

The Impact of Dropouts on the Analysis of Dose-Finding Studies with Recurrent Event Data

Mouna Akacha^{1*}, Norbert Benda²

¹ *Department of Statistics, University of Warwick, United Kingdom*

² *CLASS, Novartis Pharma AG, Basel, Switzerland*

SUMMARY

This work is motivated by dose-finding studies, where the number of events per subject within a specified study period form the primary outcome. The aim of these studies is to determine the efficacy of a new drug compared to an active control or placebo. In particular, we are interested in identifying the dose-response relationship and the target dose for which the new drug can be shown to be simultaneously safe and as effective as the control.

Given an outcome which is pain-related, we expect a considerable number of patients to drop out before the end of the study period. The impact of missingness on the analysis and diverse models for the missingness process must be carefully considered.

The recurrent events are modeled as over-dispersed Poisson process data, with dose as a regressor. Additional covariates such as age may be included. Constant and time-varying rate functions are examined. Based on these models the impact of missingness on the precision of the target dose estimation is evaluated. Diverse models for the missingness process are considered, including dependence on covariates and number of events. The performances of five different analysis methods (a complete case analysis; two analyses using different single imputation techniques; a direct likelihood analysis; and an analysis using pattern-mixture models) are assessed via simulation studies.

It is shown that the target dose estimation is robust if the same missingness process holds for the target dose group and the active control group. Furthermore, we demonstrate that this robustness is lost as soon as the missingness mechanisms for the active control and the target dose differ. Among the explored missing data handling methods it is shown that the direct-likelihood approach performs best, even when a missing not at random mechanism holds. Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS: Missing Data, Count Data, Recurrent Event Data, Dose-Finding Studies, Pattern-Mixture Models

1. INTRODUCTION

Typically, the aim of a clinical trial is to determine the efficacy of a new drug compared to placebo or an active comparator. In the special case of *dose-finding studies* the interest lies

*Correspondence to: Mouna Akacha, Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom. E-mail: M.Akacha@warwick.ac.uk

in identifying the *dose-response relationship* and the *target dose* for which the drug can be shown to be simultaneously as effective as the comparator and safe. A good understanding of the dose-response relationship is crucial in clinical drug development. A dose which is too low will hinder the proof of efficacy and a dose which is too high could result in safety issues. In fact, one of the main reasons for the high discontinuation rate of phase three clinical trials was found to lie in a poor understanding of the dose-response relationship and consequently in an inadequate dose selection, [1].

Several approaches have been proposed for the efficacious planning and analysis of a dose-finding study. A methodology that combines formal hypothesis testing for dose response with flexible modeling of the dose-response relationship and estimating a target dose, i.e. a *minimum effective dose* (MED) that produces a clinical relevant effect was proposed in [2, 3]. This concept of selecting the best model while controlling the familywise error rate and the subsequent target dose estimation is an extension of the ideas proposed in [4]. The estimation of the MED based on a given model can be regarded as a calibration problem which can be seen as a reverse process to regression, i.e. the estimation of a value for an independent variable that yields an expected outcome for the dependent variable equal to a predefined value. An overview on the classical calibration problem is described in [5]. There is also an extensive literature on calibration problems related to dose estimation, e.g. [6, 7, 8, 9, 10] and [1]. See also [11] and [12] regarding designing aspects of non-linear calibration problems.

This paper focuses on the estimation of a target dose, defined as the dose for which the expected response is equal to that of a control group. In contrast to the literature cited above, the reference value for the outcome of the target dose must be estimated from the current study, i.e. imbedded in the dose-response model, where - depending on the chosen parametrization - either the response of the comparator drug or the target dose itself represents an additional parameter to estimate.

The dose-finding studies that motivated our work seek to analyze processes which generate events repeatedly over time. Such processes are referred to as *recurrent event processes*. Examples include the number of seizures in epileptic studies, the number of hot-flushes postmenopausal women suffer from or the number of flares in gout studies.

Statisticians involved in these studies are usually interested in understanding the underlying event occurrence process. This includes the investigation of the rate at which events occur, the inter-individual variation and most importantly, the relationship between the event occurrence and explanatory variables such as treatment or dose.

Modelling of recurrent events can be approached in a number of ways. The two most common ways are through *event counts* and *gaps* or *waiting times* between two events. Models and methods based on counts are often useful when individuals frequently experience the events of interest, and the events are ‘incidental’ in the sense that their occurrence does not materially alter the process itself, either directly or through resulting interventions, [13]. In contrast, analyses based on waiting times are often relevant when either events are relatively infrequent, some type of individual renewal occurs after an event, or when prediction of the time to the next event is of interest. The applications in mind for this work are of the former nature, and

the canonical approach for the analysis of event counts is the *Poisson process*. Poisson processes can be defined in various ways (see [14]), one of which is via the *intensity function*. It gives the instantaneous probability of an event occurring at a certain point in time, given the history of the process. Modelling recurrent events on explanatory variables can be achieved by specifying the associated intensity as a function of those variables. Corresponding work and different models have been presented in [15, 16, 17] and [14].

The studies that motivated our work focus mainly on the number of events that occur by the end of the study period. Therefore, modelling the counts through a *generalized linear model* as introduced in [18] might seem sufficient. However, in the dose-finding studies of interest and more generally in settings with repeated measurements, incomplete data due to missed visits, dropouts or non-return of questionnaires are quite common. Hence, the endpoint of interest may be missing. In many situations, however, information about the counting process prior to the dropout is available, e.g. through patient diaries. This knowledge can be incorporated in a recurrent event data analysis.

The literature on the analysis of data with dropouts is extensive. References for different missing data methods with the main focus on repeated measurement studies are [19, 20] and [21]. There are three potential problems that typically arise with missing data: loss of efficiency, complication in data handling and analysis, and bias due to differences between the observed and unobserved data, [22]. Given the importance of selecting the adequate dose, the impact of missingness on the dose selection and different models for the missingness mechanism must be carefully considered.

The *missing data mechanism* concerns the relation between the *missingness process* and the outcome variable. Three missingness processes are distinguished, namely the *missing completely at random* (MCAR), the *missing at random* (MAR) and the *missing not at random* (MNAR) mechanisms, [23].

A missingness process is said to be MCAR, when missingness is not related to any measurements, observed or missing in the study. In particular, the incomplete data set can be seen as a random subsample of the complete data set, that would have been observed without missingness. Furthermore, it is not necessary to construct a model for the missingness process. A missingness process which uses less restrictive assumptions is the MAR mechanism. In this case, missingness depends on observed quantities, which include outcomes and explanatory variables, but not on the missing components. If, in addition to MAR, the parameter vectors associated with the measurement and missingness process are disjoint, in the sense that the joint parameter space is the product of the single parameter spaces (*separability* or *distinctness condition*), the missing data mechanism is termed *ignorable*. Likelihood-based or Bayesian inference for the measurement parameter of interest can then be based on the observed data likelihood while ignoring the missing data mechanism, [24]. Finally, if the missingness probability depends on unknown quantities the missingness process is termed MNAR. In the case of *non-ignorability* and MNAR, we need to model the measurement and missingness process jointly. Methods, such as *selection models*, *pattern-mixture models* and *shared parameter models* have been proposed for this case.

Regulatory aspects of missing data in clinical trials are described in the ICH E9 guideline [25], as in the EMEA Points to Consider on Missing Data [26]. Recently, a revised EMEA guideline has been published as a draft [27], giving a more detailed insight in potential issues related to the different methods of missing data handling. Whilst, from a regulatory point of view, the main focus appears to lie on confirmatory Phase III trials, the undesired effects of missing data are also relevant in dose-finding studies. As described in [27] bias is the most important concern resulting from missing data. In dose-finding, this may lead either to an over- or an underestimation of the treatment effect or a dose-response slope. An overestimation would, in general, lead to an underestimation of a target dose, and vice versa. Consequently, the assumed therapeutic window would be either too narrow or too wide. In the first case, the development program might erroneously be stopped or continued with an unnecessarily high dose that has a potentially poor safety profile, whereas in the latter case an ineffective dose might be put forward to Phase III.

The interplay between recurrent event data modeling, dose selection and different missingness mechanisms poses the main focus of this work. Section 2 is devoted to different methods to deal with the missing data problem in dose-finding studies. In Section 3 we will investigate approaches to model the dose-response relationship in the case of recurrent event data and an endpoint-analysis. Furthermore, an alternative modeling approach, which enables the incorporation of information of non-completers, is presented. In Section 4 a dose-finding study which motivated the presented work and the conducted simulation study is introduced. In Section 5 the impact of missingness on the precision of the target dose estimation and the performance of different missing data handling methods is assessed via simulation studies. Concluding remarks are given in Section 6.

2. APPROACHES TO THE ANALYSIS OF RECURRENT EVENT DATA WITH DROPOUT

It is very common for sets of repeated measurements to be incomplete. Missingness usually occurs for reasons outside of the control of the investigators and may be related to the outcome measurement of interest, hence complicating the data analysis.

