Published in final edited form as:

Clin Trials. 2015 April ; 12(2): 177. doi:10.1177/1740774515569854.

A Note on Masking in the SOX Trial

Vance W Berger

National Cancer Institute and University of Maryland Baltimore County, Biometry Research Group, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD 20850, (240) 276-7142 (voice)

Vance W Berger: vb78c@nih.gov

Editor

Houweling et al.¹ concluded that masking was successful in the SOX Trial. This conclusion does not appear to be warranted, or supported by the data they present. We note, first, that successful masking means that *all* relevant parties were masked, not only that some of them were, or even that most of them were. As illustrated in Chapter 3 of reference 2, it does not take much unmasking to completely invalidate a trial. Therefore, it is not the average we are after. Even if we can accept that those reporting uncertainty truly were uncertain, at least one pass through the data would still have to condition on these responses, and consider only the rest. Reporting the overall rates of correct responses as 10.4% for investigators, 17.8% for research coordinators, and 29.4% for patients is grossly misleading, given the large number who did not guess at all. Among those who *did* guess, these proportions of correct responses (taken from Table 2) are actually 75% (78/104), 69% (138/201), and 63% (200/318), respectively, with all three yielding two-sided p-values under 0.0001 for testing the null hypothesis $H_0: p=0.5$ (random guessing). In other words, the “guessing” was not random at all. Clearly, masking was not successful among those who ventured a guess. Ergo, masking was not successful.

But what about those who replied “don’t know”? We might be more inclined to believe the patients, who have no incentive to report this if they in fact do suspect that they know which treatment they received. This amounts to only 363 patients, as compared to 649 and 573 responses, respectively, from site investigators and research coordinators, much larger proportions even considering the larger denominators. These responses must be taken with a grain of salt, because investigators certainly recognize what will happen to their trial, and, indeed, their professional reputations, if it is found that the trial was substantially unmasked.³ Why, then, would they wish to incriminate the trial (and, indirectly, themselves) when they can instead simply report “don’t know”? Even the randomized response technique might not be sufficient to produce entirely honest answers to this question, and therefore the James and Bang indices are only partially informative. They can establish a lack of masking, but they cannot, by themselves, establish that masking was successful.

Fortunately, there is another test of the success of masking that does not require investigators to be forthcoming. Unmasking allows for selection bias, whereas successful masking eliminates it. Therefore, if selection bias is found, then the masking could not have been successful. The authors were remiss in not reporting the results of the Berger-Exner

test of selection bias,⁴ which improves upon the James and Bang indices by virtue of being objective, and independent of imperfect recall or unwillingness to divulge what was known. Of course, the Berger-Exner test is similar to the James and Bang indices in that it can establish unmasking, but not successful masking. Nevertheless, the effort to establish masking requires rigorous tests aimed at falsifying the claim. The claim becomes more plausible when multiple (and varied) challenges were met.⁵ Therefore, the Berger-Exner test must be a part of any serious effort aimed at establishing the success of the masking of a randomized trial.

References

1. Houweling AH, Shapiro S, Cohen JM, et al. Blinding strategies in the conduct and reporting of a randomized placebo-controlled device trial. *Clin Trials*. 2014; 11:547–552. [PubMed: 24902921]
2. Berger, VW. Selection bias and covariate imbalances in randomized clinical trials. Chichester: John Wiley & Sons; 2005.
3. Berger VW. On the insufficiency of reporting masking. *J Rehabil Med*. 2013; 45:221–222. [PubMed: 23303555]
4. Berger VW, Exner DV. Detecting selection bias in randomized clinical trials. *Control Clin Trials*. 1999; 20:319–327. [PubMed: 10440559]
5. Rosenbaum PR. Replicating effects and biases. *Am Stat*. 2001; 55:223–227.