Joint modelling with competing risks of dropout for longitudinal analysis of health-related quality of life in cancer clinical trials

Benjamin Cuer, Thierry Conroy, Beata Juzyna, Sophie Gourgou, Caroline Mollevi, Célia Touraine

Joint modelling with competing risks of dropout for longitudinal analysis of health-related quality of life in cancer clinical trials

Author information

Benjamin Cuer[1,2,3], Thierry Conroy[4,5], Beata Juzyna[6], Sophie Gourgou[1,2], Caroline Mollevi[1,2,7], Célia Touraine[1,2,7]

[1]Montpellier Cancer Institute (ICM), University of Montpellier, Montpellier, France

[2]French National Platform Quality of Life and Cancer, Montpellier, France

[3]Institute of Cancer Research of Montpellier (IRCM) (Inserm 1194), ICM, University of Montpellier, Montpellier, France

[4]Institut de Cancérologie de Lorraine (Department of Medical Oncology), Vandoeuvre-lès-Nancy, France

[5]APEMAC (équipe MICS), Université de Lorraine, Nancy, France

[6]R&D UNICANCER, French Federation of Comprehensive Cancer Centres, Paris, France

[7] Desbrest Institute of Epidemiology and Public Health (IDESP), UMR Inserm, University of Montpellier, Montpellier, France

Corresponding author: Benjamin Cuer, Biometrics Unit, Montpellier Cancer Institute, 208, avenue des Apothicaires, 34298 Montpellier, France.

Mail: cuerbenjamin@gmail.com; ORCID: 0000-0002-5634-0704

Abstract

Purpose

Health-related quality of life (HRQoL) is an important endpoint in cancer clinical trials. Analysis of HRQoL longitudinal data is plagued by missing data, notably due to dropout. Joint models are increasingly receiving attention for modelling longitudinal outcomes and the time-to-dropout. However, dropout can be informative or non-informative depending on the cause.

Methods

We propose using a joint model that includes a competing risks sub-model for the cause-specific time-to-dropout. We compared a competing risks joint model (CR JM) that distinguishes between two causes of dropout with a standard joint model (SJM) that treats all the dropouts equally. First, we applied the CR JM and SJM to data from 267 patients with advanced oesophageal cancer from the randomized clinical trial PRODIGE 5/ACCORD 17 to analyse HRQoL data in the presence of dropouts unrelated and related to a clinical event. Then, we compared the models using a simulation study.

Results

We showed that the CR JM performed as well as the SJM in situations where the risk of dropout was the same whatever the cause. In the presence of both informative and non-informative dropouts, only the SJM estimations were biased, impacting the HRQoL estimated parameters.

Conclusion

The systematic collection of the reasons for dropout in clinical trials would facilitate the use of CR JMs, which could be a satisfactory approach to analysing HRQoL data in presence of both informative and non-informative dropout.

Trial registration: This study is registered with ClinicalTrials.gov, number NCT00861094.

Keywords:

Joint modelling, Competing risks, Health-related quality of life, Missing data, Clinical trials, Oncology.

1. Introduction

Health-related quality of life (HRQoL) is an important endpoint in cancer clinical trials. HRQoL is a patient-reported outcome used to investigate the clinical benefit of new treatment strategies in oncology. This endpoint, reflecting the patients' perception of their health status, is measured by self-administered questionnaires collected longitudinally throughout the care process. Linear mixed modelling is classically used to analyse HRQoL longitudinal data because it is a flexible and powerful approach to assessing the impact of the new treatment strategy on quality of life [1, 2]. Nevertheless, HRQoL assessment may be incomplete due to missing or incomplete questionnaires. High rates of missing HRQoL data are common in clinical trials, and most studies ignore this missing data or do not report the method used to handle them. Notably, when the disease and treatment are aggressive and toxic, monotone missing data may occur from dropout due to toxicity, disease progression or death. This form of dropout can be informative, leading to a biased estimate of the treatment effect [3]. In order to produce valid results in the analysis of the longitudinal outcome, it is crucial to consider the missing data mechanism [4–6].

Joint modelling has received much attention in various contexts and can be used to account for informative dropout when modelling a longitudinal outcome. The idea is to define two sub-models (a linear mixed model for the longitudinal outcome and a survival model for the time-to-dropout) linked through an association structure [7]. The association between the two processes is captured using a function of the shared random effects from the linear mixed model as a covariate in the survival model.

In advanced cancer or palliative care, a patient may drop out before the end of the clinical trial due to various reasons. On one hand, the HRQoL of the patients is generally related to the occurrence of clinical events (such as death or disease progression), leading to informative dropouts that should be taking into account in the analysis [8, 9]. On the other end, HRQoL being often relegated to a secondary focus, monotone missing data could arise independently from HRQoL. In this situation, a standard joint model (SJM) that assumes the association between dropout and HRQoL to be the same whatever the causes of dropout is likely to be inappropriate. Such modelling can produce biased estimates of the HRQoL parameters, and therefore unreliable conclusions from the predicted trajectories of the mean HRQoL score in the different treatment arms and an incorrect understanding of the association between the HRQoL outcome and the dropout process.

An alternative, more satisfactory approach would be to use a time-to-dropout model that considers competing risks, allowing the study of multiple causes of dropout that have different associations with the longitudinal HRQoL outcome. Assuming different types of dropout may also shed new light on the association between dropout and HRQoL outcome. Recent developments have extended SJMs, enabling them to consider more complex data, such as considering multiple events for the time-to-event model [10, 11]. Although underused in clinical contexts, such models appear to be of particular interest to distinguish between different types of dropout [12].

The aim of this paper is to illustrate the potential of using a competing risks joint model (CR JM) to analyse HRQoL longitudinal data in the presence of different types of dropout and investigate the impact of using a SJM instead of a CR JM in various situations. Section 2 describes the two joint models. In Section 3, we apply the models to data from 267 patients with advanced oesophageal cancer from the randomized clinical trial PRODIGE 5/ACCORD 17. Section 4 presents a simulation study evaluating the performance of the two joint models according to various scenarios and highlights the situations where the SJM may not be suitable. Finally, we conclude with a discussion in Section 5.

2. Models

2.1 General presentation and notations

The joint models used consist of two sub-models: a linear mixed model for the continuous longitudinal HRQoL score and a model for the time-to-dropout, linked by an association structure involving common parameters. The sub-model for the longitudinal outcome is common to both joint models. The two processes are assumed to be conditionally independent given the random effects. In the SJM the dropout is considered the only possible event through a survival model. In the CR JM, we consider the dropout as a multiple event through a competing risks model to distinguish between two possible causes of dropout.