Data from such trials can in general be analysed in four ways:

1. Perform the analysis only on those subjects who complete the trial;
2. Analyse only the available data;
3. Use a single or multiple imputation technique to replace the missing observations with plausible values, then analyse the completed data set(s); or
4. Model the repeated data and missingness process jointly, [21].

The first option yields a *complete case analysis*. In contrast, the second option can be realised through the *direct likelihood approach*, which is the likelihood-based way of using only the available information, see [19]. Various other (mostly nonparametric) methods of using only the observed data are discussed in [24], and for *single* and *multiple imputation* techniques in [22, 24, 28, 29, 30, 31]. As mentioned in the introduction, the fourth option becomes necessary in the case of non-ignorability and MNAR. Three model families, namely selection models, pattern-mixture models and shared parameter models are the most commonly used

frameworks.

A selection model factorises the joint distribution of the measurement and response mechanism into the marginal measurement distribution and the response distribution, conditional on the measurements. In a pattern-mixture model, we use the alternative factorisation, where the joint density of the full data is factorised into the marginal response distribution and the measurement process, conditional on the missingness pattern. In a shared parameter model, the density of the full data is modelled through the incorporation of random effects, which drive both the outcome and the missingness process.

As described in the introduction, the studies that motivated this work focus mainly on the number of events that occur by the end of a specific study period. Unfortunately, in practice, the endpoint of interest is not observed for each subject. At the same time, information about the counting process before dropping out is usually available. Hence, the remaining derivations of this paper are based on this assumption.

In the following subsections the focus will lie on different methods to deal with missing data in this specific case. Five approaches will be discussed: complete case analysis; two single imputation techniques; direct likelihood approach; pattern-mixture models; and their advantages and disadvantages explored.

2.1. Notation

Suppose m independent subjects are randomized into a trial and that each subject experiences a type of recurrent event. Let $N_i(T_i) = n_i$ denote the number of events over the complete study period $[0, T_i]$ for the i -th subject. For convenience we let $T_i = T$ for all $i \in \{1, \dots, m\}$. The event times for subject i are denoted by $0 < t_{i1} < \dots < t_{in_i} \leq T$. Furthermore, let x_{ij_i} be the vector of explanatory variables for subject i and event time $j_i \in \{t_{i1}, \dots, t_{in_i}\}$. It is assumed that x_{ij_i} are time-independent explanatory variables. Therefore, $x_{ij_i} = x_i$ for all $j_i \in \{t_{i1}, \dots, t_{in_i}\}$. Explanatory variables of interest are age, treatment and dose. Let z_i denote the treatment for subject i . Assuming the study design involves $\ell \in \mathbb{N}$ doses, d_1, \dots, d_ℓ , of the new drug and one comparator (C), then $z_i \in \{d_1, \dots, d_\ell, C\}$. Furthermore, let $N_i = \{N_i(t)\}_{t \in \{t_{i1}, \dots, t_{in_i}\}}$ denote the recurrent event data information for subject i .

Now let $S_i \in (0, T]$ indicate the dropout time for subject $i \in \{1, \dots, m\}$. For clarity, in this paper we assume that once patients drop out, they do not return.

2.2. Complete Case Analysis

In a complete case analysis all participants with missing data are simply discarded and the missingness process is not explicitly incorporated. Although still used in many areas, this strategy is generally inadvisable in the case of longitudinal data. In the context of clinical trials, the FDA is currently bringing out a guideline which discourages use of complete case analysis, because the reason for missing data can be treatment-related and the remaining patients may not adequately represent the study population.[†] Further disadvantages of this approach are the loss of information, potentially severe bias if a MAR process holds and lack

[†]See the FDA draft for the guidance of industry, which can be found under <http://www.fda.gov/cder/guidance/5460dft.pdf>

of cost-effectiveness.

Using this approach in a study of interest would lead to the exclusion of all patients for whom $N_i(T) = n_i$ is missing. Based on the remaining ‘complete’ data a generalized linear model for count data can be formulated (see Section 3.1).

2.3. Single Imputation Techniques

Another common, albeit problematic, strategy is to replace a missing value by a single value which is based on the observed measurements. There are several ways in which the observed information can be used to impute the missing measurements, e.g. *last observation carried forward*, *mean imputation*, *regression imputation* and *hot deck imputation*, [23]. Most of these methods require the MCAR assumption to hold, while some even require additional assumptions. Therefore, all of these techniques have to be used with caution.

“The idea of imputation is both seductive and dangerous. It is seductive because it can lull the user into the pleasurable state of believing that the data are complete after all, and it is dangerous because it lumps together situations where the problem is sufficiently minor that it can be legitimately handled in this way and situations where the standard estimators applied to the real and imputed data have substantial biases.” [32]

The key problem with single imputations is that inferences about parameters based on the imputed data do not account for the imputation uncertainty.

“Thus standard errors computed from the filled-in data are systematically underestimated, P values of tests are too small and confidence intervals are too narrow.” [24]

In our work we will explore two single imputation techniques.

2.3.1. Last Observed Rate Carried Forward This single imputation technique makes the implicit assumption that events for a specific patient occur at the same rate before and after dropout. More explicitly, assume subject i drops out at time point $S_i \in (0, T)$ and let $N_i(S_i)$ be the number of events prior to the dropout. The missing true endpoint $N_i(T)$ is then replaced by the rounded value of $N_i(S_i) + (T - S_i) \frac{N_i(S_i)}{S_i}$, which corresponds to the assumptions of an equal average rate before and after dropout. Replacing every missing endpoint with the imputed value leads to a complete data set and inference can be based on a generalized linear model (see Section 3.1).

2.3.2. Last Count Carried Forward Using this single imputation technique, the missing outcome $N_i(T)$ is simply replaced by $N_i(S_i)$. Then a generalized linear model can be fitted to the ‘completed’ data (see Section 3.1).

2.4. Direct Likelihood Approach

As mentioned before, the direct likelihood approach is the likelihood-based way of using only the available information, while ignoring the missing data mechanism. This approach is valid under the assumption of ignorability, [23]. Using this method will enable the incorporation of

information from non-completers when fitting a specific model and estimating the target dose of interest.

Suppose patient i dropped out at time point $S_i \in (0, T]$ and let $N_{i,obs} = \{N_i(t)\}_{t \leq S_i}$ denote the observed part and $N_{i,mis} = \{N_i(t)\}_{S_i < t \leq T}$ the missing part of the recurrent event data sequence. Likelihood inference for the parameter of interest θ can then be based on the observed likelihood

$$L_{obs}(\theta|N_{obs}) = \prod_{i=1}^m f(N_{i,obs}|\theta) = \prod_{i=1}^m \sum_{N_{i,mis}=0}^{\infty} f(N_{i,obs}, N_{i,mis}|\theta),$$

where $f(\cdot|\theta)$ denotes a parametric model for the recurrent event data sequence $N_i(t_{i1}), \dots, N_i(t_{in_i})$, where $N_i(t)$ is unobserved for $t \in (S_i, T)$. The measurement parameter of interest is denoted by $\theta \in \mathbb{R}^q$, $q \in \mathbb{N}$. Specific models which enable the estimation of the target dose of interest will be discussed in Section 3.2.

2.5. Analysis using Pattern-Mixture Models

In the case of pattern-mixture models it is assumed that the distribution differs according to the underlying missingness pattern, which motivates the naming of these models.

Let $N_i = \{N_i(t)\}_{t \in \{t_{i1}, \dots, t_{in_i}\}}$ be the recurrent event data information for subject i and let p denote the missingness pattern, which in case of continuous time recurrent event data is equivalent to the dropout time. The joint distribution of N_i and the missingness pattern is then given by

$$\begin{aligned} f(N_i, p = S_i) &= f(N_i|p = S_i) f(p = S_i) \\ &= \underbrace{f(N_{i,obs}|p = S_i)}_{known} \underbrace{f(N_{i,mis}|N_{i,obs}, p = S_i)}_{unknown} \underbrace{f(p = S_i)}_{known} \end{aligned} \tag{1}$$

where the first and third factors can be modelled through the data. We can use identifying restrictions to determine the unknown conditional densities of unobserved components, given a set of observed quantities, [33, 19]. Using these restrictions is not the only way to overcome the under-identification. For alternatives, such as using pattern p as an explanatory variable, we refer to [19].

Different identifying restrictions yield different missingness mechanisms. The most commonly used ones are the *complete case missing value restrictions*, where information which is unavailable is borrowed only from the model for the completers; *neighbouring case missing value restrictions*, where unavailable information is always borrowed from the first pattern which observes the corresponding missing measurement; and the *available case missing value restrictions* where all the available information is used. It has been shown that for discrete-time longitudinal data with monotone missingness, MAR is equivalent to the available missing value case, [34]. Fitting procedures for pattern-mixture models are described in [35].

In the case of recurrent event data each pattern (except the pattern of the completers) consists of one patient only. Therefore, estimating the pattern-specific identifiable densities, as requested in the fitting procedure for the discrete time pattern-mixture models, is not feasible. However, when using the complete case identifying restrictions, only the density of the completers needs to be estimated, as all information is borrowed just from these patients. In order to be able to borrow some information using the other identifying restrictions, patients with similar dropout times have to be clustered into one group or pattern. Once the clustering

is done, identifying restrictions for the unknown conditional densities need to be specified. Although very interesting, tackling this problem is beyond the scope of this paper and we will focus on the complete case missing value restrictions.

