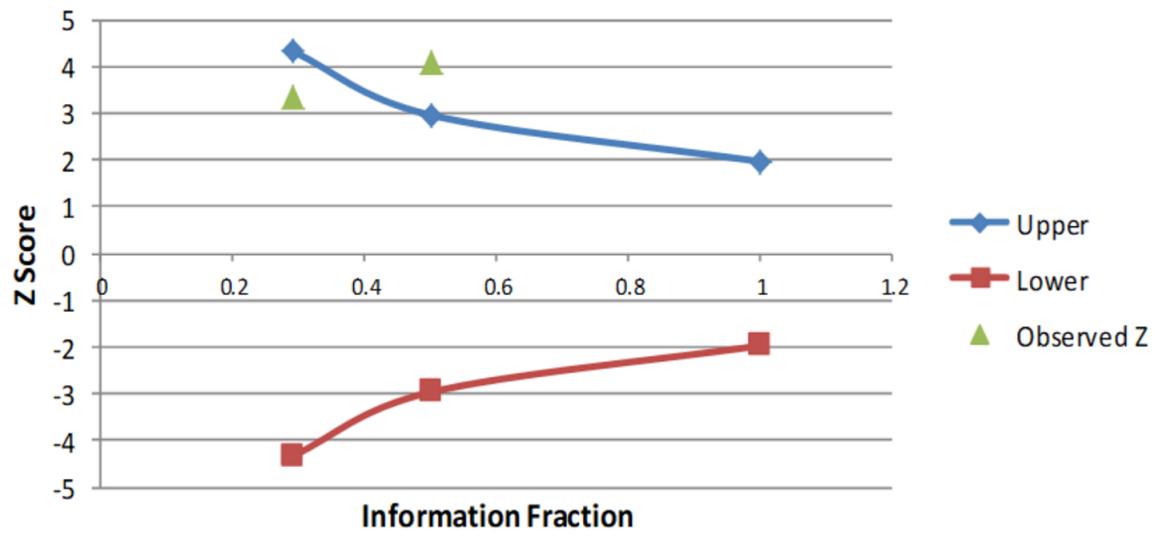
Supplemental Table 1. Randomizations by Region, Country and Clinical Centre

Region	Country	Clinical Centre	No. Enrolled
ARUBA-Europe	France	Hôpital Lariboisière	35
		Hôpital Roger Salengro, CHRU	24
		CHU Besançon	11
		Groupe Hospitalier Pellegrin	7
		Hôpital Saint Anne	1
		CHU Brest	1
	Germany	Charite Campus Benjamin Franklin	12
		Neurologische Universitätsklinik, Dresden	12
		Klinikum der J.W.v.Goethe Universität	10
		BG-Kliniken Bergmannstrost	4
		Klinikum der Universität München-Grosshadern	2
		Universitätsklinikum Freiburg	2
		Universitätsklinikum Essen	1
	Netherlands	Utrecht Medical Center	9
		UMC Groningen	3
	Switzerland	Inselspital, University of Bern	1
	United Kingdom	Western General Hospital	12
		Royal Hallamshire Hospital	8
		Walton Centre for Neurology & Neurosurgery	2
		Newcastle General Hospital	2
Salford Royal NHS		1	
ARUBA-West	Brazil	Hospital de Clinicas de Porto Alegre	14
	Canada	CHUM Notre Dame Hospital	6
		Hamilton Health Sciences	1
		University of Alberta	1
	South Korea	Seoul National University Hospital	2
	United States	Long Island Jewish Medical Center	7
		Columbia University Medical Center	6
		Winthrop University Hospital	4
		University of California at San Francisco	4
		Kaiser Permanente (LA)	4
		University of Alabama at Birmingham	3
		University of Miami	3
		Rush University	3
		Hartford Hospital	3
		Loyola University Stritch School of Medicine	2
John P. Murtha Neuroscience and Pain Institute		1	
University of South Alabama		1	
Redwood City Kaiser	1		

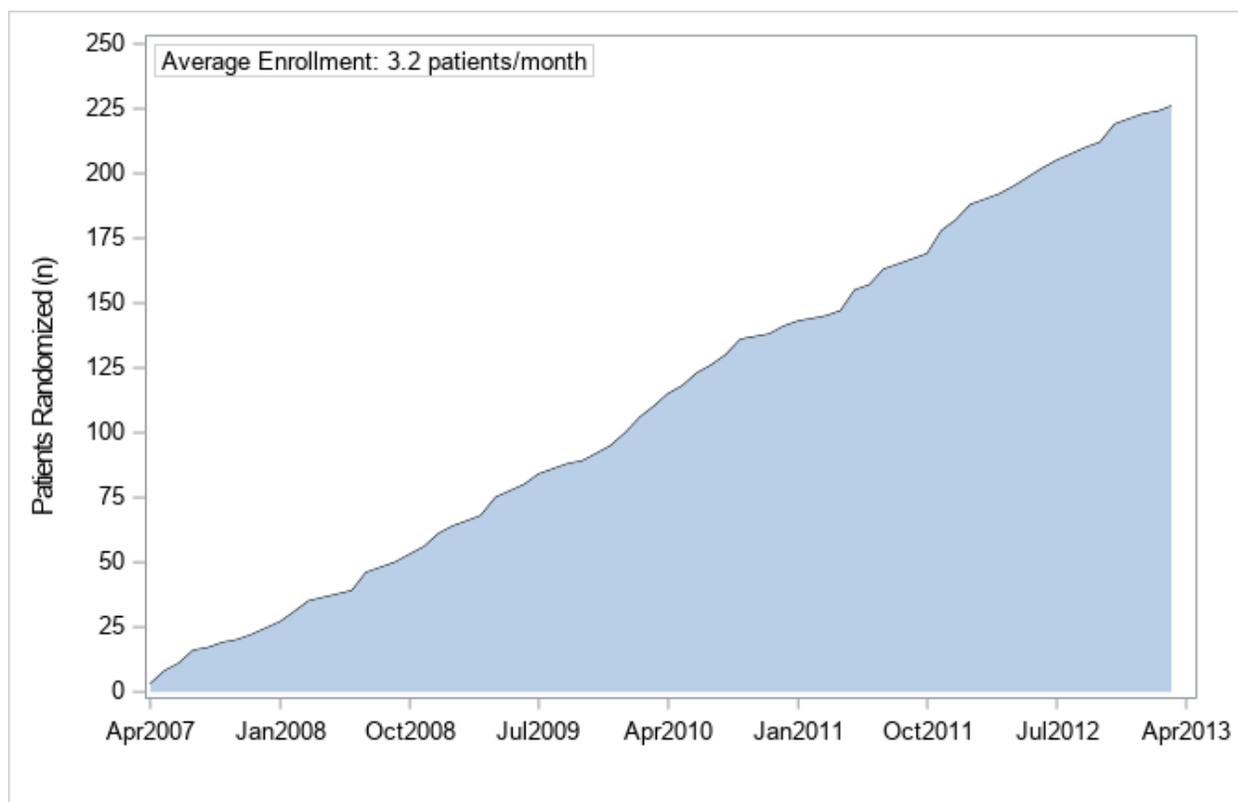
Supplemental Figure 1. Histogram of Patients Last Dates of Contact by Quarter



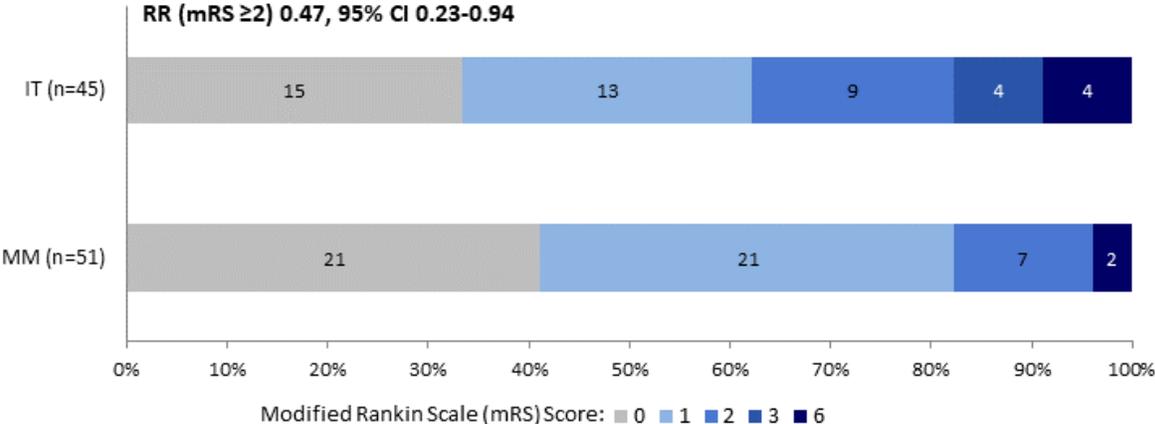
Supplemental Figure 2. Group Sequential Boundaries and Observed Z-Statistics at each Analysis



Supplemental Figure 3. Cumulative Enrolment Overtime



Supplemental Figure 4. Modified Rankin scale scores at 5 years post-randomisation*



*Modified Rankin scale scores of 4 and 5 were not observed at 5 years post-randomisation and therefore are not represented in this plot

Supplemental Table 2. Treatment strategy and primary outcome in the 106 patients that initiated interventional therapy with an unruptured brain AVM *

Type of interventional treatment	Symptomatic stroke or death	With documented AVM obliteration**
	n (row %)	n (row %)
Single modality of treatment		
Endovascular (n=28)	14 (50.0)	14 (50.0)
Surgery (n=7)	2 (28.6)	7 (100.0)
Stereotactic radiosurgery (n=33)	8 (24.2)	6 (18.2)
Multiple modalities of treatment		
Endovascular and surgery (n=14)	6 (42.9)	14 (100.0)
Endovascular and stereotactic radiosurgery (n=23)	12 (52.2)	6 (26.1)
Endovascular, surgery and stereotactic radiosurgery (n=1)	1 (100.0)	0 (0)

* Eight patients randomised to MM received IT prior to any stroke events. Among the 116 patients randomised to IT, n=15 never received therapy and n=3 suffered a stroke prior to the initiation of IT

** Documented AVM obliteration required cerebral angiography by study protocol. For n=16 (15%) patients, the AVM obliteration status was unknown due to missing imaging information, n=43 (41%) had a documented AVM remnant on last follow-up imaging.

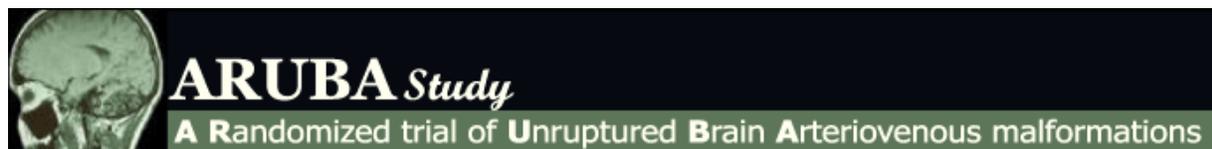


CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	3
	2b	Specific objectives or hypotheses	3
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	4
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	8-9
Participants	4a	Eligibility criteria for participants	6
	4b	Settings and locations where the data were collected	6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8
	6b	Any changes to trial outcomes after the trial commenced, with reasons	N/A
Sample size	7a	How sample size was determined	8-9
	7b	When applicable, explanation of any interim analyses and stopping guidelines	9
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	6-7

		assessing outcomes) and how	
Statistical methods	11b	If relevant, description of the similarity of interventions	N/A
	12a	Statistical methods used to compare groups for primary and secondary outcomes	9-10
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	10
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	10
	13b	For each group, losses and exclusions after randomisation, together with reasons	10, Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up	10-11
	14b	Why the trial ended or was stopped	1, 7
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	11-12
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	11, 12
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	-
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	12
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	11
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	13
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	13-14
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	13-15
Other information			
Registration	23	Registration number and name of trial registry	3, 10
Protocol	24	Where the full trial protocol can be accessed, if available	6
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	4,10

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.



A Randomized Multicenter Clinical Trial of Unruptured Brain AVMs (ARUBA)

Clinical Protocol

Sponsors National Institute of Neurological Disorders and Stroke
National Institutes of Health

Number of Patients 400

Trial Leadership J.P. Mohr, M.D. (PI, CCC)
Alan Moskowitz, M.D. (PI, DCC)
Deborah Ascheim, M.D. (Co-PI, DCC)
Annetine Gelijns, PhD, (Co-PI, DCC)
Michael Parides, PhD. (Co-PI, DCC)
Christian Stapf, M.D. (Co-PI, CCC, Europe)
Eric Vicaut, M.D. (Co-PI, DCC, Europe)
Claudia Scala Moy, PhD. (Co-PI, NINDS)

Design Randomized Multicenter Clinical Trial of Unruptured
Brain AVMs

Version Date June 2010

Version 4.0

Data Coordinating Center: International Center for Health Outcomes and Innovation Research (InCHOIR)

Mount Sinai School of Medicine, Department of Health Policy

One Gustave L. Levy Place, Box 1077 New York, NY 10029

Phone: 1-212-659-9567; Fax: 1-212-659-9894

Ellen Moquete, RN, Project Leader, Email: ellen.moquete@mounsinai.org

Alejandrina Estevez, MBA Project Manager, alejandrina.estevez@mounsinai.org

Clinical Coordinating Center: Doris & Stanley Tananbaum Stroke Center, Columbia University Medical Center, Department of Neurology

710 West 168th Street, New York, NY 10032 USA

Phone: 1+ 212 305 8033, Fax: 1+ 212 305-5796

Steve Marshall (Coordinator), Email: smm2147@columbia.edu

Coordinating Center (Europe): CHU Hôpital Lariboisière, Service de Neurologie, 2, Rue Ambroise Paré, 75475 Paris cedex 10, France

Phone: +33 1 4995 6946, Fax : +33 1 4995 2596

Christian Stapf, MD (Clinical), Email: christian.stapf@lrb.ap-hop-paris.fr

Eric Vicaut, MD (Data), Email: eric.vicaut@lrb.aphp.fr

Unite de Recherche Clinique, Hôpital Fernand Widal

200 rue du Faubourg St Denis, 75475 Paris cedex 10

Phone : +33 1 4005 4973 ; Fax : +33 1 4005 4974

PROTOCOL REVISIONS

Revision Number	Date of Revision	Text location (page)	Description of Revision
2	03/30/06	Created Appendix II	Clinical Sites Participating in the Study: Moved pages 1-3 of Rev. 1 to Appendix II. Revised to reflect update on those centers agreeing to participate and those who have declined participation.
2	03/30/06	Appendix II	Study Team Roster: Moved pages 4-5 of Rev. 1 to Appendix II. Revised to reflect personnel changes
2	03/30/06	Appendix II	Trial Organization: Moved page 6 of Rev. 1 to Appendix II.
2	03/30/06		Table of Contents: Global change to all page numbers
2	03/30/06	6; Section 4.2	Exclusion criteria #6: Change units of measurement from nl to μ L
2	03/30/06	7; Section 4.3	# Clinical sites changed from 86 to 104
2	03/30/06	8; Section 4.3.3	Dr. Deborah Gohs replaced by Steven Marshall
2	03/30/06	22; Section 10.2	# centers changed from 86 to 104
2	6/28/06	Table of Contents	Added 4.3.4 Randomization section, page 8
2	6/28/06	8	Added 4.3.4 Randomization section to reflect protocol reference page 22
2	6/28/06	8	Procedure for obtaining randomization assignment
2	6/28/06	8	Changed order of data points and added Rankin scale to entry data collection
2	6/28/06	9	AR03 now Image Study
2	6/28/06	9	AR04 now Randomization
2	6/28/06	8	Images should not be older than one year
2	6/28/06	8	De-identified CD of images will be sent to the DCC
2	6/28/06	9	Interventional therapy to begin within 3 months of assignment
2	6/28/06	15	Clinical and Laboratory Evaluations now Data Collection Schedule
2	6/28/06	15	Comments now Event Driven
2	6/28/06	15	Enrollment now Baseline
2	6/28/06	15	Rankin Scale added to screening
2	6/28/06	15	Added Delayed treatment
2	6/28/06	16	Added Form AR07 to Screening

