Initial data analysis for longitudinal data – a general framework

Lara Lusa

University of Primorska

`lara.lusa@famnit.upr.si`

Marianne Huebner

Michigan State University

`huebner@msu.edu`

Systematic initial data analysis (IDA) and clear reporting of the findings is an important step towards reproducible research. A general framework of IDA for observational studies was proposed to include data cleaning, data screening, and possible refinements of the preplanned analyses (Huebner 2018). Longitudinal studies, where participants are observed repeatedly over time, have special features that should be taken into account in the IDA steps before addressing the research question. Our aim was to propose a framework for IDA in longitudinal studies. Based on the IDA framework from Huebner et al. (2018) we refined it for use in longitudinal studies and provided guidance on how to prepare an IDA plan for longitudinal studies. The framework includes several steps that are specific to longitudinal data, or bear greater importance when data are longitudinal. Appropriate numerical and graphical tools for longitudinal data allow the researchers to conduct IDA in a reproducible manner to avoid non-transparent impact on the interpretation of model results. For example, in the framework we propose how to summarize the time frame of the study, the characteristics of the participants, the outcome and the time varying covariates, how to summarize longitudinal average trends, how to explore the variation between individuals, how to characterize the correlation and the covariance of the outcome and of selected covariates, methods for the exploration of missing values and the description of drop-out. We provide an example on how to conduct IDA in the context of a longitudinal population cohort study (Boesch-Supan et al., 2013). The paper is presented on behalf of the Topic Group "Initial Data Analysis" of the STRATOS Initiative (STRengthening Analytical Thinking for Observational Studies, http://www.stratos-initiative.org).

References

Huebner, M., le Cessie, S., Schmidt, C.O., Vach, W. A Contemporary Conceptual Framework for Initial Data Analysis. Observational Studies 4 (2018) 171-192.

Boesch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, S. Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE), International Journal of Epidemiology 42 (2013) 992—1001.