An alternative characterization of MAR in shared parameter models for incomplete longitudinal data and its utilization for sensitivity analysis

Grigorios Papageorgiou[1,2] and Dimitris Rizopoulos[1]
[1]Department of Biostatistics, Erasmus University Medical Centre, Rotterdam, The Netherlands
[2]Department of Cardiothoracic Surgery, Erasmus University Medical Centre, Rotterdam, The Netherlands

Abstract: Dropout is a common complication in longitudinal studies, especially since the distinction between missing not at random (MNAR) and missing at random (MAR) dropout is intractable. Consequently, one starts with an analysis that is valid under MAR and then performs a sensitivity analysis by considering MNAR departures from it. To this end, specific classes of joint models, such as pattern-mixture models (PMMs) and selection models (SeMs), have been proposed. On the contrary, shared-parameter models (SPMs) have received less attention, possibly because they do not embody a characterization of MAR. A few approaches to achieve MAR in SPMs exist, but are difficult to implement in existing software. In this article, we focus on SPMs for incomplete longitudinal and time-to-dropout data and propose an alternative characterization of MAR by exploiting the conditional independence assumption, under which outcome and missingness are independent given a set of random effects. By doing so, the censoring distribution can be utilized to cover a wide range of assumptions for the missing data mechanism on the subject-specific level. This approach offers substantial advantages over its counterparts and can be easily implemented in existing software. More specifically, it offers flexibility over the assumption for the missing data generating mechanism that governs dropout by allowing subject-specific perturbations of the censoring distribution, whereas in PMMs and SeMs dropout is considered MNAR strictly.

Key words: missing data, sensitivity analysis, shared parameter models, dropout, MNAR

1 Introduction

Follow-up studies that include human subjects are known to be highly susceptible to dropout. The extent to which dropout will complicate the subsequent analysis depends on how the missingness is associated with the observed and unobserved longitudinal outcomes. According to the taxonomy introduced by Rubin (1976) and Little and Rubin (2002), this may happen according to three distinct missing data mechanisms: missing completely at random (MCAR), missing at random (MAR)

and missing not at random (MNAR). Under MCAR, the missingness depends on neither observed nor unobserved outcomes. If the missingness is independent of the unobserved outcomes after conditioning on the observed outcomes, then the missing data mechanism is considered to be MAR. Conversely, under MNAR, conditioning on the observed outcomes cannot disentangle the dependence between the missingness process and the unobserved outcomes.

What separates MNAR from its counterparts in applied practice, though, is the concept of ignorability. That is under the (Bayesian) likelihood framework, one does not need to consider a model for the missingness process and hence 'ignore' it. Ignorability holds under the assumption that the process that generates the missing data is MAR and that the missingness and measurement processes depend on independent sets of parameters. While the innate convenience of ignorability popularized MAR models, one needs to consider if such an assumption holds carefully. That is especially the case since the distinction between MAR and MNAR is intractable as for every MNAR model, there exists a MAR counterpart with the same fit to the observed data (Molenberghs et al., 2008). Hence, one should start with an analysis that is valid under MAR and then perform a sensitivity analysis considering plausible MNAR departures from it.

To showcase the importance of such sensitivity analysis, consider the illustrative example in Figure 1 which is based on simulated data. The data were simulated assuming a joint model for longitudinal and survival data. To mimic MAR missingness, censoring was based on the value of previous repeated measurements, and therefore the missing longitudinal trajectories (depicted with dotted grey) consist of the subjects that fulfilled the MAR criterion or experienced the event. We then fitted a model to the complete data, whereas for the incomplete data we fitted an MAR and an MNAR model. All three models resulted in different fits to the data, which consequently lead to different subject-specific predictions as shown in the lower panel of Figure 1. Therefore, sensitivity analysis is necessary to be able to investigate the robustness of the model and the plausibility of the assumptions concerning the mechanism that generated the missing data.

To this purpose, three well-established general frameworks exist, distinguished by their approach in factoring the joint distribution of the outcome and missingness processes to conditionally independent components. In a pattern-mixture model (PMM), the joint distribution is expressed as the product of the marginal distribution of the missingness process and the conditional distribution of the outcome given the missingness process. Contrariwise, in a selection model (SeM), the joint distribution is expressed as the product of the marginal distribution of the outcome and the conditional distribution of the missingness process given the outcome. Alternatively, in a shared-parameter model (SPM), the outcome and the missingness processes are assumed to be driven by a set of unobserved variables, the so-called random effects. The use of latent variables gives rise to the conditional independence assumption under which the outcome and the missingness processes are independent, given the set of random effects.

The SeM framework naturally encompasses MAR, while Molenberghs et al. (1998) provided a characterization of MAR under the PMM setting, specifically

for the case of longitudinal data with dropout. As a result, sensitivity analysis is well-developed in the SeM and the PMM frameworks (Molenberghs and Vereke, 2005; Molenberghs and Kenward, 2007), whereas the SPM has received less attention. More specifically, Creemers et al. (2011) provided a characterization of MAR for the SPM case by introducing the general SPM (GSPM) and then proposed a sensitivity analysis approach utilizing this framework (Creemers et al., 2010). Later, Njagi et al. (2014) proposed a characterization of MAR for the case of SPMs for longitudinal and time-to-event data under the GSPM.

In this article, we introduce an alternative characterization of MAR in SPMs for longitudinal and time-to-event data by exploiting the conditional independence assumption and utilizing the censoring distribution. By doing so, we cover a wide range of assumptions for the missing data mechanism on the subject-specific level. Previous work on the topic focused on the GSPM, which includes potentially a broad set of random effects. Here, we focus on the standard SPM (Wu and Carroll, 1988; Wu and Bailey, 1988, 1989; Hogan and Laird, 1997, 1998), a particular case of the GSPM, that is most commonly encountered in practice. The benefit of doing so is that our approach is intuitively appealing and can be easily implemented in existing software. Furthermore, it allows for subject-specific perturbations of the censoring distribution, concerning the missing data generating mechanism, whereas in PMMs and SeMs dropout is considered MNAR strictly.

The rest of the article is structured as follows. In Section 2, we propose an alternative characterization of MAR for the SPM model. In Section 3, we discuss its estimation under the Bayesian framework. In Section 4, we illustrate with an application how the SPM can be used for sensitivity analysis with existing software. In Section 5, we present a simulation study to assess the behaviour of the proposed sensitivity analysis under different scenarios with respect to the amount of MAR and MNAR missingness. Discussion follows in Section 6.

2   A characterization of MAR for the shared parameter model

For each subject i out of a sample of size N from the target population, let $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^\top$ denote the vector of size $n_i \times 1$ of planned repeated measurements, with y_{ij} being the value of the longitudinal outcome at time point t_{ij}, $j = 1, \ldots, n_i$, out of a set of $\mathcal{J}$ planned measurement occasions. The vector of planned repeated measurements may further be partitioned into two vectors: $\boldsymbol{y}_i = (\boldsymbol{y}_i^o, \boldsymbol{y}_i^m)$, of observed and missing repeated measurements, respectively. Next, let T_i^* be the true dropout time, C_i the censoring time and $T_i = \min(T_i^*, C_i)$ the observed dropout or censoring time. Furthermore, let $\boldsymbol{\theta}$ and $\boldsymbol{\psi} = \{\boldsymbol{\psi}^{T^*}, \boldsymbol{\psi}^C\}$ be the vectors of parameters for the measurement, dropout and censoring processes respectively. Under these settings, the vectors of observed and missing repeated measurements can be expressed as $\boldsymbol{y}_i^o = \{y_{ij}; t_{ij} < T_i, \forall j \in \mathcal{J}\}$ and $\boldsymbol{y}_i^m = \{y_{ij}; t_{ij} > T_i, \forall j \in \mathcal{J}\}$. In other words, the vector of observed measurements includes all the observations which were recorded before
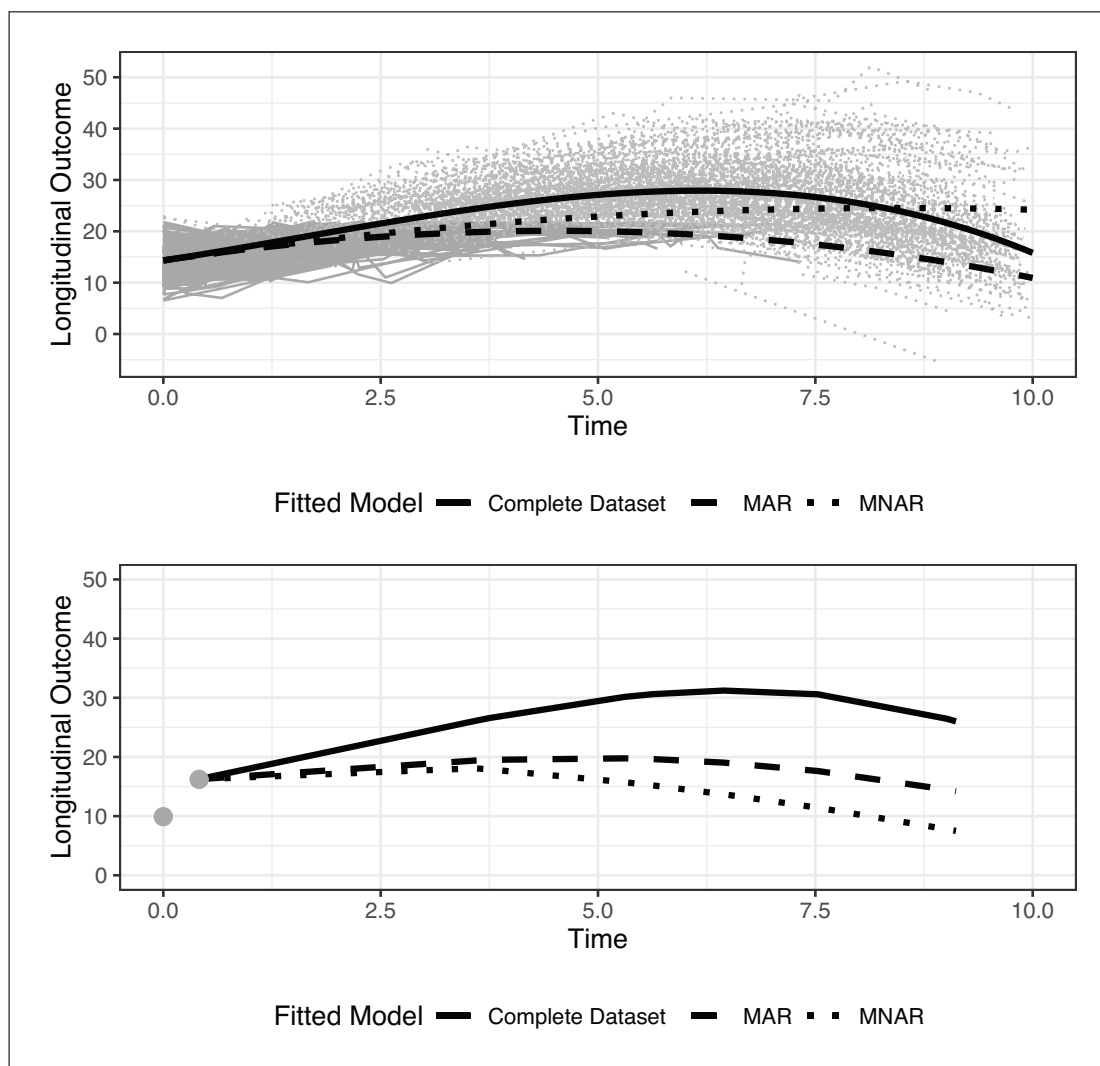
Figure 1 Illustrative Example: (Upper panel) observed longitudinal trajectories (black lines), missing longitudinal trajectories (grey), and fitted models using complete data and using incomplete data under MAR and MNAR. (Lower Panel) observed longitudinal measurements of a new subject and subject-specific predictions based on the three different models

neither dropout nor censoring occurred. On the contrary, the vector of missing measurements includes the observations which were not recorded due to either dropout or censoring. Note that we regard dropout and censoring as different causes of missingness, a distinction that we will exploit to achieve MAR characterization of SPMs.

What differentiates dropout from censoring in such a setting pertains to the causes of leaving the study. As an example, consider the case of a study, with death as the main outcome and where a longitudinal outcome is also planned to be recorded for a time period. The planned longitudinal outcome measurements will be missing for both the subjects that died during the study but also for subjects that dropped out from the study due to any other reason (e.g., they moved to another country). Depending on the information available and the research setting, one may define death as the only cause of dropout and assume all the other reasons are non-informative and therefore record them as censoring. In the case where no information is available for the cause of dropout, then one would consider only those who completed the study as censored.

Under this framework and with covariate vectors suppressed from the notation, the shared parameter model can be specified as follows:

$$p\left(y_i^o, y_i^m, T_i^*, C_i; \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \int p\left(y_i^o, y_i^m, T_i^*, C_i, \boldsymbol{b}_i; \boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{D}\right) d\boldsymbol{b}_i, \qquad (2.1)$$

where $\boldsymbol{b}_i \sim \mathcal{N}(0, \boldsymbol{D})$ is a vector of shared random effects assumed to induce the dependence between the processes involved in the joint density of the SPM. That is by conditioning on the random effects, the joint density of the SPM can be decomposed to the following product of conditional densities

$$\int p\left(T_i^*, | \boldsymbol{b}_i; \boldsymbol{\psi}^{T^*}\right) p\left(C_i \mid y_i^o, y_i^m, \boldsymbol{b}_i; \boldsymbol{\psi}^C\right) p\left(y_i^o, y_i^m \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right) p\left(\boldsymbol{b}_i; \boldsymbol{D}\right) d\boldsymbol{b}_i, \qquad (2.2)$$

which brings to light the core assumption of conditional independence in SPMs, under which the measurement and the dropout processes are independent conditional on the random effects. Note that the conditional independence assumption commonly does not imply independence of the censoring and the measurement processes. More specifically, in the general case, censoring is allowed to depend on the measurement process. Note that the conditional independence assumption may be extended to allow for independence of the censoring and the measurement processes as well. Finally, the assumption of non-informative censoring, which we discuss later, may be used to relax both these assumptions and allow for independence between the censoring and measurement processes.

