Missed opportunity: Efficient use of baseline can increase power and reduce bias in psychiatric crossover trials

Farmer[1], C, Greenstein[1], D, and Zarate Jr.[1], C
[1]Experimental Therapeutics and Pathophysiology Branch, Intramural Research Program, NIMH

Methodological question: Can baseline measurements be better leveraged to increase power and reduce bias in psychiatric crossover trials?

Introduction: Partially owing to the advent of rapid-acting antidepressants, crossover trials are increasingly used to estimate treatment effects in psychiatry. It is well-understood that the primary strength of a crossover trial is that the patient serves as her own control. It is less appreciated that because the patient is randomized to sequence, rather than to treatments individually, crossover trials share characteristics with observational studies. For example, while covarying baseline (ANCOVA) is appropriate for randomized trials, caution must be exercised in crossover trials because the treatment effect estimate is subject to cross-level bias (Kenward & Roger, 2010; referred to herein as KR). Although KR described the simple solution of including the subject-average baseline score as an additional covariate, this has been adopted by few investigators. In this poster, we will formally evaluate the extent to which crossover trials in psychiatry have adopted the KR best-practice recommendation for handling period-specific baseline measurements. Second, we will offer a plain-language explanation of the KR recommendation, and, finally, we will use archival data from the NIMH Intramural Research Program to illustrate the recommended approach and compare it with more common approaches.

Methods: We performed a systematic review of the literature, including all RCTs in psychiatry published after 2010. To explain and illustrate the KR approach, we used linear mixed models to evaluate the treatment effects from three double-blind, randomized, controlled crossover trials of rapid-acting agents. We compared these results to those using the approaches most commonly observed in our literature review, including the exclusion of baseline altogether (Model 0), use of change scores (Model Δ), and the ANCOVA approach (Model ANCOVA). We will discuss differences in the treatment effect estimates and their corresponding precision. R and SAS syntax for the calculation of these effects and their inclusion in the model of analysis will be provided.

Results: Of the 664 initial Embase search results, 285 met the inclusion/exclusion criteria. Here we present the initial results, using studies with a primary Web of Science subject of Psychiatry (n=27). No study used the KR approach. 19% did not include baseline measurements (Model 0), 41% included only study baseline, and 37% had a period specific baseline. Of the 10 with period-specific baseline, six used change scores (Model Δ) and four used Model ANCOVA. Results of the mixed models indicated that relative to Model 0, the use of both Model Δ and Model ANCOVA *decreased* the precision of the treatment effects (equivalent to sample size reductions ranging between -3% and -17%). However, compared to both Model Δ and Model ANCOVA, the use of Model KR improved precision, equivalent to sample size increases ranging from +3% to +22%.

Conclusions: The inclusion of baseline as a covariate in parallel arm trials is intended to improve the precision of the treatment effect, but in crossover trials it may introduce cross-level bias. The addition of the average baseline value can eliminate this bias and yield more accurate estimates of the treatment effect and increased power.

Disclosures: The authors report no conflicts of interest for this work.

References: Kenward, M. & Roger, J. (2010). The use of baseline covariates in crossover studies. *Biostatistics, 11(1),* 1-17.