Heumann:

Intention-to-treat with drop-out

Sonderforschungsbereich 386, Paper 199 (2000)

Online unter: http://epub.ub.uni-muenchen.de/

Intention-to-treat with drop-out

Christian Heumann

Department für Statistik

Ludwigstr. 33

80539 München

email: heumann@stat.uni-muenchen.de

SUMMARY

We consider the problem of an ITT analysis in a randomized clinical trial. Due to (study) drop-outs, standard methods are not applicable and simple imputation methods like LOCF (last observation carried forward) may lead to biased results.

Since a patient who drops out of the study often will also change or drop the assigned treatment, an "ignorable" analysis in the sense of Rubin (1976) assuming MAR (missing at random), as e. g. a propensity weighted analysis or a likelihood based MAR-analysis (Heyting, Tolboom and Essers (1992), Lavori, Dawson and Shera (1995)) is not valid. This is due to the fact that information is missing about outcomes as well as the covariate treatment after drop-out. That is, even if the drop-out process itself is ignorable, we can not treat the problem as ignorable because of the missing covariate information.

Two different proposals for treating the problem were made from Little and Yau (1996), who used a fixed effects model for modeling the treatment effect, and from Kleinman, Ibrahim and Laird (1998), who used a Bayesian approach with a random effects model.

We follow the path given by Little and Yau (1996), who created multiple imputations under various assumptions about the actual treatment after drop-out, and conduct a simulation study on the α-error and power of simple endpoint tests. This should also shed light onto the problem whether the true treatment effect can be sensibly bracketed by assumptions like zero dose or continuing dose after drop-out.

KEYWORDS: intention-to-treat, longitudinal, missing data, non-ignorable, non-compliance

1 Introduction

In a recent article, Frangakis and Rubin (1999) state: "Randomized experiments with human subjects often suffer from two major complications, namely non-compliance to treatment assignment and missing outcomes".

An intention-to-treat (ITT) analysis is the standard approach for randomized trials with non-compliance. It is a valid method for estimating the effect of assignment (to a specific treatment) on the population averages of the outcome Y. The ITT principal requires that all subjects initially randomized to one of, say, two treatments, have to be included in the analysis and that one has to analyse the data "as randomized". The latter means that even if some subjects drop their assigned treatment and switch, e.g., to the Placebo group, we treat them as if they have not dropped their treatment. In such a sense the ITT analysis is a method for dealing with *treatment drop-out* (Kleinman et al, 1998). In a Placebo-controlled trial one expects, that the ITT analysis conservatively (underestimates) the true biological effect of a new treatment, if there exists a treatment effect. It can be shown that an analysis which uses the actual received treatment ("as-treated analysis") or only the data from compliers can be biased, see Nagelkerke, Fidler, Bernsen and Borgdorff (2000).

While it is clear how to apply an ITT analysis if the (final) outcomes of all subjects are recorded, there is no standard way for an ITT analysis if some of the outcomes are missing, e.g. because of drop-outs (*study drop-out*). It also can be shown that approaches which use the assumption of missing at random (MAR) and analyse the data according to the ITT principle may lead to an overestimation of the treatment effect (Kleinman et al., 1998).

In the following we study an extension of the multiple imputation ideas of Little and Yau (1996) to the case of non-ignorable non-response using a *selection model* approach. We assume that the treatment received after a drop-out is *unknown* and that the treatment can be described by a binary treatment indicator (say, 1 for treatment, 0 for Placebo). This scenario is clearly not relevant, if a subject has no access to the treatment after drop-out. We study two main settings:

- Full compliance but missing data: non-response may be non-ignorable, that is de-

pendent on previous outcomes and previous received treatments, but the true data generating process is such that each subject stays on the assigned treatment.

- Non-compliance and missing data: the actual received treatment in the treatment group depends on previous outcomes, and: a treatment drop-out implies a study drop-out. That means, non-compliance in the treatment group automatically leads to a study drop-out (while a complier has some probability for a study drop-out).

In both settings we develop multiple imputation procedures for the missing outcomes based on the assumption of non-ignorable non-response and some *assumed* treatment after drop-out similar to Little and Yau (1996), namely the zero dose and continuouing dose assumptions. The imputation model can be considered as a conditional model while a simple end-point analysis is used for testing the ITT effect. In the following we describe the set-up of the simulation study and describe more precisely the underlying assumptions.

2 Data generating model and missing model

We restrict ourselves to the case where each subject is measured before randomization (*baseline value*) and at two distinct time points ($T = 2$) after randomization. Denote the baseline value by Z_{i0}, and the two subsequent measurements by Y_{i1} and Y_{i2}, where i is the subject index. Let further D_{i0} and D_{i1} denote the binary treatment indicators with values 1 (treatment) and 0 (Placebo). We further assume that missing responses only occur at the second time point and therefore Y_{i2} is possibly not observed. The indicator variable R_{i2} is always observed, taking values 1 (*responded*) if Y_{i2} is observed and 0 if Y_{i2} is missing. The sample sizes for the two groups were hold fixed through the simulation at $N_{\text{Placebo}} = N_{\text{Treatment}} = 60$ and thus $N = 120$.