Model (2.2) is a conventional SPM for longitudinal and time-to-event data. The term conventional here is in reference to the GSPM framework in the sense that a single common underlying random effects structure is used instead of multiple random effects structures. Another important note for model (2.1) is that by definition it allows for dependence between the dropout process T_i^* and the censoring process C_i. When there is no information from the data on the joint distribution of (T_i^*, C_i), the assumption of non-informative censoring may be utilized (Tsiatis, 1975) to further simplify the model. Under both the assumptions of non-informative censoring and conditional independence, (2.2) further simplifies to the following:

$$\int p\left(T_i^* \mid \boldsymbol{b}_i; \boldsymbol{\psi}^{T^*}\right) p\left(C_i \mid y_i^o; \boldsymbol{\psi}^C\right) p\left(y_i^o, y_i^m \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right) p\left(\boldsymbol{b}_i; \boldsymbol{D}\right) d\boldsymbol{b}_i, \qquad (2.3)$$

which, in other words, means that the censoring process no longer depends on the random effects and the missing observations. It should be noted that the assumption of non-informative censoring, while commonly applicable, may not always be a reasonable choice depending on the setting under study. Nevertheless, it is a necessary condition for achieving an MAR characterization in SPMs without introducing more latent variables as in Creemers et al. (2011).

Note that by definition, a subject may either be classified as a dropout or as censored and never both. This means that the decomposition of the joint density as described in (2.3) is, on a subject-specific level, further decomposed to the following:

$$
\begin{cases}
\int p\left(T_i^* \mid b_i; \psi^{T^*}\right) p\left(y_i^o, y_i^m \mid b_i; \theta\right) p\left(b_i; D\right) db_i, & i : \text{dropout} \\
\int p\left(C_i \mid y_i^o; \psi^C\right) p\left(y_i^o, y_i^m \mid b_i; \theta\right) p\left(b_i; D\right) db_i, & i : \text{censored}.
\end{cases}
\tag{2.4}
$$

The decomposition in (2.4) reveals that SPMs encompass MAR on the subject-specific level. That is for a subject who dropped out from the study, the model is MNAR because both the dropout, $p\left(T_i^* \mid b_i; \psi^{T^*}\right)$, and measurement, $p\left(y_i^o, y_i^m \mid b_i; \theta\right)$, processes depend on the missing observations through the random effects b_i. For a censored subject, though, the model is MAR since the random effects and the missing observations appear only in factor $p\left(y_i^o, y_i^m \mid b_i; \theta\right)$. Moreover, due to ignorability, the joint density reduces to the density of the observed measurement process $\int p\left(y_i^o \mid b_i; \theta\right) p\left(b_i\right) db_i$. Ignorability is a key consequence of MAR in SPMs. It entails that there are not any information to be gained from either the dropout or the censoring processes with respect to the measurement generating process, and therefore MAR can be explored within the SPM framework. This is achieved without broadening the definition of the conventional SPM by introducing additional random effects. Furthermore, the fact that MAR in SPMs is achieved on the subject-specific level allows for the potential of a subject- or cause-specific MAR sensitivity analysis. This is in contrast to the SeM and PMM frameworks which are MNAR or MAR for all the individuals under study. Finally, as we are going to illustrate in the upcoming section, this decomposition allows exploiting existing software for its estimation.

3 Estimation of SPM under MAR

Estimation of the SPM under MAR follows from the conventional joint model for longitudinal and time-to-event data, which exploits the decomposition of the full joint-likelihood function to conditional independent components given the random effects. More specifically, the likelihood contribution of the ith subject is given by the following:

$$
\int p\left(T_i^* \mid y_i, b_i; \psi^{T^*}\right) p\left(C_i \mid y_i; \psi^C\right) p\left(y_i \mid b_i; \theta\right) p\left(b_i; D\right) db_i,
\tag{3.1}
$$

The form of the conditional densities depends then on the model specification of each component. Let the risk of dropout for subject i depend on a function of the true underlying subject-specific trajectory $\eta_i(t)$. Then the instantaneous risk of dropout can be expressed as follows:

$$
\begin{aligned}
h_i\left(t \mid \mathcal{H}_i(t), \boldsymbol{b}_i, \boldsymbol{w}_i\right) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\left\{t \le T_i^* < t + \Delta t \mid T_i^* \ge t, \mathcal{H}_i(t), \boldsymbol{w}_i\right\} \\
&= h_0(t) \exp\left[\boldsymbol{\gamma}^\top \boldsymbol{w}_i + f\left\{\eta_i(t), \boldsymbol{b}_i, \boldsymbol{\alpha}\right\}\right], \quad t > 0,
\end{aligned} \tag{3.2}
$$

where $\mathcal{H}_i(t) = \{\eta_i(s), 0 \le s \le t\}$ denotes the history of the true underlying longitudinal process up to time t (and therefore depends on the random effects as well), $h_0(\cdot)$ is the baseline hazard function and $\boldsymbol{w}_i$ is a vector of baseline covariates with corresponding regression coefficients $\boldsymbol{\gamma}$. The association between the features of the longitudinal measurement process, $p\left(\boldsymbol{y}_i^o, \boldsymbol{y}_i^m \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right)$, and the instantaneous risk for dropout, $h_i\left(t \mid \mathcal{H}_i(t), \boldsymbol{b}_i, \boldsymbol{w}_i\right)$, is quantified by parameter vector $\boldsymbol{\alpha}$. Function $f(\cdot)$ may take various forms but for the remainder of the article we will assume that $f\left(\eta_i(t), \boldsymbol{b}_i, \alpha\right) = \alpha \eta_i(t)$, the so-called current value association, which postulates that the risk of dropout at time t depends only on the current value of the true underlying subject-specific trajectory at the same time $\eta_i(t)$. Let $\boldsymbol{u}_i$ be a vector of baseline covariates with corresponding regression coefficients $\boldsymbol{\gamma}_C$ and $g\left(\boldsymbol{y}_i^o\right)$ a function of the observed measurements. Then, analogously, the instantaneous risk to be censored can be expressed as follows:

$$
\begin{aligned}
\lambda_i\left(t \mid \boldsymbol{y}_i^o, \boldsymbol{u}_i\right) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\left\{t \le C_i < t + \Delta t \mid C_i \ge t, \boldsymbol{y}_i^o, \boldsymbol{u}_i\right\} \\
&= h_0(t) \exp\left[\boldsymbol{\gamma}_C^\top \boldsymbol{u}_i + g\left(\boldsymbol{y}_i^o\right)\right], \quad t > 0,
\end{aligned} \tag{3.3}
$$

since based on the assumption of non-informative censoring, it cannot depend on the missing observations and the random effects.

The probability of not dropping out from the study can then be obtained by the following:

$$
\begin{aligned}
S_i\left(t \mid \mathcal{H}_i(t), \boldsymbol{b}_i, \boldsymbol{w}_i\right) &= \Pr\left(T_i^* > t \mid \mathcal{H}_i(t), \boldsymbol{w}_i\right) \\
&= \exp\left(-\int_0^t h_0(s) \exp\left\{\boldsymbol{\gamma}^\top \boldsymbol{w}_i + \alpha \eta_i(s)\right\} ds\right)
\end{aligned} \tag{3.4}
$$

which means that unlike the instantaneous risk of dropout, the probability of remaining in the study depends on the whole true underlying subject-specific trajectory up to time t. Let $\mathcal{Y}_i^o(t) = \{\boldsymbol{y}_i^o(l), 0 \le l \le t\}$ denote the history of the observed repeated measurements. Then, analogously, the probability of not being

censored can be expressed as follows:

$$V_i\left(t \mid \mathcal{Y}_i^o\left(t\right), \boldsymbol{u}_i\right) = \Pr\left(C_i > t \mid \mathcal{Y}_i^o\left(t\right), \boldsymbol{u}_i\right)$$
$$= \exp\left(-\int_0^t \lambda_0\left(s\right)\exp\left\{\boldsymbol{\gamma}_C^\top\boldsymbol{u}_i + g\left(\boldsymbol{y}_i^o\right)\right\}ds\right) \tag{3.5}$$

omitting the conditioning on missing values and the random effects due to the non-informative censoring assumption. Finally, to distinguish between censored subjects and dropouts, let δ_i be an indicator variable which takes the value 1 if a subject dropped out from the study, that is, $T_i^* < C_i$ and 0 otherwise.

