

Managing Unruptured Aneurysms: The Ethical Solution to the Dilemma

Jean Raymond

Can. J. Neurol. Sci. 2009; 36: 138-142

“A man cannot learn what he thinks he already knows”
Epictetus

The management of unruptured intracranial aneurysms (UIAs) remains controversial. Should they be treated or left alone? Perhaps unsurprisingly, a caricature of current management preferences would propose that surgeons favor clipping, interventionists coiling, and neurologists observation of incidental aneurysms. How is such a diversity of practices possible? It is the bold ambition of this manuscript to claim a) that we know very little about UIAs; the unwarranted beliefs at the source of our disagreement are based on very thin error-prone studies; b) that convergence on a basic requirement of a good practice is possible, namely the demonstration that the proposed practice is truly beneficial to patients; and c) that solutions to our dilemma exist if we are ready to acknowledge that there is a clinical reality with such brutal facts that some actions lead to more deaths and disabilities than others. What we need is not more disputes but hard work. Our first duty will consist in making room for the uncertainty that will reveal essential to look at the problem with new eyes. Then we will propose the only method that can provide justification for preventive medical actions: the clinical trial. Finally we will discuss the most appropriate trial design.

1. The wrong premise

For more than a decade we were told that ‘The management of UIAs requires accurate knowledge of the natural history (NH) of these lesions’ and comparisons with ‘precise definition of the risks of repairing them’¹. This premise must now be critically appraised. First one must notice that the efficacy of treatment is not even mentioned in this argument, and if the efficacy of surgical clipping has rarely been questioned, the efficacy of coiling is more problematic. A recent report on the surgical group of International Study on Unruptured Intracranial Aneurysms (ISUIA) showed delayed hemorrhages occurring at a rate of 0.09% per year. It was 0.3% for the endovascular group². Second, no one knows what could be the appropriate method to assess this NH. There is no indication here of how groups observed and groups treated, and how two different ‘things’ as immediate operative morbidity and life time hemorrhagic risks could reliably be compared. Bias being the devil of clinical research, we can guess that selection for observation or for treatment will necessarily affect the outcome and in the absence of randomization, one may end up comparing the surgical risk of patients selected for clipping with the NH of patients we do not

want to treat, surely an invalid procedure, no matter how much you toil with statistical adjustments. Finally, assuming that the observed group would be a reliable indicator of the NH of the patients you select to treat, the argument would be acceptable in one direction only: if the risk of repairing UIAs was, on the day of surgery, more than a life time of hemorrhagic risk, then one would conclude that observation would be better; but this calculation (life-time hemorrhagic risk minus treatment risk), which may suffice to support a recommendation not to resort to a risky preventive treatment by applying a ‘*primum non nocere*’ maxim, is clearly insufficient to justify treatment. This would require much more rigorous evidence of benefit than a speculative actuarial calculation that the life time risks of hemorrhage exceeded the immediate surgical risks. Unfortunately the argument, originally introduced to argue for conservative management, has become a reason for unduly treating virtually anybody, since any incidence of hemorrhage, no matter how small the percentage/year, could exceed purported surgical risks once it was multiplied by decades of life expectancy.

If life time estimates are needed to make clinical decisions, then the precision required of our assessment of the NH becomes extreme, otherwise error will be multiplied beyond control. If we also wish to distinguish ‘kinds of aneurysms’ to feed algorithms with different risk estimates according to size, location, previous history, sex, age and their combinations, numbers just become astronomical. Nearly 10,000 observations are necessary to a single comparison of patients with a 1% risks with others at 0.5% risk. Can our NH data offer this type of precision?

2. Building pseudo-certainties on shaky ground

From 1966 to 2005, our literature reported on 4705 patients with UIAs; only 2517 have been followed prospectively³. There was no Randomized Controlled Trial (RCT). A total of 224 hemorrhagic events were recorded, for a risk estimated at 0.6 or

From the Interventional Neuroradiology Research Unit, Department of Radiology, Centre hospitalier de l'Université de Montréal, Notre-Dame Hospital, Montreal, Canada.