2.1 Data generating model

The following models were used for generating the data:

$$\begin{aligned} Z_{i0} &\sim N(0,1) \ , \\ Y_{i1} &\sim f_{i1} = N(\alpha_0 + \alpha_1(Z_{i0} - \bar{Z}_0) + \alpha_2 D_{i0}, \sigma_1^2) \ , \\ Y_{i2} &\sim f_{i2|1} = N(\beta_0 + \beta_1(Y_{i1} - \bar{Y}_1) + \beta_2 D_{i1} + \beta_3(Z_{i0} - \bar{Z}_0), \sigma_2^2) \ , \end{aligned} \tag{1}$$

where $\bar{Z}_0$ is the mean of all individual baseline values. If a subject is randomized into the Placebo group then $D_{i0} = D_{i1} = 0$ in any case. That is, a subject receiving Placebo can never receive a treatment whether it is a (potential) complier or not and whether it drops out of the study or not. If a subject is randomized into the treatment group, $D_{i0} = 1$. We assume therefore that all subjects are compliers at least until the first measurement is made. On the other hand, D_{i1} depends on whether the subject is a complier or not. Summarizing, we allow three treatment regimes: $(D_{i0}, D_{i1}) = (0,0)$ for subjects in the Placebo group, $(D_{i0}, D_{i1}) = (1,1)$ for compliant subjects in the treatment group and $(D_{i0}, D_{i1}) = (1,0)$ for non-compliant subjects in the treatment group. We further assume, that the conditional variances are equal in both groups.

We can compute the marginal expectations from (1) as

$$\begin{aligned} \mathrm{E}(Y_{i1}) &= \alpha_0 + \alpha_2 D_{i0} \ , \\ \mathrm{E}(Y_{i2}) &= (\beta_0 - 0.5\beta_1\alpha_2) + \beta_1\alpha_2 D_{i0} + \beta_2 D_{i1} \ , \end{aligned}$$

where we have used that

$$\begin{aligned} \mathrm{E}(Z_{i0}) &= \mathrm{E}(\bar{Z}_0) = 0 \ , \\ \mathrm{E}(\bar{Y}_1) &= \frac{1}{N_{\text{Placebo}} + N_{\text{Treatment}}} \left(\sum_{i=1}^{N_{\text{Placebo}}} \alpha_0 + \sum_{i=1}^{N_{\text{Treatment}}} (\alpha_0 + \alpha_2) \right) \\ &= \alpha_0 + 0.5\alpha_2 \qquad \text{since } N_{\text{Placebo}} = N_{\text{Treatment}} \end{aligned}$$

and therefore

$$\mathrm{E}(Y_{i1} - \bar{Y}_1) = \alpha_2 D_{i0} - 0.5\alpha_2 \ .$$

We can also formulate the expectations in terms of the randomization group and compliance status:

$$\mathrm{E}(Y_{i1}) = \begin{cases} \alpha_0 & \text{if subject } i \text{ is in the Placebo group} \\ \alpha_0 + \alpha_2 & \text{if subject } i \text{ is in the treatment group} \end{cases} \tag{2}$$

and

$$\mathrm{E}(Y_{i2}) = \begin{cases} (\beta_0 - 0.5\beta_1\alpha_2) & \text{if subject } i \text{ is in the Placebo group} \\ (\beta_0 - 0.5\beta_1\alpha_2) + \beta_1\alpha_2 + \beta_2 & \text{if subject } i \text{ is in the treatment group and is a complier} \\ (\beta_0 - 0.5\beta_1\alpha_2) + \beta_1\alpha_2 & \text{if subject } i \text{ is in the treatment group} \\ & \text{and is a non-complier} \end{cases} \tag{3}$$

As can be seen, even if a subject is a non-complier, there is a "carry-over" effect $\beta_1\alpha_2$.

2.2 Missing model

We used a logistic regression model for modeling the drop-out probability:

$$\log\left(\frac{pr(R_{i2}=1)}{pr(R_{i2}=0)}\right) = \log\left(\frac{\pi_{i2}}{1-\pi_{i2}}\right) = \gamma_0 + \gamma_1(Y_{i2} - \bar{Y}_2) + \gamma_2(Y_{i1} - \bar{Y}_1) + \gamma_3 D_{i1} + \gamma_4(Z_{i0} - \bar{Z}_0) \tag{4}$$

A ceterus paribus interpretation can be given as follows: if $\gamma_1 > 0$, then a subject i where $(Y_{i2} - \bar{Y}_2) > 0$ has higher probability of being observed in $T = 2$ than a subject where this difference is negative. If $\gamma_3 > 0$, a subject i which is in the treatment group and is a complier ($D_{i1} = 1$) has higher probability of being measured in $T = 2$ than a subject which is in the treatment group but is a non-complier, etc. We note, that non-compliers have a different drop-out rate than subjects in the Placebo group, if there is a treatment effect and if $\gamma_1 \neq 0$ and/or $\gamma_2 \neq 0$, since the responses Y_{i1} and Y_{i2} themselves are influenced by the treatment after randomization (D_{i0}).

