Clip Versus Coil


The International Subarachnoid Aneurysm Trial (ISAT) was a turning point in modern neurosurgical history. Despite some weaknesses that inspired many critical editorials and letters, the trial had a significant impact on practices, with a larger proportion of patients being treated with coil embolization after than before results were published. This is the normal order of things: re-treated with coil embolization after than before results on practices, with a larger proportion of patients being editorials and letters, the trial had a significant impact in the primary clinical outcome between patients treated with clip ligation and those treated with coil embolization were too large to continue with the uncertainty principle that justified randomized allocation of treatment. In other words, we could not keep on claiming that according to the reliable evidence available at that point it was still uncertain which treatment leads to better outcomes. All we have now is a generalization, not necessarily applicable to any individual: when coil embolization and clip ligation are considered valid options in a particular patient, coil embolization in general leads to better outcomes at 1 year. Now was this verified for all patient, aneurysm, group that was responsible for guiding us into the safe vicinity of the island of the Sirens (the land of speculation), all sailors having their ears unplugged. "No seaman ever sailed his black ship past this place without listening to the sweet voice that flow [sic] from our lips, and none that listened has not been delighted and gone on a wiser man … for we have foreknowledge of all that is going to happen on this fruitful earth" (The SIRENS to Odysseus, Homer, Odyssey 12.184).6

In this recent article, not only are Mitchell and colleagues proposing to look at subgroup findings that were not prespecified, with a borderline interaction test (p = 0.04), an analysis that by all standards should be considered exploratory, but they also multiply this potentially misleading finding by arbitrary and uncontrolled numbers taken from various other individuals. The list of unverified assumptions that have been included in this computation and the multiplicity of arbitrary choices (some age-specific, others not) that were involved have only been partially covered in the discussion. Why select age-specific relative risk for results at 1 year, but age-nonspecific rebleed rates? Can authors assume that younger patients will be followed up and possibly retreated in a similar fashion to older patients? Why project 7 “excess” rebleeds of the coil-treated group, but not the 12 “excess unrelated” deaths of the surgical group over the next 40–60 years of life expectancy? The authors chose reasonable assumptions, but a number of equally reasonable alternatives could lead to a variety of conclusions that contradict each other. Statements such as “the large number of possible combinations leads to graphs that are intractably complex and results that are correspondingly inaccessible” give a hint of how many attempts at modeling were explored until the authors could find a graph that would fit their preconceived opinion. We cannot even imagine the level of precision, and the consequent enormity of the sample sizes, that would have been necessary to justify such extrapolations in a preplanned controlled manner, to restrict the risks of erroneous conclusions to acceptable boundaries.

The resulting graph (Fig. 2) has been interpreted as showing that “the advantage of coil embolization cannot be assumed for patients < 40 years old,” a statement that is in itself quite bland. The absence of evidence of a benefit for younger patients is neither, however, “evidence of absence;” nor evidence in favor of clip ligation, especially when one keeps in mind that the confidence intervals never crossed the 0 line to favor clipping.

Readers of the Journal of Neurosurgery must be warned against the dangers of post-hoc explorations, unplanned subgroup findings, and mathematical extrapolations. This type of finding should always be viewed with caution, and cannot be used to guide clinical decisions.2 We hope the neurosurgical community will be as wise as Ulysses, and stay tied to the mast, despite the Sirens’ claims that one would be wiser to listen to them to have “foreknowledge of all that is going to happen.” We must be realistic regarding what kind of knowledge we can achieve, especially about the future and fate of individuals: “Reliable assessment of moderate effects on major
Neurosurgical forum

outcomes, with no data-dependent emphasis on specific parts of the overall evidence...is all that can realistically be expected for most conditions.” At best, unplanned uncontrolled subgroup findings can only be used as the hypothesis for another trial. Until such a trial is performed, “the answer to a RCT that does not confirm one’s beliefs is not the conduct of several sub-analyses until one can see what one believes. Rather the answer is to re-examine one’s beliefs carefully.”

Jean Raymond, M.D.
François Guilbert, M.D.
Alain Will, M.D.
Daniel Roy, M.D.
Université de Montréal, CHUM Notre-Dame Hospital
Montreal, Canada

References


Response: Dr. Raymond and colleagues discuss some of the limitations of data interpretation and extrapolation of results from the randomized trial to other patients. The majority of points they raise were addressed by ourselves in the discussion of the paper. They raise 2 points we did not mention. These are:

1) Why did we not use age-specific rebleed rates? The reason is that no age-dependent trend can be determined from the sparse rebleed data available in the public domain. If and when such data become available it would naturally be sensible to include them in models such as ours.

2) Dr. Raymond and colleagues accuse us of selecting one model from many because of the result it gave rather than because it had the most reasonable assumptions. In their words: “Statements such as ‘the large number of possible combinations leads to graphs that are intractably complex and results that are correspondingly inaccessible’ give a hint of how many attempts at modelling were explored until the authors could find a graph that would fit their preconceived opinion.”

This is denied.

The decision to limit the number of graphs to 6 had no bearing on the modelling assumptions. When all combinations of post–coil embolization and post–clip ligation rebleed rates are included, the area covered by the modelling functions is the same, but instead of being occupied by 6 distinct lines it is occupied by 36 lines that cross over each other at shallow angles making visual interpretation difficult.

The intention when developing the model was to help decide how long follow-up should be continued by assessing the sensitivity of the result of the trial to rebleed rates and thereby informing how much precision of rebleed rate definition was required to determine the long-term durability of the result. All parameters were chosen because they were felt to be the most representative, even-handed estimates available. The result was not expected and its sensitivity to parameters taken from outside the ISAT cohort was tested and found to be low, as described in the paper.

A randomized controlled trial is an experiment interpreted with a mathematical model that approximates a clinical issue with a simple yes or no question: Is the hypothesis true or false? In many drug trials this is a reasonable approximation. In most surgical situations, the ISAT included, it is not a fully satisfactory approximation and can only be regarded as a step toward defining the relationship between treatments and effects; nevertheless, randomized trials remain the most reliable form of surgical evidence. The interpretation of the ISAT, in common with that of many surgical trials, involves extrapolating from a reductive hypothesis to highly variable real-life situations. A further complication with respect to the ISAT is that because of the importance of as yet incompletely defined long-term efficacy, the final result will not be available for many years. It is anticipated that in the long run the question of what treatment to use for which aneurysm will be better defined by more complete data from the ISAT cohort and additional randomized data addressing issues such as subgroups and different methodological techniques. Until then we must make the best use of what evidence is available. That means extrapolating results such as those of ISAT from a particular population of patients treated over a particular historical period to patients presenting today. The less representative the ISAT is of a particular patient, the less reliable the extrapolation. We must find a compromise between on the one hand wide unreliable extrapolation and on the other narrow safe extrapolation leaving large numbers of critical decisions involving patients unsupported by randomized data. This compromise applies to our analysis. There are several assumptions that have been made, comparisons of nonidentical populations, and dependence on ISAT subgroup analysis, all of which mean that when compared with the primary ISAT result, these results are more specific in terms of patient age and life expectancy at the cost of being less reliable. They suggest the average benefit of coil embolization includes benefit that is better than the