RECEIVED SEPTEMBER 5, 2008. FINAL REVISIONS SUBMITTED OCTOBER 22, 2008.
Correspondence to: Jean Raymond, CHUM – Notre-Dame Hospital, Interventional Neuroradiology (NRI), 1560 Sherbrooke east, Pavillon Simard, room Z12909, Montreal, Quebec, H2L 4M1, Canada.

1.3%/year (0.5-1.8%). Such a small number of events necessarily limits the number of analyses that can be extracted. In contrast, Cowan⁴ estimates that 8334 patients with UIAs were treated in 2003 in the USA alone, where approximately 30,000 SAHs are expected each year. Wermer³ could identify six risk factors (RFs) for hemorrhage (age, sex, location, size, symptoms and countries), all with modest relative risks (RRs) (between 1.6 and 2.5) and borderline significance (except for symptoms). Yet RRs must be an order of magnitude larger if they are to be of any clinical value. Having performed 44 analyses, two of these RFs could have been found by chance alone. In addition, for at least four RFs (age, location, size, symptoms), treatment risks are also increased, with RRs of the same magnitude¹, forcing any prudent physician to remain agnostic regarding who should be treated or observed. All studies of this review suffer from innumerable methodological problems, including: a) systematic selection bias (in ISUIA, treated and observed cases differed for all pertinent RFs); b) excessive attrition bias (33% of the observed group ended-up operated and the ten-year unrelated mortality was close to 50% in ISUIA II²); c) incommensurable follow-up periods and intractable differences in outcome adjudication. Methodological flaws are not just violations of theory. These impact on the reliability and repeatability of findings. How risky are extrapolations of these results to patients encountered in clinical practice? For example, the initial ISUIA proposed that patients with a prior history of subarachnoid hemorrhage (SAH) had a 10-fold increased risk of rupture. Yet these same patients had a significantly ($p < 0.01$) lower risk of rupture in ISUIA II². Can we trust this sort of changing data, and can we multiply these potentially misleading numbers by years of life expectancy?

3. How “natural” is an imaging finding?

If on the one hand UIAs are structural defects that are predestined to rupture, as their name implies, it would seem helpful to attempt to determine their propensity for such a catastrophic event. On the other hand, if one is reminded that we are talking about an incidental finding in most cases, it becomes obvious that we cannot find here what we were looking for: the NH, if defined as a number of events in patients in whom the finding was not pertinent to their presentation, will heavily depend on the diagnoses that were indications for imaging studies, the impact of these diseases on the survival and follow-up of patients, the invention, availability, clinical use of and referral pattern to imaging machines, at least as much as on the presence of the UIA. This will be compounded by selection bias for observing patients with irrelevant UIAs, and for treating younger, healthier patients⁵. Aneurysms were of a very different ‘nature’ before the availability of non-invasive imaging. Only 10% of all aneurysms were unruptured, and most of them were symptomatic in 1966. Is the ‘nature’ of UIAs changing over time, or are we only witnessing the evolution of imaging?

4. Back to basic principles

The controversy ‘to treat or not to treat’ has been inflamed by recent efforts to find after the fact a size threshold that would assure reliable prediction of the future of individual patients, first 10mm, then 7mm in patients with anterior circulation aneurysms without a prior history of SAH, but only if posterior communicating aneurysms were excluded from the class of

carotid aneurysms etc... Unfortunately these post-hoc creations are typical of data torturing, a procedure that destroys any hope of seeing these findings confirmed in the future⁶. The quest for an elusive NH of UIAs was doomed to fail, but it had the merit of raising questions about the habitual recommendations of the surgical and endovascular communities. There was no need to triturate numbers to show a zero risk in certain patients, for the burden of the proof is the responsibility of those who prescribe preventive treatments. Entangled in controversies, we have lost sight of a most basic ethical requirement: while therapy only has an obligation of means, prevention has an obligation of results. Because it concerns healthy individuals, prevention can only be considered when risks of iatrogenia are low, and benefits proven by reliable studies. And the only reliable way to demonstrate the benefits of a treatment is a randomized comparison with deferral of treatment. Exceptions to this rule are very few: when the course of the disease is so uniform and the effects of an intervention so dramatic that RCTs are not necessary. Needless to say that UIAs do not fit this description. There is currently no evidence that UIAs should be treated preventively. The surgical or endovascular management of UIAs, in the face of uncertainty, can only be justified (now) by 1) suspending our current actions for patients in whom treatment is contemplated 2) giving to each patient a chance of being protected from a potential rupture (by being treated), and an equal chance of being exempted from potential complications of treatments that may ultimately be revealed inappropriate and (eventually) 3) by showing better outcomes in treated patients as compared to observed patients, the two groups being similar, for all known and unknown characteristics, except for random allocation of treatment.

5. The right approach

The risk of bleeding from UIAs is debated, but most series have reported a small annual risk, between 0-2%. Treatment can prevent ruptures but involves immediate risks. Furthermore, successful treatment does not eliminate all risks. Hence, the balance of the risks and benefits is uncertain. This calls for a direct comparison between treatment and observation, a randomized trial.

Randomization does not mean that we are bending the therapeutic obligation to current individuals to meet the scientific requirements that will provide knowledge to guide the treatment of future individuals. The research question concerns first and foremost our current patients, for whom no action has yet been proven beneficial. Thus randomization is not only a solution to the problem of bias; until we find out which option is best, it is a practical way of assuring the best possible outcome for each patient. Therefore, who should be recruited in our trial? The very patient in whom treatment is contemplated. Which treatment? The very treatment that is being offered to that patient. We are not dealing here with novel therapies. The research question is one regarding the usual management of a common condition. This calls for a ‘pragmatic trial’: a large, simple trial 1) with loose eligibility criteria based on uncertainty; 2) taking all comers; 3) retaining every admitted patient in the analysis; 4) proceeding with non-obstructive monitoring; 5) ascertaining a range of hard outcome events; 6) counting every event and charging it against intervention, looking for a practical answer to usual circumstances in real-world practices⁷. This is

the design of the Trial on Endovascular Aneurysm management (TEAM) trial.

6. The TEAM trial

TEAM is an international, randomized, multicenter, controlled trial comparing the combined mortality and morbidity (modified Rankin Scale >2) from intracranial hemorrhage or treatment in patients with UIAs treated by conservative management (or deferral of treatment for ten years or until definite indications are thought to have arisen) as compared to endovascular coiling⁸. The study will enroll 2002 patients equally divided between the two groups, a size sufficient to achieve 80% power at a 0.0167 significance to detect differences in 1) disease or treatment-related poor outcomes from 7-9% to 3-5% at ten years, as judged by an independent committee masked to treatment allocation; 2) overall mortality from 16% to 11%. The complete protocol can be found at www.TEAMstudy.org. If TEAM can eventually provide a justification for coiling UIAs, it does not address the role of surgical clipping. But a single trial cannot answer all questions.

7. Coiling and clipping are different treatments that raise different concerns

Clipping of UIAs was used long before coiling, but we are still waiting for proof that it is beneficial. While International Subarachnoid Aneurysm Trial (ISAT) has shown coiling to be preferable to clipping in patients with ruptured aneurysms eligible to both treatments, nothing allows us to extrapolate this result to UIAs. Some registries have shown a higher mortality with clipping, but this is not a valid comparison: more risky patients with UIAs may have been referred to surgery⁹. Most importantly for prevention, while the long term efficacy of coiling is questioned, the efficacy of clipping is more certain. The concern with clipping is rather in regard to initial morbidity, difficult to estimate reliably, in the range of 5 to 13%^{1,9}.

