Review of: "On the Need for Better Information from Randomized Clinical Trials in Oncology"

Ian Tannock[1]

1 University Health Network

Potential competing interests: No potential competing interests to declare.

The premise is valid that HR is a poorly understood concept, often misinterpreted by oncologists and even more by their patients, and hyped to suggest much greater effects of experimental treatments than actually occur. Measures that are more understandable, such as absolute risk or difference in probability of survival at given times, are needed, as stated by the authors. But other authors have commented on these problems, and they should refer to papers by others that have pointed out such needs.

Summarizing the reporting of 5 recent RCTs published in NEJM is reasonable, but it is not sufficient to simply summarize their results in a different way to interpret their hazard ratios. The hypothesis set by the authors is that many trials are reported poorly and are difficult to interpret, so for each trial, I expect constructive criticism of the quality of reporting. For each of them, how easy is it to extract from a quick reading of the abstract and main results data that are easy to interpret? Does the trial have other biases (suboptimal controls? Informative censoring, etc.)? Are the results presented without bias, or are they hyped?

A concern is inaccuracy in transposing results of these trials. There are several errors in the paper. Examples are: In trial 1, where 372 received everolimus, not 272; The report of the trial states "*257 participants (68.7%) in the belzutifan group and 262 (70.4%) in the everolimus group had had disease progression or had died."* The authors give numbers of 289 and 276 (perhaps derived from later analysis?), and it is unclear which groups the numbers refer to. And there is "a difference," not "an increase," in relative and absolute risk. Other errors in the description of trial 1 include a stated RR of progression of 6% with a CI of (0.960 – 1.129), so the CI doesn't include the mean? Other errors are present in the report of trial 2, where it is stated that there were 143 PFS events among 57 patients and 73 events among 63 patients, and the HR for death is given as 0.81 with a CI of 0.42-0.56, all of which is nonsensical. I have not checked for errors in the descriptions of the other trials, but the authors need to check all the data that they have transposed from the trial reports.

A minor concern is the failure to show numerical data to a meaningful number of decimal places. In trial 1, the authors state a CI of (0.960 – 1.129) with p=0.3272. If you give a number as 1.129, it implies that you believe the real number is closer to that than to 1.128 or to 1.130, which is not true. All the numbers should be truncated to no more than 2 decimal places: e.g., this CI should be 0.96-1.13, p=0.33. That should be consistent throughout the paper.

While the idea behind this paper is sound, in my opinion, it requires substantial work both to ensure its accuracy and, more important, to make it a critical appraisal of the reporting of influential RCTs, and not just a reformatting of their data.

Ian Tannock MD, PhD

Princess Margaret Cancer Centre & University of Toronto.

ian.tannock@uhn.ca