From equation (1), pattern-mixture models rely on modelling the whole recurrent event data sequence. In order to formulate the pattern-mixture model, the recurrent event model needs to be defined (see Section 3.2). Therefore, this section will be revisited in Section 3.3.

3. MODEL SPECIFICATION

The purpose of this section is to formulate a regression model that enables the estimation of the target dose and the effect of other covariates on the endpoints n_1, \dots, n_m , taking into account the missingness process. Corresponding to the missing data methods that were discussed in the last section, some approaches only need modelling of the endpoints, whereas others rely on a model formulation for the whole recurrent event sequence.

3.1. Models for Count Data

In order to analyze the effects of covariates on the number of events, n_1, \dots, n_m , that occur by the end of the study T a regression analysis has to be performed.

The probability of n_i events occurring in the interval $[0, T]$ for subject i can be modelled by a Poisson distribution with mean $\Lambda_{x_i}(\vartheta)$:

$$\mathbb{P}(N_i(T) = n_i) = \frac{\exp(-\Lambda_{x_i}(\vartheta)) \Lambda_{x_i}^{n_i}(\vartheta)}{n_i!},$$

where ϑ denotes the parameter of interest.

Following the ideas in [18] a generalized linear model will be formulated to model the mean $\Lambda_{x_i}(\vartheta)$ as a function of the dose and other explanatory variables. The canonical link function is used and the following relationship is assumed:

$$\begin{aligned} \ln \{E(N_i(T))\} &= \ln \{\Lambda_{x_i}(\vartheta)\} \\ &= \alpha_0 + \alpha_1 \log(\text{age}_i) + \alpha_2 g(z_i, \eta_{td}), \end{aligned}$$

where η_{td} is the target dose of interest and $\vartheta = (\alpha_0, \alpha_1, \alpha_2, \eta_{td})^\top$. The function $g(\cdot, \eta_{td})$ quantifies the dose-response relationship between the endpoint $N_i(T)$ and the dose, treated as a continuous variable. This work will focus on linear and log-linear dose-response relations. An extension of the models for the EMAX or logistic dose-response relation is straightforward, [36]. Now suppose subject i and j have same age. The target dose is then defined as

$$\eta_{td} = \{\eta \in \mathbb{R} : E[N_i(T) | z_i = \eta] = E[N_j(T) | z_j = C]\}. \quad (2)$$

The following parametrizations for $g(\cdot, \eta_{td})$ are proposed to enable the estimation of η_{td} :

$$\begin{aligned} \text{linear model:} & \quad g(z_i, \eta_{td}) = z_i - \eta_{td}; \text{ and} \\ \text{log-linear model:} & \quad g(z_i, \eta_{td}) = \ln(z_i) - \ln(\eta_{td}). \end{aligned}$$

For $z_i = C$ the value $g(C, \eta_{td})$ is set to zero.

As defined, the Poisson model features the constraint of equal mean and variance, but the

number of events usually varies beyond what can be explained through available covariates. In order to take this inter-individual variation and the so-called overdispersion into account, a model, in which the regression parameters vary across different patients will be considered. Thus, the following *generalized linear mixed model* is proposed:

$$\begin{aligned} N_i(T)|U_i &\stackrel{indep.}{\sim} \mathcal{P}(U_i \Lambda_{x_i}(\vartheta)); \\ U_i &\stackrel{iid.}{\sim} \Gamma(\phi^{-1}, \phi) \quad \text{and} \\ \ln \{E(N_i(T)|U_i)\} &= \ln(U_i \Lambda_{x_i}(\vartheta)) \\ &= \alpha_0 + \alpha_1 \log(age_i) + \alpha_2 g(z_i, \eta_{td}) + \ln(U_i), \end{aligned}$$

where $g(\cdot, \eta_{td})$ is defined above. Assuming the U_i are iid. gamma-distributed random variables with mean 1 and variance ϕ , implies that subjects with U_i greater than one are more likely to experience an event than those with U_i less than one. In particular, the marginal distribution of $N_i(T)$ is given by

$$\begin{aligned} N_i(T) &\stackrel{indep.}{\sim} \mathcal{NB}\left(\frac{1}{\phi}, \frac{1}{1 + \phi \Lambda_{x_i}(\vartheta)}\right) \quad \text{with} \\ \ln(\Lambda_{x_i}(\vartheta)) &= \alpha_0 + \alpha_1 \log(age_i) + \alpha_2 g(z_i, \eta_{td}); \end{aligned} \tag{3}$$

such that the overdispersion coefficient is given by $1 + \phi \Lambda_{x_i}(\vartheta)$. Based on this model we seek to estimate the target dose. The estimate obtained will depend substantially on the procedure used to deal with the missing data and the chosen dose-response model. This inference can be severely biased, as will be illustrated in Section 5.

3.2. Model for Recurrent Event Data

In Section 2, the direct likelihood method and the pattern-mixture models were presented. As illustrated, these approaches are based on different assumptions about the missingness process. Both methods enable the incorporation of information for non-completers. In order to conduct inference for these methods a model for the whole event data sequence N_i , $i \in \{1, \dots, m\}$, has to be formulated.

It is assumed that events occur in continuous time and according to a Poisson process with (time-dependent) intensity function $\lambda_{x_i}(t, \theta)$. The corresponding *cumulative* or *integrated intensity function* is given by

$$\Lambda_{x_i}(t, \theta) = \int_0^t \lambda_{x_i}(w, \theta) dw,$$

where θ quantifies the relationship between the rate and covariates. Thus, the probability of n events occurring in the interval $[0, t]$ for subject i is given by

$$\mathbb{P}(N_i(t) = n) = \frac{\exp(-\Lambda_{x_i}(t, \theta)) \Lambda_{x_i}^n(t, \theta)}{n!}.$$

Let T_{i1}, \dots, T_{in_i} denote the ‘time of occurrence’ random variables for subject $i \in \{1, \dots, m\}$. The joint distribution of the count n_i and the event times $0 \leq t_{i1}, \dots, t_{in_i} \leq T$ can be shown

to be

$$\mathbb{P}(N_i(T) = n_i, T_{i1} = t_{i1}, \dots, T_{in_i} = t_{in_i}) = \exp\{-\Lambda_{x_i}(T, \theta)\} \prod_{j=1}^{n_i} \lambda_{x_i}(t_{ij}, \theta). \quad (4)$$

In practice, however, even after conditioning on covariates, there is more inter-individual variation than is accounted for by a Poisson process. Therefore, a model, where the event rate varies across different patients is considered. Suppose

$$\lambda_{x_i}(t, \theta | U_i = u_i) = u_i \lambda_{x_i}(t, \theta), \quad (5)$$

where u_i is a realization of the gamma-distributed random variable U_i . The extended model is then given by

$$N_i(t) | U_i \stackrel{indep.}{\sim} \mathcal{P}(\Lambda_{x_i}(t, \theta | U_i)) \quad \text{and} \quad U_i \stackrel{iid}{\sim} \Gamma(\phi^{-1}, \phi), \quad (6)$$

where $\Lambda_{x_i}(t, \theta | U_i) = \int_0^t U_i \lambda_{x_i}(u, \theta) du = U_i \Lambda_{x_i}(t, \theta)$. Note that the intensity, implicitly given in equation (5), belongs to the conditional counting process $\{N_i | U_i = u_i\}$ and not to the marginal process N_i .

The contribution of the i -th subject to the joint likelihood for the outcome ' n_i events occur at times $t_{i1}, \dots, t_{i, n_i}$ ', with U_i specified as above, is then given by

$$\begin{aligned} L_i(\phi, \theta) &= \int f_{N_i, T_{i1}, \dots, T_{in_i} | U_i}(n_i, t_{i1}, \dots, t_{i, n_i}) f_{U_i}(u_i) du_i \\ &= n_i! \left(\prod_{j=1}^{n_i} \frac{\lambda_{x_i}(t_{ij}, \theta)}{\Lambda_{x_i}(T, \theta)} \right) \underbrace{\frac{\Gamma(n_i + \frac{1}{\phi})}{n_i! \Gamma(\frac{1}{\phi})} \left(\frac{\phi \Lambda_{x_i}(T, \theta)}{\phi \Lambda_{x_i}(T, \theta) + 1} \right)^{n_i} \left(\frac{1}{\phi \Lambda_{x_i}(T, \theta) + 1} \right)^{\frac{1}{\phi}}}_{f(\theta, \phi)}, \quad (7) \end{aligned}$$

where $f(\theta, \phi)$ is the density of a negative-binomial distributed random variable X with mean $\Lambda_{x_i}(T, \theta)$ and variance $\Lambda_{x_i}(T, \theta) + \phi \Lambda_{x_i}^2(T, \theta)$. In particular, the overdispersion coefficient is given by $1 + \phi \Lambda_{x_i}(T, \theta)$, and thus does depend on explanatory variables.

