SWAT 52: Statistical methods for compensating for missing longitudinal data in a cluster randomised trial

Objective of this SWAT
To investigate whether multiple imputation or maximum likelihood-based methods perform better as a statistical approach to handling missing data in a longitudinal cluster randomised trials. To use real data, calibrated by simulations, to determine the relative power and type I error of the trial, and to identify biases that arise in effect estimates.

Study area: Statistical methods
Sample type: Participants
Estimated funding level needed: Very Low

Background
In the analysis of clinical trials, maximum likelihood offers certain pragmatic advantages over multiple imputation[1] although both methods are expected to give similar results. In practice, it is not clear whether results under the two approaches converge more or less rapidly in small or modest samples, in the presence of missing-at-random (MAR) versus MCAR data, nor in situations of correlated observations (for example, due to clustering, repeated-measures, or both); where a choice must be made concerning the covariance structure of the data.

Interventions and comparators
Intervention 1: Multiple Imputation (MI)
Intervention 2: Maximum Likelihood (ML) modelling

Index Type: Method of data analysis

Method for allocating to intervention or comparator
Not applicable, because alternative analysis methods will be used on the same data sets.

Outcome measures
Primary: The standard error of the estimate of the group difference in the change in Emotional Self-Awareness (ESA) from baseline.
Secondary: Model fit statistics (information criteria, entropy, log likelihood).
The p-value (and significance at a 0.05 threshold) for the statistical test on the group difference in the change in ESA from Baseline, and subsequent post hoc contrasts.
The type I error rates under a range of missing data conditions.
The statistical power (and type II error rates) under a range of effect sizes and missing data conditions.

Analysis plans
In order to compare the two methods proposed as solutions to missing data points, the analytic strategy will match that of the CopeSmart trial: a contrast of the Group ESA scale marginal means in a linear mixed model, incorporating cross-classified random effects for School and for Person. Fixed predictors will include the baseline ESA measurement, Group, and Time, and optionally (for the ML approach) covariates which associate with the probability of a missing data point.
The primary outcome of interest will be the precision with which statistics of interest to the primary analysis of the trial are estimated, quantified by their standard error. The CopeSmart trial consists of two parallel arms with three repeated measurements over time, and therefore the primary statistic of interest will be a contrast between arms of the difference in the ESA from baseline to subsequent time-points. Model fit statistics will also be compared across the two strategies.
In order to inspect objective metrics of relative performance of the two approaches, data will be simulated using the sample size and other characteristic parameters of the CopeSmart trial, and therefore known patterns of missing data can be introduced, either MAR or not-missing-at-random (NMAR). These patterns can be made more or less dependent on measured or unmeasured covariates. Outcomes in this context will be the observed type I error rate when no group differences are permitted, and the observed type II error rate (or statistical power) when group differences of certain sizes are incorporated.

Possible problems in implementing this SWAT
Recruitment of skilled personnel; No other implementation problems anticipated as the study is purely in silico.

References
1. Allison PD. Handling Missing Data by Maximum Likelihood. SAS Global Forum 2012, paper 312-2012.

Publications or presentations of this SWAT design

Examples of the implementation of this SWAT

People to show as the source of this idea: Ricardo Segurado
Contact email address: ricardo.segurado@ucd.ie
Date of idea: 15/MAR/2016
Revisions made by:
Date of revisions: