

Comparison of multiple imputation methods in HRQoL studies with repeated measurements

Urko Aguirre¹, Inmaculada Arostegui² and José M. Quintana¹

¹*Galdakao-Usansolo Hospital, Bizkaia, Spain*

²*University of the Basque Country UPV/EHU, Bizkaia, Spain*

Longitudinal data research projects are often subject to high rates of missingness. This nonresponse may be related to the dropout before the ending of the study, or to the unavailability during one or more measurements. On the other hand, when the information comes from surveys, patients often deny to respond or are unable to answer several questions due to the lack of interest or time. This happens in Health-Related Quality of Life (HRQoL) studies. Because of that, many researchers use different imputation methods, such as simple imputation methods (mean, regression imputation, etc...) as well as more advanced methods, to avoid losing valuable information from those who are unable or unwilling to answer the questions and to conclude misleading inferences about changes in the mean response over time.

The most widely used methods of imputation are: hotdeck, propensity score and the Markov Chain Monte Carlo (MCMC). Hotdeck imputation replaces missing data using the data from other observations in the sample at hand. The propensity score technique is based on an implicit approach based on Propensity scores and an Approximate Bayesian Bootstrap (ABB) to generate imputations. The MCMC method uses a single chain to produce the imputations. The posterior mode computed from the Expectation-Maximization (EM) algorithm is used as the starting values for the MCMC process. Selection of the appropriate imputation method is a difficulty itself. Available literature about the benefits of each method exists, although it is not conclusive.

We apply these three above mentioned methods to a sample of 788 patients who had a surgical intervention of hip joint replacement. The Pain domain of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was measured before and after the hip joint replacement. The change in pain was the outcome variable and relevant covariates where included in a general linear model.

Results from data after imputation were compared to the complete analysis showing that imputation affects to the estimates and standard errors of coefficients and p-values. Different results were obtained depending on the methods of imputation. A simulation study was carried out to illustrate the performance of the mentioned methods. MCMC method shows more accurate results in the simulation study. Therefore, we recommend to apply multiple imputation techniques in the analysis of HRQoL data over time and the MCMC method is preferred to the hotdeck and propensity score methods.