Estimands, Missing Data, and Sensitivity Analysis

Geert Molenberghs

I-BioStat, Hasselt University & KU Leuven
E-mail: geert.molenberghs@uhasselt.be

Abstract: The presentation sets out by considering estimands, a very important topic in clinical trials. A connection is made with the much older use in survey sampling theory. Using an example from surrogate marker evaluation, it is discussed where information comes from: data, design, and assumptions. The latter may be unverifiable, hence the need to perform sensitivity analysis.

The setting is then broadened to various forms of enrichment; that is, every situation where the model contains more aspect than the data are able to provide information about. Subsets of the enrichment class are: (a) coarsening, where some data could have been observed but were not (e.g., missing data); (b) augmentation, where models, for convenience of interpretation, are supplemented with unobservables, such as random effects. The focus is then placed on incomplete data for the rest of the presentation.

A general framework for missing data is given, starting from Rubin's seminal work. The defining and transforming role of the National Academy of Sciences report from 2010 about the "Prevention and Treatment of Missing Data in Clinical Trials" is evocated. It is argued that the role of the patient should not be forgotten, next to academe, regulators, and industry.

It is shown that for every MAR model, there is a family of MNAR models that exhibits the same fit to the data. Hence, one cannot show that MAR holds or not, solely depending on the data. The implications for standard and sensitivity analyses are discussed.

Regarding standard analysis, the roles of ignorable likelihood and Bayesian analysis, multiple imputation, and inverse probability weighting (e.g., weighted generalized estimating equations), are discussed.

The presentation concludes with some brief illustrations of sensitivity analysis.