Revision Number	Date of Revision	Text location (page)	Description of Revision
2	6/28/06	16	AR03 is now Image Study
2	6/28/06	16	AR04 is now Randomization
2	7/12/06	1	“treatment” changed to “eradication therapy”
2	7/12/06	8	4.3.4 now Procedure for Image Interpretation and shipment
2	7/12/06	9	4.3.5 now Randomization
2	7/12/06	Table of Contents	Changed 4.3.4 to Procedure for Image interpretation and shipment
2	7/12/06	Table of Contents	Added 4.3.5 Randomization, page 9
2	7/12/06	8	Deleted “AVM Description Form”
2	7/12/06	8	Added Rankin Scale AR07
2	7/12/06	9	Added “Interventional therapy to begin within 3 months of randomization”
2	7/12/06	15	Clinical and Laboratory Evaluations now Data Collection Schedule
2	7/12/06	16	Deleted Randomization Verification
2	7/12/06	16	AR03, AR03A Image Study added
2	7/12/06	16	AR04 now Randomization
2	7/12/06	17	Added “Event Driven” to On-Study Evaluations
2	7/12/06	17	Patient Encounter schedule changed to correspond with Data Collection Schedule.
2	7/12/06	18	Added Form AR22 Adverse Events Adjudication
2	7/12/06	18	Added Form AR23 Mortality Adjudication
2	7/12/06	18	Added Form AR24 Delayed Treatment
2	7/12/06	19	AR01 now Demographics – deleted “and AVM Description”
2	7/12/06	19	Deleted Randomization Verification
2	7/12/06	19	AR03 and AR03A are now Image Study
2	7/12/06	19	Added definitions of forms AR03 and AR03A
2	7/12/06	19	AR04 now Randomization
2	7/12/06	21	Added AR22: Adverse Events Adjudication
2	7/12/06	21	Added AR23: Mortality Adjudication
2	7/12/06	21	Added AR23: Delayed Treatment
2	7/12/06	43	Added 12.4.3: Safety Monitor and definition
2	7/12/06	Table of Contents	Added 12.4.3: Safety Monitor
2	7/26/06	Table of Contents	Table of Contents: Global change to all page numbers
2	7/31/06	15	Global changes to Data Collection Schedule
2	7/31/06	32	Edited definition of Cross over
2	7/31/06	19	Added AR25 Final Study

Revision Number	Date of Revision	Text location (page)	Description of Revision
2	7/31/06	22	Added AR25 Final Study
2	7/31/06	17	AR20 changed to 60 Month Study
2	7/31/06	22	AR20 changed to 60 Month Study
2	7/31/06	15	Global changes to Data Collection Schedule
2	8/10/06	19	(4.3.3) Procedure for obtaining Intervention group assignment now Procedure for Enrollment
2	8/10/06	19	4.3.3 now Screening Log
2	8/10/06	19	4.3.4 now Procedure for Enrollment
2	8/10/06	19	4.3.5 now Procedure for Image Interpretation and Shipment
2	8/10/06	19	4.3.6 now Randomization
2	8/18/06		Global changes to page numbers
2.1	11/1/06	13	New wording for outcome assessment procedure
2.1	11/1/06	20	Added Procedure Related Nephropathy to Non-neurological Adverse Events
2.1	11/1/06	20	Added Catheter adherence to embolization material to Non-neurological Adverse Events
2.1	11/1/06	26	Added waiver request to AR02 to enroll a patient who has a value outside of the protocol defined range
2.1	11/14/06	22	Merging of Screening and Baseline on Data Collection Schedule
2.1	11/14/06	23-24	Merge Screening and Baseline Data Collection
2.1	11/14/06	23-24	AR05-AR12 now read “ <i>At baseline, prior to randomization</i> ”
2.1	11/14/06	15-16	Forms AR05-12 Deleted from post-randomization completion and added to Enrollment Procedures
3.0	6/3/08	14	Pregnancy risks updated
3.0	6/3/08	Appendix V	Updated Screening Log
3.0	6/3/08	15	Instructions for completing Screening Log updated
3.0	6/3/08	45	Updated language for monitoring
3.0	7/9/08	18	Updated language for stroke definition
3.0	7/9/08	30	Updated language for definition of intervention associated hemorrhage
3.0	7/9/08	1	Replaced Naresh Talathoti with Julie Krauchuk
3.0	10/16/08	1	InCHOIR moved to Mount Sinai School of Medicine, Department of Health Policy
3.0	10/16/08	16	Section 6.1 Added language regarding adjudication of stroke
4.0	6/25/2010	Entire Protocol	Number of Patients to be randomized changed from 800 to 400
4.0	6/25/2010	1	Updated email addresses for Ellen Moquete and Steve Marshall
4.0	6/25/2010	1	Added Alejandrina Estevez as Project Manager

4.0	6/25/2010	9	Changed 40 to 46% reduction in hazard
4.0	6/25/2010	12	Revised primary null hypothesis
4.0	6/25/2010	12	Revised non inferiority analysis
4.0	6/25/2010	15	Added South America, Asia and Australia
4.0	6/25/2010	18	Updated Completeness of Interventional Therapy
4.0	6/25/2010	23	Added Medical history on the Data Collection Schedule at month 60
4.0	6/25/2010	25	Added AR09 Medical History to on-study evaluation
	6/25/2010	28	Added antiplatelet agents and anticoagulants in Form AR10: Medications
4.0	6/25/2010	31-33	Updated Statistical Considerations
4.0	6/25/2010	35	Updated Interim Analysis
4.0	6/25/2010	49	Updated DSMB to DSMB
4.0	6/25/2010	50	Updated M&M to Event Adjudication Committee

PROTOCOL TABLE OF CONTENTS

	page
Protocol Revisions.....	2
PRÉCIS.....	8
1. SPECIFIC AIMS	9
1.1 Primary Aims.....	9
1.2 Secondary Aims	10
2. BACKGROUND.....	11
2.1 Rationale	11
2.2 Supporting Data	12
3. STUDY DESIGN	13
4. SELECTION AND ENROLLMENT OF SUBJECT.....	13
4.1 Inclusion Criteria.....	13
4.2 Exclusion Criteria	14
4.3 Study Recruitment and Enrollment Procedures.....	14
4.3.1 Minority Recruitment.....	15
4.3.2 Informed Consent Procedures	15
4.3.3 Screening Log.....	16
4.3.4 Procedure for Enrollment.....	16
4.3.5 Procedure for Image Interpretation and Shipment.....	16
4.3.6 Randomization.....	16
5. STUDY INTERVENTIONS	17
5.1 Medical Management	17
5.2 Interventional Therapy.....	17
5.2.1 Endovascular Treatment.....	17
5.2.2 Microsurgery.....	18
5.2.3 Radiotherapy.....	18
5.3 Completeness of Interventional Therapy.....	18
5.4 Handling of Study Interventions	18
5.5 Concomitant Interventions.....	18
5.5.1 Required Interventions.....	18
5.5.2 Prohibited Interventions	18
5.6 Adherence Assessment	18
6. ENDPOINTS.....	19
6.1 Primary Endpoint	19
6.2 Secondary Endpoints.....	19
6.2.1 Quality of Life and Patient Preferences.....	19
6.2.2 Adverse Events.....	19
6.2.3 Cost Endpoints.....	21
6.2.3.1 Direct Costs of Medical and Non-medical Care	21

7. CLINICAL AND LABORATORY EVALUATIONS	23
7.1 Schedule of Data Collection	23
7.2 Timing of Evaluations	24
7.3 Special Instructions and Definitions of Evaluations	27
8. MANAGEMENT OF ADVERSE EXPERIENCES	30
9. CRITERIA FOR INTERVENTION DISCONTINUATION	30
9.1 Brain Hemorrhage or Stroke related to BAVM in Medical Management Group	30
9.2 Intervention Associated Hemorrhage	31
10. STATISTICAL CONSIDERATIONS	31
10.1 General Design Issues	31
10.2 Sample Size and Accrual	31
10.3 Randomization Design and Procedure	33
10.4 Data Monitoring and Analyses	33
10.4.1 Methods of Analysis	33
10.4.2 Assessing the Proportional Hazards Assumption	34
10.4.3 Interim Analysis	35
10.4.4 Assessment of Balance of the Randomization	36
10.4.5 Analysis of Secondary Endpoints	36
10.4.6 Additional Analyses of the Primary Endpoint	38
10.4.7 Imputation Procedure for Missing Data	39
10.5 Crossovers	40
11. STUDY ORGANIZATION, DATA COLLECTION, SITE MONITORING, AND ADVERSE EXPERIENCE REPORTING	40
11.1 Study Organization	40
11.2 Training the Research Staff	41
11.3 Electronic Data Management	42
11.3.1 The Data Center	42
11.3.2 Security	42
11.3.3 Electronic Forms	43
11.3.4 Software Quality Assurance and Technical Support	44
11.3.5 Disaster Planning	44
11.3.6 HIPAA Compliance	45
11.3.7 Data Access Control for Blinded Investigators	45
11.3.8 Management of Digital Images	45
11.3.9 Management of Faxed Source Data	46
11.4 Data Monitoring and Quality Assurance	46
11.5 Adverse Experience Reporting	47
11.5.1 Adverse Event	47
11.5.2 Serious Adverse Event	48
11.5.3 Event Reporting	48
12. HUMAN SUBJECTS	48
12.1 Institutional Review Board (IRB) Review and Informed Consent	48

12.2	Potential Risks.....	48
12.3	Safety and Confidentiality.....	48
12.4	Study Modification/Discontinuation.....	49
12.4.1	Performance and Safety Monitoring Board.....	49
12.4.2	Morbidity and Mortality Committee.....	49
12.4.3	Safety Monitor.....	50
13.	PUBLICATION OF RESEARCH FINDINGS.....	50
14.	REFERENCES.....	50

APPENDICES

- I. Informed Consent Template
- II. Clinical Sites Participating in the Study
- III. Study Team Roster
- IV. Trial Organization
- V. Screening Log
- VI. Case Report Forms

PRÉCIS

Study Title

A Randomized Multicenter Clinical Trial of Unruptured Brain AVMs (ARUBA)

Objectives

Primary: To determine whether medical management improves long-term outcomes of patients with unruptured BAVMs compared to interventional therapy (with endovascular procedures, neurosurgery, or radiotherapy, alone or in combination). The trial has been designed to test whether medical management or interventional therapy will reduce the risk of death or stroke (due to hemorrhage or infarction) by at least 46% (an absolute magnitude of about 9.5% over 5 years). It will require 400 patients to detect the hypothesized 46% reduction in event rate, analyzed using the intention-to-treat principal. This sample size will support a test of non-inferiority if medical management is not superior to interventional therapy.

Secondary: To compare the impact of medical management to interventional therapy with respect to adverse events, quality of life and cost.

Design and Outcomes

The study design is a prospective, multi-center, parallel design, randomized, controlled trial. Treatment assignment will not be masked; however, clinical coordinating center personnel and outcome events committees will be blinded to treatment assignment. Interim study results will be kept confidential by the DCC. The primary outcome is the composite event of death from any cause or stroke (hemorrhage or infarction revealed by imaging). Functional outcome status will be measured by the Rankin Scale, a widely-used outcome measure for stroke. The secondary measures of outcome include adverse events, quality of life and cost.

Interventions and Duration

The interventional therapy arm of the trial involves prophylactic efforts with a plan for eradication of the observed BAVM utilizing endovascular procedures, microsurgery, or radiosurgery, alone or in combination with pharmacological therapy for existing risk factors and coexisting medical conditions. The medical management arm will involve pharmacological therapy as deemed appropriate for medical symptoms as determined by the treating investigator. Should patients in the medical management arm develop hemorrhage or infarction related to their BAVM, they would then be candidates for any single or combination of interventional therapy using endovascular procedures, microsurgery and radiosurgery. Patients will be followed for a minimum of 5 years and a maximum of 10 years (mean 7.5 years) from randomization.

Sample Size and Population

All patients with an unruptured BAVM diagnosed at a participating clinical center without prior interventional therapy to attempt eradication and with no contraindications to interventional therapy, will be candidates for this trial. A total of 400 patients will be enrolled in the ARUBA trial. Patients may be referred for enrollment by their clinical neurologist, neurosurgeon, or interventional radiologist.

1. SPECIFIC AIMS

Current interventional therapy for brain arteriovenous malformations (BAVMs) is varied and includes endovascular procedures, neurosurgery, and radiotherapy alone and in combination, largely dependent on the decisions of the local clinical team. All of these interventional therapies are administered on the assumption that they will decrease the risk of initial or subsequent hemorrhage and lead to better long-term outcomes. Despite these laudable goals, the literature contains almost no reference to the outcome for medical management before or after hemorrhage, or for intervention outcome for unruptured BAVMs. Published reports of interventional therapy outcome typically have blended the bled and non-bled cohorts together as if their risk for lesion-related morbidity and the response to intervention is expected to be the same.

Although no clinical trial data exist on the effect of interventional therapy even after BAVM hemorrhage, the most contentious issue at present is whether interventional therapy should be considered for those increasingly being discovered incidentally by brain imaging, with lesions that have not bled. Recent data from our institution on BAVM patients who presented without bleeding raises the possibility that interventional therapy may be detrimental compared with medical management. Among possible reasons may be that interventional therapy destabilizes the lesion toward hemorrhage. Furthermore, there is disappointing evidence that contradicts prior assumptions that hemorrhage associated with BAVM treatment lie in functionally-inert tissues, and, therefore, are less disabling. It appears that the disabilities associated with such events are equivalent to and possibly worse clinically than that seen with spontaneous BAVM hemorrhages, which still have a relatively low likelihood of occurring in the foreseeable future.

1.1 Primary Aims

The *primary hypothesis* of this randomized clinical trial is that medical management improves long-term outcomes of patients with unruptured BAVMs compared to interventional therapy (with endovascular procedures, neurosurgery, or radiotherapy, alone or in combination). The primary outcome is the composite event of death from any cause or stroke (hemorrhage or infarction confirmed by imaging). Functional outcome status will be measured by the Rankin Scale, a widely-used outcome measure for stroke. There are three specific aims associated with the primary hypothesis:

Specific Aim 1.1a To determine whether medical management is *superior* to interventional therapy for preventing the composite outcome of death from any cause or stroke (hemorrhage or infarction confirmed by imaging) in the treatment of unruptured BAVMs.

Specific Aim 1.1b If medical management is not superior to interventional therapy, to determine whether medical management is *not inferior* to interventional therapy for preventing the composite outcome of death from any cause or stroke (hemorrhage or infarction confirmed by imaging) in the treatment of unruptured BAVMs.

Specific Aim 1.2 To determine whether treatment of unruptured BAVMs by medical management decreases the risk of death or clinical impairment (Rankin Score ≥ 2) at 5 years post-randomization compared to interventional therapy.