We need approximations regarding treatment and rupture risks, not to proceed with clinical decisions (the methodology is exploratory and the data too uncertain¹⁰), but to generate credible hypotheses for trials. Trial designs should take into account that clipping may be more risky initially but more effective, while coiling may be safer but less effective. Finally, ISUIA provides information on 'clinical judgement': treated and untreated, as well as coiled and clipped patients differed markedly, with $P < .0001$ for most alleged RFs, forbidding any comparison, but also pointing out that surgical or endovascular trials would not address the same 'types' of patients.

8. Various trial designs and their appropriateness

A randomized comparison between coiling and clipping is clearly inappropriate, for it would implicitly justify prevention, while the crucial question now is if any treatment is indicated.

The 'supreme trial' could theoretically consist in randomization to three groups (observation, coiling, clipping), with a ratio of 2-1-1, in order not to imply that we are 2/3 certain that treatment is better. This necessitates restricting entry to patients in whom three different options, with a range of immediate risks, degrees of invasiveness, and efficacy, are felt to be equally appropriate for a variety of patients by a diversity of

specialists not sharing the same set of beliefs. By definition the trial would address the intersection of three diverging categories of patients, leaving the majority outside the realm of a rigorous evaluation, and proscribing, at the end of a long trial, generalization of results to patients physicians want to clip, but not coil, or coil but not clip.

A trial comparing conservative management with 'Treatment' (endovascular, surgical, or both) could come in various flavours:

a) A predetermined set of criteria could be used to select endovascular or surgical options. However it is impossible to provide consensual criteria agreeable to a majority of experts of different backgrounds. In addition this procedure is like putting the cart in the front of the horse: what would be the credibility of deciding a priori, before the onset of the trial in whom this or that treatment is beneficial?

b) The decision regarding 'Treatment' could be left to the local investigators. We are aware many physicians are seduced by this design, so it must be examined carefully. By lumping two treatments, determination of sample size and observation period, monitoring, stopping rules, relevance and significance of results become impossible to understand. The 'meaning' of the trial and its conduct become ambiguous and murky. More importantly, we risk losing on all scenarios for two main reasons. One is statistical: by lumping two options, we attenuate the potential benefits of each and deprive them of the opportunity (power) to prove their respective value. The initial safety of coiling may be lost by the added morbidity of surgery; the long term efficacy of surgery may be spoiled by recurrences after coiling.

An illustration may clarify this point. The hypotheses of the TEAM trial are shown in the Table. Two thousand participants are equally allocated to observation versus coiling. Observed patients suffer no initial morbidity, but present a poor outcome from ruptures at a rate of 0.8%/year (1.2%/year with 70% morbidity/mortality)³, for a total of 76 at 10 years. The endovascular group has an initial poor outcome (mRS >2) of 3%¹¹ and a rate of yearly events of 0.1% (0-0.2%); this is the main test for coiling, whether it holds on the long run), for a total of

Table: Hypotheses for TEAM, a surgical trial, and a trial combining both treatments

	Observation	TEAM	Clipping	Combination
Sample	981	981	981	70/30%
Initial morbidity/mortality (M/M)	0	3% (1.9-4.1%)	8% (6.3-9.7%)	4.6% Endo: (1.7-4.3%) Surg: (4.9-11.1%)
Hemorrhages/yr M/M	0.8%	0.1%	0.01%	0.1%
Acceptable range	---	0.1-0.2%	0.01-0.1%	0.1-0.2%
Events 5 yrs	39	34	79	50
Events 10 yrs	76	39	80	54
<i>P value 10 yrs</i>	---	0.0005*	0.8	0.06
Events 15 yrs	111	44	81	59