Let $\{y_i(t), t \geq 0\}$ be the process for the HRQoL score of patient i at time t. We can observe this process as long as $t < T_i$, where T_i is the time-to-dropout or the right-censored time (in cases of no dropout).

2.2 Linear mixed sub-model

We assume that the observed longitudinal HRQoL score of patient i measured at time t, $y_i(t)$ is composed of the unobserved true value $y_i^\star(t)$ and a residual error term $\varepsilon_i(t)$. More specifically, we use a linear mixed model that

includes fixed effects and random subject-specific effects to take into account the repeated measurements on the same patient, that is:

$$\begin{aligned} y_i(t) &= y_i^\star(t) + \varepsilon_i(t) \\ &= X_i^\top(t)\beta + Z_i^\top(t)b_i + \varepsilon_i(t) \end{aligned} \quad (1)$$

where $X_i(t)$ and $Z_i(t)$ denote the design matrices of the fixed effects β and random effects b_i, respectively. The random effects are assumed to be normally distributed with mean zero and variance-covariance matrix D, independent of the measurement errors which are assumed to be independent and normally distributed with mean zero and variance σ^2.

In our context of clinical trials, we assumed a linear time effect of the HRQoL score which depends on the treatment arm and random intercept and slope; that is, the following linear mixed model:

$$y_i(t) = \beta_0 + \beta_1 t + \beta_2\{arm_i \times t\} + b_{0i} + b_{1i}t + \varepsilon_i(t) \quad (2)$$

where arm_i is the arm indicator for patient i taking a value of 0 or 1 for the control or experimental treatment arm, respectively; $\beta = (\beta_0, \beta_1, \beta_2)$ is the vector of fixed effects, with β_0 the intercept that represents the mean score at inclusion ($t = 0$), β_1 the slope that represents the score change by unit of time in the control arm and β_2 the interaction parameter that represents the difference between the slopes of the experimental and control arms; and $b_i = (b_{0i}, b_{1i})$ is the vector of random effects, with b_{0i} the random intercept that represents the individual deviations from the fixed intercept, b_{1i} the random slope that represents the individual deviations from the fixed slope, and $D = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_{01}} \\ \sigma_{b_{01}} & \sigma_{b_1}^2 \end{pmatrix}$.

2.3 Standard joint model (SJM)

In the SJM, we use a proportional hazards model as the sub-model for the time-to-dropout, with a hazard function expressed as:

$$h_i(t) = h_0(t)\exp\{\alpha y_i^\star(t) + W_i^\top\gamma\} \quad (3)$$

where h_0 is the baseline hazard function, α is the parameter that characterizes the association between the risk of dropout and $y_i^\star(t)$ is the current true value of the HRQoL score, W_i denotes the vector of covariates with γ the corresponding vector of coefficients. Note that if the association coefficient $\alpha = 0$, the two sub-models are no longer linked. Thus, an estimated association coefficient which is not significantly different from zero would traduce a non-informative dropout. In this model, the instantaneous risk of dropout is multiplied by the hazard ratio $\exp(\alpha)$ for one unit increase of the current HRQoL score.

In our context of clinical trials, we included only the treatment arm as a covariate. The hazard function reduces to:

$$h_i(t) = h_0(t)\exp\{\alpha y_i^\star(t) + \gamma arm_i\} \quad (4)$$

where the baseline hazard h_0 is assumed to be parametric, and γ is the arm effect on the risk of dropout.

2.4 Competing risks joint model (CR JM)

In the CR JM, we use a competing risks model as the sub-model for the time-to-dropout to consider two possible causes of dropout. We use a proportional hazards model for each competing cause k of dropout, $k = 1,2$, that is, cause-specific hazard given by:

$$h_{ik}(t) = h_{0k}(t)\exp\{\alpha_k y_i^\star(t) + \gamma_k arm_i\} \quad (5)$$

where the baseline hazard functions h_{0k} are assumed to be parametric, and α_k and γ_k denote the association parameters and cause-specific arm effects, respectively.

In our context of clinical trials, we considered two types of dropout: dropouts unrelated (k=1) or related (k=2) to a clinical event such as death or disease progression. The underlying idea is that the latter are likely to be informative, in contrast with the former.

3. Application

3.1 Data description

The clinical trial PRODIGE 5/ACCORD 17 was a randomized phase 2/3 trial that included 267 patients with advanced oesophageal cancer and compared the chemoradiotherapy regimens of FOLFOX (n=134) and fluorouracil–cisplatin (n=133). The analysis for the primary efficacy endpoint of the phase 3 in the intent-to-treat population (n=267) showed no difference in progression-free survival between the two treatment arms [13].

The secondary endpoints were overall survival, endoscopic complete response, time-to-treatment failure, occurrence of grade 3 or 4 toxicities and quality of life. HRQoL was assessed using the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ-C30 version 3.0) [14]. The QLQ-C30 is a self-administered questionnaire composed of 30 items evaluating five functional scales, nine symptomatic scales and the global health status/quality of life (GHS/QoL) scale. For each scale, a standardized score from 0 to 100 can be calculated according to the EORTC QLQ-C30 scoring manual [15]. A high score for the functional and GHS/QoL scales reflects high functional capacities and a good level of HRQoL. Conversely, a high score for the symptom scales reflects high levels of symptoms and a poor level of HRQoL. The questionnaire was completed at eight evaluations defined in the protocol: at inclusion, during treatment at months 1.25, 3 and 4 (first evaluation of treatment efficacy) and during follow-up at months 6, 12, 24 and 36.

Four dimensions of the QLQ-C30 were pre-specified in the protocol as targeted dimensions. First, the objective was to compare GHS/QoL levels (QL scale) between treatment arms; this was followed by comparisons of physical functioning (PF scale), pain (PA scale) and fatigue (FA scale). Analysis of the HRQoL scores using linear mixed models showed no difference between the two regimens [16]. However, this analysis did not account for potentially informative dropout.

3.2 Application of the joint models

All analyses were performed in the evaluable intent-to-treat population: the evaluable population for a given scale included patients with at least one available HRQoL score from any measurement time. We analysed the four scales of interest included as secondary endpoints for the analysis of HRQoL, namely QL, PF, PA and FA. The evaluable intent-to-treat population reduces to 252 for the QL scale (experimental arm: n=130; control arm: n=122) and 254 for the PF, PA and FA scales (experimental arm: n=131; control arm: n=123).

Intermittent missing data of score were assumed to be missing at random. Monotone missing data of score were assumed to be missing not at random, that is, assimilated to a dropout event at the last HRQoL evaluation which was taken into account in a joint model analysis. Specifically, the purpose of this application was to analyse the evolution of the HRQoL scores taking into account two possible causes of dropout and their association with HRQoL.

We applied the two joint models described in Section 2 to each of the four scales of interest: first, we applied the SJM in Equation (4), for which dropout is considered the only possible event; then, we applied the CR JM in Equation (5) to distinguish between two causes of dropout – i.e., whether the dropouts were unrelated (cause 1) or related (cause 2) to death or disease progression (assessed by RECIST version 1.0) [17]. In the two models, the longitudinal scores followed Equation (2).

Dropouts were attributed to cause 2 if death or progression occurred between the last evaluation visit with an available score and the subsequent planned visit, and attributed to cause 1 otherwise.

The estimates of the joint models were obtained by likelihood maximization using the function jointModel() from the R package JM (R software, version 3.4.0) [18, 19]. The argument CompRisk was set to TRUE to fit a CR JM instead of a SJM. In both joint models, we used a cubic B-splines approximation for the log baseline hazard function with five internal knots placed at the quantiles of the observed event times (QL: 1.23, 2.98, 3.99, 9.43

and 16.14 months; versus PF, PA and FA: 1.35, 3.22, 4.07, 9.66 and 16.20 months). The integrals over the random effects were approximated using a nine-point adaptive Gauss–Hermite quadrature.

3.3 Results

No patient had a complete longitudinal profile (i.e., 8 available scores out of the 8 planned assessment visits) in the QL scale, and only 1 in the PF, PA, and FA scales. The amount of HRQoL evaluations by scale is more fully described in Supplementary Table 1. Overall, 95% of patients (QL: 240/252; PF, PA and FA: 242/254) dropped out before the last planned questionnaire and 5% of patients (QL: 12/252; PF, PA, and FA: 12/254) completed the last planned questionnaire, generating right-censored times-to-dropout. Among the total dropouts, 57% (137/240 for QL and 139/242 for PF, PA and FA) were attributed to cause 1 and 43% (103/240 for QL and 103/242 for PF, PA and FA) were attributed to cause 2. The results of the SJM and CR JM are summarized in table 1.

[insert Table1]

Results of the sub-model for the time-to-dropout

Both models found that there was no effect of the treatment arm on the risk of dropout (non-significant γ for the SJM and non-significant γ_1 and γ_2 for the CR JM) for the four analysed dimensions. The SJM detected a significant association between the current HRQoL score and the risk of dropout for three scales: QL ($\hat{\alpha} = -0.016,\ p = 0.050$), PF ($\hat{\alpha} = -0.019,\ p = 0.001$) and FA ($\hat{\alpha} = 0.013,\ p = 0.007$), suggesting an informative dropout. For example, a decrease of 10 points in the current PF score corresponded to an increase of $\exp(10 \times 0.019) = 1.21$ (95% CI = 1.11 – 1.35) in the instantaneous risk of dropout. The CR JM found a significant association between the current HRQoL score and the risk of dropout only for the PF and FA scales, where only dropout due to cause 1 was associated with the current HRQoL score: PF ($\hat{\alpha}_1 = -0.024, p = 0.001$) and FA ($\hat{\alpha}_1 = 0.016, p = 0.011$). For example, a decrease of 10 points in the current PF score corresponded to an increase of $\exp(10 \times 0.024) = 1.27$ (95% CI = 1.11 – 1.49) in the instantaneous risk of dropout unrelated to clinical event. In fact, these results suggest that for the PF and FA scales, dropout from cause 1 (unrelated to death or progression) was informative whereas dropout from cause 2 (related to death or progression) was non-informative. By contrast, the CR JM did not detect any association between the risk of dropout and the HRQoL for the QL or PA scales. Note that because of a smaller number of events when dividing the dropout events into two categories, the width of the confidence intervals of the cause-specific parameters from the CR JM were larger than the corresponding parameters from the SJM.

Results of the sub-model for the longitudinal outcome

Both the SJM and CR JM found a similar and significant increase in the score over time for the QL scale (SJM: $\hat{\beta}_1 = 0.311$, $p = 0.034$; CR JM: $\hat{\beta}_1 = 0.313$, $p = 0.031$) and a decrease in the score over time for the PF scale (SJM: $\hat{\beta}_1 = -0.512$, $p = 0.017$; CR JM: $\hat{\beta}_1 = -0.505$, $p = 0.017$). In addition, neither of the two models detected a significant difference between the HRQoL score trajectories of the two treatment arms. Globally, the estimates of the two models were very similar, even in the PF and FA scales where the dropouts from only cause 1 had been found informative. As an illustration, one can see that the predicted trajectories from the SJM and the CR JM were almost superimposed (see Figure 1).

[insert Fig. 1]

Fig. 1 Predicted HRQoL trajectories for the four dimensions of the EORTC QLQ-C30 (QL, PF, PA and FA) according to the SJM and CR JM

4. Simulation study

We conducted a simulation study to evaluate the performance of the joint model that considers competing risks of dropout. We compared the estimated parameters of the CR JM with those of the SJM (which does not account for different causes of dropout) using bias and variance criteria derived from a series of 1000 generated datasets of 500 patients equally assigned to two treatment arms.

4.1 Data generation

The data generation was based on the clinical trial PRODIGE 5/ACCORD 17 (see Section 3).

We considered seven planned HRQoL assessment times $t_j, j = 1, \dots, 7$, at: inclusion ($t_j = 0$), 1.5, 3, 6, 12, 24 and 36 months. For each patient i, the actual HRQoL assessment times t_{ij} were uniformly distributed around the planned ones (i.e., more or less a value proportional to the difference between the last and the next planned assessment): $t_{i0}= 0$ and $t_{ij} = t_j \pm x$ where $x = k * (t_{j-1} - t_j) * U$ with $U \sim U(-1, 1)$ a random variable distributed on $(-1, 1)$ and $k = 0.5$ to concentrate the simulated value t_{ij} around t_j without overlapping with other simulated values for the same patient i.

At each HRQoL assessment time for patient i, t_{ij}, we generated the longitudinal outcome following the linear mixed model (1) with, $\beta_1 = -0.5$, $\beta_2 = 1$, $\sigma = 13$, $\sigma_{b_0} = 12$, $\sigma_{b_1} = 1.5$ and $\rho = -0.3$ where $\rho = \frac{\sigma_{b_{01}}}{\sigma_{b_0}\sigma_{b_1}}$ is the correlation coefficient between the two random effects. We furthermore set the fixed intercept β_0 to 0 in order to avoid numerical convergence issues. This does not affect the other parameters and corresponds to a situation where

the considered outcome is the centred HRQoL score variable, calculated as the HRQoL score at time t minus the mean HRQoL score at inclusion.

For each patient i, we generated the observed time-to-dropout T_i such as $T_i = min(T_{1i}, T_{2i}, C_i)$ with T_{ki} the times-to-dropout for each cause $k = 1, 2$ and C_i the censoring time. The censoring time C_i was t_{i7}, the last evaluation of patient i. The two cause-specific event times were generated under the CR JM (5) following the method developed to simulate complex survival data, such as time-to-event data from a joint model, and implemented in the R package simsurv [20, 21]. We assumed Weibull baseline cause-specific hazard functions expressed as $h_{01}(t) = h_{02}(t) = \phi t^{\phi-1}\exp(\gamma_0)$ with $\phi = 3$ (shape parameter) and $\gamma_0 = -9$ (log of the scale parameter).

4.2 Estimation

The estimates of the joint models were obtained by likelihood maximization using the function jointModel() from the R package JM (R software, version 3.4.0) [18, 19]. The log baseline hazards were approximated using a cubic B-splines with one internal knot placed at the median of the observed event times. The integrals over the random effects were approximated using a five-point adaptive Gauss–Hermite quadrature.

4.3 Evaluations criteria

For each parameter θ, we calculated the following evaluations criteria: the mean estimated parameter $\bar{\hat{\theta}}$, the relative bias $(\bar{\hat{\theta}} - \theta)/\theta$, the root mean square error (RMSE) $\sqrt{\overline{(\hat{\theta} - \theta)^2}}$ and the coverage rate defined as the percentage of times where θ falls within the confidence interval of $\hat{\theta}$.

4.4 Scenarios

We considered five scenarios by varying the strength of the association α_k and the intensity of the arm effect γ_k on the cause-specific risk of dropout, k=1, 2. In order to facilitate comparison between scenarios, the parameters describing the HRQoL trajectories remained unchanged under all scenarios.

[insert Fig.2]

Fig. 2 Representation of the mean longitudinal outcome from the linear mixed sub-model (a) and hazard functions from the competing risks sub-model given the mean longitudinal outcome (b) used in the five scenarios of the simulation study where Scenario 1: $\alpha_1 = -0.05$, $\alpha_2 = -0.05$, $\gamma_1 = -0.70$, $\gamma_2 = -0.70$; Scenario 2: $\alpha_1 = 0.00$, $\alpha_2 = -0.05$, $\gamma_1 = 0.00$, $\gamma_2 = -0.70$; Scenario 3: $\alpha_1 = 0.05$, $\alpha_2 = -0.05$, $\gamma_1 = 0.00$, $\gamma_2 = 0.00$; Scenario 4:

$\alpha_1 = 0.05$, $\alpha_2 = -0.05$, $\gamma_1 = 0.70$, $\gamma_2 = -0.70$; and Scenario 5: $\alpha_1 = 0.00$, $\alpha_2 = -0.05$, $\gamma_1 = 0.70$, $\gamma_2 = -0.70$

Figure 2 shows the mean trajectories of the longitudinal HRQoL outcomes and the mean hazard functions from the CR JM used to generate the data in the five scenarios.

In Scenario 1, the two causes of dropout were equivalent. The cause-specific risks of dropout were equally and negatively associated with the current true HRQoL outcome, where a decrease of 10 points in the HRQoL outcome corresponded to an increase of 1.65 in the instantaneous risk of dropout ($\alpha_1 = \alpha_2 = -0.05$).
The arm effect was equal whatever the cause and corresponded to lower risks in the experimental arm ($\gamma_1 = \gamma_2 = -0.70$).
Scenario 2 was identical to Scenario 1 for cause 2 ($\alpha_2 = -0.05$ and $\gamma_2 = -0.70$), whereas for cause 1 the association was null ($\alpha_1 = 0$), corresponding to non-informative dropout, and the risk of dropout was the same in both arms ($\gamma_1 = 0$).
In Scenario 3, dropouts from causes 1 and 2 were considered to be informative, with cause-specific risks of dropout oppositely associated with the current true HRQoL outcome ($\alpha_1 = 0.05$ and $\alpha_2 = -0.05$). There was no arm effect on the risk of dropout whatever the cause ($\gamma_1 = \gamma_2 = 0$).
In Scenario 4, the dropouts were considered to be inversely informative, as in Scenario 3 ($\alpha_1 = 0.05$ and $\alpha_2 = -0.05$). There were opposite arm effects on the risk of dropout for the two causes ($\gamma_1 = 0.70$ and $\gamma_2 = -0.70$).
In Scenario 5, dropout from cause 2 was also negatively associated with the current true HRQoL outcome ($\alpha_2 = -0.05$) but dropout from cause 1 was treated as non-informative ($\alpha_1 = 0$). Again, there were opposite arm effects on the risk of dropout for the two causes of dropout ($\gamma_1 = 0.70$ and $\gamma_2 = -0.70$).

4.5 Simulation results

Table 2 describes dropout occurrence in the five different scenarios.

[insert Table 2]

In Scenario 1, we observed all planned measures of the longitudinal outcome (right-censored time-to-dropout) in 12.7% of the patients, whereas in the other scenarios all patients (or almost all) dropped out. Dropouts due to causes 1 and 2 were equally distributed in Scenarios 1 and 3, and more unbalanced in Scenarios 2, 4 and 5, where more dropouts from cause 1 occurred. Table 3 provides the mean results of the 1000 simulations.

[insert Table 3]

Figure 3 gives a graphical representation of the distribution of the estimated β_1, β_2 and association parameters.

[insert Fig. 3]

Fig. 3 Boxplots of the distributions of the association parameters, $\hat{\beta}_1$ and $\hat{\beta}_2$ for the standard joint model (SJM) and the competing risks joint model (CR JM) in the presence of dropouts from two causes, C1 and C2, for the five scenarios considered in the simulation study

In Scenario 1, the parameters of the hazard functions were the same whatever the cause. Thus, the CR JM used to generate the data can reduce to a SJM with, in particular, $\alpha = \alpha_1 = \alpha_2$ and $\gamma = \gamma_1 = \gamma_2$. There were almost zero biases for the estimates of the hazard parameters with both the SJM and CR JM ($\bar{\hat{\alpha}} = -0.050$, $\bar{\hat{\alpha}}_1 = -0.050$, $\bar{\hat{\alpha}}_2 = -0.050$ and $\bar{\hat{\gamma}} = -0.696$, $\bar{\hat{\gamma}}_1 = -0.691$, $\bar{\hat{\gamma}}_2 = -0.699$). The slope ($\beta_1$) and interaction ($\beta_2$) parameters of the longitudinal sub-model were also with very small or minimum bias and with coverage rates close to 95% for both the SJM and CR JM. The RMSE were similar in both models. Finally, in Scenario 1, for which a competing risks sub-model was not necessarily useful, both models performed similarly.