1.2 Secondary Aims

A number of Secondary Aims are planned in support of the primary hypothesis to answer the following questions:

- 1) Is there a difference in quality of life between interventional therapy and medical management?
- 2) Is there a difference in mortality between interventional therapy and medical management?
- 3) Is there a difference in quality-adjusted survival between medical management and interventional therapy?
- 4) Is there a difference in the incidence of adverse events, such as cerebral hemorrhage and infarction, between interventional therapy and medical management?
- 5) What are the costs associated with each treatment (medical management and interventional therapy); and if medical treatment is not superior, but also not inferior to interventional therapy what are the cost-effectiveness implication of choosing one therapy over another?
- 6) Does any benefit of medical management or interventional therapy depend on BAVM size?
- 7) Does any benefit of medical management or interventional therapy depend on BAVM location?
- 8) Does any benefit of medical management or interventional therapy depend on venous drainage pattern?
- 9) Does any benefit of medical management or interventional therapy depend on age at randomization?
- 10) Does any benefit of medical management or interventional therapy depend upon the length of time the AVM was known?
- 11) Is there a difference in the risk of the composite event of death from any cause or stroke between prophylactic treatment modalities (i.e. endovascular procedures, neurosurgery, and radiotherapy)?
- 12) Among patients treated by interventional therapy, is there a relationship between the completeness of eradication of the BAVM and the composite event of death from any cause or stroke?
- 13) Among patients treated by interventional therapy, is there a relationship between the Spetzler-Martin grading scale and the composite event of death from any cause or stroke?

The *primary null hypothesis* is that there is no difference between medical management and interventional therapy in the time to stroke or death from any cause. The null hypothesis will be tested against the alternative hypothesis that there is a difference between treatments with a two-sided 0.05 level log-rank test. With a plan to enroll 400 patients, the test will have 80% power to detect a risk reduction of 46% (hazard ratio of 0.54) This hazard ratio corresponds to an absolute decrease in 5-year event rates of 9.5% for medical management, from an assumed 5-year event rate of 22% for interventional therapy.

If the null hypothesis is not rejected, a test of non-inferiority of medical management compared to interventional therapy will be performed. The null hypothesis for the test of non-inferiority is that the hazard ratio for the composite event of death from any cause or stroke for interventional therapy compared to medical management is less than 0.87 (an 13% reduction in risk for interventional therapy). Thus, the null hypothesis that medical management is inferior will be rejected, and non-inferiority claimed, if the reduced risk of interventional therapy compared to medical management is less than the non-inferiority margin of 13% (hazard ratio ≥ 0.87) based on a one-tailed 0.05 level test. An 13% reduction in risk corresponds to an absolute difference in 5-year event rates of 2.5%.

The *secondary hypothesis* to be tested is that early intervention decreases the risk of death or clinical impairment at 5 years post-randomization. Death in this young, and otherwise healthy, population is a rare event. The primary hypothesis has been constructed to be inclusive of all strokes that occur during the course of the trial (thereby averting judgment about severity), while the secondary hypothesis concentrates only on those events associated with impairment.

2. BACKGROUND

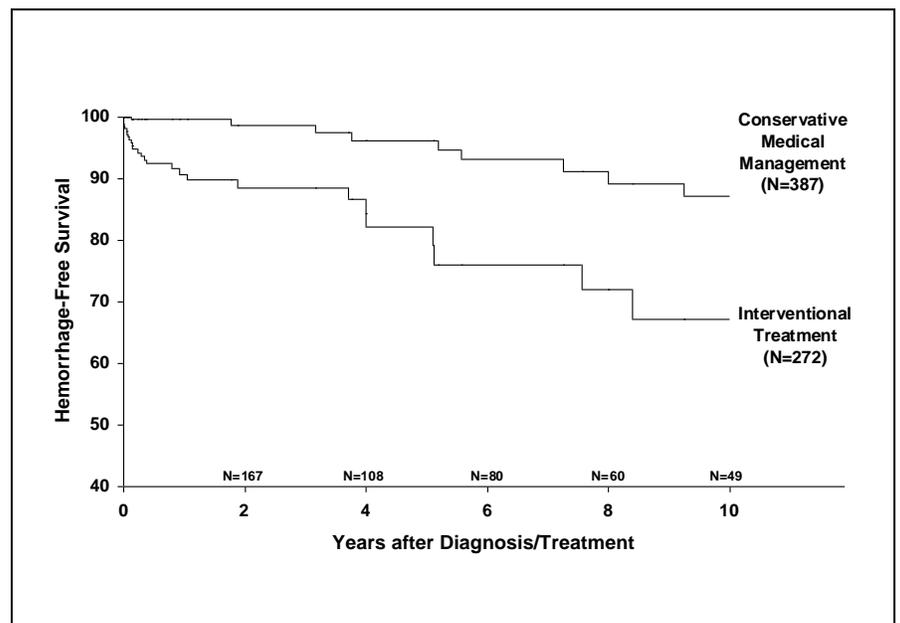
2.1 Rationale

With the emergence of new non-invasive imaging techniques, there has been a substantial increase in the incidental detection of non-ruptured BAVMs. These BAVMs are being treated in a variety of ways, including medical management, endovascular procedures, neurosurgery, or radiotherapy. The widespread diffusion of these various treatment approaches is partially driven by the existence of variations in the perception about the risks of rupture and how devastating such events would be. The increased treatment rate of non-ruptured BAVMs consumes a considerable amount of health resources. With an annual incidence in the US of nearly 3000 cases, and treatment costs in the range of \$50,000 to \$100,000 per patient, widespread utilization of early intervention would amount to an expenditure of between \$150 million and \$300 million per year. Thus, the choice between early interventional therapy and medical management involves making a critical trade-off between avoiding the upfront risks and cost of an early intervention and possibly mitigating the long-term risks and costs associated with medical management. These trade-offs have not been adequately addressed in the clinical literature.

2.2 Supporting Data

There have not been any randomized trials comparing any of the forms of interventional therapy for BAVMs among themselves or with medical management. This is the case despite the enormous resources committed to the treatment of patients with BAVMs. Some data indicate

that interventional therapy is superior to medical management for BAVMs, but many of these studies do not distinguish between AVMs that have previously bled and those that have not. Other data suggests that there is a spectrum of risk for medical management of BAVMs and those that are unruptured have a much lower risk for future hemorrhage than those that have previously bled. The currently available published data on both medical management and treatment-related morbidity and mortality do not separate outcomes by pre-treatment status (bled or unbled), and show little



consistency for mode and number of treatments or for clinical severity.

An important source of data that we have relied upon in planning this study is the Columbia AVM Databank project, which has prospectively enrolled 622 consecutive AVM patients clinically encountered at Columbia University Medical Center since 1989. The mean age of these patients is 34 years with a standard deviation of 15 years. Three hundred and twenty-two of the patients, or 53%, are female. Of the 622 study subjects, 282 (45%) presented with hemorrhage and 340 (55%) had unruptured AVMs.

A recent analysis of these data favors early treatment intervention in patients who have bled, showing little additional clinical injury for the extirpation of the lesion, particularly in those harboring additional morphological risk factors. Of concern, however, is the low risk of spontaneous rupture in as yet unbled AVMs and the mild clinical syndrome from such rupture. As shown in the figure, interventional treatment was associated with an increased risk of hemorrhage ($p < 0.0001$; hazard ratio (HR) = 5.53, 95% CI 2.91 to 10.49). In this figure the value on the abscissa for the subgroup that underwent interventional treatment was defined as time-since-treatment-was-initiated in order to mimic the result that would be obtained in a clinical trial. The actual analysis utilized time-dependent covariates that classify treatment status at each time point of follow-up on the basis of its relationship to the time at which treatment began. Interventional treatment was also associated with an increased risk of clinical impairment as assessed by a Rankin score ≥ 2 (HR = 11.04, 95% CI 7.21 to 16.90, $p < 0.0001$). These observational data suggest that for AVM patients who have not yet bled, treatment may increase the risk of both hemorrhage and an acute, disabling persisting clinical syndrome.

Comparing our own data to those taken from the literature (not stratified by AVM rupture status) offers two extreme cases for comparing the benefits of early interventional treatment versus watchful waiting. Comparing the worst 5-year risk of stroke or death with medical management of 20%, and the best 5-year risk with early intervention of 5%, supports the strategy of early intervention, while the best 5-year natural history outcome of 5% and the worst 5-year early intervention outcome of 19% support medical management. Thus, there is considerable uncertainty in the existing clinical literature, which does not provide conclusive evidence about

optimal treatment approaches for this vexing clinical problem. The Columbia database was collected prospectively, but like other clinical series, is not a randomized trial of treatment versus medical management, or of various modes of treatment. The findings in a randomized clinical trial could well be different from that in this one-center clinical cohort.

3. STUDY DESIGN

The overall purpose of this multi-center RCT is to evaluate the effectiveness and safety (in terms of survival, clinical impairment, adverse events and quality of life), and costs of medical management compared to interventional therapy of patients with unruptured BAVMs. While the nature of the treatments precludes blinding of patients and their treating clinicians, outcome evaluations should be done by an experienced person who is not directly involved in providing the interventional procedure. Therefore a neurologist at each site who is certified to perform the Rankin assessment will do so for all outcome assessments at that center. A parallel groups design with random assignment of patients to interventional therapy or medical management with equal probability will be performed. A total of 400 patients will be randomized. Patients will be followed for a minimum of 5 years and a maximum of 9.5 years (mean 7.5 years) from randomization.

4. SELECTION AND ENROLLMENT OF SUBJECTS

The patient population for this trial consists of patients with unruptured BAVMs. All patients who meet eligibility criteria may be included in the study regardless of gender, race, or ethnicity.

4.1 Inclusion Criteria

1. Patient must have unruptured BAVM diagnosed by MRI/MRA, CTA and/or angiogram
2. Patient must be 18 years of age or older
3. Patient must have signed Informed Consent, Release of Medical Information, and Health Insurance Portability and Accountability Act (HIPAA/U.S. only) Forms

4.2 Exclusion Criteria

1. Patient has BAVM presenting with evidence of recent or prior hemorrhage
2. Patient has received prior BAVM therapy (endovascular, surgical, radiotherapy)
3. Patient has BAVM deemed untreatable by local team, or has concomitant vascular or brain disease that interferes with/or contraindicates any interventional therapy type (stenosis/occlusion of neck artery, prior brain surgery/radiation for other reasons)
4. Patient has baseline Rankin ≥ 2
5. Patient has concomitant disease reducing life expectancy to less than 10 years
6. Patient has thrombocytopenia ($< 100,000/\mu\text{L}$),
7. Patient has uncorrectable coagulopathy (INR >1.5)
8. Patient is pregnant or lactating
9. Patient has known allergy against iodine contrast agents
10. Patient has multiple-foci BAVMs
11. Patient has any form of arteriovenous or spinal fistulas

Previous diagnosis of any of the following:

12. Patient has a diagnosed Vein of Galen type malformation
13. Patient has a diagnosed cavernous malformation
14. Patient has a diagnosed dural arteriovenous fistula
15. Patient has a diagnosed venous malformation
16. Patient has a diagnosed neurocutaneous syndrome such as cerebro-retinal angiomas (von Hippel-Lindau), encephalo-trigeminal syndrome (Sturge-Weber), or Wyburn-Mason syndrome
17. Patient has diagnosed BAVMs in context of moyo-moya-type changes
18. Patient has diagnosed hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber)

Pregnancy Risks:

This study involves treatments or procedures which could be harmful to a fetus or breastfed baby. Women of childbearing age should be aware of the potential risks associated with the diagnostic and interventional treatments that are standard of care for the diagnosis and treatment of AVMs. This information must be discussed by the investigator at the time of enrollment. Women who are pregnant or nursing at the time of enrollment may not participate. Women of childbearing age who are randomized to interventional therapy are encouraged to use an effective form of birth control during the course of treatment. For women of childbearing age, a serum or urine HCG should be recorded as part of the source documentation.

4.3 Study Recruitment and Enrollment Procedures

There are 104 clinical sites in the U.S., Europe, South America, Asia, and Australia proposed as participating clinical centers in the ARUBA trial. These centers have extensive clinical and research experience with the management of BAVMs. Combined, these centers have an annual volume between 650 and 1000 patients, who would meet the eligibility criteria for the trial.

Mailings will be sent out, with IRB approved flyers prepared for posting announcement of the study. Our recruitment efforts will target the front-line physicians, local neurosurgeons, and neuro-radiologists to make them aware of the trial, so that when they evaluate a patient with an unruptured BAVM, they have the option of referring them to the clinical investigators for consideration of enrollment in the trial. We will conduct ARUBA seminars for the staff of the local neurology practices to inform them of the trial requirements. A set of Power Point ARUBA slides will be prepared by the DCC and made available to the site investigators so that they can meet and present the trial to physicians who practice in local communities. A pocket size laminated eligibility criteria list will be sent to all investigators to be distributed to all referring physicians. The DCC will develop a template ARUBA informational packet directed at referring physicians, which can be adapted by the sites. All such publicity materials targeted to patients will require IRB approval. Through this method they will be able to identify potential candidates, and make appropriate referrals to the ARUBA team.

All patients who are diagnosed with an unruptured BAVM are potential candidates for this trial. There are three main pathways that patients with unruptured BAVMs may be referred for evaluation for the ARUBA trial. They may be referred by their clinical neurologist, neurosurgeon, or interventional radiologist. If the examination and work-up confirms that a patient has met the eligibility criteria, the trial will be presented. When a patient expresses interest, they can be referred to an ARUBA investigator who will evaluate the medical records

and initiate the consent process. All referring physicians will be encouraged to present this trial to all their patients, including women and minority patients.

An ARUBA web site has been developed to allow physicians to have access to up-to-date trial information. Once the web site is approved by the IRB it will also be available to patients and their families. The ARUBA web site will be linked to other medical and clinical trial sites, including the NIH, CDC, WebMD, and Center Watch sites. Key words such as brain aneurysm and cerebral aneurysm will be included. Lay terms will be used to make the information accessible to patients, their families and friends.

4.3.1 Minority Recruitment

Recruitment will not discriminate on the basis of age, gender, race, or socioeconomic status. The proposed clinical trials pose no scientific justification to exclude any gender or ethnic group. Given the international nature of this trial, a wide spectrum of ethnic backgrounds is expected. A special effort will be made to ensure that no opportunity for recruitment of eligible women is overlooked and development of the recruitment database in cooperation with the clinical sites will place a special emphasis on effective recruitment of women from the general population.

4.3.2 Informed Consent Procedures

Only adults (those ≥ 18 years) with unruptured BAVMs will be considered for enrollment in the ARUBA trial. The site clinical investigator will discuss the trial with the patient's primary care physician who will ascertain from the prospective enrollee whether or not they wish to be approached by the investigator. The clinical investigator or a designated member of the investigative team will provide a thorough explanation of the objectives, patient responsibilities, risks and benefits of the study, and will fully address all the concerns raised by the patient and/or family. After all issues have been adequately resolved, and the investigator confirms that the patient has fully consented to participate, the patient will be asked to sign the informed consent. All patients will be given a signed copy of the informed consent for future reference. Patients who decline to be in the trial will receive the same quality of care.

4.3.3 Screening Log

Patients who are screened for enrollment in ARUBA who are not enrolled should be recorded on the patient screening log. (see Appendix V). This document is located in the Electronic Data Capture system and should be completed each time a patient is screened. This will include all AVMs seen at a participating institution.

4.3.4. Procedure for Enrollment.

The site clinical investigator or clinical coordinator will log into the Electronic Data Center and complete the following data collection forms:

- A. Demographics (AR01) which includes verification of signed Informed Consent, Release of Medical Information, and Health Insurance Portability and Accountability Act (HIPAA) Clinical Research Authorization (U.S. only) forms
- B. Eligibility Evaluation form (AR02)
- C. Imaging data (AR03) or (AR03A)
- D. Presentation history (AR05)
- E. AVM morphology (AR06)
- F. Rankin Scale (AR07)
- G. NIH stroke scale (AR08)

- H. Medical history (AR09)
- I. Medications (AR10)
- J. Quality of Life: SF-36 (AR11)
- K. Quality of Life: EuroQol (AR12).

4.3.5 Procedure for Image Interpretation and Shipment

Relevant image(s) chosen by the local investigator should not be older than one year.

If an image study is older than one year, a waiver must be obtained from the CCC PI who will decide whether that image may be used for enrollment and randomization. The decision will be based upon further conversation with the site clinical investigator. The CCC PI will document reasons for acceptance/rejection of images. On the EDC, a waiver request box will appear when the date of the images exceeds one year from enrollment date.

