Multiple imputation models should incorporate the outcome in the model of interest

Jonathan W. Bartlett, Chris Frost and James R. Carpenter

Sir, In a recent publication in Brain, Jack Jr et al. (2010) reported on the value of hippocampal atrophy and amyloid-β measures in predicting conversion from mild cognitive impairment to Alzheimer’s disease. The authors used data from 218 subjects in the Alzheimer’s Disease Neuroimaging Initiative with mild cognitive impairment, who had a measure of amyloid-β either through CSF amyloid-β42 or Pittsburgh compound B positron emission tomography imaging (PIB-PET). Of the 218 subjects, only 53 (24%) had PIB-PET available, and so Jack Jr et al. (2010) used multiple imputation for measurement error correction (following Cole et al., 2006) to impute the missing PIB-PET values, based on each subject’s CSF amyloid-β42 and Apolipoprotein E (APOE) ε4 status. The imputation model was fitted using data from a calibration data set of 41 subjects who had both PIB-PET and CSF amyloid-β42 data available. The fitted imputation model was then used to impute 100 ‘completed’ data sets, each with no missing PIB-PET values. In line with standard multiple imputation methodology, a Cox proportional hazards model was then fitted to each imputed data set, relating time to conversion to Alzheimer’s disease, to amyloid-β load (as measured by PIB-PET) and atrophy, and the results combined using Rubin’s rules for final inference.

Our concern focuses on the imputation model used by Jack Jr et al. (2010) that may be mis-specified since it did not include variables representing conversion status and time to conversion or last follow-up (the outcome of interest). In general, omitting the outcome from the imputation model results in biased estimates (Moons et al., 2006; Sterne et al., 2009). Indeed, Cole et al. (2006) included the censoring indicator and logarithm of time to event as covariates in their imputation model (Appendix 2).

Recently, it has been shown that a more accurate approach is to use an estimate of the baseline cumulative hazard function as covariate rather than log time (White and Royston, 2009).

In the present context, the extent of the bias induced by omitting the outcome from the imputation model depends on the percentage of variation ($R^2$) in PIB-PET explained by CSF amyloid-β42 and APOE ε4. Given the imputation model coefficients and residual standard deviation reported by Jack Jr et al. (2010), and the variance and correlation of CSF amyloid-β42 and APOE ε4 in mild cognitive impairment subjects in the Alzheimer’s Disease Neuroimaging Initiative, we estimate that $R^2$ was ~80% and hence the extent of the bias was likely to be relatively small. Nevertheless, an analysis that included the outcome is equally feasible computationally, is likely to be less biased and is thus preferable.

More generally, the effect of omitting the outcome variable from the imputation model is not always small (Sterne et al., 2009). A striking example of the dangers occurred in the development of the UK QRISK cardiovascular risk score (Hippisley-Cox et al., 2007b), in which missing cholesterol values were imputed using multiple imputation. Surprisingly, serum cholesterol ratio was found to have no independent effect on risk of cardiovascular disease. The authors subsequently clarified that the censoring indicator had been inadvertently omitted from the imputation model, and a re-analysis using an improved imputation model did result in an independent effect of cholesterol (Hippisley-Cox et al., 2007a).

In summary, we emphasize that multiple imputation is a powerful statistical tool for the analysis of partially observed data that can alleviate biases and recover information. However, the validity...
of estimates and inferences relies critically on appropriate specification of the imputation model. In general, one should always include the outcome variable of the final model of interest in the imputation model, and failure to do so may result in biased estimates of associations between the variable being imputed and the outcome of interest.

Funding

JWB was supported by an Economic and Social Research Council Follow-On Funding grant (RES-189-25-0103). CF’s work is funded by grants from the CHDI Foundation, the Euro-HD Network, the EU FP7 Health programme, the Institute of Neurology (University College London), the Multiple Sclerosis Trials Collaboration and from Effective Intervention (Centre for Economic Performance, London School of Economics). JRC is supported by an Economic and Social Research Council Research Fellowship (RES-063-27-0257).

References