Under these assumptions and suppressing covariates $\boldsymbol{w}_i$ and $\boldsymbol{u}_i$ from the notation, the conditional density of the dropout process can be expressed as follows:

$$p\left(T_i^* \mid \boldsymbol{b}_i; \boldsymbol{\psi}^{T^*}\right) = \begin{cases} h_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right)S_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right), & \delta_i = 1 \\ S_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right), & \delta_i = 0, \end{cases} \tag{3.6}$$

while the conditional density of the censoring process is given by the following:

$$p\left(C_i \mid \boldsymbol{y}_i^o; \boldsymbol{\psi}^C\right) = \begin{cases} \lambda_i\left(t\right)V_i\left(t\right), & \delta_i = 0 \\ V_i\left(t\right), & \delta_i = 1, \end{cases} \tag{3.7}$$

Depending on the dropout and censoring status of the ith individual, the joint-likelihood in (3.1) can be written as follows:

$$\begin{cases} V_i\left(t\right)\int h_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right)S_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right)p\left(\boldsymbol{y}_i \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right)p\left(\boldsymbol{b}_i; \boldsymbol{D}\right)d\boldsymbol{b}_i, & \delta_i = 1 \\ \\ V_i\left(t\right)\lambda_i\left(t\right)\int S_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right)p\left(\boldsymbol{y}_i \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right)p\left(\boldsymbol{b}_i; \boldsymbol{D}\right)d\boldsymbol{b}_i, & \delta_i = 0. \end{cases} \tag{3.8}$$

Under (3.8), the model is MNAR for both uncensored and censored subjects. The latter is because even for censored subjects, the term $S_i\left(t \mid \mathcal{H}_i\left(t\right), \boldsymbol{b}_i\right)$ still appears in the likelihood decomposition. However, for the case that a subject is censored, this term equals to 1 reducing (3.8) to the following:

$$V_i\left(t\right)\lambda_i\left(t\right)\int p\left(\boldsymbol{y}_i \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right)p\left(\boldsymbol{b}_i; \boldsymbol{D}\right)d\boldsymbol{b}_i, \tag{3.9}$$

which due to ignorability further reduces to the following:

$$\int p\left(\boldsymbol{y}_i \mid \boldsymbol{b}_i; \boldsymbol{\theta}\right)p\left(\boldsymbol{b}_i; \boldsymbol{D}\right)d\boldsymbol{b}_i, \tag{3.10}$$

an MAR model for the observed longitudinal measurements. The transition from equation (3.9) to (3.10) holds due to the fact that the instantaneous risk to be censored $\lambda_i\left(t\right)$ and the probability of not being censored $V_i\left(t\right)$ are ignorable, since they strictly depend on the history of the observed repeated measurements and not the missing

ones (see equations (3.3) and (3.5)). Then when a subject is censored, instead of dropout, he/she is considered MAR allowing for straightforward sensitivity analysis since it can be achieved with existing software just by doing a data manipulation step and changing all the dropout indicators to zero.

As far as the true underlying longitudinal trajectory is concerned, let $y_i(t)$ be the contaminated, with measurement error ϵ_i, observed values of the true longitudinal trajectory, η_i. Then the longitudinal process can be expressed using a mixed-effects model defined as follows:

$$
\begin{cases}
y_i(t) = \eta_i + \epsilon_i(t), \\
\eta_i(t) = E\left[y_i(t) \mid b_i\right] = x_i^\top(t)\,\beta + z_i^\top(t)\,b_i, \\
\epsilon_i \sim \mathcal{N}(0, \Sigma_i), \\
b_i \sim \mathcal{N}(0, D), \\
b_i \perp\!\!\!\perp \epsilon_i,
\end{cases}
\tag{3.11}
$$

where the true longitudinal trajectory η_i is assumed to be a function of some general time-dependent design vectors $x_i^\top(t)$ and $z_i^\top(t)$ which are associated with a set of fixed-effects β and a set of subject-specific random effects b_i, respectively. The errors are typically assumed to be normally distributed with mean zero and a general variance—covariance matrix Σ_i which is commonly specified as $I \times \sigma^2$. The random effects b_i are also assumed to be normally distributed with mean zero and a general variance—covariance matrix D, and they are also assumed to be independent from the errors.

The conditional density for the longitudinal responses is then given by the following:

$$
p(y_i \mid b_i; \theta) = \left(2\pi\sigma^2\right)^{n_i/2} \exp\left\{-\|y_i - x_i\beta - z_i b_i\|^2/2\sigma^2\right\}
$$

with $\|x\| = \sum_i \left\{x_i^2\right\}^{1/2}$ denoting the Euclidean vector norm. The density of the random effects is, then, given by the following:

$$
p(b_i; D) = (2\pi)^{-q/2} \det(D)^{-1/2} \exp\left(-b_i^\top D^{-1} b_i/2\right)
\tag{3.12}
$$

Under the Bayesian framework, estimation of the parameters of the SPM is achieved using Markov chain Monte Carlo (MCMC) algorithms. Let $\Theta = \left\{\beta, \Sigma_i, \psi^{T^*}, \psi^C, D\right\}$ be a vector of unknown parameters, then the posterior distribution is proportional to the following:

$$
p(\Theta, b) \propto \prod_i p(y_i \mid b_i; \theta)\, p\left(T_i^* \mid b_i; \psi^{T^*}\right) p\left(C_i \mid y_i^o; \psi^C\right) p(b_i; D)\, p(\Theta)
\tag{3.13}
$$

The computation of the integrals involved in $p\left(T_i^* \mid b_i; \psi^{T^*}\right)$ and $p\left(C_i \mid y_i^o; \psi^C\right)$ is achieved using standard Gauss–Kronrod and Gauss–Legendre quadrature rules.

For the parameters in $\boldsymbol{\Theta}$, standard prior distributions are used. More specifically, for the vector of regression coefficients $\boldsymbol{\beta}$ of the longitudinal submodel and for the vector of regression coefficients of the dropout submodel $\boldsymbol{\gamma}$ we use independent univariate diffuse normal priors. For the covariance matrix of the random effects an inverse-Wishart prior is used, while for the variance of the error terms σ^2 an inverse-Gamma prior is used. For the baseline hazard, we use a B-spline approximation expressed as follows:

$$h_0\left(t\right) = \exp\left(\gamma_{h_0,0} + \sum_{q=1}^{Q} \gamma_{h_0,q}B_q\left(t,k\right)\right), \tag{3.14}$$

with $B_q\left(t,k\right)$ denoting the qth basis function of a B-spline with knots $k_1, \ldots, k_q$ and $\boldsymbol{\gamma}_{h_0}$ the vector of coefficient corresponding to the spline terms. To achieve the B-spline approximation to the baseline hazard under the Bayesian framework, the following prior is specified:

$$p\left(\boldsymbol{\gamma}_{h_0} \mid \tau_h\right) \propto \tau_h^{\rho(K)/2} \exp\left(-\frac{\tau_h}{2}\boldsymbol{\gamma}_{h_0}^{\top}K\boldsymbol{\gamma}_{h_0}\right) \tag{3.15}$$

In the above specification, τ_h is a smoothing parameter with a Gamma $(1, 0.005)$ hyper-prior, K is a penalty matrix and $\rho(K)$ is its rank. For more details, the reader is referred to Rizopoulos (2016).

4 The HIV CD4 data: A sensitivity analysis

As an illustrative case study, we will use data from a randomized clinical trial designed to compare the efficacy and safety of Didanosine (ddI) versus Zalcitabine (ddC) in HIV patients (Abrams et al., 1994; Goldman et al., 1996). In total, 467 advanced HIV patients were included in the study and were randomized to ddI (230; 49.2%) or ddC (237; 50.8%). The primary goal of the trial was to compare survival between the two treatment arms, while a secondary goal of the study was to investigate the association between CD4 cell count (a marker for the strength of the immune system) and the risk of death. For the latter goal, measurements of the CD4 cell count were recorded at scheduled visits at baseline, 2, 6, 12 and 18 months after the start of the study. Figure 3 shows the observed trajectories of the square root CD4 cell count. Until the end of the study, 184 (39.4%) died while only 1 405 (60.1%) out of the 2 335 planned visits were recorded, leading to 930 (39.9%) missing observations. In total, only 24(5.1%) patients were present for all five planned visits, 382(81.8%) dropped out from the study due to death or other reasons, and 61(13.0%) missed at least one planned visit without dropping out though from the remainder of the study. Table 1 shows the number of available measurements per dropout pattern per treatment arm, excluding the 61 cases of intermittent missingness.
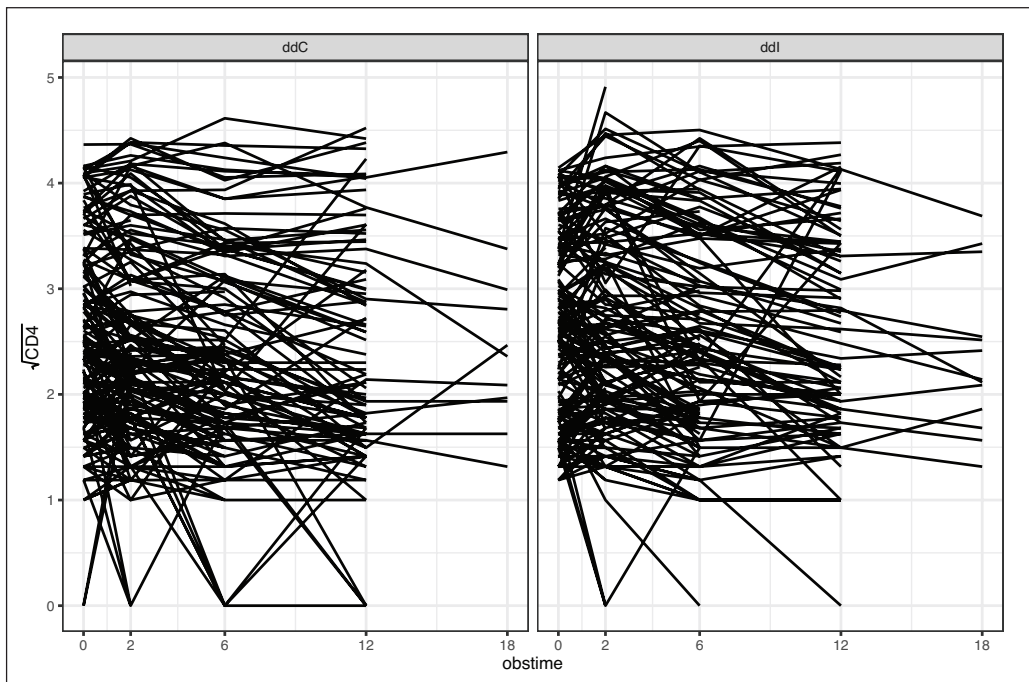
Figure 2 Observed longitudinal trajectories of $\sqrt{CD4}$ by treatment arm

We observe that dropout rates are similar between the treatment groups, the closer we look at the beginning of the study, but they slightly start to deviate towards the end of the study. Nevertheless, it is not possible to distinguish if the missing data are missing completely at random, MAR or MNAR. This highlights the importance of being able to explore different scenarios with respect to missing data in the SPM framework via sensitivity analysis.

To model the longitudinal measurements of CD4 cell count, we used a linear mixed-effects model with an interaction between time and ddI in the fixed-effects

Table 1 Number of observed CD4 measurements per dropout pattern and per treatment arm. 'OXXXX' denotes a dropout pattern where only the first planned measurement was obtained (denoted with 'O') while the rest were not (denoted with 'X'), whereas 'OOOOO' denotes a pattern for a completer (all planned measurements were obtained)

	ddC		ddI	
Dropout Pattern	N	%	N	%
OXXXX	29	14.4	32	15.6
OOXXX	35	17.4	37	18.0
OOOXX	41	20.4	47	22.9
OOOOX	85	42.3	76	37.1
OOOOO	11	5.5	13	6.3
Total	201	100	205	100

Table 2 Posterior means (standard errors) for an MAR and MNAR analyses

Effect	Parameter	Scenario 1	Scenario 2	
		MNAR	MNAR	MAR
Intercept	β_0	2.4671 (0.0640)	2.4680 (0.0629)	2.4423 (0.0647)
Time	β_1	−0.0655 (0.1701)	−0.0674 (0.1715)	−0.0399 (0.0050)
ddI	β_2	0.1038 (0.0908)	0.1067 (0.0879)	0.1188 (0.0946)
ddI × Time	β_3	0.0219 (0.2332)	0.0290 (0.2376)	0.0089 (0.0070)

structure. We also used random intercepts and random slopes to allow for subject-specific deviations. The square root of the CD4 cell count was used in order to meet the normality and homoscedasticity assumptions of the model, which is defined as follows:

$$\sqrt{y_{ij}^o} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1}) t_{ij} + \beta_2 \mathrm{ddI} + \beta_3 \left(t_{ij} \times \mathrm{ddI} \right) + \epsilon_{ij}, \qquad (4.1)$$

where $\boldsymbol{b}_i \sim \mathcal{N}(0, \boldsymbol{D})$, independent of $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$.

For the time-to-dropout T_i, we assumed a relative risk model of the form:

$$h_i \left(t \mid \mathcal{H}_i(t), \boldsymbol{b}_i \right) = h_0(t) \exp \left[\gamma_1 \mathrm{ddI} + \boldsymbol{\alpha} \eta_i(t) \right] \qquad (4.2)$$

where the baseline hazard $h_0(t)$ was approximated using penalized B-splines with three internal knots placed at quantiles of the observed event times (Eilers and Marx, 1996). The number of knots was kept small due to the low number of time points per subject.