In the following, different regression models for the intensity function $\lambda_{x_i}(t, \theta)$ will be discussed. It is commonly assumed that covariates affect the intensity through a *multiplicative model* of the form

$$\lambda_{x_i}(t, \theta) = \lambda_0(t, \delta) h(x_i, \beta), \quad (8)$$

where $\theta = (\beta, \delta)$ is the parameter vector of interest, $\lambda_0(\cdot)$ is a *baseline intensity function* and $h(\cdot)$ is a positive-valued function of x_i and β , [14, 37]. Then, the corresponding integrated intensity function is given by

$$\Lambda_{x_i}(t, \theta) = \Lambda_0(t, \delta) h(x_i, \beta), \quad \text{where} \quad \Lambda_0(t, \delta) = \int_0^t \lambda_0(u, \delta) du.$$

The case where $\lambda_0(\cdot)$ is left arbitrary (semi-parametric model) is distinguished from the case where the baseline function is specified up to a parameter vector δ (full parametric model). As implied by the notation in equation (8), the focus of this paper lies on fully parametric models. Further discussion on semiparametric models can be found in [15, 38].

Different models for the baseline function and the function $h(\cdot)$ may be proposed. It is often convenient to choose the function $h(\cdot)$ according to the *Andersen-Gill-Model*, [39, 16]:

$$h(x_i, \beta) = \exp(x_i^\top \beta)$$

since no restrictions on the values of the regression parameter β are needed. The regression parameters can be interpreted as log-relative intensities. Note that model (8) implies that the rate functions associated with any two sets of covariate values are proportional over time, i.e. using a fully parametric Andersen-Gill model leads to a time dependence solely through the baseline intensity function. This restriction may be too strong in practice, [38]. To overcome this drawback *proportional odds models* and other transformation models were studied, [17, 40, 41]. Moreover, the model (8) can be extended such that the baseline rates depend on explanatory variables: $\lambda_{x_i}(t, \theta) = \lambda_0(t, x_i, \delta) h(x_i, \beta)$. Furthermore, random effects can be incorporated as in model (5).

For convenience this paper will focus on a slight extension of the Andersen-Gill-Model. The following relation is assumed

$$h(x_i, \beta) = \exp(k(x_i, \beta)),$$

where $k(\cdot)$ is a (not-necessarily linear) function of x_i and β . One possible model with $x_i = (age_i, z_i)^\top$ is

$$k(x_i, \beta) = \beta_0 \log(age_i) + \beta_1 g(z_i, \eta_{td}). \tag{9}$$

Different choices for $g(\cdot, \eta_{td})$ were discussed in the last subsection. In the following, various models for the baseline intensity function are proposed.

If the rate of events is expected to be the same over the whole study period, it is sensible to choose $\lambda_0(\cdot)$ as a constant function. This yields a *homogeneous Poisson process*. Alternatively, a monotone decreasing or increasing rate function can be realistic and among others yields a *Weibull rate*, see [15]. In the case of a seasonal illness, a periodic rate function might be considered.

The corresponding baseline functions for the constant and Weibull rate assumptions are given by $\lambda_0(t, \delta) = \delta$, $\delta \in \mathbb{R}_+$, and $\lambda_0(t, \delta) = \delta_0 \delta_1 t^{\delta_1 - 1}$ with $\delta = (\delta_0, \delta_1)^\top \in \mathbb{R}_+^2$, respectively. Depending on the chosen rate function the likelihood in equation (7) will become more or less cumbersome; with a constant rate function the likelihood simplifies substantially.

In summary, the following model for the recurrent event data sequence $\{N_i(t)\}_{t \in \{t_{i1}, \dots, t_{in_i}\}}$ is considered:

$$\begin{aligned} N_i(t)|U_i &\stackrel{indep.}{\sim} \mathcal{P}(U_i \Lambda_{x_i}(t, \theta)); \\ U_i &\stackrel{iid.}{\sim} \Gamma(\phi^{-1}, \phi) \quad \text{and} \\ \ln \{E(N_i(t)|U_i)\} &= \ln(U_i \Lambda_{x_i}(t, \theta)) \\ &= \beta_0 \log(age_i) + \beta_1 g(z_i, \eta_{td}) + \ln(\Lambda_0(t, \delta)) + \ln(U_i). \end{aligned} \tag{10}$$

The constant and Weibull rate function will be considered in the simulation studies presented in Section 5.

3.3. Pattern-Mixture Models: Revisited

Having formulated a model for the recurrent event data sequence, we now revisit the pattern-mixture models described in Section 2.5.

We aim to identify the conditional model

$$f(N_{i,mis} | N_{i,obs}, p = S_i),$$

where $N_i = \{N_i(t)\}_{t \in \{t_{i1}, \dots, t_{in_i}\}}$. As we are observing an overdispersed Poisson process we incorporate an unobservable subject-specific random effect $U_i \sim \Gamma(\phi^{-1}, \phi)$ in the model framework proposed in Section 3.2. Conditional on this variable the counts were assumed to follow a Poisson process. However, as here the interest lies in the unconditional counting process N_i , we note that this is no longer a Poisson process with rate function $u_i \lambda_{x_i}(t, \theta)$. Instead, using the notation in Section 3.2, the rate function of N_i is given by

$$\tilde{\lambda}_{x_i}(t, \phi, \theta) = \frac{1 + \phi N_i(t-)}{1 + \phi \Lambda_{x_i}(t, \theta)} \lambda_{x_i}(t, \theta),$$

where $N_i(t-)$ is the number of events that occur in the time interval $[0, t)$, [13].

In order to identify $f(N_{i,mis} | N_{i,obs}, p = S_i)$ we use the CCMV identifying restrictions, where we only have to distinguish between completers (set I) and non-completers (set I^c). Let all parameters of interest in the density (7) be summarized through $\varphi = (\theta, \phi)$. Assume that φ was estimated through $\hat{\varphi}_{comp}$ for the complete cases, using the model in (10). Then, the rate function of N_i for the completers $i \in I$ is determined through

$$\tilde{\lambda}_{x_i}(t, \hat{\varphi}_{comp}, \hat{\theta}_{comp}) = \frac{1 + \hat{\varphi}_{comp} N_i(t-)}{1 + \hat{\varphi}_{comp} \Lambda_{x_i}(t, \hat{\theta}_{comp})} \lambda_{x_i}(t, \hat{\theta}_{comp}).$$

We then identify

$$N_i(T) = N_i(S_i) + N_{comp}\{[S_i, T]\} | N_i(S_i), \quad (11)$$

where $N_{comp}\{[S_i, T]\} | N_i(S_i)$ is a draw from the distribution of the completers, that is, from a counting process with rate function $\tilde{\lambda}_{x_i}(t, \hat{\varphi}_{comp}, \hat{\theta}_{comp})$, $t \in [S_i, T]$.

As it is cumbersome to draw from the corresponding distribution, we choose to draw the event times and to count the event occurrences by time point T . Given an event occurred at time point t_{j-1} , the waiting time w_j for event t_j can be calculated based on

$$B_j = \int_{t_{j-1}}^{t_{j-1}+w_j} \tilde{\lambda}_{x_i}(t, \hat{\varphi}_{comp}, \hat{\theta}_{comp}) dt.$$

Solving this integral for the constant and Weibull rate function $\lambda_{x_i}(t, \theta)$ is straightforward. Now, for each patient from the set I^c , multiple draws of $N_i(T)$ are created according to the identifying restriction defined in equation (11). Then, the missing endpoints for subjects $i \in I^c$ are replaced by these multiple draws and the generalized linear mixed model, introduced in Section 3.1, is fitted to these ‘completed’ data sets. The same model is fitted separately to the data of all completers. For each parameter the final estimate is a weighted average of the pattern specific estimates. The weights are estimated as the pattern probabilities (see [35] for details about fitting procedures for pattern-mixture models).

4. EXAMPLE: CHRONIC GOUT STUDY

Throughout the remaining part of the paper, we illustrate our ideas with simulated data which mimic a Phase II trial. An active-controlled dose-finding study is planned to evaluate a new compound for the prophylaxis of signs and symptoms of acute flares in chronic gout patients. For similar studies see [42, 43].

The primary objective of this study is to determine the target single dose of the new compound that leads to the same efficacy as an active-control, with respect to the mean number of gout flares occurring during 16 weeks after randomization.

Since the primary outcome is the mean number of flares for each patient after 16 weeks an endpoint-analysis is to be performed. However, given the painful nature of gout flares a considerable number of patients is expected to drop out throughout the study. Safety reasons for high doses and adverse events are further expected reasons for dropout.

Performing an endpoint-analysis without adjustment for missingness would imply no inclusion of information about non-completers. This practice would lead to the underlying assumption of an identical evolution for completers and non-completers. Given the reasons for dropout, this is rather unrealistic.

All patients are asked to record the date they suffered from a gout flare. Using this information, the number of flares until the time of dropout will be known and can be included in the modelling framework, as was shown in Section 3.2.

5. SIMULATION STUDY

One of the key goals of this work is to investigate the performance of several missing data handling methods for the estimation of the target dose in the context of a recurrent event data trial. To allow a direct quantitative comparison of the proposed methods in Section 2 under the same conditions and using the same performance metrics, a simulation study, motivated by the setting of the example in Section 4, was undertaken.

5.1. Scenarios

Different scenarios will be explored using different:

- rate functions: constant and (decreasing) Weibull;
- dose-response profiles: linear and log-linear;
- missingness mechanisms;
- dropout-rates: 20% – 75%;
- overdispersion coefficients: dependent on covariates $1 + \Lambda_{x_i}(T, \theta) \in [1.1, 1.4]$.

It is assumed that besides the comparator, five different doses ($\{0.25, 0.5, 1, 2, 3\}$) of the new drug are used in the trial. Furthermore, every simulated data set has a sample size of 3500 patients. We choose a relatively large sample size in order to reduce the bias due to finite sampling. The comparator is assigned to 1000 patients and each of the different doses of the new drug is assigned to 500 patients. The study period is 112 days. Moreover, the following missing data mechanisms were used to introduce missingness:

MAR:

Person i drops out after $t_{ij_i} \sim \text{Bernoulli}(p(t_{ij_i}))$, where

$$\text{logit}(p(t_{ij_i})) = \gamma_{0,z_i} + \gamma_{1,z_i} \log(\text{age}_i) + \gamma_{2,z_i} z_i + \gamma_{3,z_i} \frac{N_i(t_{ij_i})}{t_{ij_i}} \quad (12)$$

where $\gamma_{j,z_i} < 0$ for $j \in \{0, 1\}$ and $\gamma_{j,z_i} > 0$ for all $j \in \{2, 3\}$, i.e. the probability of dropping out decreases with age but increases with the assigned dose (due to adverse events) and number of events.

MNAR:

Person i drops out after $t_{ij_i} \sim \text{Bernoulli}(p(t_{ij_i}))$, where

$$\begin{aligned} \text{logit}(p(t_{ij_i})) = & \gamma_{0,z_i} + \gamma_{1,z_i} \log(\text{age}_i) + \gamma_{2,z_i} z_i + \gamma_{3,z_i} \frac{N_i(t_{ij_i})}{t_{ij_i}} \quad (13) \\ & + \gamma_{4,z_i} \frac{N_i(T) - N_i(t_{ij_i})}{T - t_{ij_i}}, \end{aligned}$$

where $\gamma_{j,z_i} < 0$ for $j \in \{0, 1\}$ and $\gamma_{j,z_i} > 0$ for all $j \in \{2, 3, 4\}$. Here, in addition to the MAR case, the probability of dropping out increases with the potentially unobserved number of events by the end of the trial.

Moreover, we distinguish two cases. Firstly, for $j \in \{0, 1, 3, 4\}$

$$\gamma_{j,z_i} = \gamma_j \quad \text{for } z_i \in \{0.25, \dots, 3, C\};$$

and

$$\gamma_{2,z_i} z_i = \begin{cases} \gamma_2 z_i & \text{for } z_i \in \{0.25, \dots, 3\}; \\ \gamma_2 \eta_{td} & \text{for } z_i = C; \end{cases}$$

i.e. the missingness processes for the comparator and the target dose group are identical. Secondly, for $j \in \{0, 1, 3, 4\}$

$$\gamma_{j,z_i} = \begin{cases} \gamma_j & \text{for } z_i \in \{0.25, \dots, 3\}; \\ \gamma_j^* \neq \gamma_{j,\eta_{td}} & \text{for } z_i = C; \end{cases}$$

and

$$\gamma_{2,z_i} z_i = \begin{cases} \gamma_2 z_i & \text{for } z_i \in \{0.25, \dots, 3\}; \\ \gamma_2^* \neq \gamma_2 \eta_{td} & \text{for } z_i = C; \end{cases}$$

i.e. the missingness processes for the comparator and the target dose group differ. For each scenario 1000 data sets are simulated according to the data model given in Section 3.2, equation (10), and assuming the target dose $\eta_{td} = 2$. Patients are assumed to suffer from three events on average.

In total, 32 scenarios were investigated. The different compositions are given in Tables I, VI, VIII and IX. The simulations and the analyses were performed using the software SAS (PROC IML, NLMIXED and created macros).

Simulation scenarios for constant rate and linear dose-response relation

Scenario	Missingness		Contrast		Dropout-Rate
	MAR	MNAR	same miss.	diff. miss.	
1	X		X		22.10%
2		X	X		21.69%
3	X		X		46.06%
4		X	X		45.51%
5	X			X	25.88/36.71/21.55%
6		X		X	26.27/35.13/22.72%
7	X			X	54.18/71.04/47.43%
8		X		X	56.53/70.74/50.84%

Table I. Overview of the different scenarios for the simulation study using a constant rate function with a linear dose-response relationship. In case of different missingness rates for the new drug and the comparator, the column 'Dropout-Rate' consists of the dropout rates for the complete data set/ only for the comparator/ only for the experimental drug.

5.2. Results

The results of the simulation study, using the mean bias of the estimated target dose as our performance metric, are given in Tables II, VII, X and XI. Next to the mean and median of the estimated target doses, the standard errors and the 90% range for the estimates are recorded for the proposed methods for dealing with missingness (see Section 2). The pattern-mixture model was fitted using the CCMV identifying restrictions and 5 imputations.

The results for the scenarios presented in Table I are given in Table II.

For these eight scenarios a homogeneous Poisson process with a linear dose-response relationship was assumed. Different missingness rates are considered. Regarding the results for the first four scenarios, which assume the same missingness rate and process for the target dose group and comparator, it becomes apparent that the estimates are very close to the true value $\eta_{td} = 2$. Obviously, the estimates obtained by analyzing the complete data sets are the most accurate ones. However, in the case of MAR, it is remarkable how close these are to the direct likelihood (DL) estimates (see Scenario 1 and Scenario 3). Due to the reduced information, standard errors are just slightly higher than those obtained from analyzing the complete data. Even in the case of MNAR, where ignoring the missingness process, i.e. applying the DL approach, is usually not valid, we obtain acceptable estimates. This observation even holds for Scenario 4, where half of the endpoints went missing. Also the estimates obtained by the complete case analysis (CC), the last count carried forward technique (LCCF) and the pattern-mixture models (PMM) are within a 2.5% range of the true value. The target dose is usually slightly underestimated. The standard errors are marginally higher than those of the complete data. In the case of the last observed rate carried forward (LORCF) approach the mean of the estimates is highly biased due to outliers. We observe two different reasons for outlying estimates. Firstly, imputing the missing endpoint through the rounded value of $N_i(S_i) + (T - S_i) \frac{N_i(S_i)}{S_i}$ can yield very large numbers of events. Take for example a subject which drops out after two days, but already suffered from three events. The imputed value based on LORCF is then given by 167 and will affect the estimation of the target dose. The second issue that leads to outliers is of numerical nature: the dual quasi-Newton algorithm used

Simulation results for constant rate and linear dose-response relation

Scenario	Parameter	CD	CC	LORCF	LCCF	DL	PMM
1	mean(η_{td})	2.002	1.985	-84.60	1.997	2.003	1.997
1	median($t\eta_{td}$)	1.998	1.983	1.968	1.995	2.000	1.994
1	$\hat{\sigma}_{\eta_{td}}$	0.07	0.07	2382.1	0.07	0.08	0.07
1	$\hat{\gamma}$	[1.89,2.12]	[1.86,2.12]	[-11.78,6.41]	[1.88,2.13]	[1.88,2.13]	[1.88,2.12]
2	mean(η_{td})	2.002	1.987	-84.45	1.997	2.003	1.997
2	median(η_{td})	1.998	1.986	1.966	1.999	2.002	1.996
2	$\hat{\sigma}_{\eta_{td}}$	0.07	0.08	2417.8	0.08	0.08	0.08
2	$\hat{\gamma}$	[1.89,2.12]	[1.86,2.12]	[-11.75,7.18]	[1.87,2.12]	[1.88,2.13]	[1.87,2.13]
3	mean(η_{td})	1.999	1.981	-881.9	1.973	1.999	2.006
3	median(η_{td})	1.999	1.978	1.967	1.977	1.999	2.005
3	$\hat{\sigma}_{\eta_{td}}$	0.07	0.09	27576	0.09	0.08	0.09
3	$\hat{\gamma}$	[1.89,2.10]	[1.84,2.13]	[-3.44, 6.56]	[1.83,2.11]	[1.88,2.19]	[1.87,2.15]
4	mean(η_{td})	1.999	1.972	-989.9	1.921	1.992	1.992
4	median(η_{td})	1.999	1.970	1.959	1.919	1.991	1.994
4	$\hat{\sigma}_{\eta_{td}}$	0.07	0.09	30962	0.10	0.08	0.11
4	$\hat{\gamma}$	[1.89,2.10]	[1.82,2.12]	[-3.79, 6.38]	[1.76,2.09]	[1.87,2.12]	[1.83,2.15]
5	mean(η_{td})	2.000	2.403	-0.982	2.494	2.000	2.404
5	median(η_{td})	1.999	2.403	1.720	2.492	1.999	2.403
5	$\hat{\sigma}_{\eta_{td}}$	0.07	0.10	72.69	0.10	0.08	0.09
5	$\hat{\gamma}$	[1.88,2.12]	[2.25,2.57]	[-2.77,5.90]	[2.35,2.65]	[1.87,2.14]	[2.26,2.56]
6	mean(η_{td})	2.000	2.473	-4.915	2.626	2.083	2.481
6	median(η_{td})	1.999	2.474	1.496	2.623	2.081	2.478
6	$\hat{\sigma}_{\eta_{td}}$	0.07	0.10	144.51	0.10	0.08	0.10
6	$\hat{\gamma}$	[1.88,2.12]	[2.32,2.63]	[-4.85,5.38]	[2.47,2.78]	[1.96,2.22]	[2.32,2.64]
7	mean(η_{td})	2.002	3.279	-561.21	3.082	2.000	3.237
7	median(η_{td})	2.001	3.274	1.821	3.079	2.001	3.235
7	$\hat{\sigma}_{\eta_{td}}$	0.07	0.15	15117	0.11	0.09	0.13
7	$\hat{\gamma}$	[1.89,2.11]	[3.03,3.54]	[-10.56,6.09]	[2.91,3.26]	[1.85,2.14]	[3.03,3.46]
8	mean(η_{td})	2.002	3.591	-614.04	3.602	2.085	3.586
8	median(η_{td})	2.001	3.586	1.758	3.604	2.08	3.584
8	$\hat{\sigma}_{\eta_{td}}$	0.07	0.18	15187	0.20	0.09	0.20
8	$\hat{\gamma}$	[1.89,2.11]	[3.30,3.91]	[-9.92,4.83]	[3.34,3.88]	[1.93,2.24]	[3.30,3.89]

Table II. Simulation results for a constant rate with a linear dose-response relationship. The amount of missingness and the missingness process vary according to Table I. In the column 'CD' (complete data) the data sets were analyzed before introducing missingness. Moreover, 'CC' refers to the complete case analysis, 'LORCF' to last observed rate carried forward, 'LCCF' to the last count carried forward imputation and 'DL' to the direct likelihood approach. The estimate $\hat{\sigma}_{\eta_{td}}$ denotes the standard error of $\hat{\eta}_{td}$ and $\hat{\gamma}$ the 90% range of the η_{td} -estimates.

to fit the generalized linear model defined in Section 3.1 did not converge for every simulated data set. However, it seems reasonable to use the more outlier-robust median, in order to compare the performance of this imputation techniques to the other approaches. Using this point estimate also shows that the LORCF estimates range within a 2.5% range of the true value.

At first glance these results seem very surprising, in particular the good results using the LORCF, LCCF and CC approach which were not expected. Hence, the results of Scenario 3 and Scenario 4 suggest that estimating the target dose in the given setting is very robust. But which component of the setting does lead to these results? Having a look at the remaining parameter estimates of the model shows that some of these estimates in turn are severely biased for LORCF, CC, LCCF and PMM.

Using the linear dose-response relation and the count model introduced in Section 3.1, yields

$$\ln \{E(Y_i(T)|U_i)\} = \alpha_0 + \alpha_1 \log(\text{age}_i) + \alpha_2 (z_i - \eta_{td}) + \ln(U_i).$$

For fitting purposes we reparametrized the parameter α_0 as follows:

$$\alpha_0 = \ln(\rho T).$$

In comparison, the model for the whole recurrent event data sequence implies

$$\ln \{E(Y_i(T)|U_i)\} = \beta_0 \log(\text{age}_i) + \beta_1 (z_i - \eta_{td}) + \ln(\delta T) + \ln(U_i).$$

In particular, both parametrisations imply $\rho = \delta$, $\alpha_1 = \beta_0$, $\alpha_2 = \beta_1$ and the same subject-specific parameter ϕ and target dose η_{td} in both models. The estimates for these estimates and Scenario 3 and Scenario 4 are given in Table III. Note that the estimates for these parameters are more sensitive to dropout. Especially the estimates for $\alpha_1 = \beta_0$, $\delta = \rho$ and ϕ are biased for some of the proposed data handling methods. In particular, we observe a higher bias under the assumption of MNAR. The parameter $\alpha_1 = \beta_0$ is severely biased in the case of LORCF, but also for the CC, LCCF and the PMM approach. So are the estimates for ρ and the target dose using the LORCF method. Very likely, this is due to outlying measurements. Moreover, the estimation of the overdispersion parameter ϕ seems to be biased in all approaches. This is in line with the results presented in [44]. However, the estimation of ϕ could be hindered due to the finite sample size and its small value we chose for the simulation study ($\phi = 1/65 \approx 0.0154$). In order to rule out the influence of ϕ on the results obtained, simulations were conducted with the value $\phi = 1$. The parameter estimates for ϕ and the corresponding effect on the target dose estimation are shown in Table IV and Table V respectively.

Note that all approaches except the direct likelihood approach and the LORCF approach underestimate the parameter ϕ . This is due to the dropout probabilities, which depend on explanatory variables and the number of events by a certain point in time, thus reducing the heterogeneity for the complete cases or imputed data sets. In contrast, the overdispersion parameter ϕ is highly overestimated in the case of LORCF which is related to the outlying imputations mentioned earlier. Solely the results yielded by the direct likelihood approach are very accurate. The results in Table V suggest that the bias of the target dose estimation increases with higher overdispersion.

Overall the direct likelihood estimates for Scenario 3, that is MAR, are satisfactory and near the true values. In the case of MNAR a higher bias is observable.

Interpreting these results, leads to the hypothesis that the estimation of the target dose is very

Scenario	Parameter	True	CD		CC		LORCF	
			EST	SE	EST	SE	EST	SE
3	$\beta_0 = \alpha_1$	0.01	0.011	0.02	0.022	0.04	-0.051	0.61
3	$\beta_1 = \alpha_2$	-0.4	-0.401	0.01	-0.445	0.02	-0.451	0.34
3	$\delta = \rho$	0.02	0.020	0.002	0.014	0.002	101.74*	2909*
3	ϕ	0.015	0.015	0.01	0.0003	0.002	3.130	0.50
3	η_{td}	2	1.999	0.07	1.981	0.09	-881.907*	27577
4	$\beta_0 = \alpha_1$	0.01	0.011	0.02	0.019	0.04	-0.055	0.63
4	$\beta_1 = \alpha_2$	-0.4	-0.401	0.01	-0.396	0.02	-0.462	0.35
4	$\delta = \rho$	0.02	0.020	0.002	0.014	0.003	120.40*	3473*
4	ϕ	0.015	0.015	0.01	$2.94 \cdot 10^{-6}$	$1.9 \cdot 10^{-5}$	3.239	0.51
4	η_{td}	2	1.999	0.07	1.972	0.09	-989.9*	30962*

Scenario	Parameter	True	LCCF		DL		PMM	
			EST	SE	EST	SE	EST	SE
3	$\beta_0 = \alpha_1$	0.01	0.018	0.03	0.012	0.03	0.016	0.04
3	$\beta_1 = \alpha_2$	-0.4	-0.312	0.01	-0.400	0.02	-0.380	0.02
3	$\delta = \rho$	0.02	0.014	0.002	0.020	0.002	0.019	0.01
3	ϕ	0.015	0.0005	0.002	0.016	0.01	0.0002	0.001
3	η_{td}	2	1.974	0.09	1.999	0.08	2.006	0.09
4	$\beta_0 = \alpha_1$	0.01	0.015	0.04	0.011	0.03	0.013	0.05
4	$\beta_1 = \alpha_2$	-0.4	-0.263	0.01	-0.379	0.01	-0.335	0.02
4	δ	0.02	0.014	0.003	0.019	0.003	0.019	0.02
4	ϕ	0.015	$2.74 \cdot 10^{-6}$	$2.1 \cdot 10^{-5}$	$7.98 \cdot 10^{-6}$	$6.1 \cdot 10^{-5}$	$5.50 \cdot 10^{-6}$	$2.1 \cdot 10^{-4}$
4	η_{td}	2	1.921	0.10	1.992	0.08	1.992	0.11

Table III. Mean estimates and standard errors for the parameters involved in the models of Scenario 3 and Scenario 4, given in Table I.

	$\phi = \frac{1}{65} \approx 0.0154$				$\phi = 1$			
	same miss		diff miss		same miss		diff miss	
	Est	SD	Est	SD	Est	SD	Est	SD
CD	0.015	0.01	0.015	0.01	0.98	0.03	0.98	0.03
CC	0.0003	0.002	$2.94 \cdot 10^{-6}$	$1.9 \cdot 10^{-5}$	0.88	0.04	0.83	0.04
LORCF	3.130	0.50	3.239	0.51	4.86	0.64	4.76	0.59
LCCF	0.0005	0.002	$2.74 \cdot 10^{-6}$	$2.1 \cdot 10^{-5}$	0.68	0.03	0.60	0.03
DL	0.016	0.01	$7.98 \cdot 10^{-6}$	$6.1 \cdot 10^{-5}$	0.99	0.04	0.99	0.04
PMM	0.0002	0.001	$5.50 \cdot 10^{-6}$	$2.1 \cdot 10^{-4}$	0.74	0.03	0.67	0.03

Table IV. Mean estimate for the overdispersion parameter ϕ and the corresponding standard error in case of constant rate, linear dose response and MAR.