3 A multiple imputation procedure for non-ignorable non-response

In the following we describe a multiple imputation procedure applicable to the case of non-ignorable non-response. Although the basic ideas can be found in Shafer (1997) for the MAR case, we have to develop a method where the missing model has to be incorporated in the estimation and imputation tasks. The procedure is a Bayesian approach and repeats the following steps:

- P(robability)-Step: draw $\theta^{(p)}$ from the conditional distribution of the parameters given observed and imputed data. $\theta^{(p)}$ stands for the vector of all parameters used in the described models above.
- I(mputation)-Step: draw $Y_{i2}^{(p)}$ from the conditional distribution of the missing data given $\theta^{(p)}$ and the observed data for all i, where $R_{i2} = 0$.

In the following we describe in detail the Bayesian procedure.

3.1 Complete data likelihood

As we use a selection model, the joint distribution of the response Y and the indicator variable R can be factorized as

$$[Y, R] = [R|Y][Y] \ .$$

To be precise, let Y_{i2}^* be the *potential* outcome in $T = 2$ with $Y_{i2}^* = Y_{i2}$, if the response of subject i is observed and $Y_{i2}^* = Y_{i2}^{mis}$, if the response of subject i is missing. Then the contribution of subject i to the *complete data likelihood* is (in the following we condition on the baseline values Z_{i0})

$$\begin{aligned} L_i &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{1}{2\sigma_1^2}(y_{i1} - \alpha_0 - \alpha_1(z_{i0} - \bar{z}_0) - \alpha_2 D_{i0})^2\right\} \\ &\quad \times \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{-\frac{1}{2\sigma_2^2}(y_{i2}^* - \beta_0 - \beta_1(y_{i1} - \bar{y}_1) - \beta_2 D_{i1} - \beta_3(z_{i0} - \bar{z}_0))^2\right\} \end{aligned}$$

$$\times \left[\frac{\exp\{\gamma_0 + \gamma_1(y^*_{i2} - \bar{y}^*_2) + \gamma_2(y_{i1} - \bar{y}_1) + \gamma_3 D_{i1} + \gamma_4(z_{i0} - \bar{z}_0)\}}{1 + \exp\{\gamma_0 + \gamma_1(y^*_{i2} - \bar{y}^*_2) + \gamma_2(y_{i1} - \bar{y}_1) + \gamma_3 D_{i1} + \gamma_4(z_{i0} - \bar{z}_0)\}} \right]^{r_{i2}}$$

$$\times \left[\frac{1}{1 + \exp\{\gamma_0 + \gamma_1(y^*_{i2} - \bar{y}^*_2) + \gamma_2(y_{i1} - \bar{y}_1) + \gamma_3 D_{i1} + \gamma_4(z_{i0} - \bar{z}_0)\}} \right]^{(1-r_{i2})}$$

$$= \quad f_{i1} f^*_{i2} \pi^{r_{i2}}_{i2} (1 - \pi_{i2})^{1-r_{i2}} \tag{5}$$

We note that the *observed data likelihood* contribution of subject i is

$$L^{obs}_i = f_{i1} f^*_{i2} \pi_{i2} \ ,$$

if the response in $T = 2$ is observed, and

$$L^{obs}_i = f_{i1} \int f^*_{i2} (1 - \pi_{i2}) dy^*_{i2}$$

if the response in $T = 2$ is missing.

3.2 Prior distributions

Since we implement a Bayesian procedure, we need to choose some prior distributions for the model parameters.

Parameter $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$: In the complete data case, $\boldsymbol{\alpha}$ can be estimated by a regression of the y_{i1}'s on the baseline value and treatment indicator. We choose a diffuse prior for $\boldsymbol{\alpha}$:

$$P(\boldsymbol{\alpha}) \propto \text{const} \ .$$

Parameter σ^2_1: We choose a highly dispersed Gamma distribution with parameters 1 and 0.001 for the *precision* $\tau_1 = 1/\sigma^2_1$:

$$P(\tau_1) \sim G(1, 0.001) \ .$$

Parameter $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)$: As for $\boldsymbol{\alpha}$, we choose a diffuse prior

$$P(\boldsymbol{\beta}) \propto \text{const} \ .$$

Parameter σ_2^2: As for σ_1^2 we choose a highly dispersed Gamma distribution with parameters 1 and 0.001 for the precision $\tau_2 = 1/\sigma_2^2$:

$$P(\tau_2) \sim G(1, 0.001) \ .$$

Parameter $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)$: For the parameters of the missing model we choose independent normal and informative priors:

$$\begin{aligned} P(\gamma_0) &= P(\gamma_1) = P(\gamma_2) = P(\gamma_3) = P(\gamma_4) \sim N(0,1) \\ P(\boldsymbol{\gamma}) &= \prod_{j=0}^{4} P(\gamma_j) \ . \end{aligned}$$

3.3 Creating multiple imputations by data augmentation

Let Y_{obs} and Y_{mis} denote all observed and missing responses and (Y_1, Y_2^*) all potential outcomes. Further let R_2 be the vector of all missing indicators and $\theta = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_1, \tau_2)$. Having defined the complete data likelihood and the priors, a data augmentation procedure (Shafer, 1997) allows us to create *proper* imputations for the missing values. The procedure has much similarity with the EM-algorithm: the E-step is replaced by a P-step, a draw from the *complete data posterior* distribution

$$P(\theta | Y_{obs}, Y_{mis}, R_2) = P(\theta | Y_1, Y_2^*, R_2) \ ,$$

and the M-step is replaced by an I-step, a draw from the *posterior predictive* distribution

$$Y_{mis} \sim P(Y_{mis} | Y_{obs}, R_2, \theta)$$

The complete data posteriori distribution is with (5) and the priors in section (3.2) given by

$$P(\theta | Y_1, Y_2^*, R_2) \propto \left(\prod_{i=1}^{N} L_i \right) P(\boldsymbol{\gamma}) P(\tau_1) P(\tau_2) \ . \tag{6}$$

The P- and I-steps are then repeatedly applied. In the following, the details are given.

P-step: The P-step is built by the following MCMC updating scheme. We assume that the missing values Y_{mis} have been imputed such that we have generated a complete matrix of potential outcomes (Y_1, Y_2^*). Since the full conditionals are proportional to the *joint posterior*, we get

$$P(\tau_1|\boldsymbol{\alpha}, \tau_2, \boldsymbol{\beta}, \boldsymbol{\gamma}, Y_1, Y_2^*) \propto \prod_{i=1}^{N} \left[(\tau_1)^{\frac{1}{2}} \exp\left\{ -\frac{\tau_1}{2}(y_{i1} - \alpha_0 - \alpha_1(z_{i0} - \bar{z}_0) - \alpha_2 D_{i0})^2 \right\} \right] P(\tau_1) \tag{7}$$

With the chosen Gamma-prior for τ_1 it can be shown that the full conditional is also a Gamma distribution with parameters

$$G\left(1 + \frac{N}{2}, 0.001 + \frac{1}{2}\sum_{i=1}^{N}(y_{i1} - \alpha_0 - \alpha_1(z_{i0} - \bar{z}_0) - \alpha_2 D_{i0})^2\right) \ . \tag{8}$$

Further,

$$P(\boldsymbol{\alpha}|\tau_1, \tau_2, \boldsymbol{\beta}, \boldsymbol{\gamma}, Y_1, Y_2^*) \propto \prod_{i=1}^{N} \left[\exp\left\{ -\frac{\tau_1}{2}(y_{i1} - \alpha_0 - \alpha_1(z_{i0} - \bar{z}_0) - \alpha_2 D_{i0})^2 \right\} \right] \tag{9}$$

Let $x_{i1}' = (1, (z_{i0} - \bar{z}_0), D_{i0})$ and $X_1 = (x_{11}, \ldots, x_{N1})'$. Then (9) may be written as

$$\begin{aligned} & P(\boldsymbol{\alpha}|\tau_1, \tau_2, \boldsymbol{\beta}, \boldsymbol{\gamma}, Y_1, Y_2^*) \\ & \propto \exp\left\{ -\frac{\tau_1}{2}(Y_1 - X_1\boldsymbol{\alpha})'(Y_1 - X_1\boldsymbol{\alpha}) \right\} \\ & \propto \exp\left\{ -\frac{\tau_1}{2}[\boldsymbol{\alpha}'(X_1'X_1)\boldsymbol{\alpha} - Y_1'X_1\boldsymbol{\alpha} - \boldsymbol{\alpha}'X_1'Y_1 + Y_1'X_1(X_1'X_1)^{-1}X_1'Y_1] \right\} \\ & = \exp\left\{ -\frac{\tau_1}{2}[\boldsymbol{\alpha} - (X_1'X_1)^{-1}X_1'Y_1]'(X_1'X_1)[\boldsymbol{\alpha} - (X_1'X_1)^{-1}X_1'Y_1] \ , \right\} \end{aligned} \tag{10}$$

which is the kernel of a multivariate normal distribution with parameters

$$MVN(\hat{\boldsymbol{\alpha}}, \tau_1^{-1}(X_1'X_1)^{-1}) \ , \tag{11}$$

where $\hat{\boldsymbol{\alpha}} = (X_1'X_1)^{-1}X_1'Y_1$ is the least square estimate of the regression of Y_1 on X_1. Random variates can be generated by using "independent" standard normal pseudo random variates and a Cholesky decomposition of $(X_1'X_1)^{-1}$. Analogously, the full conditional for τ_2 is