We will now comment on the results for Scenarios 2, 3, 4 and 5 obtained from the SJM for the association parameter, the parameter of the arm effect on the risk of dropout, and the longitudinal sub-model parameters. In these scenarios, the estimates from the CR JM were close to the simulated values, with a coverage rate close to 95%, and had smaller biases and RMSE than the SJM. In Scenarios 2, 3, 4 and 5, the SJM gave an estimate for the association parameter α that was between the two simulated values for each cause, α_1 and α_2. In Scenarios 3 and 4, where there were opposite cause-specific associations, the SJM estimated an association that was between α_1 and α_2, suggesting no association with the longitudinal outcome. In Scenarios 2 and 5, where there was no association for cause 1 and a negative association of -0.05 for cause 2, the SJM estimated an association closer to α_1 than to α_2: $\bar{\hat{\alpha}} = -0.017$ in Scenario 2 and $\bar{\hat{\alpha}} = 0.007$ in Scenario 4.

In a similar way to that of the association parameter, the SJM estimated a γ parameter between the simulated values for each cause, γ_1 and γ_2. However, in all the scenarios where the arm had an effect on at least one cause of dropout – that is, in Scenarios 2, 4 and 5 – the estimated value was not in the middle of γ_1 and γ_2 but was closer to γ_1, even when the cause-specific effects were opposite. In the latter cases, the SJM suggested a risk of dropout that was a little more important in the experimental arm than in the control arm (Scenario 4: $\bar{\hat{\gamma}} = 0.152$, Scenario

5: $\bar{\gamma} = 0.213$), whereas this risk was much more important for dropouts due to cause 1 ($\gamma_1 = 0.7$), and, in contrast, less important for dropouts due to cause 2 ($\gamma_2 = -0.7$).

When the effect parameters of the time-to-dropout sub-model differed according to the dropout causes of dropout, SJM tended to incorrectly estimate the estimations of the longitudinal sub-model (β_1 and β_2).

Using the SJM, the relative bias for β_1 varied from -11.5% in Scenario 3 to -18.9% in Scenario 4, and for β_2 from -8.3% in Scenario 2 to -24.6% in Scenario 4. By comparison, the maximal relative bias using the CR JM was -0.5% (for both β_1 and β_2) in Scenario 3. In the worst case – that is, in Scenario 4 – the coverage rates from the SJM were 87% for β_1 and 71% for β_2. The bias of β_1 corresponded in Scenarios 2, 3, 4 and 5 to an overestimation of the HRQoL level in the control arm and the bias for β_2 to an underestimation of the HRQoL improvement in the experimental arm compared with the control arm. The consequence of the misestimation of β_1 and β_2 on the predicted HRQoL score trajectories in the two arms can be appreciated in figure 4.

[insert Fig. 4]

Fig. 4 Simulated trajectories of the health-related quality of life (HRQoL) score and predicted trajectories from the mean estimates of the standard joint model (SJM) and competing risks joint model (CR JM) by treatment arm for the five scenarios considered in the simulation study, the CR JM-predicted trajectories (blue lines) are superimposed on the simulated trajectories (black lines)

The SJM-predicted trajectory is above the simulated trajectory in the control arm and below the simulated trajectory in the experimental arm. The CR JM-predicted trajectories are superimposed on the simulated trajectories.

5. Discussion

This article focuses on the analysis of longitudinal HRQoL data in clinical trials where patients are likely to drop out during treatment or follow-up. In a previous work, we had compared the use of a SJM taking into account dropout with a linear mixed model on the HRQoL data from the clinical trial PRODIGE 5/ACCORD 17. In the present work, we first compared on the same data the use of a CR JM that distinguishes between two causes of dropout with a SJM that treats all the dropouts equally. We then compared the CR JM and the SJM on simulated data.

In the simulation study, we showed that the CR JM performed as well as the SJM in situations where the risk of dropout was the same whatever the cause. However, in situations where two different risks of dropout coexisted, the SJM provided biased estimates. The estimates of the effect parameter on the risk of dropout were between the values of the cause-specific parameters. In particular, when the two types of dropout were informative and oppositely associated with the current true HRQoL outcome, the SJM estimated a null association parameter, suggesting non-informative dropout. The misspecification of the time-to-dropout sub-model that did not distinguish between the dropouts resulted in biased estimates of the HRQoL parameters. In particular, the difference between the slopes of the HRQoL trajectories of the two arms was underestimated. In contrast with the SJM, the CR JM performed well in all scenarios; these results being expected since the CR JM served to generate the data.

In our application, we found the same pattern as in the simulation study: the parameters estimated by the SJM were between the two cause-specific parameters estimated by the CR JM. The CR JM captured supplementary information about the dropout and revealed that only one cause of dropout was associated with the current HRQoL score. Unexpectedly, it was not dropout related to death or progression that was associated with the HRQoL outcome. This may be due to a lack of reliability concerning the cause of dropout. Indeed, the cause of dropout was constructed *a posteriori*, since the protocol did not include collecting the reasons for dropout. Another explanation is that a non-negligible proportion of dropout unrelated to death or progression could actually be informative. Death and progression are generally well documented, but other events, not documented, could be related to HRQoL. The systematic collection of the reasons for missing HRQoL data in clinical trials would allow a better attribution of the cause of dropout and facilitate the use and the interpretation of CR JMs. Recommendations exist to limit the occurrence of missing data and some authors encourage documenting why the data are missing [22, 23]. This could guide the choice of analysis strategy while only 50% of clinical trials use adequate methods for handling missing data [24, 25].

As in previous analyses using linear mixed models, no significant difference between the HRQoL score trajectories of the two treatment arms was detected [6, 16]. Note that we assumed a linear trend for the HRQoL trajectories for readability and comparability with the simulation study. However, flexible trajectories could be considered, for example using B-splines [26].

In conclusion, the competing risks joint modelling is a satisfactory approach to distinguishing between informative and non-informative dropouts in order to obtain valid results for the longitudinal analysis of HRQoL in clinical trials. Moreover, it offers a better understanding of the association between the HRQoL outcome and the dropout process than a standard joint modelling. Nevertheless, the cause of dropout can be unknown or unreliable in practice. Collecting the reasons for missing data in clinical trials would allow a better attribution of the cause of dropout and facilitate the use and interpretation of competing risks joint models.