Parameter	CD	CC	LORCF	LCCF	DL	PMM
same missingness: 21.14%						
mean(td)	1.992	1.971	-4.399	1.984	1.999	1.980
median(td)	1.994	1.964	1.821	1.982	2.001	1.980
$\hat{\sigma}_{td}$	0.12	0.14	72.61	0.13	0.14	0.13
$\hat{\gamma}$	[1.79,2.19]	[1.75,2.20]	[-10.17,5.30]	[1.78,2.21]	[1.78,2.22]	[1.77,2.20]
different missingness: 24.24% (total) - 32.44% ($z_i = C$) - 20.95% ($z_i \in \{0.25, 0.5, 1, 2, 3\}$)						
mean(td)	1.992	3.092	-28.90	2.911	2.004	3.098
median(td)	1.987	3.085	1.682	2.906	2.004	3.090
$\hat{\sigma}_{td}$	0.12	0.20	542.90	0.17	0.15	0.19
$\hat{\gamma}$	[1.79,2.19]	[2.78,3.45]	[-8.93,4.94]	[2.64,3.21]	[1.75,2.26]	[2.80,3.43]

Table V. Mean and median estimates for the target dose, the corresponding standard error and the 90% range in case of MAR.

robust, because the same missingness process was applied for patients receiving the comparator and the target dose. The responses for both treatment groups are equally biased and will therefore still yield the correct estimate for the target dose (see Section 3.1 and Section 3.2). In order to investigate this hypothesis, data sets according to Scenario 5 - Scenario 8 in Table I were simulated and analyzed. The dropout mechanisms for the comparator and the different doses of the new drug have the same functional form but different parameters (see Section 5), which lead to a higher dropout rate in the comparator group. Comparing the resulting target dose estimates for these scenarios, a remarkable bias for the CC, LCCF and PMM approaches is observed. In the case of Scenario 5 and Scenario 6, where we have an overall dropout rate of about 26%, these approaches overestimate η_{td} by approximately 25%. In the case of 50% missingness this bias increases to approximately 75%.

The overestimation of the target dose in the case of the complete case analysis is due to the nature of the simulated missingness processes. Patients who suffer from many events are more likely to drop out than those with only a few events. As the dropout rate in the comparator group is higher than for the other dose groups, it appears that the CC analysis discards relatively more patients assigned to the comparator group. The remaining patients in the comparator group present a subsample with relatively less events, thus leading to an overestimation of the target dose. Now, as we used the CCMV restrictions for the pattern-mixture models, this bias is also incorporated in the PMM estimates.

In the case of LCCF the last observed count is carried forward. Therefore, the number of events is underestimated for all non-completers. As the dropout rate for the comparator group is higher than for the other dose groups, LCCF underestimates more event counts for the comparator group, and therefore overestimates the target dose.

In contrast, the target dose is underestimated in the case of LORCF by maximal 25%. The substantial bias is slightly surprising, because the data were simulated via a homogeneous Poisson process and the LORCF approach imputes the missing values based on this assumption.

Note that the true target dose η_{td} is not included in the 90% range of estimates for either

Simulation scenarios for constant rate and log-linear dose-response relation

Scenario	Missingness		Contrast		Dropout-Rate
	MAR	MNAR	same miss.	diff. miss.	
1	X		X		22.86%
2		X	X		22.49%
3	X		X		47.08%
4		X	X		46.58%
5	X			X	26.50/36.69/22.41%
6		X		X	27.04/35.09/23.82%
7	X			X	55.18/71.02/48.85%
8		X		X	57.55/70.69/52.29%

Table VI. Overview of the different scenarios for the simulation study using a constant rate function with a log-linear dose-response relationship. In case of different missingness rates for the new drug and the comparator, the column 'Dropout-Rate' consists of the dropout rates for the complete data set/ only for the comparator/ only for the experimental drug.

CC, LCCF or PMM. Furthermore, we observe that all parameters involved in the model are highly biased. The results for the overdispersion parameter ϕ are shown in Table IV. Note that the bias in the target dose estimation increases with ϕ in the case of different missingness processes too (see Table V).

In Table VII, the results for all scenarios summarized in Table VI, are given.

Again, the data were simulated according to a homogeneous Poisson process but with a log-linear dose-response relation. The results are very similar in spirit to those discussed previously. A negligible bias can be observed for the first four scenarios, where the same missingness process is assumed for the comparator and the target dose group. Moving to Scenario 5 - Scenario 8 the bias increases greatly. While in Table II a bias of maximal 75% was observed, here we are dealing with up to 600% bias. The DL approach performs very well in the MAR case but leads to a bias of around 10% in the case of MNAR. This bias is nevertheless considerably smaller than that of the remaining techniques. CC, LCCF and PMM overestimate the target dose. The bias of the LCCF is the highest, but nearly followed by the CC and PMM estimates, which are very similar. This is not surprising as we are borrowing information for non-completers from the complete case. That is, if the target dose is overestimated using the CC approach then the estimate of the PMM method will be overestimated as well.

These results show that the impact of missingness on the estimated target dose does not only depend on the assumed missingness process and the dropout rate, but also on the assumed dose-response relationship.

In Appendix A we show the findings for data based on a Poisson process with Weibull rate. Linear and log-linear dose-response relations are assumed. From the qualitative point of view the results are identical to those obtained for the previous scenarios and will therefore not be discussed in more detail.

Simulation results for constant rate and log-linear dose-response relation

Scenario	Parameter	CD	CC	LORCF	LCCF	DL	PMM
1	mean(η_{td})	2.007	1.971	$7.3 \cdot 10^{10}$	1.989	2.011	1.974
1	median(η_{td})	1.993	1.963	1.868	1.977	2.003	1.966
1	$\hat{\sigma}_{\eta_{td}}$	0.14	0.15	$8.1 \cdot 10^{11}$	0.15	0.16	0.15
1	$\hat{\gamma}$	[1.80,2.26]	[1.75,2.22]	$[1.5 \cdot 10^{-4}, 987.5]$	[1.77,2.25]	[1.78,2.28]	[1.76,2.23]
2	mean(η_{td})	2.007	1.970	$6.2 \cdot 10^{10}$	1.984	2.008	1.970
2	median(η_{td})	1.993	1.961	1.839	1.973	1.995	1.963
2	$\hat{\sigma}_{\eta_{td}}$	0.14	0.15	$7.1 \cdot 10^{11}$	0.15	0.15	0.15
2	$\hat{\gamma}$	[1.80,2.26]	[1.75,2.23]	$[1.1 \cdot 10^{-5}, 594.7]$	[1.75,2.26]	[1.78,2.28]	[1.74,2.24]
3	mean(η_{td})	2.009	1.970	$4.4 \cdot 10^{10}$	2.002	2.015	1.927
3	median(η_{td})	2.006	1.949	2.029	1.983	2.006	1.942
3	$\hat{\sigma}_{\eta_{td}}$	0.14	0.19	$3.9 \cdot 10^{11}$	0.20	0.17	0.25
3	$\hat{\gamma}$	[1.79,2.24]	[1.70,2.29]	$[2.1 \cdot 10^{-3}, 197.3]$	[1.69,2.36]	[1.76,2.30]	[1.59,2.27]
4	mean(η_{td})	2.009	1.962	$3.1 \cdot 10^{10}$	2.009	1.969	1.953
4	median(η_{td})	2.006	1.956	1.967	1.994	1.999	1.950
4	$\hat{\sigma}_{\eta_{td}}$	0.14	0.21	$3.7 \cdot 10^{11}$	0.26	0.33	0.22
4	$\hat{\gamma}$	[1.79,2.24]	[1.66,2.31]	$[2.5 \cdot 10^{-3}, 143.7]$	[1.64,2.45]	[1.73,2.29]	[1.65,2.29]
5	mean(η_{td})	2.008	3.012	$4.3 \cdot 10^{10}$	3.372	2.010	2.968
5	median(η_{td})	2.005	3.004	1.605	3.347	2.007	2.953
5	$\hat{\sigma}_{\eta_{td}}$	0.13	0.28	$5.4 \cdot 10^{11}$	0.33	0.16	0.28
5	$\hat{\gamma}$	[1.79,2.24]	[2.58,3.51]	$[9.3 \cdot 10^{-3}, 44.05]$	[2.87,3.95]	[1.76,2.29]	[2.54,3.46]
6	mean(η_{td})	2.008	3.278	$4.5 \cdot 10^{10}$	4.135	2.201	3.267
6	median(η_{td})	2.005	3.261	1.299	4.095	1.999	3.239
6	$\hat{\sigma}_{\eta_{td}}$	0.13	0.33	$5.8 \cdot 10^{11}$	0.46	0.18	0.34
6	$\hat{\gamma}$	[1.79,2.24]	[2.78,3.90]	$[6.4 \cdot 10^{-3}, 24.19]$	[3.46,4.98]	[1.93,2.51]	[2.76,3.86]
7	mean(η_{td})	2.001	7.322	$3.1 \cdot 10^{10}$	6.442	2.001	6.542
7	median(η_{td})	1.991	7.178	1.686	6.375	1.990	6.464
7	$\hat{\sigma}_{\eta_{td}}$	0.14	1.25	$3.2 \cdot 10^{11}$	0.89	0.19	0.97
7	$\hat{\gamma}$	[1.79,2.25]	[5.50,9.61]	$[1.4 \cdot 10^{-2}, 103.3]$	[5.17,8.02]	[1.71,2.34]	[5.14,8.25]
8	mean(η_{td})	2.001	10.210	$3.5 \cdot 10^{10}$	14.812	2.185	9.530
8	median(η_{td})	1.991	9.974	1.543	14.656	2.169	9.365
8	$\hat{\sigma}_{\eta_{td}}$	0.14	2.17	$3.6 \cdot 10^{11}$	4.24	0.25	1.85
8	$\hat{\gamma}$	[1.79,2.25]	[7.14,13.97]	$[1.1 \cdot 10^{-2}, 100.8]$	[6.64,21.61]	[1.85,2.59]	[6.88,12.86]