$$G\left(1 + \frac{N}{2}, 0.001 + \frac{1}{2}\sum_{i=1}^{N}(y_{i2}^* - \beta_0 - \beta_1(y_{i1} - \bar{y}_1) - \beta_2 D_{i1} - \beta_3(z_{i0} - \bar{z}_0))^2\right) \tag{12}$$

and the full conditional for $\boldsymbol{\beta}$ is

$$MVN(\hat{\boldsymbol{\beta}}, \tau_2^{-1}(X_2'X_2)^{-1}) \tag{13}$$

where $x_{i2}' = (1, (y_{i1} - \bar{y}_1), D_{i1}, (z_{i0} - \bar{z}_0))$, $X_2 = (x_{12}, \ldots, x_{N2})'$ and $\hat{\boldsymbol{\beta}} = (X_2'X_2)^{-1}X_2'Y_2^*$. The last full conditional we have to look at, is

$$\begin{aligned} & P(\boldsymbol{\gamma}|\tau_1, \boldsymbol{\alpha}, \tau_2, \boldsymbol{\beta}, Y_1, Y_2^*) \\ & \quad \propto \prod_{i=1}^{N} \pi_{i2}^{r_{i2}}(1-\pi_{i2})^{1-r_{i2}} P(\boldsymbol{\gamma}) \\ & \quad \propto \prod_{i=1}^{N} \pi_{i2}^{r_{i2}}(1-\pi_{i2})^{1-r_{i2}} \prod_{j=0}^{4} \exp(-\frac{1}{2}\gamma_j^2) \ , \end{aligned} \tag{14}$$

from which can be sampled by a Metropolis-Hastings (MH) step. For simplicity we update each parameter separately using a normal proposal centered at the current value. The proposal variance is automatically chosen in a first trial run such that the acceptance rate lies in the interval $[0.2, 0.6]$ for all parameters. Let $\phi(\cdot)$ denote the standard normal density function and, e.g. γ_1' a new proposal for γ_1. Further, let $\pi_{i2} = \pi_{i2}(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)$ and $\pi_{i2}' = \pi_{i2}(\gamma_0, \gamma_1', \gamma_2, \gamma_3, \gamma_4)$. Since we have chosen standard normals as priors for the parameter vector γ and the proposal distribution is symmetric, one computes the ratio

$$q \ = \ \frac{\prod_{i=1}^{N} \pi_{i2}'^{r_{i2}}(1-\pi_{i2}')^{1-r_{i2}}}{\prod_{i=1}^{N} \pi_{i2}^{r_{i2}}(1-\pi_{i2})^{1-r_{i2}}} \cdot \frac{\phi(\gamma_1')}{\phi(\gamma_1)} \tag{15}$$

and accepts the proposal γ_1' with probability $\min(1, q)$. A modification to prior variances other than 1 is straight forward.

I-step: In the I-step, we have to draw from the posterior predictive distribution of the missing values given the actual parameter values and the observed data. Since we assume independence among subjects, we get for subjects with missing $y_{i2,mis}$ (and therefore $r_{i2} = 0$):

$$\begin{aligned} & P(y_{i2,mis}|\tau_1, \boldsymbol{\alpha}, \tau_2, \boldsymbol{\beta}, \boldsymbol{\gamma}, Y_{obs}, Y_{mis,-i}) \\ & \quad \propto \ \exp\left\{-\frac{\tau_2}{2}(y_{i2,mis} - \beta_0 - \beta_1(y_{i1} - \bar{y}_1) - \beta_2 D_{i1} - \beta_3(z_{i0} - \bar{z}_0))^2\right\} \\ & \qquad \times \frac{1}{1 + \exp\left\{\gamma_0 + \gamma_1(y_{i2,mis} - \bar{y}_2^*) + \gamma_2(y_{i1} - \bar{y}_1) + \gamma_3 D_{i1} + \gamma_4(z_{i0} - \bar{z}_0)\right\}} \ , \end{aligned} \tag{16}$$

where we assume, that the missing values of subjects other than i are actually fixed or imputed (therefore $Y_{mis,-i}$ in the condition). Note, that $\bar{y}_2^*$ contains $y_{i2,mis}$. To get draws from this conditional distribution we again apply MH-steps independently for each subject with a missing $y_{i2,mis}$ using normal proposals centered at the actual value of $y_{i2,mis}$. As can be seen one has to recompute the mean $\bar{y}_2^*$ to compute the ratio q as in (15) in each step.

