Incomplete Data in Clinical Studies: Analysis, Sensitivity, and Sensitivity Analysis Rejoinder

Incomplete Data in Clinical Studies: Analysis, Sensitivity, and Sensitivity Analysis

Rejoinder

Geert Molenberghs

Center for Statistics, Universiteit Hasselt, B-3590 Diepenbeek, Belgium

Biostatistical Centre, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium

I owe a sincere word of thanks to Dr Flyer, Professor Ibrahim, Professor Scharfstein, and Professor Tchetgen for their well-informed discussion and for raising some and underscoring other points. Let me come back to these in a systematic way.

1 Simple Methods

All discussants agree that we should be extremely careful with simplistic methods, and arguably abandon them completely. This includes, in particular, LOCF. In this respect, Dr Flyer makes a crucial distinction between the intention-to-treat (ITT) paradigm, which justifiably plays a key role in a large class of clinical trials, and LOCF, which is often erroneously coupled with ITT. Professor Scharfstein also emphasizes the role of ITT.

2 Parametric or Semi-parametric?

Professor Scharfstein and Professor Tchetgen supplement my own contribution by re-drawing the balance between parametric and semi-parametric methodology. It is my conviction that both parametric (likelihood-based or Bayesian) analyses and semi-parametric methods, in particular the seminal work by Robins, Rotnitkzy, Scharfstein, and a variety of co-workers, are both very valuable and find their place in the standard toolkit of the analyst. Professor Scharfstein states repeatedly and quite accurately what the relative merits of the various approaches are. Indeed, I believe it is scientifically difficult to justify an absolute choice for or against one of the strands. There is the bias–precision trade-off, for one. Also, while a semi-parametric approach makes less assumptions about the joint distribution of the outcomes, assumptions do come in for the specification of the weights in inverse-probability-weighting-based methods, also in view of the curse of dimensionality. Of course, doubly robust methods alleviate this by the mild requirement of its necessitating only one of the two dimension-reducing models to be correct: for the weights or for the prediction of the unobserved outcomes given the observed ones and predictor variables.

This notwithstanding, all methods make assumptions; they are all assuming certain scenarios, an important point also made by Dr Flyer. This is fine, as long as one is very explicit about such assumptions, allowing the reader to decide on whether s/he agrees or disagrees with such assumptions and scenarios, and allowing one to supplement the analyses done with one's own collection of analyses. In this sense, every conclusion reached based at least partially on the missing data, should be accompanied by the assumptions on which it rests, be it MAR or otherwise. In this sense, clinical-trial data, like any other data, can provide estimates of what would happen if subjects were complete contrary to fact; but again, these are conditional on the assumptions holding true.

Professor Tchetgen rightly draws our attention to doubly robust or multiply robust estimators. Although slightly technical, it is a theory of key importance and great potential that should be embraced by the community, be it on the industry side, at the regulatory authorities, or in academia.

Related to this, it is crucial to reflect up front on the scientific question actually posed, and then to select the analyses methodology deemed most appropriate. This will naturally lead the researcher to stay close to practice, in particular to stay close to the patient.

3 Sensitivity Analysis

Professor Scharfstein eloquently restates the role of sensitivity analysis. I fully endorse his assessment: "Regardless of the approach (i.e., likelihood- or non-likelihood-based) adopted, we feel that formal, coherent and scientifically driven sensitivity analysis strategies must play a critical role in the regulatory approval process." At this point, a variety of sensitivity-analysis methods are available, branching out from the likelihood-based, Bayesian, or frequentist models initially considered. The field is in full development; likely and hopefully many more methods are to come in years ahead of us. Other than embarking on a quest for the optimal sensitivity analysis, it is more important for the researcher to consider a number of well chosen sensitivity analysis routes and then to assess stability of the conclusions reached across these.

Professor Ibrahim restates that, while there is value to assess a model's quality of fit, one should be very careful when making formal statements regarding the nature of the missing data mechanism. In the spirit of the MAR companion to MNAR models, this means that a model can be assessed in as far as it fits to the observed data, but that the part describing the incomplete data given the observed data is unverifiable. The latter part is intimately linked to, though not exactly the same as, the missing data mechanism.

The unverifiable parts of a model come to surface, in a pattern-mixture model for example, through the identifying restrictions. Professor Ibrahim notes that such explicit restrictions may be less than appealing in practice. As he correctly remarks, the same is true for (13) in our manuscript, referring to the MAR counterpart of an MNAR model. While this is true, it is arguably at least equally problematic that some models contain identifications so implicit (and admittedly easy to use) that their mere existence might well go unnoticed. Bringing the effect of assumptions and restrictions, whether implicit or explicit, to surface, is one of the main roles of sensitivity analysis.

4 Aggressive Data Collection

The discussants, in particular Dr Flyer and Professor Scharfstein, quite rightly draw our attention to the fact that every effort should be made to collect as much data as possible, before and after treatment discontinuation. This is ever so important: No method, regardless how sophisticated though it may be, can replace the luxury of actually collecting all data.

In this sense, I fully subscribe to Professor Scharfstein's words: "Regulatory authorities should use their power to encourage pharmaceutical companies to design and implement studies which minimize missing data. We believe that any extra costs incurred by companies will be worth the reduction in the uncertainty caused by missing data. The best that can happen to missing data is not to have them.

5 Software

I fully subscribe to Professor Ibrahim's point of view that standard and advanced methods badly need to be implemented in standard statistical software. Otherwise, regardless of how sophisticated our insights and fancy our models may become, we are bound to remain stuck with such museum pieces as last observation carried forward.