Each clinical site will have a credentialed radiologist/neuroradiologist who will read the images and attest to the presence of an unruptured BAVM. In the US, the radiologist will be board certified and all non-US radiologists will have the appropriate clinical privileges at the academic institution affiliated with ARUBA. After completing the Image Data form (AR03), the investigator or clinical coordinator can proceed with the randomization process. A de-identified CD of the images will subsequently be sent to the DCC along with a copy of the radiologist/neuroradiologist's written report.

In the event that a credentialed radiologist/neuroradiologist is not available at the local site, the images can be uploaded to the Imaging data form (AR03A) to be reviewed by the coordinating center radiologist within 24 hours.

4.3.6. Randomization

The randomization process will assign the patient to either medical management or interventional therapy. When the site investigator or clinical coordinator has completed the data collection forms required for enrollment, a randomization button will appear in the top left hand corner of the EDC. After clicking the button, the randomization form (AR04) will be automatically completed with the patient's randomization assignment. The coordinator or investigator will then sign the form electronically.

5. STUDY INTERVENTIONS

5.1 Medical Management (Refer to Manual of Procedures)

Patients participating in the trial will receive the best medical management possible for the disorder being tested in the trial and for any general medical illnesses they are demonstrated to have. One important consideration in the medical management of patients in this trial is stroke risk factor reduction.

An additional consideration for the medical management group is that an angiogram is not required for randomization for those unruptured BAVMs for whom the diagnosis can be made by non-invasive imaging alone. The purpose of this planned limitation of data source is for patient safety. If a patient has a successful diagnosis of BAVM without conventional angiogram and is randomized to the non-intervention arm, there is no management reason for the risks, however

small, of a diagnostic angiogram. That risk (and whatever subsequent angiogram or procedure risk exists) will remain in the interventional therapy arm of the study. If an angiogram exists, performed for reasons decided by the local center or its referring clinical team, the data is to be included with the screening data forms.

5.2 Interventional Therapy (Refer to Manual of Procedures)

A patient randomized to interventional therapy is expected to begin interventional therapy within 3 months following randomization. Interventional therapy consists of endovascular attempts at occlusion of the nidus and feeding vessels, coiling or microsurgery for feeding artery aneurysms, microsurgery for BAVM itself, and radiosurgery, these alone or in various combinations and timings.

5.2.1. Endovascular treatment

Endovascular treatment may include AVM embolization, coiling of aneurysms in the vascular territories feeding the BAVM (BAVM-related aneurysm), or coiling of aneurysms unrelated to the BAVM. The embolization materials used for those who undergo embolization as part of the treatment plan will be limited to those agents approved by the FDA or by the approval agency applicable to the country in which the patient receives treatment at the time of the procedure. This plan allows for the introduction of new agents during the course of the study. The name of the agent, the amount, and the frequency of use during each treatment will be recorded on the Interventional Therapy form (AR13).

5.2.2. Microsurgery

Microsurgery may include AVM resection, aneurysm clipping related to AVM, and aneurysm clipping unrelated to AVM.

5.2.3 Radiotherapy

Radiotherapy involves the targeting of the BAVM nidus and adjacent vessels intended to induce a reduction, and possible obliteration, of the BAVM. Based on local patterns of practice, variations exist in the exact equipment used, the methods of measurement used to assess the location and size of the BAVM chosen for therapy, the individual doses and numbers of treatments, and whether radiosurgery is used before or after embolization or microsurgery. The modality, energy, number of isocenters, collimator size, Gamma angle, prescription and duration of treatment will be recorded on the Interventional Therapy form (AR13).

5.3 Completeness of Interventional Therapy

The goal of randomization into the interventional therapy arm is to achieve eradication of the BAVM. The eradication plan may include any or a combination of endovascular, surgical, or radiotherapy treatments. Following interventional therapy, using a diagnostically relevant image study, treatment outcome will be documented as: technically complete AVM removal based on catheter angiography, technically complete AVM removal based on other than catheter angiography, technically incomplete AVM removal, technically complete aneurysm treatment, or technically incomplete aneurysm treatment

5.4 Handling of Study Interventions

Not Applicable.

5.5 Concomitant Interventions

5.5.1 Required Interventions

The local Investigator will make these decisions for the extent of the treatment.

5.5.2. Prohibited Interventions

Medications and materials not approved by the U.S. FDA for American subjects or those not approved by the local country equivalent of the U.S. FDA are prohibited while the subject is on study.

5.6 Adherence Assessment

Compliance of the subjects with the study will be assessed by adherence to the follow-up visit schedule. If a patient is unable to return for follow-up before the closure of a study visit window, the coordinator will make every attempt to contact the patient and complete the Patient Encounter form (AR14). If unable to contact the patient, a Missed Visit form (AR18) will then be submitted.

6. ENDPOINTS

6.1 Primary Endpoint

The *primary outcome* is the composite event of death or stroke. *Stroke* is defined as a symptomatic event (presenting with a new focal neurological deficit, seizure, or new onset headache) that is associated with brain imaging indicating hemorrhage (defined as fresh intracranial blood on head CT and/or MRI or in the cerebrospinal fluid, the primary bleeding location further classified as parenchymatous, subarachnoid, intraventricular, or any combination) or infarction, also defined as a clinically-related new CT (low density) or MRI (DWI, FLAIR, or T2) lesion. The severity of the resulting clinical impairment from stroke will be analyzed. Clinical impairment will be determined by a score of 2 or greater on the Rankin Disability scale. This scale will be measured at baseline, every 6 months to study completion, at every intervention, and at every neurological adverse event.

For the purposes of adjudication, the Event Adjudication Committee will use the following guidelines:

Diagnosis may be based on imaging features alone or with supporting clinical symptoms up to a year after the clinical event. In the absence of supporting clinical symptoms (neurological deficit, unusual headaches, epileptic seizure), diagnosis of stroke should not be coded as being “symptomatic”. In the absence of supporting images, diagnosis of stroke should not include a specification of type (ischemic versus hemorrhagic).

6.2 Secondary Endpoints

6.2.1 Quality of Life and Patient Preferences

This clinical trial will employ a combined approach to assessing the health-related quality of life of participants by using two broad types of measurements: those that capture health status

through the description of functional capabilities, symptoms, and general health perceptions and those that generate global utility measures, which reflect both the health status and value placed on the health status by the individual. Patient utility measures will be used as quality adjustment factors to derive quality adjusted life years for the cost-effectiveness study.

The SF-36 is a 36 item generic self-report QoL instrument which provides measures on 8 dimensions of quality of life: physical functioning, role limitations due to physical factors, mental health, general health, role limitations due to emotional factors, social functioning, bodily pain and vitality. The analysis of quality of life as a secondary endpoint will include both the physical and mental composite scores of the SF-36.

We will use the EuroQoL questionnaire to derive patient preferences. This instrument examines five quality of life dimensions (mobility, self-care, usual activities (work, study, housework, family, or leisure), pain/discomfort and anxiety/depression). In addition, respondents record their perception of their overall health on a visual analog scale (0, worst, 100, best). The visual analog scale score directly reflects the respondents' view of their own health status. A societal view of the health states can be derived from population-based valuations of the 243 unique states of health described by the 5 quality of life dimensions.

6.2.2 Adverse Events

The incidence of all protocol defined adverse events will be evaluated, regardless of whether they are anticipated. Serious adverse events are defined as those that cause death or permanent disability, are life threatening or require a hospitalization, or prolong an existing hospitalization.

Protocol-defined events will include:

I. Neurological Adverse Events:

1. *Stroke* is defined as a clinically symptomatic event (revealed by a new focal neurological deficit, seizure, or new onset headache) when associated with brain imaging indicating hemorrhage (defined as fresh intracranial blood on head CT and/or MRI or in the cerebrospinal fluid, the primary bleeding location further classified as parenchymatous, subarachnoid, intraventricular, or any combination) or infarction, also defined as a clinically-related new CT (low density) or MRI (DWI, FLAIR, or T2) lesion.

Stroke presentation will be classified by the following subtypes:

- A. *Intracranial hemorrhage*: Revealed by imaging showing subarachnoid, parenchymatous or intraventricular fresh blood, or by spinal tap.
- B. *Brain infarction*: Signs of infarction on brain CT or MR imaging by DWI, T2, or FLAIR imaging.

Stroke symptoms will be classified by:

- A. *New focal neurological deficit*: A functional deficit on examination, stratified as to whether the deficit was persistent, progressive or reversible.

- B. *New onset headache*: Patient complaint of new onset headache.
- C. New onset seizures: Newly observed seizure activity.
- 2. *Seizure (unrelated to stroke)*: Clinically suspected epileptic activity without signs of recent intracranial hemorrhage or cerebral infarction on brain imaging (CT and/or MRI).
- 3. *Focal neurological deficit (unrelated to stroke)*: Focal neurological deficit on clinical exam without signs of recent intracranial hemorrhage or cerebral infarction on brain imaging (CT and/or MRI).
- 4. *Headache (unrelated to stroke)*: Patient complaint of new onset headache without signs of recent intracranial hemorrhage or cerebral infarction on brain imaging (CT and/or MRI).
- 5. *Other Neurological Event*: Any new, temporary or permanent, focal or global neurological deficit ascertained by standard neurological exam and appropriate diagnostic tests that is not a stroke, seizure, focal neurological deficit, or headache.

II. Non-Neurological Adverse Events

- 1. *Acute renal failure*: An episode of acute renal failure requiring peritoneal dialysis, hemodialysis or hemofiltration (excluding hemofiltration for fluid management alone).
- 2. *Procedure related nephropathy*—a rise in the plasma creatinine concentration of more than 50 percent above baseline or of more than 1 mg/dL (88 μmol/L), whichever is smaller within 7 days following a procedure.
- 3. *Contrast reaction*: Anaphylactic reaction in the context of intravenous or intra-arterial contrast dye injection.
- 4. *Infection related to BAVM invasive therapy*: Clinical or paraclinical signs of local or systemic infection related to invasive therapy.
- 5. *Peri-procedure bleeding (other than intracranial)*: Bleeding that results in death or transfusion of packed red blood cells during the 24 hour period following an invasive therapy for an AVM.
- 6. *Systemic (non-brain) embolization*: Unintended dislocation of embolic material into non-cerebral arteries or veins.
- 7. *Vascular injury related to BAVM invasive therapy*: Mechanical injury to any arterial or venous structures during the course of the intervention without stroke.
- 8. *Catheter adherence to embolization material*: Unintended adherence of a catheter delivering embolization material to the BAVM and the inability to remove the catheter without causing damage to the vessel and/or requiring a surgical procedure to correct it.
- 9. *Other non-neurological Adverse Event*: An event that causes clinically relevant changes in the patient's health or any event that is life-threatening, results in a fatality, results in

permanent disability, requires hospitalization, or prolongs an existing hospital stay.

6.2.3. Cost Endpoints

We will employ a health care perspective in this RCT and calculate the costs of all services associated with care, regardless of who bears the cost. These costs will include the direct costs of medical care, the costs of non-medical care and indirect health care costs. The Investigators will identify those costs that are related to the research protocol and are not part of usual care. We will conduct the economic analysis in the cohort of U.S. patients, with an expected sample size of 500 150 patients.

6.2.3.1 Direct Costs of Medical and Non-Medical Care

We will derive costs by using the clinical dataset to identify the resources that patients use during the course of the trial, and then assign payments/prices for each resource used. There are a multitude of payers in the U.S. that reimburse for services at different rates. We propose to use the Medicare payments as representative rates. For inpatient hospital days, we will use the Medicare reimbursement for the DRG codes assigned on the patient's discharge. We will not include physician time in our costing, as it is a much smaller part of the overall costs and it requires substantial data collection efforts to capture. For those patients who need nursing facilities or long-term institutional care, we will use the National Medicare average allowed daily rate to impute payments. The use of services outside the study hospitals, such as emergency room visits, out of network hospitalizations, nursing home care, and rehabilitative facility care will be determined by a structured questionnaire administered by site coordinators to all enrolled patients.

In seeking medical treatment, patients may also incur significant non-medical care costs. These costs may include the value of unpaid care provided by family members and friends, the costs of uncompensated home health care and the "costs" of time dedicated to care by the patient. We will focus on obtaining the value of unpaid care provided by family and friends with the following question administered at 6 month intervals by site coordinators on the Patient Encounter form (AR14): Has your illness required any members of your family or friends to restrict their work or social activities? If yes, about how many hours per week have friends or family spent in helping with your care? Each hour of care will be valued at an average hourly total compensation rate for civilian workers as reported in the base year by the Bureau of Labor Statistics. The value of home health care will be determined by asking patients directly if they had a home health aid or home nurse and the number of hours per week that they are employed. The hourly wage rate will be determined by the average Medicare reimbursement rate. We will not collect data on travel costs or the amount of time patients must spend seeking treatment (i.e. the opportunity cost of lost leisure time as measured by the wage rate), because of the substantial burden involved in such data collection.

7.1 DATA COLLECTION SCHEDULE

Study Visit (months after randomization) (V)=Visit, (P)=Phone Call	Screening / Baseline	6	12	18	24	30	36	42	48	54	60	66	72	78	84	End of Study (V or P)		Event Driven
		(V)	(V)	(V)	(V)	(P)	(V)	(P)	(V)	(P)	Month Visit (V)	(P)	(V)	(P)	(V)	(P)	(V)	
Window (days)	0	±30	±30	±30	±30	±30	±30	±30	±30	±30	±30	±30	±30	±30	±30	-30		
Informed Consent	X																	
Release of Medical Information	X																	
HIPAA Authorization (US only)	X																	
Demographics	X																	
Eligibility Evaluation	X																	
Randomization	X																	
Image Study	X																	
Presentation History	X																	
AVM Morphology	X																	
Rankin Scale	X [#]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X*
NIH Stroke Scale	X	X	X	X	X		X		X		X		X		X	X	X	X*
Medical History	X										X							
Medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Quality of Life: SF-36	X	X	X	X	X		X		X		X		X		X	X	X	
Quality of Life: Euroqol	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Interventional Therapy																		X
Delayed Treatment																		X
Patient Encounter		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X&
Adverse Event																		X
Hospitalization																		X
Mortality																		X
Missed Visit																		X
Voluntary Withdrawal																		X
60 Month Visit											X [^]							
End of Study																	X	
Investigators Statement																	X	

* collected following each BAVM interventional therapy, all neurological adverse events and hospitalization

& collected following each BAVM interventional therapy

[^] window of -6 months for image study

[#] if Rankin scale collected more than 6 months prior to randomization, assessment should be repeated

7.2 Timing of Evaluations

Enrollment: Screening and Baseline

Consent:

Prior to chart review, screening data and protocol defined procedures.

Release of Medical Information Form (May be combined with consent):

Prior to chart review, screening data and protocol defined procedures.

Health Insurance Portability and Accountability Act (HIPAA) Clinical Research Authorization (US only)

Prior to chart review, screening data and protocol defined procedures.

Form AR01: Demographics

At initiation of screening.

Form AR02: Eligibility Evaluation

At initiation of screening.

Screening Log

Documentation of patients screened by not enrolled in ARUBA.

Form AR03 or Form AR03A: Image Study

At initiation of screening.

Form AR05: Presentation History

At baseline, prior to randomization.

Form AR06: AVM Morphology

At baseline, prior to randomization.

Form AR07: Rankin Scale

At baseline, prior to randomization. If Rankin scale is collected more than 6 months prior to randomization, assessment should be repeated.

Form AR08: NIH Stroke Scale

At baseline, prior to randomization

Form AR09: Medical History

At baseline, prior to randomization.

Form AR10: Medications

At baseline, prior to randomization.

Form AR11: Quality of Life - SF-36

At baseline, prior to randomization.

Form AR12: Quality of Life - EuroQOL

At baseline, prior to randomization.