4 Measuring the ITT effect

Usually, the data augmentation algorithm is run for a pre-specified number of iterations where the parameters and the imputed missing values are stored after an initial burn-in phase. Since we are mainly interested in the ITT effect, that is, the difference of the *marginal* means in the two groups at the endpoint in $T = 2$, we looked at only m completed data sets. In our simulation study we chose $m = 10$, which should be sufficient. Since the run length was 2500 for each parameter setting (after a burn-in of length 2500), we simply take every 250-th completed dataset for the anylysis of the ITT effect where we hope that the draws from the posterior predictive distribution with lag 250 are approximately independent. That is, for measuring the ITT effect, we take each completed dataset $k = 1, \ldots, m$, and compute the regression

$$y_{i2} = \kappa + \psi D_{i,assigned} + \text{error} \ , \tag{17}$$

where y_{i2} is an observed or imputed value and $D_{i,assigned}$ is the assigned (not the actual received) treatment. This results in m least square estimates $\hat{\psi}_k$ with associated estimated variances U_k. According to a rule for combining the estimates given by Rubin (1987), we can construct a confidence interval for ψ from the m completed data sets as follows under the assumption of asymptotic normality of $U_k^{-1/2}(\psi - \psi_k)$:

$$\begin{aligned} \bar{\psi} &= \frac{1}{m}\sum_{k=1}^{m} \hat{\psi}_k \\ \bar{U} &= \frac{1}{m}\sum_{k=1}^{m} U_k \\ B &= \frac{1}{m-1}\sum_{k=1}^{m} (\psi_k - \bar{\psi})^2 \end{aligned}$$

$$V \;=\; \bar{U} + (1 + m^{-1})B \;. \tag{18}$$

V is called the total variance, which has as components the *within imputation variance* $\bar{U}$ and the *between imputation variance* B. A $100\%(1-\alpha)$ CI is given by

$$\bar{\psi} \pm t_\nu \left(1 - \frac{\alpha}{2}\right) V^{\frac{1}{2}} \tag{19}$$

with $\nu = (m-1)(1+l^{-1})^2$ and $l = (1+m^{-1})B/U$. Analogously a test for the hypothesis $H_0 : \psi = 0$ versus the alternative $H_1 : \psi \neq 0$ can be constructed.

5 Simulation results

In the following we give the simulation results splitted into two parts. In the first part, the true data are generated according to the assumption, that all subjects stay on their assigned treatment, that means either $D_{i0} = D_{i1} = 0$ or $D_{i0} = D_{i1} = 1$ (no non-compliers). In the second part we drop that assumption. We compare the multiple imputation procedure with complete case analysis (CC) and last observation carried forward (LOCF) under different assumptions about the missing model: missing completely at random (MCAR), missing at random (MAR), non-ignorable non-response (NI). Each result is based on 1000 simulations where each simulation consists of 2500 repeated applications of the P- and I-step after a 2500 burn-in phase. Every 250-th imputation is used to produce a 95%-CI for the ITT effect.

5.1 Full compliance case

In the full compliance case, the true treatment difference is given by $\beta_1\alpha_2 + \beta_2$. We tried a number of different parameter settings: no treatment effect, increasing treatment effect, increasing and decreasing treatment effect, etc. The results are given in tables 1, 2 and 3.

The first two columns in table 1 show a parameter setting, where the missing processes are MCAR and MAR and where the treatment effect is zero. The responses at both time points depend on the baseline value ($\alpha_1 = \beta_3 = 1$). As can be seen, all methods work sufficiently well. The row denoted by "full" (F) takes the "true" response value known in the simulation

for measuring the ITT effect. The MI procedure is slightly conservative, especially in the MAR case. The last two columns in table 1 describe a case where the treatment effect is (linearly) increasing over the two time points. The missing processes are MAR and NI. In the MAR case the response probability is higher for individuals with a response exceeding the mean response in $t = 1$ and is higher in the treatment group ($\gamma_3 > 0$). In the NI case, the response probability depends on the (possible unknown) response in $t = 2$ in a similar way. CC has the lowest power, F has the highest power and LOCF and MI are comparable and lie between CC and F.

The first two columns in table 2 describe a situation where the treatment effect is first increasing and then decreasing in a way that the marginal treatment effect in $t = 2$ is zero. MAR and NI situations were simulated. In both situations, LOCF exceeds the nominal value with a very high value of 0.094 in the NI case. That is, LOCF in nearly 10% of the 1000 simulations wrongly detects a treatment effect. That can be explained by the fact the LOCF procedure simply takes the observed values at $t = 1$ where some effect exists and thus the imputed values in the treatment group are in general too high. The last two columns describe a case similar to the last two columns of table 1, but where the response probability is lower for individuals exceeding the mean response in $t = 1$ (MAR) and $t = 2$ (NI). The results are as expected with CC having the lowest power and F the highest.

Finally, the last column in table 3 gives the results for a case where the response probability is lower for subjects exceeding the mean in $t = 2$ *and* higher drop-out probability for subjects in the treatment group. The treatment effect is again linearly increasing. The power of CC completely breaks down in that case.