These two sub models define the MNAR joint model for the CD4 cell count and the dropout. We then assumed two different scenarios with respect to the cause of dropout. Under Scenario 1, we assume no information concerning the cause of dropout, which means we treat dropout due to death or any other reason the same. Contrariwise in Scenario 2, we consider as dropouts only the subjects who died during the study. Finally, to achieve MAR, we assume that all the subjects are censored instead of dropping out. As shown in Section 3, this reduces the model to the mixed-effects sub model, which is considered MAR. All models were fitted using **R** version **3.6.1** and package **JMbayes** (Rizopoulos, 2016). Table 2 shows the parameter estimates after fitting all three different models.

While the results of all the models are close, there are quantitative differences which indicate that the findings might not be stable. This is especially the case when comparing the MAR model with both MNAR models. In Figures 3 and 4, the posterior means of the random effects estimated from the MAR model against the respective estimated random effects from each MNAR model are shown. Similarly, in Figure 5, the posterior means of the random effects estimated from the MNAR model under the 1st Scenario are plotted against the respective estimated random effects from the MNAR model under the 2nd Scenario. These plots give insight into the differences between the three models. We see that especially for the slope components, the differences in the random effects' estimates are more intense between the MAR
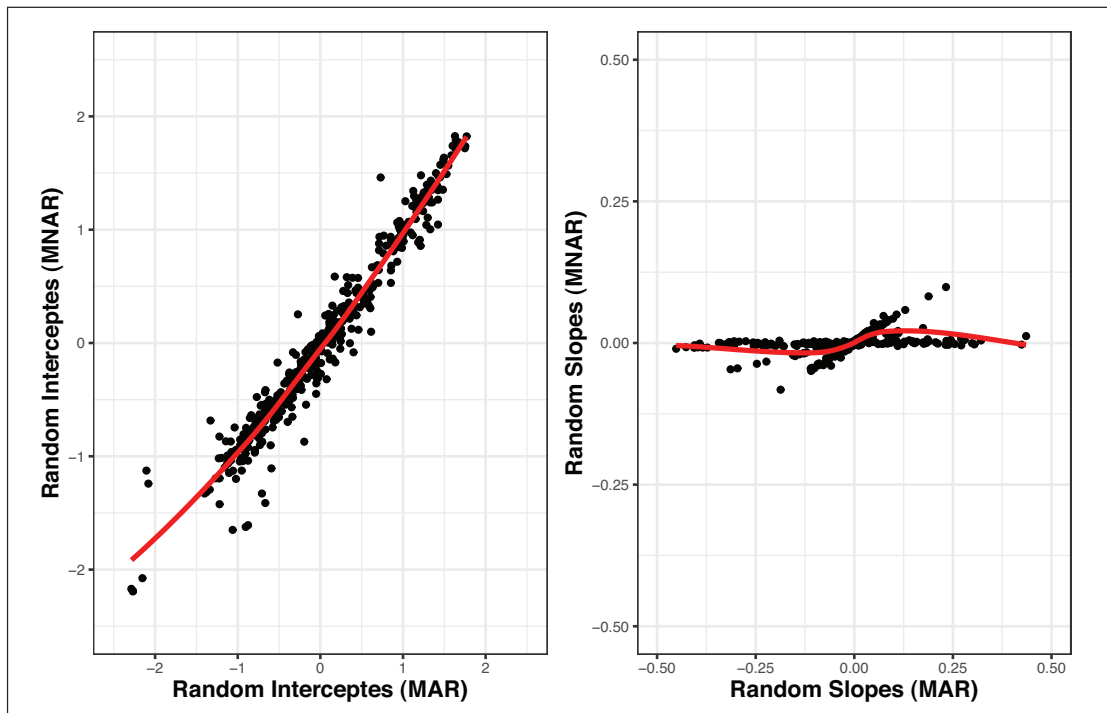
Figure 3 Scenario 1: Posterior means of the random effects under MNAR and MAR

model and the MNAR models under both scenarios. These differences become weaker when we compare the MNAR model under the 1st Scenario and the MANR model under the 2nd Scenario, which utilizes the information and distinguishes between dropout due to an event and dropout due to any other reason. This highlights the importance of sensitivity analysis in this context since the differences in the random effects estimates would translate to differences in the subject-specific predictions based on each of the models. The results are stable under different MNAR scenarios, but substantial differences are observed when using an MAR model.

5 Simulation study

We conducted a simulation study to evaluate the performance of the following three models (used in the analysis of the HIV CD4 data):

- Model 1: all dropout cases considered MNAR (as the MNAR model used under Scenario 1 MNAR in the analysis of the HIV dataset),
- Model 2: all dropout cases considered MAR,
- Model 3: dropout cases considered MNAR if dropout due to the event or MAR otherwise (as the MNAR model used under Scenario 1 MNAR in the analysis of the HIV dataset),
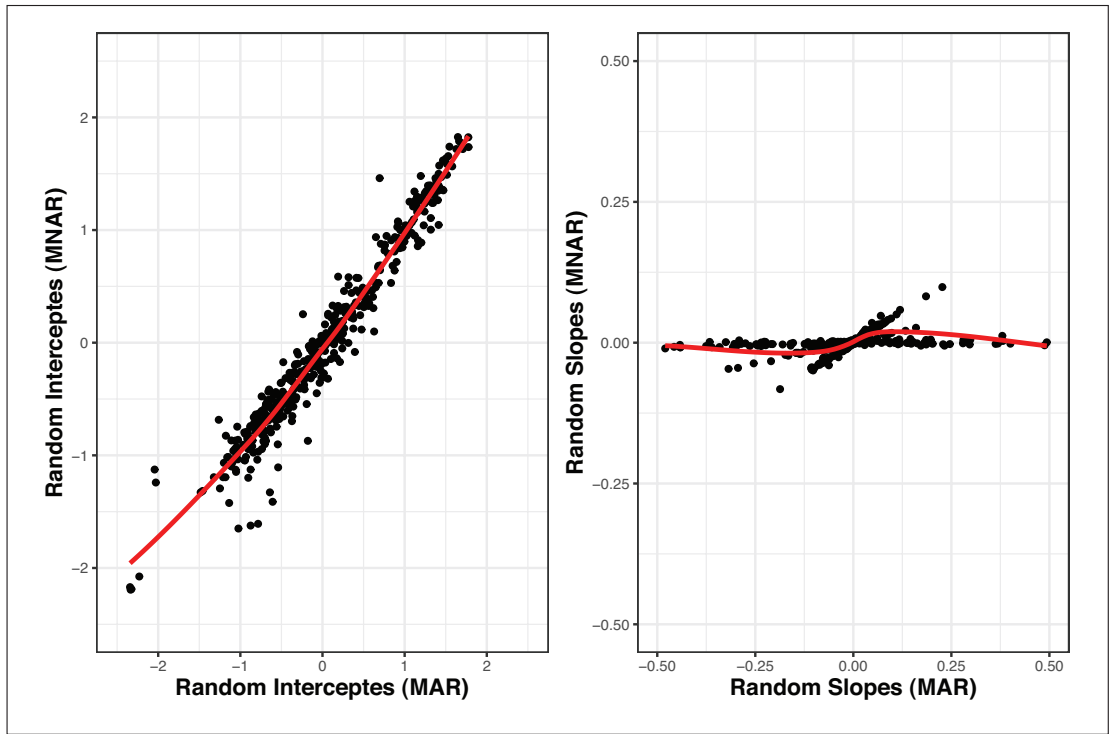
Figure 4 Scenario 2: Posterior means of the random effects under MNAR and MAR

under different scenarios for the amount of MAR and MNAR dropouts. More specifically, assuming a total dropout rate of 50%, we considered three different scenarios concerning the amount of dropout type: (a) 5% MAR dropouts versus 45% MNAR dropouts, (b) 25% MAR dropouts versus 25% MNAR dropouts and (c) 45% MAR dropouts versus 5% MNAR dropouts. Our goal is to investigate the behaviour between these models under each scenario.