Randomization:

Form AR04: Randomization

The investigator/coordinator who receives the randomization assignment from the DCC must electronically sign the form at which point they may proceed with the treatment assignment.

On-Study Evaluations

Form AR14: Patient Encounter

Patients who meet the eligibility criteria and are randomized into the study will follow the same in person study visit and telephone communication schedule. Patient study visits will be scheduled at 6, 12, 18 and 24 months for the first two years with a 30 day window (± 30 days). After two years, an in person annual visit (± 30 days) for years 3, 4 and 5 will be scheduled with telephone communication scheduled at 6 month (± 30 days) intervals in between. For patients who are randomized to interventional therapy, an interventional therapy visit which falls within the ± 30 day window of a scheduled visit may be counted as a protocol defined scheduled visit. After 5 years (60 months) an annual in person visit (± 30 days) will be scheduled with telephone communication at 6 month ($+30$ days) intervals in between until the end of the study.

Form AR07: Rankin Scale

At the time of all protocol defined scheduled visits. A Neurologist who has completed the ARUBA Rankin training must complete this scale.

Form AR08: NIH Stroke Scale

At the time of all protocol defined in person visits. A Neurologist or certified coordinator must complete this scale.

Form AR09: Medical History

At the 60 month point following randomization, this form must be completed

Form AR10: Medications

At the time of all protocol defined scheduled visits.

Form AR11: Quality of Life - SF-36

At the time of all protocol defined in person scheduled visits.

Form AR12: Quality of Life – EuroQOL

At the time of all protocol defined scheduled visits.

Form AR20: 60 Month Visit

Once a patient reaches the 60 month point following randomization, this form must be completed

Event Driven

Form AR13: Interventional Therapy

At time following each BAVM treatment.

Form AR14: Patient Encounter

At time following each BAVM treatment a Patient Encounter needs to be documented.

Form AR07: Rankin Scale

Within 48 hours of each BAVM treatment, following each neurological adverse event and hospitalization, a Rankin Scale needs to be documented.

Form AR08: NIH Stroke Scale

Within 48 hours of each BAVM treatment, following each neurological adverse event and hospitalization, an NIH Stroke Scale needs to be documented.

Form AR15: Adverse Events

Event driven.

Form AR16: Hospitalization

Event driven.

Form AR17: Mortality

Event Driven-within 24 hours of knowledge of event.

FormAR18: Missed Visit

Event Driven.

Form AR19: Voluntary Withdrawal

Event Driven.

Form AR22: Adverse Events Adjudication

Event Driven.

Form AR23: Mortality Adjudication

Event Driven.

Form AR24: Delayed Treatment

Event Driven.

Final Evaluations

Form AR 25: End of Study

This form will be completed on all patients currently enrolled in the study 5 years (60 months) following the randomization of the last patient.

Form AR21: Investigator Statement

The investigator will sign this form electronically after all patient forms have been submitted at the completion of the trial.

7.3 Special Instructions and Definitions of Evaluations

Consent:

The method of obtaining informed consent involves a discussion with the investigator, as a result of which, the patient accurately understands the study, the risks and benefits of participation, and has had all their questions answered about the study prior to making a decision whether or not to participate in the study (Refer to the protocol Appendix I. Informed Consent Template).

Release of Medical Information Form:

The patient must sign the Release of Medical Information form that authorizes release of medical records to the study investigators, monitors, NINDS, and the DCC.

HIPAA Clinical Research Authorization

The HIPAA Form approved by the IRB or Privacy Board allows site investigators to approach, screen, and enroll patients into the study (U.S. only).

(Refer to the Manual of Procedures Appendix VI. for a copy of all Case Report Forms)

Form AR01: Demographics

A screened patient is an individual who was referred to, or identified at a clinical center for consideration in the study, and for whom some preliminary data (i.e. medical records) have been reviewed. Demographic information including the patient's first, middle and last name initials, date of birth, gender, ethnic category, racial category, and handedness. The electronic data capture system (EDC) will generate a sequential screening number to identify each patient.

Form AR02: Eligibility Evaluation

A complete checklist of inclusion and exclusion criteria will be documented. A waiver must be requested to enroll a patient who has a value outside of the protocol defined range.

Form AR03: Image Study

This form includes the date of the image study, the type(s) of the images, the presence of an unbled BAVM, the initials of the radiologist and the signature of the site investigator/clinical coordinator to confirm the information. The form also documents that the de-identified CD of the images and the radiologist's final report were sent to the DCC. In the event that an image is older than one year, a waiver must be obtained from the CCC PI who will document acceptance or rejection of the images.

Form AR03A: Image Study

This form is to be used by sites which do not have a credentialed radiologist available. The form includes the date of the image study, the type(s) of the images, the uploaded images and the radiologist's review and approval, rejection or request for more images.

Form AR04: Randomization

The form includes the name of the clinical center, the date and time of randomization, the randomization assignment and the patient ID. The form will be signed electronically by the clinical center by the clinical coordinator or the clinical investigator at the respective site.

Form AR05: Presentation History

This form documents the date of the diagnostic event, mode of presentation, incidental diagnosis, and clinical presentation.

Form AR06: AVM Morphology

This form documents the specific imaging source, location, and size of the AVM, arterial supply, presence of intranidial aneurysm(s) and the presence of unrelated aneurysm(s). Questions 5-13 can only be answered if an angiogram was performed.

Form AR07: Rankin Scale

This standardized scale includes 5 components to assess functional status.

Form AR08: NIH Stroke Scale

This standardized scale includes 14 components to assess for neurologic deficits. (See Appendix III. of the Manual of Operations for instructions on administration.)

Form AR09: Medical History

This form captures the information pertaining to the patient's baseline vital signs and medical history including vascular risk factors, and other concomitant diseases.

Form AR10: Medications

This form captures prescribed medications the patient has taken over the last seven days in the following categories: antiepileptic medications, headache medications, antihypertensive medications, lipid lowering agents, anti-diabetic medications, anti-platelet agents and anticoagulants.

Form AR11: Quality of Life - SF-36

The Short Form 36 Health Questionnaire (SF-36) will be completed by the patient and used to assess quality of life.

Form AR12: Quality of Life - EuroQOL

The Euroqol is a 6 item questionnaire completed by the patient and is used to assess the patient's perception of their overall health status.

Form AR13: Interventional Therapy

This form captures the details about the interventional therapy that enrolled patients receive, including the reason for interventional therapy, the type of interventional therapy, and the result of the BAVM therapy. This form will be completed for all patients regardless of treatment arm when AVM eradication therapy is performed.

Form AR14: Patient Encounter

This form will capture any diagnostic procedures, interventional therapy, and/or clinical events the patient has experienced in between communication assessments and follow-up visits.

Form AR15: Adverse Events

Detailed information regarding adverse events will be recorded at the time an adverse event takes place. Events will be sub-categorized as neurologic or non-neurologic. Investigators will be asked to make a judgment as to the seriousness of the event. The relationship of the event to the natural history of the index AVM or AVM-related interventional therapy will be ascertained. Interventions performed as a result of acute interventional therapy will be collected if an adverse event occurred.

Form AR16: Hospitalization

Information regarding all hospitalizations will be reported and include information regarding date of hospitalization, number of days in intensive care unit setting (i.e. Neuro ICU, MICU, CCU, SICU), length of stay, medical and surgical procedures performed, a clinical narrative, and disposition at time of discharge (home, skilled nursing facility, rehabilitation facility, death).

Form AR17: Mortality

The mortality form must include the primary cause of death and what the immediate cause of death was attributed to. A clinical narrative will be instrumental in the adjudication of mortality classification. Supporting source documentation should be collected, copied, and filed in the Case Report Binder.

Form AR18: Missed Visit

If a patient is unable to return for follow-up and unable to be contacted via telephone before the closure of a study visit window, a missed visit form must be completed.

Form AR19: Voluntary Withdrawal

The Voluntary Withdrawal form must be completed if the patient chooses to withdraw from this study. The only anticipated withdrawal from this study is patient request.

Form AR20: 60 Month Visit

Once a patient reaches the 60 month point following randomization, the following data points must be obtained:

- a. A diagnostically relevant imaging study is suggested for all patients at the 60 month post-randomization point. The imaging study should be performed no earlier than 6 months prior to

the 60 month point. If the image is performed before month 54, a waiver should be obtained from the CCC PI. A de-identified CD and the radiologist's report will be sent to the DCC.

b. A reminder that scheduled study visit forms for month 60 have to be completed. These include: Rankin scale (AR07), NIH stroke scale (AR08), Medications (AR10), SF-36 (AR11), Euroqol (AR12), Patient encounter (AR14). Adverse Events (AR15) and Hospitalization (AR16) must be submitted as applicable.

Form AR21: Investigator Statement

After a complete review of the electronic CRFs and patient summaries, the investigator will sign this form to attest to the accuracy and completeness of the data collected.

Form AR22: Adverse Events Adjudication

The adjudication committee will review all adverse events both neurological and non-neurological and adjudicate the seriousness and relatedness of the event.

Form AR23: Mortality Adjudication

The adjudication committee will review all mortality adverse events and adjudicate the proximate and primary underlying cause of death.

Form AR24: Delayed Treatment

Any patient randomized to interventional therapy who does not begin interventional therapy within 3 months after randomization must have a reason documented.

Form AR25: End of Study

5 years (60 months) after the last patient is randomized, all patients currently enrolled will be seen either in person or contacted by phone. A diagnostically relevant image study is suggested at this time. If the imaging study is performed, a de-identified CD and the radiologist's report will be sent to the DCC. The form contains a reminder that scheduled study visit forms for end of study have to be completed. These include: Rankin scale (AR07), NIH stroke scale (AR08)(in person visit only), Medications (AR10), SF-36 (AR11) (in person visit only), Euroqol (AR12), Patient encounter (AR14). Adverse Events (AR15) and Hospitalization (AR16) must be submitted as applicable.

8. MANAGEMENT OF ADVERSE EXPERIENCES

It is anticipated that any complications suffered by the patient will be brought directly to the attention of the local PI. The information from the local investigator will separately document the accuracy of the initial classification and assess its severity. These will be classified into neurological and non-neurological events (as defined in section 6.2.2). Adverse experiences will be sought and reported according to their relationship to index AVM and/or interventional therapy and seriousness. The investigators will be trained to identify and document these events in the EDC. Non-serious adverse events must be entered into the EDC within 72 hours. Serious adverse events must be reported to the DCC within 72 hours of knowledge of the event.

9. CRITERIA FOR INTERVENTION DISCONTINUATION

9.1 Brain Hemorrhage or Stroke Related to BAVM in Medical Management Group

Any patient suffering hemorrhage in the course of the medical management arm is censored at that point and is eligible for intervention at the discretion of the patient and local center, but remains in the original randomization arm for the purposes of intent-to-treat analysis. The occurrence of hemorrhage will be documented in the Adverse Event form (AR15), the coordinating center notified within 72 hours, the patient seen within 48 hours of the event by the neurologist blinded as to the nature of the event, and examined for the NIH Stroke Scale and Rankin Scale.

9.2 Intervention-associated Hemorrhage

Any occurrence of intervention-associated intracranial hemorrhage that occurs with evidence of the onset of new symptom(s) (new focal neurological deficit, seizure or new onset headache) is to be documented in the Intervention form (AR13) and applicable Adverse Event form (AR15) at the time of the event. The coordinating center must be notified within 72 hours, the patient seen within 48 hours of the event by the neurologist blinded as to the nature of the event, and examined for the NIH Stroke Scale and Rankin Scale. The decision(s) to continue with treatment and plan(s) for the types of treatment, are to be made by the treating team.

A serious adverse experience in the course of interventional therapy will be sufficient for discontinuation of interventional therapy. The range of adverse events is noted in section 6.2.2.

10. STATISTICAL CONSIDERATIONS

10.1 General Design Issues

The study design is a prospective, multi-center, parallel design, randomized, controlled trial. Treatment assignment will not be masked; however, clinical coordinating center personnel and outcome events committees will be blinded to treatment assignment. The primary outcome is the composite event of death from any cause or stroke (hemorrhage or infarction revealed by imaging). Clinical outcome status will be measured by the Rankin Scale, a widely-used outcome measure for stroke. The secondary measures of outcome include adverse events, quality of life and cost.

10.2 Sample Size and Accrual

Sample size calculations are based upon both previously published studies and preliminary data obtained from our institution. Columbia data yielded Kaplan-Meier estimates of event rates at 5 years of 4% for medical management and 19% for interventional therapy. Published estimates of natural history risk range from 2-4% while quoted estimates of treatment risk vary but have been assumed to be 5% or lower for the purposes of cost-benefit analysis.

Our sample size calculations are based on the following assumptions: (1) time-to-event is exponentially distributed with a constant hazard, (2) the five-year event rate for patients assigned to interventional therapy is 22% (3) patient accrual will occur uniformly for 60 months (5 years), and follow-up will continue for an additional 60 months (5 years) after the last patient is randomized. A total of 400 patients, randomized with equal allocation to medical management or to interventional therapy, assures 80% power to detect a 46% reduction (hazard ratio of 0.54) in the risk of death or stroke for the medical management arm compared to interventional therapy. Assuming that the 5-year event rate for patients treated with interventional therapy is 22%, this corresponds to an absolute reduction of 9.5% the in 5-year event rate for treatment by medical management.

With 400 randomized patients, the potential test of non-inferiority has 80% power to declare medical management non-inferior to intervention with a one-sided 0.05 level test, assuming: (1) a non-inferiority margin of 13% risk reduction (hazard ratio of 0.87) for interventional therapy compared to medical management (2) medical management reduces the risk of death or stroke by 30% (hazard ratio of 0.70) compared to intervention. One hundred and five events are expected under these assumptions. The rationale for the test of non-inferiority is provided below

Four hundred patients will also assure sufficient power to assess group differences in clinical impairment, the metric of interest for the third aim associated with the primary hypothesis (Aim 1.1b). Based on our preliminary data, the proportion of patients clinically impaired (Rankin Disability Score ≥ 2) at five years after discovery of their AVM is expected to be in the range of 10-20% for those treated by medical management. If the proportion of impaired patients is as low as 10% in the medical management arm, then power is approximately 90% to detect a relative risk of impairment of 0.45 compared to the interventional therapy group (i.e., 10% impaired versus 22% impaired). If the proportion impaired is as high as 20% in the medical management group, then power is approximately 90% to detect a relative risk of 0.57 (i.e., 20% impaired versus 35% impaired). These results are based on two-sided 0.05 level exact tests. For 80% power, the detectable relative risks are 0.49 (10% versus 20.5%) and 0.61 (20% versus 32.8%) respectively.

The power of this trial is adequate to detect a fairly large, but we believe, reasonable effect. There is 80% power to detect a 46% relative reduction in risk (absolute 9.5% reduction in five-year events rates). We note that these effects are substantially smaller than the effects observed in our single center non-randomized series of 387 patients presenting with unruptured BAVMS between 6/87 and 7/03. Those data indicate a large benefit of medical management with an estimated hazard ratio of 0.18 (a relative reduction in risk of 82% for medical management) and 95% confidence interval of (0.095, 0.34) (a relative reduction in risk between 66% and 90.5% for medical management).

Ideally, the trial is powered to detect a smaller effect, in particular, the smallest effect that would likely change clinical practice. However, we believe that 400 patients enrolled from 104 centers for a 10-year study conservatively represents the largest and longest practicable trial possible to assess this important question. To address this concern, we propose performing a non-inferiority test if the primary null hypothesis of no treatment difference is not rejected, to establish whether or not medical management is at least as good as interventional therapy.

Medical management would be at least as good as interventional therapy, if the relative benefit of interventional therapy could be clearly established to be less than 13% compared to medical management. This non-inferiority margin represents an absolute difference in five-year event rates of approximately 2.5%. With 400 patients, there is 80% power to reject inferiority of medical management using this non-inferiority margin (13%), if the relative reduction in risk with medical management is at least 30%.

10.3 Randomization Design and Procedure

The ARUBA trial will use a 1:1 ratio in randomizing patients to the two treatment arms. Pre-stratification (stratification in the design) seeks to ensure that treatment groups are balanced with respect to factors that are likely to affect the outcome. We will stratify by clinical center using a random permuted block design.

Randomization will be implemented as described in section 4.3.5.

10.4 Data Monitoring and Analysis

10.4.1 Methods of Analysis

The primary outcome of this RCT is the composite event of stroke or death. The null hypothesis is that there is no difference in the time to occurrence of this outcome between patients randomized to receive medical management and those randomized to receive interventional therapy. The primary null hypothesis will be tested in an intent-to-treat analysis using the log-rank statistic to test for differences between survival curves. The analysis will be based on a two-tailed 0.05 level test. Due to the large number of centers, the primary analysis will not be stratified by center even though the randomization is stratified by center.

Failure to reject the primary null hypothesis will not necessarily imply that the two treatments are equivalent. To conclude equivalence or that one treatment is at least as good as another, a statistical test or confidence interval procedure must rule out clinical inferiority with a high probability. If the primary null hypothesis is not rejected, we will test whether medical management is inferior to intervention with a one-sided 0.05 level test (taking account of two interim analyses for superiority). No inflation of Type I error is associated with performing a test of inferiority after a test of superiority has not rejected the null hypothesis, since this strategy represents a simple closed test procedure. The test of non-inferiority will be based on a confidence interval approach. We will compute a 95% lower confidence bound for the hazard ratio for interventional therapy compared to medical management. If this lower bound is greater than 0.87 (i.e. the relative benefit of intervention compared to medical management is less than 13%) the null hypothesis of inferiority will be rejected. Details of the approach follow.

Cox proportional hazards regression will be used to obtain an estimate of the (natural) log of the hazard ratio, $\log(\hat{\theta})$, and its asymptotic standard error, $se(\log(\hat{\theta}))$. The hazard ratio θ represents the relative risk of the composite event of death from any cause or stroke for interventional therapy compared to medical management. The lower bound to assess non-inferiority will be computed as $\exp\{\log(\hat{\theta}) - 1.695 \text{ se}(\log(\hat{\theta}))\}$, where $\exp(x) = e^x$. The Cox model will contain a single indicator for randomization group. The log hazard will be estimated as the maximum

partial likelihood estimator. The variance (squared standard error) of the estimate will be based on the inverse information matrix evaluated at the estimated log hazard ratio.

We recognize that, in general, tests of non-inferiority are not conservative. Flaws including violations in entry criteria, noncompliance, losses to follow-up, missing data, and protocol deviations tend to bias toward a conclusion of non-inferiority. This trial's straightforward treatment protocol, gate keeping strategies, minimal cross-over between treatments and likelihood for minimal losses to follow-up (detailed below in sections on Missing Data and Crossovers) will tend to minimize this kind of bias. Nevertheless, we propose to test non-inferiority using two analysis sets; the intention-to-treat set, considering all patients as randomized regardless of whether they received the randomized treatment, and the "per protocol" analysis set. Criteria for determining the "per protocol" group assignment would be established by the Steering Committee and approved by the DSMB before the trial begins. Given our expectation that very few patients will crossover or be lost to follow-up, these analyses should agree very closely. We propose declaring medical management non-inferior to interventional therapy, only if shown to be non-inferior using both the "intention to treat" and "per protocol" analysis sets.

The time to event experience of each randomization group will be described by survival curves estimated by the Kaplan-Meier method. Of particular interest is the point estimate and associated 95% confidence interval of the five-year actuarial event rates in each randomization group. Cox proportional hazards regression will be used to estimate the hazard ratio and its associated 95% confidence interval.

The primary analysis of the trial is the comparison of event-free survival between randomization groups described above. Several additional analyses will be performed to support the assessment of the trial's first specific aim. These analyses focus on the clinical impairment of patients. The principal measure of clinical impairment is the composite event of death or Rankin Score ≥ 2 .

The principal assessment of clinical impairment will be a comparison of the proportion of patients dead or clinically impaired (Rankin Score ≥ 2), at five years between randomization groups using an exact binomial test. Additionally, we will (1) compare the time to the first occurrence of death or clinical impairment (Rankin Score ≥ 2), between randomization groups using the log-rank test, (2) compare the rate with which the composite event of death or clinical impairment (Rankin score is ≥ 2) occurs over the course of the study between randomization groups using a Poisson regression model (the Poisson model will have the number of times the composite event occurs as the dependent variable and include the natural log of follow-up time as an offset), and (3) execute two longitudinal analyses of all available Rankin data for each patient (collected every 6 months) the first analysis will use logistic regression with parameter estimation via generalized estimating equations (GEE) to estimate the odds ratio of having a Rankin score ≥ 2 over the course of the study, and a second analysis using a linear mixed effects model considering the Rankin score as a continuous variable.

10.4.2. Assessing the proportional hazards assumption

The validity of the log-rank test of the equality of event-free survival depends on the appropriateness of the proportional hazards assumption. This assumption will be assessed both

graphically and by a formal statistical test. Graphical assessments will be based on two plots: (1) a “log-negative-log plot”, i.e., a plot of $\log(-\log(S(t)))$ versus $\log t$ for each treatment group and (2) a plot of the “scaled Schoenfeld residuals”²⁰ versus $\log t$ for each treatment group (where by “log” we mean the natural logarithm and by “t” we mean time in months). A formal test for the appropriateness of the proportional hazards assumption will also be performed if there is strong evidence of non-proportional hazards that could bias the result of the test of the null hypothesis (e.g., the survival curves cross). Note that we are concerned about crossing hazards as might be expected if there were an early benefit to medical management and a later benefit to interventional therapy. We do not plan to deviate from the proposed log-rank analysis if the non-proportionality stems from diverging hazards resulting from a monotonic accelerated benefit for one arm compared to the other.

The formal test will assess the significance of the interaction between the indicator for treatment group and $\log(t)$ in a Cox proportional hazards regression model that also includes a main effect for the randomization group. Statistical significance of the interaction term (based on a two-tailed 0.05 level test) would indicate a violation of the proportional hazards assumption. In that case, a comparison of five-year survival estimates based on a Kaplan-Meier analysis would be more appropriate. Therefore, if the proportional hazards assumption is not valid due to crossing survival functions, the primary null hypothesis will be tested using a confidence interval approach based on the log-log survival function, as suggested by Kalbfleish and Prentice.

10.4.3 Interim Analysis

We will conduct formal interim analyses with respect to the primary endpoint to give us the option of stopping early should results strongly favor one arm or the other. As the decision to terminate early would likely occur after all patients were randomized, the principal benefit of early termination would be prompt dissemination of results and the possibility of cross-over from the medical arm, should interventional therapy prove to be superior. A group sequential procedure allows for flexibility in the number and timing of interim analyses should the DSMB choose to modify the proposed plan. We will use the Lan-DeMets approach, implementing an O'Brien-Fleming-type spending function that allots most of the type I error to the final look. We plan for two formal interim looks. Given the assumptions underlying sample size calculation, approximately 87 events are expected to occur during the study. We propose to perform the two interim analyses at approximately equally spaced intervals with respect to the number of expected events, that is, after observing 29 and 58 events. The resulting critical values to be used for each analysis are 3.7103 at the first interim analysis, 2.5114 at the second interim analysis, and 1.9930 at the terminal analysis

In addition to the ethical concern of continuing a trial that shows a clear benefit in favor of one treatment, there is usually a corresponding ethical concern of continuing a trial that has little chance of ever showing a benefit of one treatment compared to the other. However, since a major goal of this trial is to establish noninferiority of medical management to intervention if neither treatment is superior, we believe that an assessment of futility (i.e., the conditional probability of finding a statistically significant result if the trial were to continue to the planned maximum sample size) is not relevant. We intend to perform the test of noninferiority at the final analysis only (not at any interim analysis).

We do not propose any a priori stopping criteria based on adverse events. The treatments in this trial are not experimental, and have well known adverse event profiles. Mortality is expected to be rare. Moreover, we believe that incident rates of adverse events and mortality must be interpreted along with information about the consistency of related measures, consistency across centers, data completeness, and any external factors including scientific developments that might impact patient safety. In addition to considering the data generated by this trial, the DSMB will consider all relevant background knowledge about the treatment of BAVMs. The DSMB would be capable, and uniquely suited, to determine decisions for convening outside the schedule of meetings, and to determine decisions to suspend or terminate the trial. These decisions should be at the discretion of the DSMB alone. We therefore recommend that the DSMB should be responsible for defining its deliberative processes, including event triggers that would call for an unscheduled review. We propose that the DSMB meet every six months to review data prepared by the Coordinating Center.

10.4.4. Assessment of Balance of the Randomization

The success of the randomization procedure in balancing important covariates (e.g. age, AVM location, presence of aneurysms) between randomization groups will be assessed at each interim analysis and at the final analysis. Continuous measures such as age and AVM size will be compared using t-tests, while chi-squared tests will be used to compare categorical variables such as age and ethnicity. As four hundred patients will be randomized, no substantial imbalances are expected. However, should any covariate differ significantly between treatment groups at the 0.01 level, and be substantively large, we will adjust for those covariates in all analyses. For example, in the survival analyses described above, such covariates would be included as stratification variables.

10.4.5 Analyses of Secondary Endpoints

Aim 2.1 Quality of Life. The SF-36 will be used to measure health-related quality of life (QoL). One approach to analyzing such data is to estimate longitudinal linear models, as in the Proc Mixed procedure in the SAS System. Our models will predict outcome from treatment group and time. While we expect few drop-outs in this otherwise young and healthy group some missing QOL data are possible, if only from deaths. The mixed modeling approach requires an assumption that the dropout is ignorable in that the probability of dropping out at any time is related only to previously observed data items. Of course, this assumption may not hold, and moreover it is impossible to test it robustly from the data at hand. An alternative approach we will also use, not subject to this criticism, will be to separate the data into strata defined by the time of death or dropout. We will then estimate a separate linear model, including a treatment effect, for the data in each stratum. This method, known as pattern-mixture modeling is not sensitive to untestable assumptions about the dropout mechanism because it models the data directly in strata defined by dropout time. The method of Wu and Bailey is an instance of pattern-mixture modeling.

Aim 2.2 Mortality. Differences in time to death between randomization arms will be tested using a log-rank test in the same manner as the primary analysis of the composite event of death or stroke.

Aim 2.3 Quality-Adjusted Survival. To measure net health outcomes for the cost-effectiveness analysis, we will adjust survival for associated quality of life to derive quality-adjusted life years. This will allow us to capture the pertinent aspects of each of these individual measures of outcome in a single value and, thus, enhance our ability to make overall comparisons between the two treatment arms, as well as to facilitate exploring the trade-offs between quality of life, survival, and cost that are inherent in these therapeutic decisions. QALYs are a general enough measure of outcome to support comparisons of disparate medical interventions, and, thus, will allow us to put our observations about the rates of health care resources expended in achieving quality survival with the management of brain AVMs into a more global health economic perspective.

Aim 2.4 Adverse Events. Differences in the incidence of individual adverse events will be compared between randomization arms using Poisson regression. Exact 95% confidence intervals (based on the Poisson distribution) for the risk ratios for individual adverse events for medical management compared to interventional therapy will be computed.

Aim 2.5 Costs and Cost-Effectiveness.

Costs

The differences in average costs will be compared between the two treatment approaches using a t-test. The log transformation of costs will be utilized if distributions do not meet the assumption of normality. Results will be expressed using 95% confidence intervals.

Cost Effectiveness Analysis

Cost-effectiveness is measured as the difference in the average costs of conservative medical management as compared to interventional therapy, relative to the difference in effectiveness of these two treatment approaches. The difference in costs over the difference in effectiveness of medical management versus interventional therapy is known as the incremental cost-effectiveness ratio, the economic parameter we will use in this trial. We will also compute net health benefits (NHB) as an alternative way of looking at cost-effectiveness. This parameter compares the incremental effectiveness of an intervention with the minimum health effect that society would demand in return for the investment, i.e., with the health produced by investing at the societal ceiling cost-effectiveness ratio (CR).

Our main measure of effectiveness for economic analysis will be quality-adjusted life years (QALYs), or survival weighted by the QoL experienced by trial patients (as measured by the EuroQol). Because we anticipate that few patients will die or be lost to follow-up, censoring should not be an issue here, and the data can be analyzed using straightforward statistical methods. As with the QALY endpoint, we will measure health care costs (see above) incurred during the trial period. Except for the few patients whom we will lose to follow-up, we expect to have complete cost data on all U.S. patients. Consequently, our cost-effectiveness analysis should require relatively straightforward methods. We will express uncertainty in the cost-effectiveness ratio using the Bayesian methods (probability intervals) that we have pioneered.[26-28] We anticipate that the distribution of costs will be skewed to the right. If this violates the assumption of normality, we will modify the method using the nonparametric Bayesian bootstrap.[29] We will employ traditional sensitivity analyses to garner insight into the effect of specific factors on the CE ratio, such as the location of care, the type of modality of

therapy used, and potential modifications in specific interventions over time. We will use standard discount rates for both QALYs and costs.

International Applicability of the Cost-Effectiveness Analysis

When we publish the results of the cost-effectiveness analysis, we will document carefully the amount of resources used in both treatment arms. We will also examine if there are any differences in patterns of hospitalization or other clinical resource use among the North American, European, and Australian centers due to potential differences in practice patterns. By inserting country-specific payment rates and the specific resource use, the cost-effectiveness analysis can easily be tailored by our international collaborators for their specific national context.

10.4.6 Additional Analyses of the Primary Endpoint

We propose several secondary analyses of the primary objective, which address differences between prophylactic treatment modalities and differences in risk according to general patient characteristics and specific BAVM characteristics.

Aim 2.6: This aim addresses whether BAVM size acts as an effect modifier for the relationship between treatment and the primary composite outcome of death or stroke. To test for effect modification, we will use a Cox model including an indicator for randomization group, BAVM size, and the interaction (product) of BAVM size and randomization group indicator as covariates. BAVM size will be considered as a continuous covariate to maximize power. The null hypothesis that the treatment effect does not depend on BAVM size will be rejected if the treatment-by-BAVM size coefficient is significant at the 0.05 level. Note that we will preliminarily assess the assumption that the effect of BAVM size on the hazard is linear by fitting a confirmatory model in each treatment arm separately in which BAVM is entered categorically in three levels defined by observed tertiles. The models would include two indicators for the three-level categorical version of BAVM size and the original linear BAVM size term. A formal test of linearity is obtained by calculating the difference between the partial log-likelihoods from the models with and without the BAVM size category indicators. If there is no significant departure from linearity at the 0.05 level, the analysis will be as described above. If there is a significant departure from linearity the Cox model used will include the two indicator variables for BAVM size and the two treatment-by-indicator interaction terms. The test of interaction would then be a two degree of freedom log partial likelihood test of the joint significance of the two treatment-by-BAVM category coefficients.

Aim 2.7: This aim addresses whether BAVM location (deep versus other) is an effect modifier for the relationship between treatment and the primary composite outcome of death or stroke. As for Aim 2.6, the test for effect modification will be based on a Cox model that contains indicators for randomization group, BAVM location (deep versus other) and a treatment-by-BAVM location term. The null hypothesis that the treatment effect does not depend on BAVM location will be rejected if the treatment-by-BAVM location coefficient is significant at the 0.05 level.

Aim 2.8: This aim will be addressed as is Aim 2.7 with appropriate substitution of venous drainage pattern (exclusively deep versus other) for BAVM location.

Aim 2.9: This aim will be addressed as is Aim 2.6 with appropriate substitution of age for BAVM size.