5.2 Non-compliance case

We have also made two simulations for the non-compliance case, where the actual treatment D_{i1} in the treatment group is dependent on the response in $t = 1$. That is there are some individuals having a true treatment regime of $(1, 0)$. We have simulated the "true" actual received treatment before $t = 2$ in the following way: let U a uniform random number on

$[0,1]$. For all $i = 1, \ldots, N$

$$D_{i1} = \begin{cases} 1, & \text{if } U < 0.5 + 0.5(Y_{i1} - \bar{Y}_1) \ (*) \\ 0, & \text{else } (**) \end{cases}$$

[1] [2] Further, treatment drop-out implies study drop-out:

$$P(R_{i2} = 0 | D_{i1} = 0, D_{i0} = 1) = 1 \ .$$

The imputations were then applied under two models:

- zero dose model: the assumption is, that *any* subject that drops out, has no access to the treatment and thus only has some "carry-over" effect.
- continuing dose model: the assumption is, that *any* subject that drops out, continuoues its assigned treatment.

For both assumptions confidence intervals were calculated. The results are given in table 4.

The first column shows the results for a linearly increasing treatment effect. Since some subjects are non-compliers, the overall mean effect is smaller (< 1) than if all subjects would comply to their assigned treatment. The methods produce the expected results: zero-dose leads to an under-estimation of the true treatment effect (the true parameter lies above the upper bound of the CI in all cases where it isn't covered by the CI), continuouing dose leads to an over-estimation of the true treatment effect (in nearly all cases the true parameter lies below the lower bound of the CI if the parameter is not covered by the CI).

The second column shows the results for the case of an increasing and then decreasing treatment effect. Now the results are contrary to the results of the former case. Under the continuous dose assumption, the treatment effect is under-estimated. The true parameter is covered by the CI in only 60% of the simulations. In all other cases it lies above the upper bound of the CI. An explanation could be the following: the missing model correctly identifies the dependence on D_{i1} (more drop-outs in the treatment arm) but incorrectly estimates the

[1] $(*)$ is always fulfilled, if $0.5 + 0.5(Y_{i1} - \bar{Y}_1) > 1$.

[2] $(**)$ always fulfilled, if $0.5 + 0.5(Y_{i1} - \bar{Y}_1) < 0$.

dependence on $(Y_{i2} - \bar{Y}_2)$ (the estimate $\hat{\gamma}_1$ is in most cases greater than zero) and thus incorrectly assumes that individuals with low values of $(Y_{i2} - \bar{Y}_2)$ have higher probability of drop-out. The imputed values are therefore too low and the estimate of the treatment effect is biased towards zero. On the other hand, under the zero dose assumption, the missing model correctly identifies the dependence on $(Y_{i2} - \bar{Y}_2)$ but incorrectly estimates the dependence on D_{i1} (the estimate $\hat{\gamma}_3$ is in most cases greater than zero) and thus incorrectly assumes a lower drop-out rate in the treatment group. This leads to an over-estimation of the treatment effect.

6 Discussion

We have developed a method creating multiple imputations under a selection model for the missing process. The first set of simulations show that MI is a relatively stable procedure, while CC may break down concerning the power of detecting a true treatment effect and LOCF may falsely detect a treatment effect to often. In the case of non-compliance, new issues arise, when using the zero or continuous dose assumption. A combination of both methods may give bounds for the treatment effect, but the resulting interval may be too long to draw a definite conclusion about the effectiveness of the treatment, especially if the sample sizes are as small as in the simulation study. It seems that the progress of the treatment effect has an influence on the quality of the ad hoc methods as well as our extended MI procedure. As a cautionary note, we remind that selection models are known to be very sensitive. Their value lies in the possibility of a sensitivity analysis, especially using a Bayesian approach. We do not claim that each simulated situation is realistic but the examples show how the ad hoc methods can fail even in simple set-ups. To handle more realistic situations, methods for dealing with different causes for drop-out have to be developed. An approach, where a random effect influences the compliance as well as the response probability (representing e.g. the health status) could perhaps be an interesting alternative to this approach.

	MCAR	MAR	MAR	NI
α_0	0	0	0	0
α_1 $(Z_{i0} - \bar{Z}_0)$	1	1	1	1
α_2 (D_{i0})	0	0	0.5	0.5
β_0	0	0	0	0
β_1 $(Y_{i1} - \bar{Y}_1)$	0	0	1	1
β_2 (D_{i1})	0	0	0.5	0.5
β_3 $(Z_{i0} - \bar{Z}_0)$	1	1	0	0
γ_0	0	0	0.5	0.5
γ_1 $(Y_{i2} - \bar{Y}_2)$	0	0	0	0.5
γ_2 $(Y_{i1} - \bar{Y}_1)$	0	0.5	0.5	0
γ_3 (D_{i1})	0	0	0.2	0.2
γ_4 $(Z_{i0} - \bar{Z}_0)$	0	0	0	0
Δ $(\beta_1\alpha_2 + \beta_2)$	**0**	**0**	**1**	**1**
Sig.	0.05	0.05	0.05	0.05
Full	0.053	0.054	0.87	0.89
CC	0.057	0.05	0.62	0.58
LOCF	0.057	0.054	0.75	0.78
MI	0.040	0.032	0.75	0.74

Table 1: Full compliance case; left two columns: treatment effect zero; right two columns: increasing treatment effect