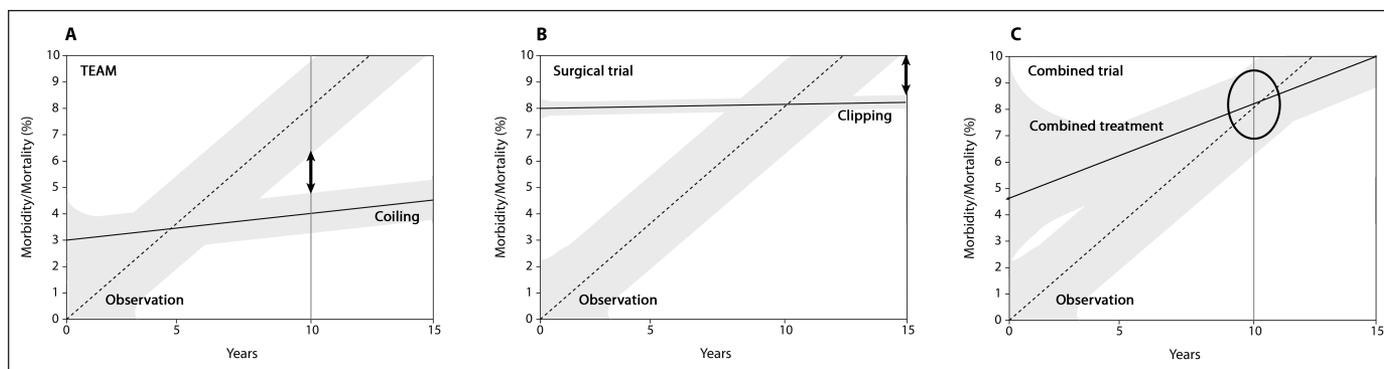


Figure: Artistic representation illustrating the primary outcome (Disease or treatment related morbidity and mortality; $mRS > 2$) as a function of time in three different trials. A) TEAM; B) surgical trial; C) a trial comparing 'treatment, surgical or endovascular, and conservative management.'

39 at 10 years, significantly less than the conservative group (Figure, A). If we imagine a surgical trial with an initial morbidity of 8% and a better efficacy (0.01% (0-0.1%)), curves cross later and a benefit can only be shown at 15 years (Figure, B). However, imagine an interim analysis at five years. The data monitoring committee (DMC) would be in a position to show that 5 versus 39 deficits caused by ruptures occurred in treated versus observed patients. This could be judged sufficient evidence that 'clipping provides >99% protection'. Projected over 10 more years, the curves would cross, if initial complications are < 10%. This evidence would come at a time when the M&M of treated patients would still exceed that of observed patients (83 versus 39; $P < 0.001$). Now let us examine combining treatments (Figure, C), assuming 30% clipping/70% coiling: we have lost the demonstration of a better outcome at 10 years ($P = .06$) and we have too small a surgical group ($n = 294$), with no matching controls, to conclude about the value of surgery.

Worse, a trial lumping clipping and coiling forbids any comparison between clipping and observation, or coiling and observation. We have lost the benefit of randomization. Results cannot be dissected and analysed separately, since the control group is no longer a randomized match of individuals with equal chances of being in one group or the other (unless stratification is used).

But the main reason for rejecting a trial comparing 'Treatment' and conservative management is logical: such a trial, according to the logic of medicine, can only show superiority of observation, the 'pure' arm of the trial, since a verdict such as '(coiling or clipping) is better cannot be used to justify either clipping or coiling. In medicine, a claim of the type 'treatments A or B are good' necessitates a proof that (A is good) AND a proof that (B is good), otherwise anything goes (if A is good, any treatment B, no matter how esoteric or dangerous, would be accepted). Therefore a combined trial would need to be powered to give an answer for clipping AND an answer for coiling. Someone may claim that a combined trial would be useful to convincingly show that conservative management is better than 'Treatment', whatever 'Treatment' means. However, besides the

fact that a trial constrained to provide a one-sided answer is not a trial, starting a trial with a hypothesis favouring conservative management is ethically inappropriate. The burden of the proof is on active treatment. We cannot tell our patients 'we believe this treatment is toxic, but let us prove it'. Like any trial, we need to say something like 'we believe this treatment is good, but we are uncertain. Immediate risks are involved; therefore we must prove that treatment is beneficial before we can recommend it'. The principle of prudence forces them to admit they are uncertain, that they should submit their beliefs to the verdict of experience, that as things may turn out they may be wrong, but the trial must be designed to allow a positive perspective to be confirmed or refuted by credible evidence.