Aim 2.10: We will estimate the time to the composite event of death or stroke via Kaplan-Meier survival functions for each prophylactic treatment modality, and test for differences between groups using a Cox model including two indicator variables to represent the three treatments. We will also estimate the proportion of patients dead or clinically impaired at 5 years after randomization, and the associated exact 95% confidence interval for each prophylactic treatment modality. Logistic regression will be used to test for treatment differences. These analyses will include only the sub-group of patients randomized to receive interventional therapy.

Aim 2.11: We will attempt to consider completeness of eradication of the AVM as a continuous variable and proceed analogously to the analysis of BAVM size described for Aim 2.6. If this is not reasonable due to the distribution of values, we will treat completeness of eradication as categorical. The categories will be defined based on an examination of the distribution of values, but prior to analysis of the outcome.

Aim 2.12: To estimate via a Cox proportional hazards regression model the hazard ratios and associated 95% confidence intervals comparing Speltzer-Martin grading scales (grades of 1-5 with 5 being most severe) for death or stroke among patients randomized to receive interventional therapy. The hazard ratios will reflect the risk of grades 2, 3, 4 and 5 compared to the reference grade of 1. A corresponding logistic regression model will estimate the odds ratios and associated 95% confidence intervals for relating these risk factors to the composite event of death or clinical impairment (Rankin Score ≥ 2) at five years after randomization. These models will also be fit for each prophylactic treatment modality separately.

10.4.7 Imputation Procedure for Missing Data

While the analysis of the primary endpoint (death or stroke) will be based on a log-rank test and, therefore, not affected by patient withdrawals (as they will be censored) provided that dropping out is unrelated to prognosis; other outcomes, such as the Rankin Score at five years post-randomization, could be missing for patients who withdraw from the trial. We will report reasons for withdrawal for each randomization group and compare the reasons qualitatively. Given the relatively young age and overall good health of this population, we believe the severity of potential events and associated anxiety will limit patient withdrawal from the study. The effect that any missing data might have on results will be assessed via sensitivity analysis of augmented data sets. Dropouts (essentially, participants who withdraw consent for continued follow-up) will be included in the analysis by modern imputation methods for missing data.

The main feature of the approach is the creation of a set of clinically reasonable imputations for the respective outcome for each dropout. This will be accomplished using a set of repeated imputations created by predictive models based on the majority of participants with complete data. The imputation models will reflect uncertainty in the modeling process and inherent variability in patient outcomes, as reflected in the complete data.

After the imputations are completed, all of the data (complete and imputed) will be combined and the analysis performed for each imputed-and-completed dataset. Rubin's method of multiple (i.e., repeated) imputation will be used to estimate treatment effect. We propose to use 15 datasets (an odd number to allow use of one of the datasets to represent the median analytic result).

These methods are preferable to simple mean imputation, or simple "best-worst" or "worst-worst" imputation, because the categorization of patients into clinically meaningful subgroups, and the imputation of their missing data by appropriately different models, accords well with best clinical judgment concerning the likely outcomes of the dropouts, and therefore will enhance the trial's results.

10.5 Crossovers

By design, crossovers (patients who after randomization switch from the allocated treatment to the non-allocated treatment) are expected to be few in this trial. Patients randomized to interventional therapy who do not receive it during the trial can be considered crossovers. In addition, patients who are randomized to medical therapy and subsequently receive interventional therapy are considered to have crossed over if the reason for intervention was other than AVM rupture. Though we expect the rate of crossover to interventional therapy to be low, if medical management was actually superior, a crossover to interventional therapy would bias the study in favor of the null hypothesis.

11. STUDY ORGANIZATION, DATA COLLECTION, SITE MONITORING, AND ADVERSE EXPERIENCE REPORTING

11.1 Study Organization

The trial has separate Clinical and Data Coordinating Centers. The CCC and DCC jointly compose the operations committee. This committee will direct the day-to-day operations of the trial and oversee the overall conduct throughout the course of the trial. The Operations Committee will consist of the principal investigators and co-principal investigators of the CCC and DCC, the international coordinators and trial monitors.

The CCC and DCC will conduct a series of annual meetings during the course of the trial. These meetings will stimulate enthusiasm for the trial, enhance the synergy of the research team, and train the investigators and coordinators in specifics of the protocol design. Moreover, they may result in updating the guidelines prepared by the sub-committees, protocol revisions, or investigator-generated sub-studies.

The CCC will maintain regular contact with all the clinical sites, and address questions concerning the eligibility of patients, definitions of clinical factors, including endpoints, and on issues of managing patients in the ARUBA trial. A contact log for all interactions will be kept for DSMB inspection to assure that all rules of the trial have been followed. The CCC staff includes the trial's gatekeeper, who will be involved in the clearance of the randomization process. The P.I. will provide overall scientific leadership to the trial, chair the Steering Committee, and ensure that all contacts of centers are directed to the appropriate sources of information, including the CCC co-investigators with expertise in clinical neurology,

endovascular procedures, microsurgical procedures, and radiosurgery. The CCC will also communicate on a regular basis with the investigators to ensure that enrollment targets are being met, and will discuss any barriers to enrollment and opportunities to increase enrollment.

The major committee of the DCC is the data committee. This committee will meet every week during the first year of the trial, and subsequently on a bi-weekly basis. This committee will oversee data flow, quality, and completeness. In addition, the DCC (U.S. and Europe-based) will maintain open lines of communication with clinical center collaborators. Query generation and site response will guarantee regular communication. At least one member from the DCC will personally contact (by telephone or email) the clinical centers on a weekly basis to discuss enrollment and screening activity, resolve data entry issues, clarify protocol requirements, discuss adverse events, IRB status, other regulatory issues, and to troubleshoot when necessary. The Database Manager will be available for on-line support, report generation and technical help. Two coordinators (1 U.S., 1 Europe-based) will be available by beeper on a 24-hour basis for troubleshooting.

The underlying purpose of frequent and open communication is to ensure that the sites fully understand the protocol, and to provide support during the start-up phase. After the first patient has been enrolled, issues surrounding data entry may arise requiring additional tutorials, or modification of the system. This communication model will ensure that this trial will remain in the forefront of our collaborators concerns, enhance enrollment, and ensure we capture the highest quality data.

11.2 Training the Research Staff

We will employ several methods for training the investigators. Firstly, the trial's initial investigator meeting will be dedicated primarily to training. In the plenary session, we will discuss the scientific rationale, hypotheses, specific aims, adverse events, and data collection schedule. During the second part of the meeting, there will be breakout sessions for clinical sub-committees: (1) neurosurgeons, (2) neuroradiologists, (3) neurologists and (4) coordinators. The neurosurgical and neuroradiology sub-committees of the CCC will review guidelines for procedural techniques and post-procedural management, and the neurologists will do the same regarding short and long-term clinical care. The coordinators will be trained to use the web-based data entry system and to administer the quality of life instruments (refer to section 11.4 Data Monitoring and Quality Assurance for more explicate detail).

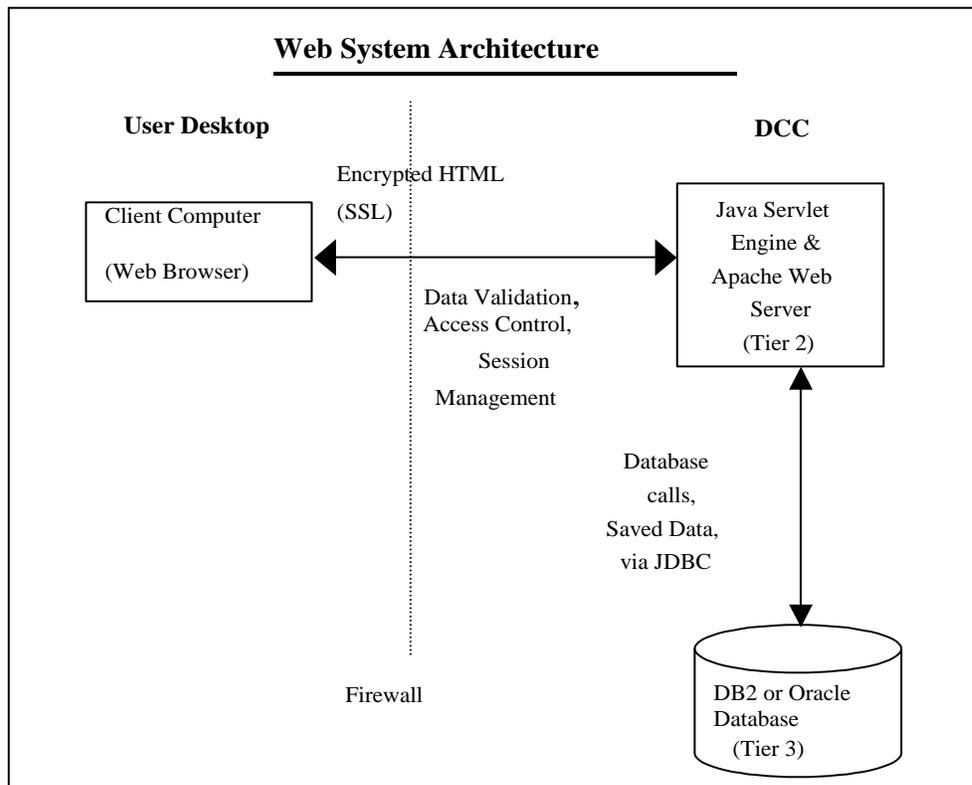
In addition, the DCC will conduct a site initiation for each clinical center, using a combination of conference call and web based demonstration, which we have successfully used in other trials. Prior to the site initiation, the DCC will send each clinical center an ARUBA binder that contains all the documents needed to conduct the study. This binder will include the protocol, a detailed, comprehensive operations manual, and blank case report forms (CRFs). A section of the ARUBA binder will be dedicated to regulatory documents. This section will include updated IRB approval letters and all other communications with the IRB, approved informed consent form, signature verification pages, responsibilities of the investigators, the monitoring visit log, the investigators' curriculum vitae, and the signed investigator agreement. There will be a communication section containing a telephone/email log on which all communication with the DCC will be recorded. During the site visit, all aspects of the protocol, operations manual and

case report forms, including IRB requirements and regulations, will be reviewed with the complete clinical center staff. In-person site initiations in Europe will be organized by the European Coordinating Center (Paris)

11.3 Electronic Data Management

11.3.1 The Data Center

The InCHOIR Data Center will provide a centralized data storage and reporting facility and computing systems support for the trial. The center will be responsible for the development and implementation of consistent standard operating procedures for the management of all trial data to ensure appropriate standards for software quality, data quality, access control, security, and physical protection of study data.



11.3.2 Security

The Data Center provides a strong network security infrastructure that treats the rest of the campus network as a potentially hostile environment. The use of a private firewall to create a “network within a network,” a private email server with an aggressive email attachment filtering policy, and network monitoring hardware have all helped to create an exceptionally secure networking environment for clinical data management. The Center has been completely unaffected by the large number of worldwide network security incidents over the past two years, including the most recent widely publicized Internet worm and email attacks.

The Center's firewall automatically blocks all Windows networking and database server traffic, which have been frequent sources of intrusions at other locations. All web-based data management sessions will be encrypted with the 128-bit SSL standard. All file transfers of project data use either secure FTP (SFTP) or WebDAV over SSL. The Center maintains several Virtual Private Networking (VPN) connections that extend the secure networking environment to branch offices outside of our main location, and a network intrusion detection system monitors and logs all network traffic at the Center.

11.3.3 Electronic Forms

The primary data collection tool for forms-based data to be collected in the trial will be a web-based forms management system that has been developed by the Center's system development team. The system design is based on an industry-standard three-tier architecture consisting of the following components:

- Tier 1: A client computer equipped with an ordinary web browser such as Netscape Navigator or MS Internet Explorer
- Tier 2: An SSL-enabled application server consisting of the Apache web server and a Java servlet engine on the middle tier. The servlet engine executes custom programs written in Java that process HTTP requests from the user's browser, access the relational database server (see below), and generate HTML response pages to be sent back to the user.
- Tier 3: A relational database server (IBM DB2) on the back end. The database can be queried directly by Data Center personnel using the SQL query language, and data can be automatically exported to a variety of machine-readable formats including SAS, ASCII, and Microsoft Access.

The system has a number of features that facilitate the management of multi-center clinical trials-
Rapid Application Development: The initial development and subsequent modification of the electronic form designs are both done using an in-house development system that enables the rapid development of low-cost web forms with a rich feature set for clinical data management. The system can generate database, validation and display components of the three-tier system, which can reduce both the overall development cost of the system and ensure a rapid implementation schedule.

Form revision tracking: The electronic system can accommodate any number of modifications to the form designs, including additions, deletions and rewording of questions. A form revision identifier is stored with each record so that data can be displayed with the version of the form with which it was captured. If desired, different form versions can be automatically assigned to different centers. Our implementation supports database queries that access either distinct revisions of a form or data that was collected using all versions.

Pass-through authentication: The database system can securely pass a user's login to an existing institutional email or Kerberos server for authentication to avoid the need for a separate database account and local password storage. Users need only remember their institutional login credentials.

Audit Trail and FDA 21 CFR Part 11 compliance: The database server stores each submitted set of changes to form data or a form design as a separate transaction in a log file, along with the identifier of the user who made the changes and the date and time of the edit session. This design enables the database manager to identify the source and time of each change in the database, and the database can be “rolled back” to recreate its state at any point in time. All archived transactions are stored on-line so that they are accessible without the need to restore data from a backup tape. An interactive audit trail is available on-line that can enable a suitably authorized user to view, print, and restore any copy of a form as it was originally saved during an edit session.

A reliable audit trail critically depends on an accurate system clock. All servers at the DCC maintain synchronization with a stratum 2 time server at Mount Sinai Medical Center using the NTP protocol.

11.3.4 Software Quality Assurance and Technical Support

The electronic data systems development group at the DCC has an ongoing quality assurance program, which is based on their experience as developers of the FDA-regulated REMATCH Trial data collection software and their current work on the development of hospital medical error reporting systems.

They have developed a software testing system that uses a clinical project’s data dictionary to generate a series of automated tests for a web-based database to verify that every form variable in the system can be correctly entered, saved and restored over its expected range of values. Values outside the expected range are also programmed to validate the appropriate error responses. The program generates scripts that simulate the actions of an end user by opening a web form in a browser window, entering data, saving, recalling the form, and comparing the results to a set of expected values. The test runs are timed to ensure that the database provides acceptable response times under the anticipated user loads. Additional scripts are generated to verify the system’s access controls, ensuring that each simulated user can only access authorized functionality.

In the event that a user experiences technical difficulty, direct 24-hour telephone access to the Center’s development team is provided via cell phone. Reported problems are logged and tracked on-line with the Center’s web-based issue tracking system to ensure that any discovered system problems are recorded and addressed. System change control is managed with a concurrent versioning system (CVS) server at the Center that enables the development team to document and retrieve all changes to system source code.

11.3.5 Disaster Planning

All servers at InCHOIR use battery-backed UPS’s to ensure a controlled shutdown in the event of a power disruption. All servers, UPS’s and auxiliary networking equipment are housed in enclosures that are raised at least three inches from the floor, and they are protected from overhead leaks by a secondary waterproof covering between the enclosures and the ceiling. Both on and off-site data backups will be maintained according to a schedule specified in the Data Center’s operations manual. All data entry locations will have blank hard copies of the forms, and manual alternatives will be provided for any automated functions such as patient

randomization or visit schedule display where a delay caused by a network disruption would adversely affect data collection activity.

11.3.6 HIPAA Compliance

No patient identifiers will be permanently stored in the study database. During the course of the trial the study coordinator at each site will maintain a hard-copy list of participating patients, their contact information, and their study identifiers. This list will be used by study coordinators to maintain patient follow-up and by trial monitors during monitoring visits to link actual patients to records in the study database in order to verify data entry with information from the patient chart. This paper copy containing patient contact information will be destroyed at the end of the trial as part of the center closeout procedure.

