

Diploma thesis:

## **Sample Size Calculation for Phase II Multiple Sclerosis Studies with Magnetic Resonance Imaging Based Outcome Variables**

Andrea Graml

In cooperation with the Ludwig-Maximilians-Universität, Munich, Prof. Dr. Helge Toutenburg and advisors from the SLC, Prof. Dr. Ludwig Kappos and Dr. Ulrike Held.

### **Background:**

This diploma thesis dealt with two major topics, namely the replacement of missing values and sample size estimation.

In longitudinal multiple sclerosis (MS) trials with MRI based outcome measures, techniques for replacing missing values are used without thorough validation. In most studies the method of choice (if at all explicitly mentioned) is "last observation carried forward". The use of imputation methods provides larger datasets for sample size calculations, which can make these more exact and reliable.

Two different approaches for sample size estimation in MS trials with the number of new enhancing lesions in frequent magnetic resonance imaging (MRI) as primary outcome have been proposed in the past: a non-parametric bootstrapping approach introduced by Nauta et al. (1994) and a parametric approach after Sormani et al. (1999). In the parametric approach a negative binomial distribution is assumed for the cumulative number of lesions in a certain time interval. Both approaches make specific assumptions about how treatment influences the distribution of the data.

The techniques for the replacement of missing values and the two approaches for sample size estimation were compared separately using placebo data from the Sylvia Lawry Centre for Multiple Sclerosis Research (SLCMSR) dataset, that currently includes the placebo arms from 27 MS trials with monthly measured MRI data.

### **Methods:**

Different methods for imputation have been compared. For that, missing values were generated at random in a complete dataset of 67 patients observed monthly for a period of 6 months. Each observation of each patient had the same probability to be chosen.

The deleted values were replaced by 10 different methods and compared with the original data using the mean squared error and other measures of goodness of fit. As a next step, another group of 35 patients, each with one missing observation was examined for the proposal of being "missing at random" (i.e. missing independently of the value itself and of the value of any covariate).

The sample size calculation was done for the sum of six consecutive monthly MRI Gadolinium enhancing lesion counts as the primary outcome. From the open portion of the SLCMSR database, 67 patients with complete observations (33 relapsing remitting (RR), 34 secondary progressive (SP)) and 35 patients (21 RR, 14 SP) with one missing value in the observation period were available. Simulations for both approaches were carried out on the 67 patients with complete observations, with and without separation for disease course.

Additionally five partially overlapping subsamples from this group, each containing 23 patients, were drawn without replacement and their power estimates compared to those presented in the article by Nauta, where the dataset was of the same size.

### **Results:**

Imputing the mean of the neighbouring observations of a missing value was found to be the most valuable method for replacement. The values of the 35 patients mentioned above seem to be "missing completely at random", in the way Simonoff (1988) and Toutenburg (1998) have indicated.

For that reason these observations could be merged with the 67 complete patients to a larger dataset of 102 patients with up to one missing observation per patient. In the sample size estimation the RR and SP subgroups lead to significantly different power estimates. The power was influenced even more by characteristics of the datasets, in particular by the proportion of inactive patients, than by their size. In some cases as expected from Sormani et al. (1999) there were also significant differences between the algorithms, but unexpectedly these were not always in the same direction.

### **Conclusion:**

Sample size calculations for future phase II trials (Sormani et al., 2001) can be based on larger datasets than usual using the SLC database which can be further enlarged using replacement methods such as “mean of neighbours”.

It could also be thought of using this imputation method in the analysis of clinical trials.

Sample size estimation in future phase II trials in MS should be conducted separately for different disease courses, as already recommended by Tubridy et al. (1998). Further studies on the influence of dataset characteristics and different algorithms on sample size estimation is necessary.

### **References:**

**Nauta, J.J.P.** , Thompson, A.J., Barkhof, F. and Miller, D.H. (1994). Magnetic resonance imaging in monitoring the treatment of multiple sclerosis patients: statistical power of parallel-groups and crossover designs. *J. Neurol. Sci.* 122, 6-14.

**Simonoff, J.** (1988). Regression Diagnostics to Detect Nonrandom Missingness in Linear Regression. *Technometrics* 30, No. 2, 205-214.

**Sormani, M.P.**, Bruzzi, P., Miller, D.H., Gasperini, C., Barkhof, F. and Filippi, M. (1999). Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. *J. Neurol. Sci.* 163, 74-80.

**Sormani, M.P.**, Miller, D.H., Comi, G., Barkhof, F., Rovaris, M., Bruzzi, P. and Filippi, M. (2001). Clinical trials of multiple sclerosis monitored with enhanced MRI: new sample size calculations based on large datasets. *J. Neurol. Neurosurg. Psych.* 70, 494-499.

**Toutenburg, H.** (1992). *Lineare Modelle*. Physica-Verlag, Heidelberg.

**Tubridy N.**, Ader, H.J., Barkhof, F., Thompson, A.J. and Miller, D.H. (1998). Exploratory treatment trials in MS: sample size calculations for relapsing-remitting and secondary-progressive subgroups using placebo controlled parallel groups. *J. Neurol. Neurosurg. Psych.* 64, 50-55.