We assumed 600 subjects and then randomly selected follow-up visits, t_{ij} from a uniform distribution between 0 and 10. We set the total number of measurements per subject to 10, but the final number of measurements may vary depending on whether a subject was a dropout or not. The type of dropout was then determined as follows: if the value of the longitudinal trajectory of a subject exceeded a pre-specified value, then the next time point was selected as the candidate MAR dropout time, which was then compared to the candidate MNAR dropout time. We simulated the MNAR dropout time using a joint model. Depending on which of these time points occurred sooner (if occurred), the subject was classified as MAR dropout, MNAR dropout or completer. The threshold value for MAR varied between the three scenarios in order to attain the target percentages of dropout per type.
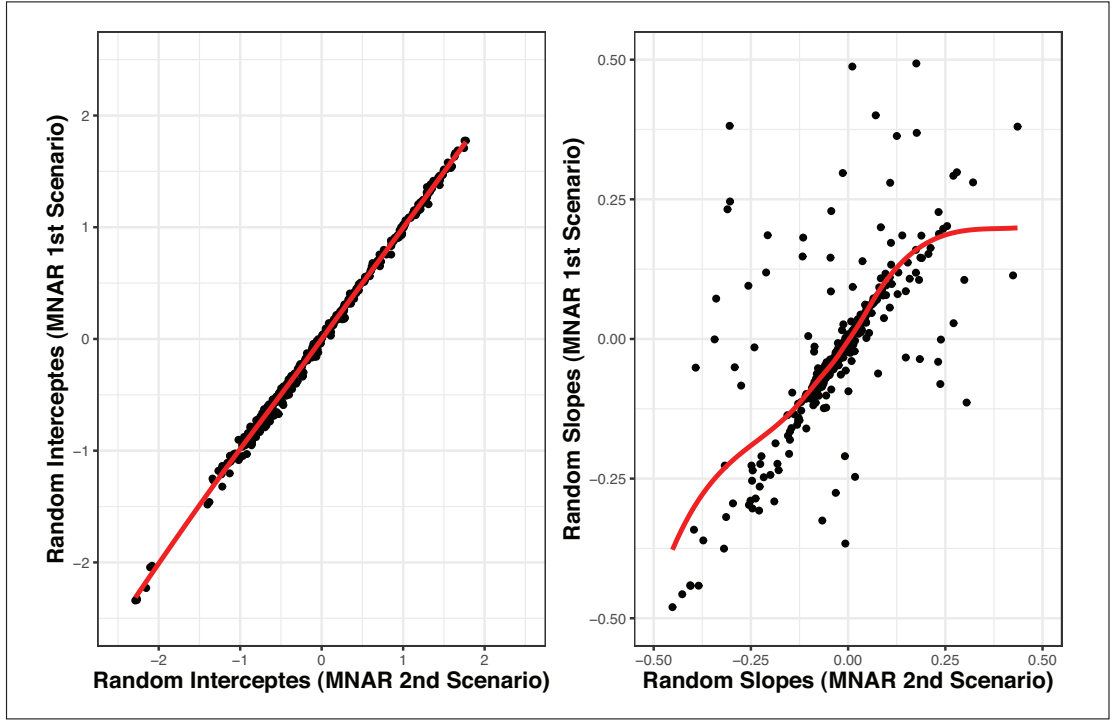
Figure 5 Scenario 2: Posterior means of the random effects under MNAR (1st Scenario) and MNAR (2nd Scenario)

For the continuous longitudinal outcome, the data were simulated from a linear mixed-effects model as follows:

$$y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1}) t_{ij} + \beta_2 \text{Xcov} + \beta_3 (\text{Xcov} \times t_{ij}) + \epsilon_{ij}, \qquad (5.1)$$

where Xcov is a categorical variable with two groups, $\epsilon_i(t) \sim \mathcal{N}\left(0, \sigma^2 I_{n_i}\right)$ and $b \sim \mathcal{N}(0, D)$. The parameter values used are as follows: $\beta_0 = 2.47$, $\beta_1 = -0.067$, $\beta_2 = 0.107$, $\beta_3 = 0.029$, $\sigma = 1.431$ and for the variance—covariance matrix of the random effects:

$$D = \begin{bmatrix} d_{11} = 0.765 & d_{12} = -0.038 \\ d_{21} = -0.038 & d_{22} = 4.910 \end{bmatrix}$$

The values of the parameters were based on the analysis of the HIV CD4 data. For time-to-dropout, we assumed a relative risk model of the form:

$$h_i(t \mid \mathcal{H}_i(t), b_i) = h_0(t) \exp\left[\gamma_1 \text{Xcov} + \boldsymbol{\alpha}\eta_i(t)\right] \qquad (5.2)$$

where the baseline hazard was simulated from a Weibull distribution $h0\,(t) = \xi t^{\xi-1}$ with $\xi = 3.765$.

We simulated 500 datasets per scenario, and we fitted each of the three models on each dataset. To assess the behaviour of the three models, we calculated the Bias and Root Mean Squared Error (RMSE) for both the regression coefficients, and the variance parameter estimates. Table 3 summarizes the results for the regression coefficients, while Table 4 summarizes the results for the variance parameters. The results suggest that when the major cause of missingness is MAR, Models 2 and 3 perform similarly and slightly better than Model 1, as far as the regression coefficients are concerned. As the amount of MNAR dropout increases, the three models seem to have similar performance in the estimation of the regression coefficients with only slight differences.

On the other hand, when looking at the Bias and RMSE for the estimates of variance parameters for the random effects, there are more distinct differences between the models. More specifically, the results suggest that Models 1 and 3 perform better than Model 2. Moreover, for the variances of the random effects, under Model 2, the Bias seems to increase as the balance between dropout types moves from MAR dominance to MNAR dominance, whereas the case is the opposite for Model 3. The differences in the variance parameters' estimates of the random effects between the models also imply that there are considerable differences in the estimated random effects. The difference in the random effects means that the subject-specific predictions derived from these models will differ.

6 Discussion

In this article, we have proposed an alternative characterization of MAR for the conventional SPM and proposed its application as a sensitivity analysis tool towards MNAR deviations from the MAR assumption. In doing so, we did not broaden the definition of the SPM by adding additional random effects structures and hence retaining the computational feasibility of the model.

Furthermore, we argued how the subject-specific nature of the MAR characterization in SPM comes with the advantage of more flexible comparisons regarding the causes of missingness. This is an important advantage over other MNAR frameworks such as the PMM and SeM since it allows distinguishing groups of subjects as MAR or MNAR depending on the information available for the causes of missingness. This was illustrated in the application to the HIV data, where we were able to consider different MNAR scenarios with respect to the information available on the causes of missingness.

