Reviewer's report

Title: A Perfect Correlate Does Not A Surrogate Make

Version: 1 Date: 22 August 2003

Reviewer: Jeremy M Taylor

Reviewer's report:

General
The purpose of this article is educational, it is designed to reiterate that there are significant caveats associated with the use of surrogate endpoints in clinical trials. These caveats have been written about by many other authors, so it doesn’t represent new ideas or new research. The one novel aspect of the paper is the simple graphical representation, which will make it understandable to a very wide audience. This appears to be the aim of the paper.

I have comments in three areas (i) terminology (ii) balanced presentation of the plausibility of figure 1 and (iii) missing reference.

Terminology. Many researchers have a different concept of what is meant by the words “surrogate endpoint”. To some it means any biomarker or intermediate endpoint which you might consider as an endpoint in a clinical trial (randomized and non-randomized), to others you can only attach the label “surrogate” if it is actually valid to replace the true endpoint by the intermediate endpoint in a randomized trial. It seems to me that the use of the word “surrogate endpoint” needs additional adjectives, such as “potential” or “valid”. The use of the phrase “surrogate endpoint” in the first sentence is confusing, perhaps replace it by “biomarker” or “intermediate endpoint” or “potential surrogate endpoint”. There is a lot of haphazard use of the phrase “surrogate endpoint” in the literature. This article might be a good place to try to reduce this confusion.

A second terminology issue occurs on page 6, when the authors discuss using a surrogate endpoint to predict the true endpoint. When inference is going to be based on the true endpoint, and the surrogate endpoint is just additional data which is designed to help in this inference, then I thought the established terminology was “auxiliary variable”. The term surrogate is reserved for the situation where the true endpoint is being replaced by the intermediate endpoint. The authors might take this opportunity to clarify the terminology, or at the very least be precise as to what they mean by surrogate endpoint.

Plausibility of figure 1. While it is true that the scenario in figure 1 is possible, it is probably not something that happens very often. For a surrogate which is likely to be affected by the intervention and central to the “pathway” of the disease, it seems much more likely that the group E line will be below the group C line. Or if it is greater, it won’t be greater by very much, and tE would still be less than tC. While your figure requires the within group correlation to be +1, it also requires the between group correlation to be negative, how common is this? A little more balanced presentation of the plausibility of figure 1 is required.

Missing reference. An important missing reference is Wang & Taylor, Biostatistics, 2002. They have a figure very similar to figure 1, and develop these ideas in a precise way.

Specific Comments

1. You should clarify that you are talking about perfect within group correlation, because the pooled
correlation is not perfect.

2. Page 4, second paragraph of results. Both endpoints here are binary, so what do you mean by perfect correlation between surrogate and true here?

3. Page 5, line 1. It is a bit more complicated than just “different slope”. It is really that the experimental line has to be higher than the central line, and sufficiently higher than the order of tE and tC are switched from what is expected. A small difference in slope won’t cause the contradiction you are discussing.

4. Page 4, first paragraph. I don’t find this “bad” and “innocent” adenomas example very convincing. Think you should give the reader more details of what has to transpire, so they can judge whether they think it is plausible. Specifically, the intervention needs to cause a reduction in the total number (bad + innocent) of adenomas, yet an increase in the number of bad adenomas and thus a substantial decrease in the number of innocent adenomas. This would be a very strong interaction. Although clearly not impossible, is it very plausible?

5. Another example to consider is the finasteride trial, which reduced prostate cancer, but increased high grade prostate cancer. What will happen with the true survival endpoint here?

6. Page 2. The last sentence of the conclusion paragraph, doesn’t follow from the paper, since the paper hasn’t been about observational studies or early (single arm) studies. What conclusion are you talking about when you say “should not base conclusions on surrogate endpoints? The conclusion of the observational study?

Discretionary Revisions (which the author can choose to ignore)

Minor Compulsory Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

**Advice on publication:** Accept after minor compulsory revisions

**Level of interest:** A paper whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Declaration of competing interests:**

None