References

[1] Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; 963–974.
[2] Chen H, Cohen P. Using individual growth model to analyze the change in quality of life from adolescence to adulthood. *Health and Quality of Life Outcomes* 2006; 4: 10.
[3] Bell ML, Kenward MG, Fairclough DL, et al. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. *Bmj* 2013; 346: e8668.
[4] Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics* 2000; 1: 465–480.
[5] Ibrahim JG, Chu H, Chen LM. Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. *Journal of Clinical Oncology* 2010; 28: 2796–2801.
[6] Cuer B, Mollevi C, Anota A, et al. Handling informative dropout in longitudinal analysis of health-related quality of life: application of three approaches to data from the esophageal cancer clinical trial PRODIGE 5/ACCORD 17. *BMC Medical Research Methodology* 2020; 20: 223.
[7] Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; 809–834.
[8] Fielding S, Fayers PM, Loge JH, et al. Methods for handling missing data in palliative care research. *Palliat Med* 2006; 20: 791–798.
[9] Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Statistics in Medicine* 1998; 17: 517–532.
[10] Elashoff R, li G, Li N. *Joint Modeling of Longitudinal and Time-to-Event Data*. CRC Press, 2016.
[11] Williamson PR, Kolamunnage-Dona R, Philipson P, et al. Joint modelling of longitudinal and competing risks data. *Statistics in Medicine* 2008; 27: 6426–6438.
[12] Kolamunnage-Dona R, Powell C, Williamson PR. Modelling variable dropout in randomised controlled trials with longitudinal outcomes: application to the MAGNETIC study. *Trials* 2016; 17: 222.
[13] Conroy T, Galais M-P, Raoul J-L, et al. Definitive chemoradiotherapy with FOLFOX versus fluorouracil and cisplatin in patients with oesophageal cancer (PRODIGE5/ACCORD17): final results of a randomised, phase 2/3 trial. *The Lancet Oncology* 2014; 15: 305–314.
[14] Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *JNCI: Journal of the National Cancer Institute* 1993; 85: 365–376.
[15] Fayers P, Aaronson NK, Bjordal K, et al. *EORTC QLQ-C30 scoring manual*. 3rd ed. European Organisation for Research and Treatment of Cancer, 2001.
[16] Bascoul-Mollevi C, Gourgou S, Galais M-P, et al. Health-related quality of life results from the PRODIGE 5/ACCORD 17 randomised trial of FOLFOX versus fluorouracil–cisplatin regimen in oesophageal cancer. *European Journal of Cancer* 2017; 84: 239–249.
[17] Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000; 92: 205–216.
[18] Rizopoulos D. JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software, Articles* 2010; 35: 1–33.
[19] Rizopoulos D. *Joint models for longitudinal and time-to-event data: With applications in R*. Chapman and Hall/CRC, 2012.
[20] Crowther MJ, Lambert PC. Simulating biologically plausible complex survival data. *Statist Med* 2013; 32: 4118–4134.
[21] Crowther MJ, Lambert PC. Simulating complex survival data. *The Stata Journal* 2012; 12: 674–687.
[22] Little RJ, D'agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine* 2012; 367: 1355–1360.
[23] Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA* 2013; 309: 814–822.
[24] Stockler MR, Hilpert F, Friedlander M, et al. Patient-reported outcome results from the open-label phase III AURELIA trial evaluating bevacizumab-containing therapy for platinum-resistant ovarian cancer. *Journal of Clinical Oncology* 2014; 32: 1309.
[25] Powney M, Williamson P, Kirkham J, et al. A review of the handling of missing longitudinal outcome data in clinical trials. *Trials* 2014; 15: 237.
[26] Ediebah DE, Galindo-Garre F, Uitdehaag BMJ, et al. Joint modeling of longitudinal health-related quality of life data and survival. *Quality of Life Research* 2015; 24: 795–804.

Figure 1

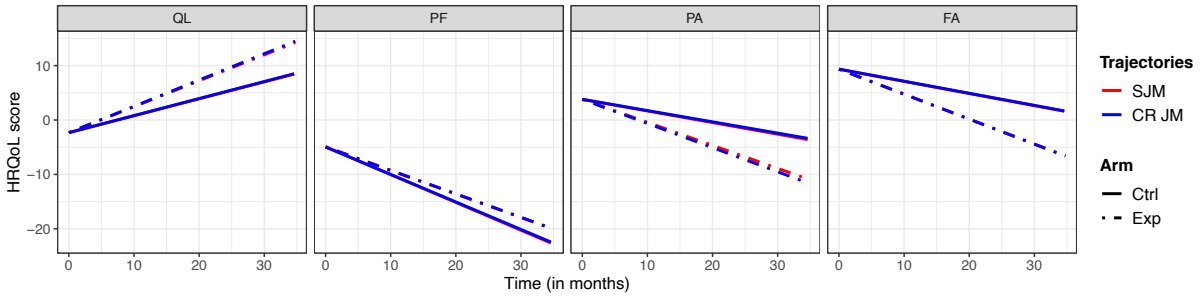
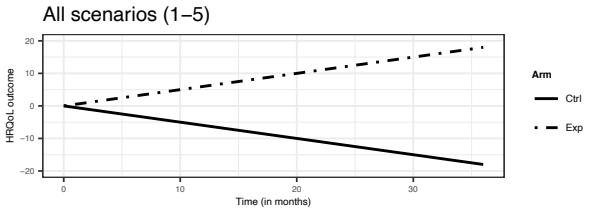

QL
PF
PA
FA
HRQoL score
10
0
−10
−20
0
10
20
30
Time (in months)
Trajectories
SJM
CR JM
Arm
Ctrl
Exp

