Dealing with missing data in randomized and cluster randomized trials

Chapter 8 - Scientific and social impact of the thesis

As explained in this thesis, randomized and cluster randomized trials are gold standard procedures for assessing the effectiveness of interventions (i.e. treatments) since randomized allocation of study participants to different treatment groups guarantees unbiased comparability between the treatment groups and, so, allows unbiased assessment of the effectiveness of treatments. RCTs are predominantly used, for example, in phase 3 confirmatory clinical trials for evaluating the safety and effectiveness of drugs, biologic products, and some medical devices, which require the highest standard of scientific rigor (Little et al. 2012). CRTs are often encountered in public and mental health, and in family medicine, among others. When correctly designed, run and reported, RCTs and CRTs can produce an immediate beneficial impact on clinical practice and patient care, as well as inform policy decisions about them. However, missing data in such trials, if not treated appropriately in statistical analysis, can seriously compromise inferences from clinical trials. This may in turn mislead, for example, health practitioners or public health authorities in their decision making on the safety and effectiveness of an intervention (e.g. a drug or vaccine) that should be used as cure or prevention of certain diseases, which would badly affect individuals or public health safety (see e.g. Lagakos 2006; Little et al. 2012; Lurie and Levine 2010; Molnar et al. 2009; Raboud et al. 1996). Therefore, the investigations conducted in the present thesis, and the findings and recommendations thereof, can help avoid or mitigate the risk of having invalid results and misleading information from statistical analysis of RCTs and CRTs with missing data, and thus positively impact the social wellbeing of individuals or the community at large.

As a research with scientific impact, this thesis is a useful and practical guideline for researchers and practitioners primarily interested in statistical analysis of scientific clinical

trials about how to adequately deal with missing data in statistical analysis of RCTs and CRTs, in terms of bias and efficiency in treatment effect estimation. In this thesis, we raise awareness about the detrimental impact that not accounting for missingness in statistical analysis of RCTs and CRTs may have on statistical inferences from such trials, and then prescribe the statistical strategies that can be used to appropriately handle the missing data under various realistic missingness scenarios in RCTs and CRTs. This thesis contributes to the existing work on methods for handling missing data in RCTs and CRTs, as it mostly covers scenarios not yet covered in existing work and particularly emphasizes (unlike other work) the usefulness of simple missing data methods such as mean imputation and the missing-indicator method in the context of RCTs and CRTs with missing covariates, which (methods) are known to be inappropriate for handling missing data in the context of observational studies. See for example chapters 2 (published in "pharmaceutical statistics") and 3 where these two methods were found to be performing at least as good as advanced methods such as multiple imputation (MI) and maximum likelihood (ML)-based methods in dealing with missing covariates in RCTs. Moreover, the findings of this thesis are mostly based on extensive, detailed and carefully designed simulation studies, as described in chapters 2-5. The present thesis will, therefore, also be useful for researchers and lecturers as didactic material about simulation studies in statistics. This thesis is not exclusively aimed at statisticians or researchers with a strong technical background on statistical analysis, but also at researchers with a non-technical background, since we also explain in simple wording all technical expressions contained in this report, and the merits and demerits of each missing data method are considered. This methodology will thus hopefully contribute to the dissemination, understanding and correct use of statistical analyses of clinical trials with missing data amongst medical and health researchers.

In order to make accessible the findings of this thesis to the scientific community, chapter 2 has been published in an international scientific journal at the interface of statistics and pharmaceutical RCTs, chapter 3 is under review (after revision) and will hopefully be published soon in a similar journal, and chapters 4 and 5 have been submitted for publication. Moreover, chapter 2 has been presented in a webinar, and 3 and 4 have been presented at two international scientific conferences. The data and R-code that support the findings of this thesis are available from the author or supervisors upon request.

8.1 References

Kaushal, S. (2014). Missing data in clinical trials: Pitfalls and remedies. *International Journal of Applied and Basic Medical Research*, *4*(3), 6.

Lagakos, S. W. (2006). Time-to-event analyses for long-term treatments-the APPROVe trial. *New England Journal of Medicine*, *355*(2), 113.

Little, R. J., D'Agostino, R., Cohen, M. L., Dickersin, K., Emerson, S. S., Farrar, J. T., ... & Stern, H. (2012). The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, *367*(14), 1355-1360..

Lurie, I., & Levine, S. Z. (2010). Meta-analysis of dropout rates in SSRIs versus placebo in randomized clinical trials of PTSD. *The Journal of nervous and mental disease*, *198*(2), 116-124.

Molnar, F. J., Man-Son-Hing, M., Hutton, B., & Fergusson, D. A. (2009). Have last-observation-carried-forward analyses caused us to favour more toxic dementia therapies over less toxic alternatives? A systematic review. *Open Medicine*, *3*(2), e31.

Raboud, J. M., Montaner, J. S. G., Thorne, A., Singer, J., Schechter, M. T., & Canadian HIV Trials Network A002 Study Group. (1996). Impact of missing data due to dropouts on estimates of the treatment effect in a randomized trial of antiretroviral therapy for HIV-infected individuals. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *12*(1), 46-55.