Finally, it should be mentioned that this characterization of MAR in SPMs allows for potential extensions to account for multiple causes of missingness by using for example a multi-state model and/or generalized mixed-effects models for categorical longitudinal outcomes. Dropout is present in most longitudinal studies, and our approach showed how one could perform a sensitivity analysis using standard joint

Table 3 Mean, Bias and RMSE for the regression coefficients from the three models under different scenarios for the amount of dropout caused by each missing data generating mechanism

Parameter	Scenario 1: 5% MNAR - 45% MAR			Scenario 2: 25% MNAR - 25% MAR			Scenario 3: 45% MNAR - 5% MAR		
	Mean	Bias	RMSE	Mean	Bias	RMSE	Mean	Bias	RMSE
Model 1									
β_0	2.4532	−0.0147	0.0753	2.4667	−0.0012	0.0715	2.4705	0.0025	0.0711
β_1	0.0166	0.0840	0.1477	−0.0088	0.0585	0.1339	−0.0129	0.0544	0.1322
β_2	−0.0091	−0.1158	0.1819	−0.0109	−0.1176	0.1827	−0.0114	−0.1181	0.1821
β_3	0.1056	0.0766	0.1045	0.1045	0.0755	0.1002	0.1048	0.0758	0.0993
Model 2									
β_0	2.4688	0.0009	0.0738	2.4707	0.0028	0.0716	2.4723	0.0043	0.0712
β_1	−0.0136	0.0538	0.1322	−0.0127	0.0547	0.1322	−0.0139	0.0535	0.1319
β_2	−0.0106	−0.1174	0.1808	−0.0107	−0.1174	0.1825	−0.0116	−0.1183	0.1823
β_3	0.1060	0.0770	0.1043	0.1051	0.0761	0.1002	0.1046	0.0756	0.0994
Model 3									
β_0	2.4680	0.0001	0.0738	2.4687	0.0007	0.0717	2.4694	0.0014	0.0713
β_1	−0.0129	0.0545	0.1325	−0.0115	0.0558	0.1328	−0.0121	0.0553	0.1323
β_2	−0.0107	−0.1174	0.1810	−0.0105	−0.1172	0.1827	−0.0114	−0.1181	0.1820
β_3	0.1060	0.0770	0.1042	0.1051	0.0761	0.1005	0.1047	0.0757	0.0995

Table 4 Mean, Bias and RMSE for the variance parameters from the three models under different scenarios for the amount of dropout caused by each missing data generating mechanism

Parameter	Scenario 1: 5% MNAR - 45% MAR			Scenario 2: 25% MNAR - 25% MAR			Scenario 3: 45% MNAR - 5% MAR		
	Mean	Bias	RMSE	Mean	Bias	RMSE	Mean	Bias	RMSE
Model 1									
d_{11}	0.7577	−0.0074	0.0910	0.7397	−0.0254	0.0848	0.7394	−0.0257	0.0824
d_{12}	−0.0514	−0.0134	0.1206	−0.0419	−0.0040	0.1132	−0.0265	0.0115	0.1107
d_{22}	4.7665	−0.1439	0.3248	4.6891	−0.2214	0.3543	4.6768	−0.2337	0.3600
σ	1.4318	0.0003	0.0182	1.4314	0.0000	0.0160	1.4293	−0.0022	0.0157
Model 2									
d_{11}	0.1406	−0.6245	0.6258	0.1381	−0.6270	0.6279	0.1381	−0.6270	0.6278
d_{12}	−0.0222	0.0158	0.0742	−0.0208	0.0171	0.0708	−0.0202	0.0178	0.0694
d_{22}	5.4293	0.5188	0.8668	5.4355	0.5251	0.8401	5.4484	0.5380	0.8477
σ	1.4296	−0.0018	0.0180	1.4302	−0.0013	0.0159	1.4301	−0.0014	0.0155
Model 3									
d_{11}	0.7469	−0.0182	0.0908	0.7427	−0.0224	0.0841	0.7392	−0.0259	0.0829
d_{12}	−0.0271	0.0108	0.1184	−0.0321	0.0059	0.1133	−0.0343	0.0037	0.1098
d_{22}	4.6621	−0.2484	0.3817	4.6715	−0.2389	0.3643	4.6798	−0.2306	0.3576
σ	1.4299	−0.0015	0.0181	1.4300	−0.0014	0.0160	1.4299	−0.0016	0.0156

modelling software to explore different causes of missingness. By doing so, we do hope that such sensitivity analyses will become standard procedure and will be routinely reported in longitudinal studies with dropout.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

References

Abrams DI, Goldman AI, Launer C, Korvick JA, Neaton JD, Crane LR, Grodesky M, Wakeeld S, Muth K, Kornegay S, Cohn DL, Harris A, Luskin-Hawk R, Markowitz N, Sampson JH, Thompson M and Deyton L (1994) A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine*, **330**, 657–62.

Creemers A, Hens N, Aerts M, Molenberghs G, Verbeke G and Kenward MG (2010) A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes. *Biometrical Journal*, **52**, 111–25.

——— (2011) Generalized shared-parameter models and missingness at random. *Statistical Modelling*, **11**, 279–310.

Eilers P and Marx B (1996) Flexible smoothing with B-splines and penalties. *Statistical Science*, **11**, 89–121.

Goldman A, Carlin B, Crane L, Launer C, Korvick J, Deyton L and D A (1996) Response of CD4+ and clinical consequences to treatment using ddI or ddC in patients with advanced HIV infection. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, **11**, 161–69.

Hogan JW and Laird NM (1997) Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine*, **16**, 239–57.

——— (1998) Increasing efficiency from censored survival data by using random effects to model longitudinal covariates. *Statistical Methods in Medical Research*, **7**, 28–48.

Little RJA and Rubin DB (2002) *Statistical Analysis with Missing Data*. New York, NY: Wiley.

Molenberghs G and Kenward MG (2007) *Missing Data in Clinical Studies*. Chichester: Wiley.

Molenberghs G and Vereke G (2005). *Models for Discrete Longitudinal Data*. New York, NY: Springer.

Molenberghs G, Michiels B, Kenward MG and Diggle PJ (1998) Monotone missing data and pattern-mixture models. *Statistica Neerlandica*, **52**, 153–61.

Molenberghs G, Beunckens C, Sotto C and Kenward MG (2008) Every missingness not at random model has a missingness at random counterpart with equal fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**, 371—88.

Njagi EN, Molenberghs G, Kenward MG, Verbeke G and Rizopoulos D (2014) A characterization of missingness at random in a generalized shared-parameter joint modeling framework for longitudinal and time-to-event data, and sensitivity analysis. *Biometrical Journal*, **56**, 1001–15.

Rizopoulos D (2016) The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. *Journal of Statistical Software*, **72**, 1–46.

Rubin DB (1976) Inference and missing data. *Biometrika*, **63**, 581–92.

Tsiatis AA (1975) A characterization of missingness at random in a generalized shared-parameter joint modeling framework for longitudinal and time-to-event data, and sensitivity analysis. *Proceedings of the National Academy of Sciences*, **72**, 20–22.

Wu MC and Bailey K (1988) Analysing changes in the presence of informative right censoring caused by death and withdrawal. *Statistics in Medicine*, **7**, 337–46.

——— (1989) Estimation and comparison of changes in the presence of informative right censoring: Conditional linear model. *Biometrics*, **45**, 939–55.

Wu MC and Carroll RJ (1988) Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, **44**, 175–88.