Figure 2

a. Longitudinal sub–model



b. Competing risks sub–model

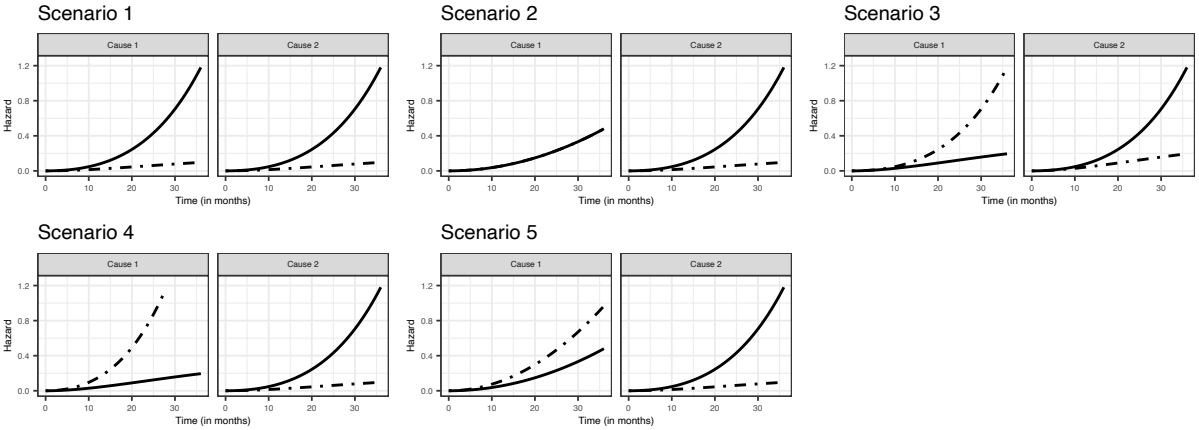


Figure 3

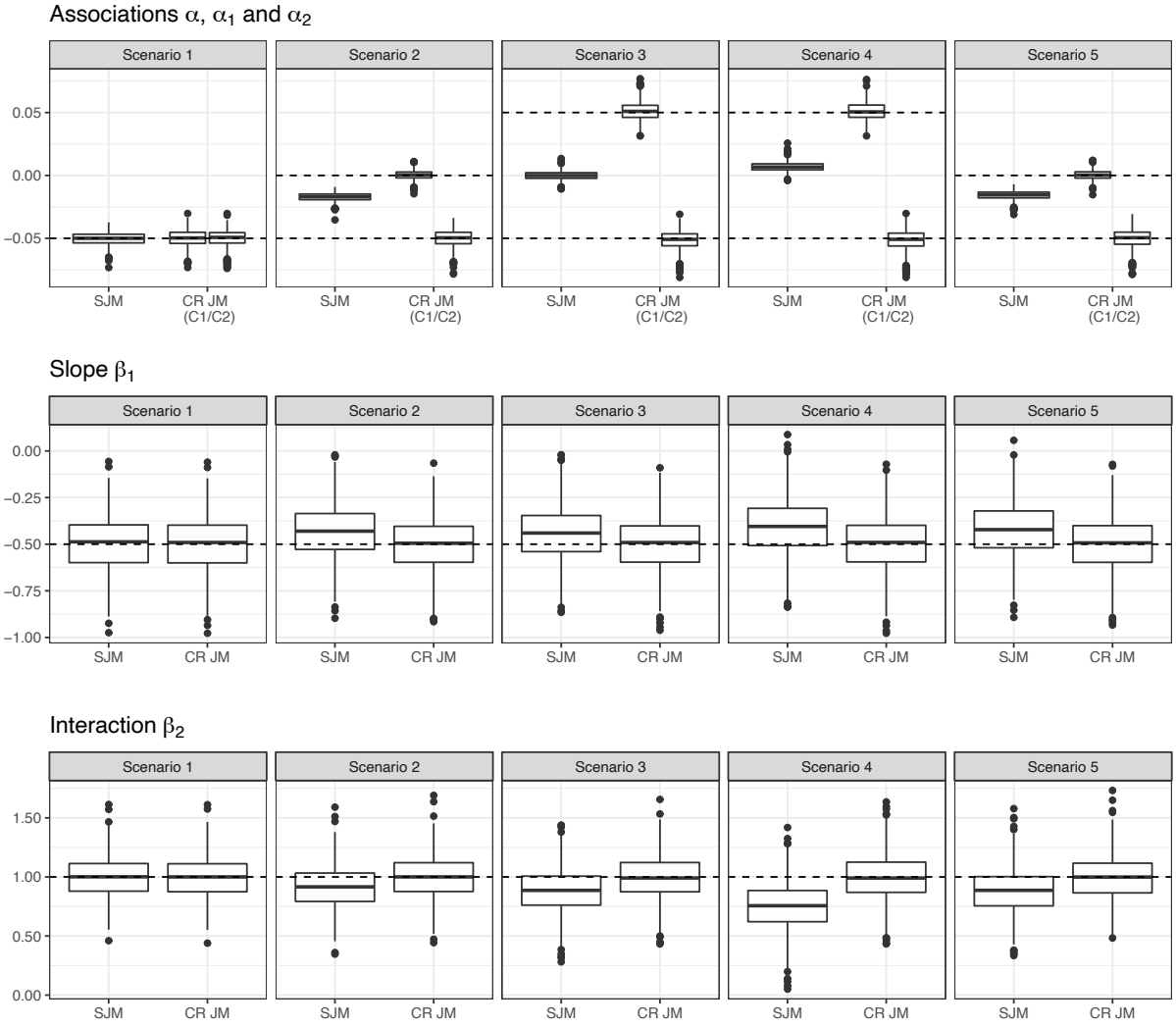

Associations α, α1 and α2
Scenario 1
Scenario 2
Scenario 3
Scenario 4
Scenario 5
0.05
0.00
−0.05
SJM
CR JM (C1/C2)
Slope β1
0.00
−0.25
−0.50
−0.75
−1.00
SJM
CR JM
Interaction β2
1.50
1.00
0.50
0.00

Figure 4

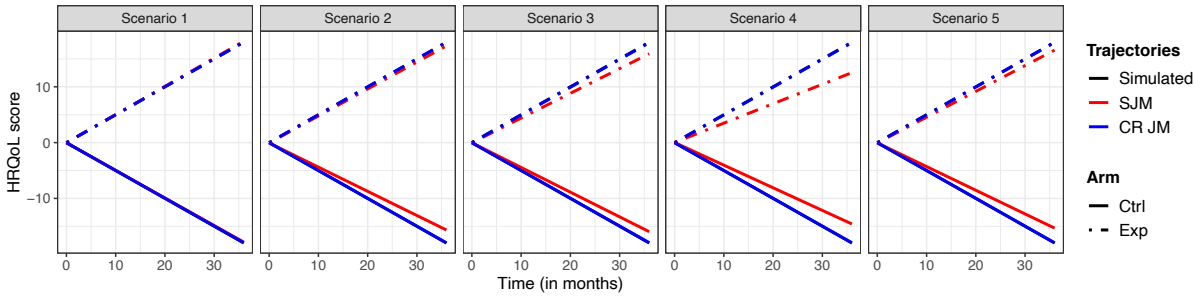

Scenario 1
Scenario 2
Scenario 3
Scenario 4
Scenario 5
HRQoL score
10
0
−10
0
10
20
30
Time (in months)
Trajectories
Simulated
SJM
CR JM
Arm
Ctrl
Exp