Some diagnostic images received at the DCC may have been routinely marked with patient identifies by the local hospital imaging system. All images received at the DCC will be inspected for identifiers, and they will be manually removed with digital photo editing software (Adobe Photoshop) prior to inclusion in the study database.

11.3.7 Data Access Control for Blinded Investigators

During the course of the trial no investigators other than unblinded users who are designated as system managers will have direct access to the study database. Other investigators who need access to trial data may log on to the web system to view a version of the database that is filtered according to the access control rules that apply to their roles in the study. More complex database queries will be forwarded through the Operations Committee to the statistical group for manual processing, in which case the study statisticians will be responsible for ensuring that blinding and study confidentiality are maintained. A data dictionary will be developed prior to the commencement of the trial. This document will specify the metadata for all variables in the trial will define data access restrictions for each user role.

11.3.8 Management of Digital Images

Representative diagnostic images will be collected for each patient at baseline, at 5 years and/or at study completion, if not previously performed. The baseline images will be interpreted by the site credentialed radiologist and a de-identified CD will be sent to the DCC along with a copy of the radiologist's report. Images at 5 years and/or at study completion will be sent as a de-identified CD and radiologist report to the DCC.

In the event that a credentialed radiologist is not present during enrollment at the clinical site the following procedure will be followed:

After logging in to the web database system, the local study coordinator will upload these images in digital form to the DCC via an SSL-encrypted connection. The system will automatically catalog each image with a timestamp and patient study ID using software that we have previously developed for a separate X-ray data management project.

To upload the images the coordinator at each site will log in to the DCC web site, open the patient screening form (or new procedure page for an existing patient) and click the "upload"

button. The browser will open a file selector window that will enable the user to select image files to be uploaded. After selecting the files and clicking the “OK” button the files will be automatically uploaded to the DCC server where the files will be cataloged. The server will respond with a page that displays the images along with the patient identifier so that the user can verify that the correct images have been uploaded.

In order to gather information about the technical capabilities and experience at each clinical center and to confirm the feasibility of using web-based image management at our anticipated set of centers we have developed an ARUBA clinical center registration site (see screenshot below). This web site gathers data about AVM experience and displays a set of 14 images for evaluation. We have also used a network diagnostic tool (tracert) to measure network latency across the Internet links to some of the more distant center locations from the DCC, and we have found average delays that would barely be noticeable to the user (92 milliseconds round trip to Paris, 252 ms to Melbourne and 280 ms to Perth).

11.3.9 Management of Faxed Source Data

In order to reduce monitoring costs and to facilitate the rapid verification of screening data during patient enrollment, source documents will be faxed to a fax server at the DCC using the following procedure:

At the clinical site the data coordinator affixes a printed patient identifier label to each page and faxes the stack of pages to the DCC fax server. The server automatically converts the faxed pages to both html format for web viewing and PDF format. The web pages are integrated to the ARUBA data management server so that the appropriate site monitor anywhere in the world can view the documents, move them into the appropriate patient folder and check them off on the patient’s source document checklist.

11.4 Data Monitoring and Quality Assurance

Through the combination of our web-based, instantaneous electronic validation, the DCC’s daily visual cross-validation of the data for complex errors, and regular on-site monitoring, the quality and completeness of the data will be reflective of the state of the art in clinical trials.

Both the European and US DCCs will conduct monitoring of source documents via fax at all enrolling ARUBA sites and will conduct at least one onsite monitoring visit per year over the course of the study at 100% of clinical sites (with repeat visits to sites where performance is a concern). Monitoring of European study sites will be assured by the European Coordinating Center (Paris). The primary objectives of the DCC during the on-site visits are to educate, support and solve problems. The monitors will discuss the protocol in detail and identify and clarify any areas of weakness. At the start of the trial, the monitors will conduct a tutorial on the web-based data entry system. The coordinators will practice entering data so that the monitors can confirm that the coordinators are proficient in all aspects of data entry, query response, and communication with the DCC. They will audit the overall quality and completeness of the data, examine source documents, interview investigators and coordinators, and confirm that the clinical center has complied with the requirements of the protocol. The monitors will verify that all adverse events were documented in the correct format, and are consistent with protocol definition.

The monitors will review the source documents as needed, to determine whether the data reported in the Web-based system are complete and accurate. Source documents are defined as medical charts, associated reports and records including initial hospital admission report, operative procedure record, anesthesia record, discharge and re-admission reports, consult notes, diffusion-weighted MRI reports, radiology reports, lab reports, clinic records, and other study related notes. Copies of all of these records must be kept in a binder with the patient's study code.

The monitors will confirm that the regulatory binder is complete and that all associated documents are up to date. The regulatory binder should include the protocol and informed consent (all revisions), IRB approvals for all of the above documents, IRB correspondence, case report forms, investigator's agreements, IRB roster, signature verification page, investigators', coordinator's and credentialed radiologist's curriculum vitae, monitor site visit log, telephone contact log, and correspondence with the DCC.

Scheduling monitoring visits will be a function of patient enrollment, site status and other commitments. The DCC will notify the site in writing at least three weeks prior to a scheduled visit. The investigators must be available to meet with the monitors. Although notification of the visits will include the list of patients scheduled to be reviewed, the monitors reserve the right to review additional ARUBA patients.

If a problem is identified during the visit (i.e., poor communication with the DCC, inadequate or insufficient staff to conduct the study, missing study documents) the monitor will assist the site in resolving the issues. Some issues may require input from the Operations Committee, Steering Committee or one of the principal investigators.

The focus of the visit/electronic monitoring will be on source document review and confirmation of adverse events. The monitor will verify the following variables for all patients: initials, date of birth, sex, signed informed consent, eligibility criteria, date of randomization, treatment assignment, adverse events, and endpoints including mortality, stroke, and completeness of the functional health status tools and quality of life questionnaires.

11.5 Adverse Experience Reporting

11.5.1 Adverse Event

The endpoints for safety will be reported as the frequencies of occurrence of each adverse event, and time to each event. Safety data will be collected throughout this study and the incidence of each event will be computed along with associated 95% confidence intervals.

An adverse event (AE) is defined as any undesirable clinical occurrence in a study patient.. Any condition that was recorded as pre-existing is not an AE unless there is a change in the nature, severity, or degree of the condition.

Detailed information regarding adverse events will be recorded at the time of their occurrence. Investigators will be asked to classify the seriousness of the event and if the event was related to the index AVM or interventional therapy.

11.5.2 Serious Adverse Event

A serious adverse event (by FDA regulation) is one that results in a fatality; is life-threatening; results in permanent disability; requires hospitalization or prolongs a hospital stay. If an event is classified as 'Other' it must meet the FDA definition of serious.

11.5.3 Event Reporting

Serious, protocol defined or interventional therapy related adverse events must be reported to the DCC and captured in the electronic data capture system within 72 hours of knowledge of the event.

Any serious 'Other' adverse events must be reported directly to the individual IRB within 10 working days of knowledge of the event, or as dictated by the individual IRB.

Non-serious events must be entered into the electronic data capture system within 72 hours.

The cause of death, the categorization, and the severity of all adverse events will be determined by the site investigators, then monitored by the DCC, and finally adjudicated by an independent Morbidity and Mortality Committee. The charge of this committee is to ensure that the categorization of adverse events and mortality meet the protocol definitions. Published mortality and adverse event data will be based on the data adjudicated by the Morbidity and Mortality Committee.

12. HUMAN SUBJECTS

12.1 Institutional Review Board (IRB) Review and Informed Consent

The Human Subjects Committee/Institutional Review Board of each clinical center will approve the informed consent form. A copy of the letter of approval from the IRB and a copy of the consent form will be filed with the Project Office and reviewed and approved by the NINDS before a clinical center will be allowed to initiate enrollment. The informed consent will include the objectives of the study, a description of the screening process, the potential risks and benefits, the cost to the patient, alternatives to participation and liabilities of the particular participating center. It will be made clear to patients that both treatment options are available, even if they decline to participate. The European CCC will ensure certified translation of informed consent forms and IRB-related materials as needed for participating sites in 8 different language zones. The signed informed consent will be faxed to the DCC prior to randomization. A copy of the consent form will be given to the patient (or legal guardian) and this fact will be documented in the subject's record.

12.2 Potential Risks

Medical management and interventional therapy are both well-recognized clinical options, and are not considered to be experimental in nature. Whether one approach offers a relative advantage over the other in terms of risk is not known, and is expected to emerge from this trial.

12.3 Safety and Confidentiality

Patients will undergo either medical management or interventional therapy. All precautions to avoid untoward effects of interventional therapy will be taken, and defined in treatment guidelines. Patients will undergo regular follow-up and receive care for any adverse events by their own physicians.

We will follow rigorous procedures to protect patients' and clinicians' confidentiality. Access to identifying information will be limited to those whose project roles demand it, and only for the period of time in which they need it. Physical safeguards, such as locked file cabinets, will be used to protect the data and prevent unauthorized access. We will use the following additional measures: identifying information will be physically separated from data collection instruments and only code numbers will identify individual participants or facilities; contact sheets with identifying information will be stored in locked cabinets; access to the database will be limited to project staff. Patient contact information will only be accessible to the clinical site research nurse for a given site, the data monitor and the database system manager.

12.4 Study Modification/Discontinuation

12.4.1 Data and Safety Monitoring Board (DSMB)

To meet the trial's ethical responsibility to its patients, an independent group will monitor the results during the trial. This board will have no formal involvement with the patients or the investigators. The clinical centers will have no contact with the DSMB. The DSMB will act in a senior advisory capacity to the NIH on data matters throughout the duration of the study. The DSMB will communicate directly only with the NIH. In addition, it will periodically review study results by treatment group and evaluate the treatment for beneficial and adverse effects. NINDS will appoint members of the Data and Safety Monitoring Board. The Board will ideally include neurologists, neurosurgeons, interventional radiologists, and statisticians. Board meetings will be attended by senior representatives of the CCC, DCC and NIH, as well as by the Chair of the Steering Committee. No voting member of the Board may participate in the study as an investigator.

Specific functions of the DSMB are: (a) to review the protocol before it is implemented, and any subsequent changes; (b) to examine outcome and adverse experience data by treatment group; (c) to make recommendations to the NINDS on any proposed extension of the study or study arm because of beneficial or harmful effects; (d) to monitor the performance of the clinical centers and the DCC; (e) to advise the NINDS about policies related to confidentiality and conflict of interest. The members of the DSMB will review the interim analyses of the primary endpoint, and adverse event data, as adjudicated by the Morbidity and Mortality Committee. The DSMB will approve stopping guidelines developed by investigators, analyze the interim results and recommend an early termination because of safety issues or because of evidence of efficacy, and will also develop guidelines for recommending that the trial be extended if the assumptions that

went into the power calculations are found to be incorrect. We anticipate 2 meetings of the DSMB annually (either by conference call or in person) to review interim analyses of the data, and one closing meeting to review the final results.

12.4.2 Event Adjudication Committee

The Event Adjudication Committee will classify the cause of mortality for all cases and review and classify all adverse events. The individuals who serve on the committee will not be investigators in the trial.

12.4.3 Safety Monitor

The safety monitor is an independent medical advisor to the Data and Safety Monitoring Board of the NIH concerned with the safety of patients enrolled in the ARUBA study. The safety monitor will communicate directly with the DCC regarding patient outcomes and adverse events. The safety monitor will treat all study data as confidential and subscribe to the protocol defined confidentiality guidelines.

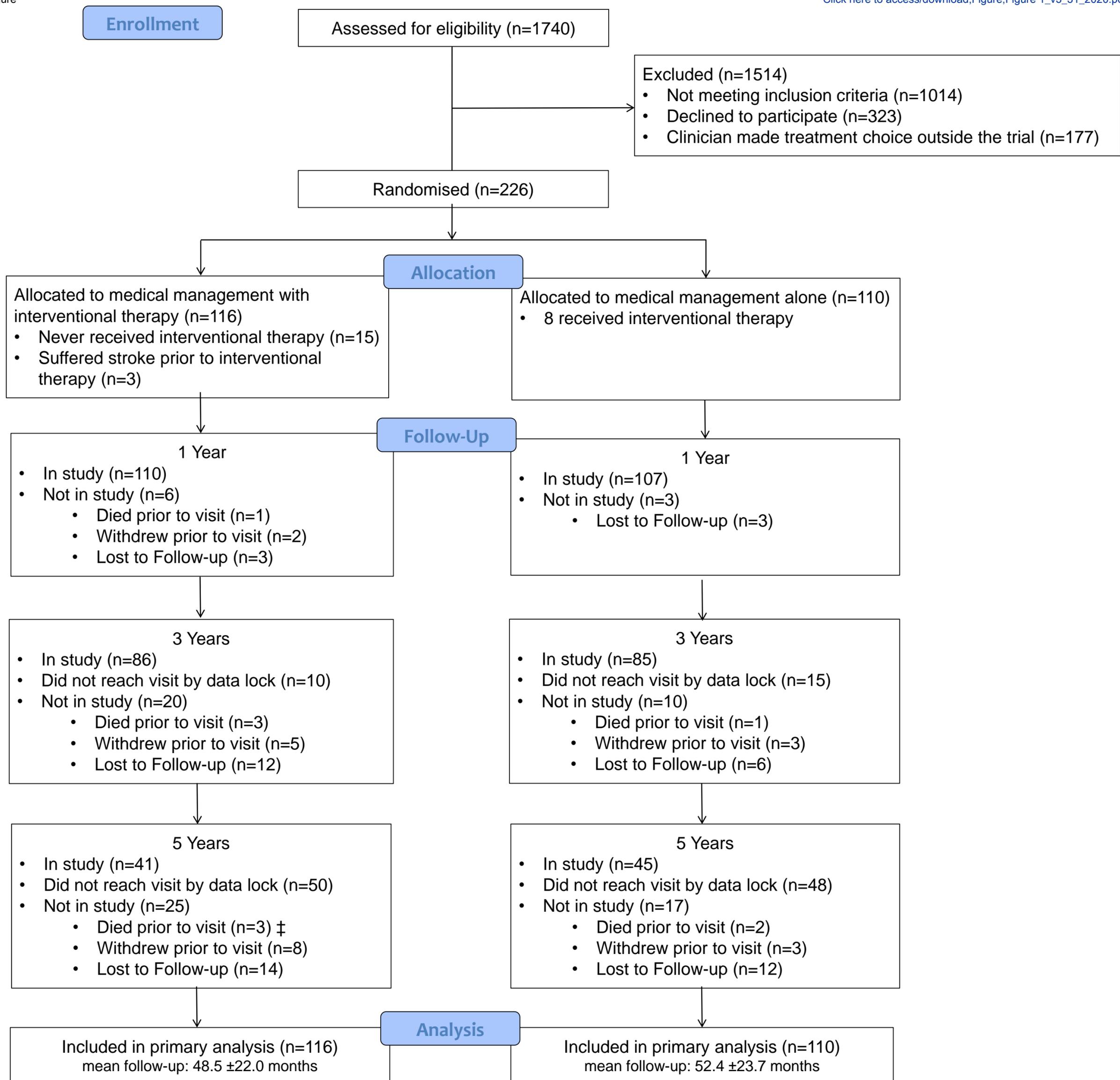
The safety monitor will receive data on all serious and protocol defined adverse events and primary endpoint events on an occurrence basis. Every two months the safety monitor will receive a summary report of all serious adverse events and primary endpoint events. If necessary, the safety monitor may request more information from the DCC. After review of the data summary provided by the DCC, the safety monitor will provide the DSMB with an interim report and will contact the chair of the DSMB directly in the event of any safety concerns.

13. PUBLICATION OF RESEARCH FINDINGS

Publication of the results of this trial will be governed by the policies and procedures developed by the Executive Committee. Any presentation, abstract, or manuscript will be made available for review by the CCC, DCC, and the NINDS prior to submission.

14. REFERENCES

Please refer to Manual of Procedures for complete Reference list.



‡ One additional patient in the IT group died ~1 month after the 5 year visit