	MAR	NI	MAR	NI
α_0	0	0	0	0
α_1 $(Z_{i0} - \bar{Z}_0)$	1	1	1	1
α_2 (D_{i0})	0.5	0.5	0.5	0.5
β_0	0	0	0	0
β_1 $(Y_{i1} - \bar{Y}_1)$	1	1	1	1
β_2 (D_{i1})	-0.5	-0.5	0.5	0.5
β_3 $(Z_{i0} - \bar{Z}_0)$	0	0	0	0
γ_0	0.5	0.5	0.8	0.8
γ_1 $(Y_{i2} - \bar{Y}_2)$	0	0.5	0	-0.5
γ_2 $(Y_{i1} - \bar{Y}_1)$	0.5	0	-0.5	0
γ_3 (D_{i1})	0.2	0.2	0.2	0.2
γ_4 $(Z_{i0} - \bar{Z}_0)$	0	0	0	0
Δ $(\beta_1\alpha_2 + \beta_2)$	**0**	**0**	**1**	**1**
Sig.	0.05	0.05	0.05	0.05
Full	0.047	0.046	0.87	0.88
CC	0.041	0.049	0.75	0.72
LOCF	0.071	0.094	0.79	0.79
MI	0.031	0.036	0.78	0.79

Table 2: Full compliance case; left 2 columns: first increasing than decreasing treatment effect; right 2 columns: linearly increasing treatment effect

	NI	NI
α_0	0	0
α_1 $(Z_{i0} - \bar{Z}_0)$	1	1
α_2 (D_{i0})	0.5	0.5
β_0	0	0
β_1 $(Y_{i1} - \bar{Y}_1)$	1	1
β_2 (D_{i1})	0.5	0.5
β_3 $(Z_{i0} - \bar{Z}_0)$	0	0
γ_0	0.5	1.0
γ_1 $(Y_{i2} - \bar{Y}_2)$	-1.5	-1.5
γ_2 $(Y_{i1} - \bar{Y}_1)$	0	0
γ_3 (D_{i1})	1.0	-1.0
γ_4 $(Z_{i0} - \bar{Z}_0)$	0	0
Δ $(\beta_1\alpha_2 + \beta_2)$	**1**	**1**
Sig.	0.05	0.05 (0.1)
Full	0.88	0.89 (0.93)
CC	0.77	0.15 (0.24)
LOCF	0.79	0.76 (0.84)
MI	0.79	0.47 (0.70)

Table 3: Full compliance case; higher versus lower response rates in the treatment group (than in the Placebo group) in the case of a linearly increasing treatment effect

	NI	NI
α_0	0	0
α_1 $(Z_{i0} - \bar{Z}_0)$	1	1
α_2 (D_{i0})	0.5	1
β_0	0	0
β_1 $(Y_{i1} - \bar{Y}_1)$	1	1
β_2 (D_{i1})	0.5	-1
β_3 $(Z_{i0} - \bar{Z}_0)$	0	0
γ_0	0.5	1.0
γ_1 $(Y_{i2} - \bar{Y}_2)$	-1.5	-1.5
γ_2 $(Y_{i1} - \bar{Y}_1)$	0	0
γ_3 (D_{i1})	1.0	-1.0
γ_4 $(Z_{i0} - \bar{Z}_0)$	0	0
Δ $(\neq \beta_1\alpha_2 + \beta_2)$	$<$**1**	$>$**0**
Level of confidence	0.95	0.95
Zero dose	0.94^a	1.0
Cont. dose	0.95^b	0.6^d

Table 4: Non-compliance case; ^{a}) In all other cases the true parameter lies above the upper bound of the CI. ^{b}) In nearly all other cases the true parameter lies below the lower bound of the CI ^{d}) In all other cases the true parameter lies above the upper bound of the CI.

References

Frangakis, C. and Rubin, D. (1999). Addressing complicationsof intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes, *Biometrika* **86**: 365–379.

Heyting, A., Tolboom, J. T. B. M. and Essers, J. G. A. (1992). Statistical handling of drop-outs in longitudinal clinical trials, *Statistics in Medicine* **11**: 2043–2061.

Kleinman, K., Ibrahim, J. and Laird, N. (1998). A bayesian framework for intent-to-treat analysis with missing data, *Biometrics* **54**: 265–278.

Lavori, P. W., Dawson, R. and Shera, D. (1995). A multiple imputation strategy for clinical trials with truncation of patient data, *Statistics in Medicine* **14**: 1913–1925.

Little, R. and Yau, L. (1996). Intent-to-treat analysis for longitudinal studies with drop-outs, *Biometrics* **52**: 1324–1333.

Nagelkerke, N., Fidler, V., Bernsen, R. and Borgdorff, M. (2000). Estimating treatment effects in randomized clinical trials in the presence of non-compliance, *Statistics in Medicine* **19**: 1849–1864.

Rubin, D. B. (1976). Inference and missing data, *Biometrika* **63**: 581–592.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Sample Surveys*, Wiley, New York.

Shafer, J. (1997). *Analysis of incomplete multivariate data*, Chapman & Hall, London.