Table 1

		SJM Estimate [CI 95%]	*p*	CR JM Estimate [CI 95%]	*p*
QL (n=252)					
Linear mixed sub-model for the longitudinal outcome					
	β_1	0.311 [0.02; 0.60]	0.034	0.313 [0.03; 0.60]	0.031
	β_2	0.166 [-0.21; 0.54]	0.383	0.170 [-0.20; 0.54]	0.363
Survival sub-model for the time-to-dropout					
	γ	-0.019 [-0.28; 0.24]	0.884		
	γ_1			-0.050 [-0.39; 0.29]	0.772
	γ_2			0.079 [-0.43; 0.59]	0.763
	α	-0.016 [-0.03; 0.00]	0.050		
	α_1			-0.016 [-0.04; 0.00]	0.126
	α_2			0.001 [-0.03; 0.03]	0.951
PF (n=254)					
Linear mixed sub-model for the longitudinal outcome					
	β_1	-0.512 [-0.93;-0.09]	0.017	-0.505 [-0.92;-0.09]	0.017
	β_2	0.078 [-0.41; 0.56]	0.753	0.075 [-0.41; 0.56]	0.763
Survival sub-model for the time-to-dropout					
	γ	-0.050 [-0.31; 0.21]	0.699		
	γ_1			-0.087 [-0.42; 0.25]	0.614
	γ_2			0.083 [-0.43; 0.59]	0.751
	α	-0.019 [-0.03;-0.01]	0.001		
	α_1			-0.024 [-0.04;-0.01]	0.001
	α_2			0.012 [-0.01; 0.03]	0.258
PA (n=254)					
Linear mixed sub-model for the longitudinal outcome					
	β_1	-0.214 [-0.58; 0.15]	0.247	-0.207 [-0.57; 0.16]	0.262
	β_2	-0.210 [-0.69; 0.27]	0.392	-0.237 [-0.72; 0.24]	0.334
Survival sub-model for the time-to-dropout					
	γ	-0.069 [-0.32; 0.19]	0.599		
	γ_1			-0.087 [-0.42; 0.25]	0.610
	γ_2			0.071 [-0.44; 0.59]	0.787
	α	0.008 [0.00; 0.02]	0.185		
	α_1			0.008 [-0.01; 0.02]	0.342
	α_2			0.000 [-0.02; 0.02]	0.992
FA (n=254)					
Linear mixed sub-model for the longitudinal outcome					
	β_1	-0.223 [-0.56; 0.12]	0.200	-0.222 [-0.57; 0.13]	0.215
	β_2	-0.235 [-0.70; 0.22]	0.316	-0.238 [-0.70; 0.22]	0.313
Survival sub-model for the time-to-dropout					
	γ	-0.033 [-0.29; 0.22]	0.797		
	γ_1			-0.068 [-0.40; 0.27]	0.693
	γ_2			0.073 [-0.44; 0.58]	0.778
	α	0.013 [0.00; 0.02]	0.007		
	α_1			0.016 [0.00; 0.03]	0.011
	α_2			-0.007 [-0.03; 0.01]	0.471

Table 2

	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5	
	%	Median time	%	Median time	%	Median time	%	Median time	%	Median time
No dropout	12.70		0.18		0.00		0.00		0.04	
Dropout		14.05		14.18		12.80		12.28		13.06
cause 1	43.71	14.06	57.59	15.13	49.96	12.80	59.00	11.84	62.92	13.39
cause 2	43.59	14.05	42.23	13.15	50.04	12.80	40.99	12.93	37.04	12.56

Table 3

			SJM				CR JM			
Scenario		**Value**	**Mean**	**Rel. Bias (%)**	**RMSE**	**Cov. Rate (%)**	**Mean**	**Rel. bias (%)**	**RMSE**	**Cov. Rate (%)**
1	β_1	-0.5	-0.494	-1.1	0.147	94	-0.497	-0.6	0.147	94
	β_2	1	1.000	0.0	0.167	96	0.999	-0.1	0.167	96
	α	-	-0.050	0.7	0.005	95	-	-	-	-
	α_1	-0.05	-	-	-	-	-0.050	0.0	0.007	94
	α_2	-0.05	-	-	-	-	-0.050	-0.3	0.006	94
	γ	-	-0.696	-0.6	0.126	94	-	-	-	-
	γ_1	-0.7	-	-	-	-	-0.691	-1.3	0.161	95
	γ_2	-0.7	-	-	-	-	-0.699	-0.1	0.159	95
2	β_1	-0.5	-0.434	-13.3	0.160	92	-0.499	-0.3	0.144	94
	β_2	1	0.917	-8.3	0.194	93	0.999	-0.1	0.177	95
	α	-	-0.017	-	-	-	-	-	-	-
	α_1	0	-	-	-	-	0.000	0.0	0.004	95
	α_2	-0.05	-	-	-	-	-0.050	0.2	0.007	96
	γ	-	-0.298	-	-	-	-	-	-	-
	γ_1	0	-	-	-	-	0.004	0.4	0.132	94
	γ_2	-0.7	-	-	-	-	-0.699	-0.2	0.174	94
3	β_1	-0.5	-0.443	-11.5	0.157	92	-0.497	-0.5	0.148	94
	β_2	1	0.886	-11.4	0.215	89	0.995	-0.5	0.184	96
	α	-	0.000	-	-	-	-	-	-	-
	α_1	0.05	-	-	-	-	0.051	2.6	0.007	96
	α_2	-0.05	-	-	-	-	-0.051	2.7	0.008	96
	γ	-	0.005	0.5	0.101	94	-	-	-	-
	γ_1	0	-	-	-	-	0.003	0.3	0.165	94
	γ_2	0	-	-	-	-	0.009	0.9	0.162	95
4	β_1	-0.5	-0.406	-18.9	0.178	87	-0.498	-0.4	0.149	93
	β_2	1	0.754	-24.6	0.319	71	0.997	-0.3	0.194	95
	α	-	0.007	-	-	-	-	-	-	-
	α_1	0.05	-	-	-	-	0.051	2.2	0.007	95
	α_2	-0.05	-	-	-	-	-0.051	2.7	0.008	95
	γ	-	0.152	-	-	-	-	-	-	-
	γ_1	0.7	-	-	-	-	0.709	1.3	0.153	94
	γ_2	-0.7	-	-	-	-	-0.702	0.2	0.182	95
5	β_1	-0.5	-0.424	-15.1	0.165	91	-0.498	-0.4	0.146	94
	β_2	1	0.886	-11.4	0.218	90	0.998	-0.2	0.185	95
	α	-	-0.016	-	-	-	-	-	-	-
	α_1	0	-	-	-	-	0.000	0.0	0.004	95
	α_2	-0.05	-	-	-	-	-0.050	0.1	0.008	94
	γ	-	0.213	-	-	-	-	-	-	-
	γ_1	0.7	-	-	-	-	0.713	1.8	0.126	94
	γ_2	-0.7	-	-	-	-	-0.715	2.2	0.200	95