9. We need two trials

If we accept the above principles we need: a) a sample size sufficient for each treatment to show its value with a reasonable power; b) for each treatment, a matched control group, since controls may differ substantially; c) a follow-up period tailored to each treatment, with provision for tolerating a higher initial morbidity if perspectives of a long-term efficacy are good; d) a research hypothesis that the proposed treatment is beneficial to participants. This in effect is planning two separate trials, TEAM and a surgical trial.

A rigorous examination of the literature shows there is little reliable knowledge that could guide clinical decisions. Recent publications have suggested that clipping or coiling of UIAs is rarely justified¹. Concerns regarding treatments are distinct, and deserve dedicated trials. Either we stop performing risky treatments of unknown benefit, or we do it with full disclosure of the uncertainty, within the context of valid trials.

DISCLOSURE

JR is the principal investigator of the TEAM trial. In such he is the recipient of a research grant from the CIHR Canadian Institutes of Health Research (MCT-80799) (<http://www.cihr-irsc.qc.ca>). Trial registration: Current controlled trials ISRCTN62758344 (www.controlled-trials.com).

REFERENCES

1. Wiebers DO, Whisnant JP, Huston J, Meissner I, Brown RD, Piepgras DG, et al. International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet*. 2003;362:103-10.
2. Torner JC, Brown RD, Meissner I, Piepgras DG, Huston J, Whisnant J, et al. Long-term follow-up of patients with unruptured intracranial aneurysms. Presented at the International Stroke Conference, February 20-22, 2008, New Orleans, USA. *Stroke*. 2008;39(2):535.
3. Wermer MJ, van der Schaaf IC, Algra A, Rinkel GJ. Risk of rupture of unruptured intracranial aneurysms in relation to patient and aneurysm characteristics: an updated meta-analysis. *Stroke*. 2007;38(4):1404-10.
4. Cowan JA, Ziewacz J, Dimick JB, Upchurch GR, Thompson BG. Use of endovascular coil embolization and surgical clip occlusion for cerebral artery aneurysms. *J Neurosurg*. 2007;107(3):530-5.
5. Britz GW, Salem L, Newell DW, Eskridge J, Flum DR. Impact of surgical clipping on survival in unruptured and ruptured cerebral aneurysms: a population-based study. *Stroke*. 2004;35(6):1399-403.
6. Mills JS. Data torturing. *N Engl J Med*. 1993;329(16):1196-9.
7. Haynes BR, Sackett DL, Guyatt GH, Tugwell P. *Clinical epidemiology: how to do clinical practice research*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2006.
8. Raymond J, Molyneux AJ, Fox AJ, Johnston SC, Collet JP, Rouleau I. The TEAM trial: safety and efficacy of endovascular treatment of unruptured intracranial aneurysms in the prevention of aneurysmal hemorrhages: a randomized comparison with indefinite deferral of treatment in 2002 patients followed for ten years. *Trials*. 2008;9(1):43-53.
9. Higashida RT, Lahue BJ, Torbey MT, Hopkins LN, Leip E, Hanley DF. Treatment of unruptured intracranial aneurysms: a nationwide assessment of effectiveness. *AJNR Am J Neuroradiol*. 2007;28(1):146-51.
10. Lee T, Baytton M, Sciacca R, Mohr JP, Pile-Spellman J. Aggregate analysis of the literature for unruptured intracranial aneurysm treatment. *AJNR Am J Neuroradiol*. 2005 Sep;26(8):1902-8.
11. Pierot L, Spelle L, Vitry F, ATENA Investigators. Immediate clinical outcome of patients harboring unruptured intracranial aneurysms treated by endovascular approach: results of the ATENA Study. *Stroke*. 2008;39(9):2497-504.