Letters to the editor

Neurosurgical forum

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: retreated 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 patients, aneurysm, and center characteristics, such as sex, age, race, location, size, experience? Of course not. By chance phenomena and sample diversity, this could never be achieved, and one has to expect heterogeneous results, some apparently increasing, others decreasing treatment effects, when dissecting the data in so many post-hoc categories. Planning for a study that would be powered to provide evidence for so many “kinds of patients” would lead to eternal trials, the evidence becoming non-convincing each time the data are split into other interesting subgroups.

There is a multitude of rules and norms that regiment clinical trials. An important one is to prespecify the research question with precision, defining the primary outcome, the error rates, the sample size, and any subgroup we wish to study, in order to minimize the risks of jumping to wrong conclusions. This is not theoretical; medical history is replete with examples of misguided guidelines driven by poorly designed or wrongly interpreted trials. Norms exist to put some restraint to our irresistible urge to project beyond the data, and somehow pretend we have more knowledge than can be justified by evidence.

It is with dismay that we now witness the leading group that was responsible for guiding us into the safe waters of evidential medicine propose that we sail in the vicinity of the island of the Sirens (the land of speculations), 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). 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. 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
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 WIEIL, 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 inaccen-