Table VII. Simulation results for a constant rate and a log-linear dose-response relationship. The amount of missingness and the missingness process vary according to Table VI. The estimate $\hat{\sigma}_{\eta_{td}}$ denotes the standard error of $\hat{\eta}_{td}$ and $\hat{\gamma}$ the 90% range of the η_{td} -estimates.

6. CONCLUSIONS

In the context of our work, we were challenged by the analysis of dose-finding studies, in which patients are expected to drop out throughout the study. The reasons for dropping out are assumed to be related to the randomized dose, other explanatory variables and the outcome of interest. Hence, the investigation of the impact of dropouts on the target dose selection has to be carefully considered. Here, the target dose is defined as the dose which leads to the same efficacy as an active comparator.

The specific study which motivated our work explores a recurrent event data process over a time period of several weeks. In practice, investigators are not interested in modelling the whole recurrent event sequence. Instead, they focus on the number of events occurring in a specific time interval. Due to missingness this endpoint is not observed for all patients and the classical approach of investigators is to perform a complete case analysis where all non-completers are discarded from the analysis. In our study of interest, however, dropout is expected to be outcome related and the underlying assumption justifying a complete case analysis, i.e. MCAR, is violated. Therefore, other techniques to handle missing data and their performances have to be investigated. Some classical techniques for dealing with missing data in the case of repeated measurements were discussed in Section 2. Some of these methods require models for the whole recurrent event process, whereas others enable an endpoint analysis for the number of events occurring by the end of the study. Regression models which enable the target dose selection and the investigation of the effect of other explanatory variables were laid out in Section 3. In order to compare the performances of the proposed missing data methods a simulation study was performed.

The results suggest that the estimation of the target is very robust if the same missingness mechanisms apply for the comparator and the target dose group. However, care should be taken because the other model parameters can be substantially biased. This holds in particular for the parameters which estimate subject-specific characteristics.

Further simulation studies showed that this robustness is lost as soon as the missingness processes differ for the comparator and the target dose group. A dropout rate as little as 25% was able to severely bias the target dose when performing the ‘traditional’ complete case analysis. In the case of a log-linear dose-response and a homogenous Poisson process the target dose was overestimated by more than 50%. The pattern-mixture models perform similarly badly. This is not surprising as all unavailable information is borrowed from completers. Moreover, both single imputation techniques perform poorly in this case. The bias introduced by the last count carried forward approach is comparable to that of a complete case analysis. The target dose is always overestimated and dependent on the rate function, dose-response relationship, dropout rate and mechanism the bias varies between 25% and 600%. In contrast, the target dose is usually underestimated in the case of the last observed rate carried forward approach. The bias ranges between 25% and 75%. Outliers and numerical problems hinder an adequate dose selection. In the case of the direct likelihood approach we observe a bias of maximal 5% in the case of a MAR process and of up to 10% in the case of a MNAR mechanism. The good performance in the case of MAR is not surprising. However, these results also suggest that the direct likelihood approach provides a certain protection against bias for the MNAR process used.

In light of these results, we remark that the direct likelihood performed best among the

missing data handling approach proposed in Section 2. Hence, we recommend the inclusion of this method in the study protocols for dose-finding studies. Furthermore, we discourage the use of complete case analysis and single imputation techniques. The poor performance of the pattern-mixture models is simply due to the identifying restrictions used. Ideally, we would like to consider other possibilities to borrow available information. Moreover, we observe that the magnitude of bias increases with the overdispersion. Furthermore, the bias highly depends on the modelling framework, i.e. the underlying rate function and the assumed dose-response model.

ACKNOWLEDGEMENT

Mouna Akacha is very thankful for the continuous help and support given by her supervisor Prof. Jane L. Hutton, University of Warwick, United Kingdom. This work was funded by CRISM and the Novartis Pharma AG.

REFERENCES

1. Bretz F, Hsu J, Pinheiro JC, Liu Y. Dose finding - a challenge in statistics. *Biometrical Journal* 2008; **50**:480–504.
2. Bretz F, Pinheiro JC, Branson M. *Dose finding in drug development*, chap. Analysis of dose-response studies. Springer: New York, 2006; 146–171.
3. Pinheiro JC, Bornkamp B, Bretz F. Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. *Journal of Biopharmaceutical Statistics* 2006; **16**:639–656.
4. Tukey J, Ciminera J, Heyse J. Testing the statistical certainty of a response to increasing doses of a drug. *Biometrics* 1985; **41**:295–301.
5. Osborne C. Statistical calibration: A review. *International Statistical Review* 1991; **59**:309–336.
6. Filloon T. Estimating the minimum therapeutically effective dose a compound via regression modelling and percentile estimation. *Statistics in Medicine* 1995; **14**:925–932.
7. Hsu J, Berger R. Stepwise confidence intervals without multiplicity adjustment for dose-response and toxicity studies. *Journal of the American Statistical Association* 1999; **94**:468–482.
8. Tamhane A, Logan B. Multiple test procedures for identifying the minimum effective and safe dose of drugs. *Journal of the American Statistical Association* 2002; **97**:293–301.
9. Morales K, Ibrahim J, Chen C, Ryan L. Bayesian model averaging with applications to benchmark dose estimation for arsenic in drinking water. *Journal of the American Statistical Association* 2006; **101**:9–17.
10. Budtz-Jørgensen E. Estimation of the benchmark dose by structural equation models. *Biostatistics* 2007; **8**:675–688.
11. Forkman J. A method for designing nonlinear univariate calibration. *Technometrics* 2008; **50**:479–486.
12. Dette H, Bretz F, Pepelyshev A, Pinheiro JC. Optimal design for dose-finding studies. *Journal of the American Statistical Association* 2008; **103**:1225–1237.
13. Cook J, Lawless J. *The statistical analysis of recurrent events*. Springer, 2007.
14. Andersen P, Borgan O, Gill R, Keiding N. *Statistical models based on counting processes*. Springer Series in Statistics, 1993.
15. Lawless J. Regression methods for poisson process data. *Journal of the American Statistical Association - Theory and Methods* 1987; **82**:808–815.
16. Cook R, Lawless J. Analysis of repeated events. *Statistical Methods in Medical Research* 2002; **11**:141–166.
17. Lindsey J. *Modelling Frequency and Count Data*. Oxford University Press, 1995.
18. McCullagh P, Nelder J. *Generalized Linear Models*. Chapman & Hall, 1989.
19. Molenberghs G, Kenward MG. *Missing data in clinical studies*. Wiley, 2007.
20. Diggle P, Farewell D, Henderson R. Analysis of longitudinal data with drop-out: objectives, assumptions and a proposal. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2007; **56**:1–31.
21. Carpenter J, Pocock S, Lamm C. Coping with missing data in clinical trials: A model-based approach applied to asthma trials. *Statistics in Medicine* 2002; **21**:1043–1066.
22. Horton N, Lipsitz S. Multiple imputation in practice: Comparison of software packages for regression models with missing data. *The American Statistician* 2001; **55** (3):244–254.

23. Rubin D. Inference and missing data. *Biometrika* 1976; **63**:581–592.
24. Little R, Rubin D. *Statistical analysis with missing data*. Wiley Interscience, 2002.
25. ICH-E1. *ICH Topic E 9. Statistical principles of clinical trials*, 1998. Available at <http://www.emea.eu.int>.
26. EMEA. *Points to consider on missing data. Doc. Ref. EMEA/CPMP/EWP/1776/99, Committee for Medicinal Products in Human Use (CHMP)*. London, 15 November 2001.
27. EMEA. *Guideline on missing data in confirmatory clinical trials. Doc. Ref. EMEA/CPMP/EWP/1776/99 Rev.1 Corr, Committee for Medicinal Products for Human Use (CHMP)*. London, 23 April 2009.
28. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. Wiley, 1987.
29. Rubin D. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 1996; **91**:473–489.
30. Schafer J. Multiple imputation: a primer. *Statistical methods in medical research* 1999; **8**:3–15.
31. Schafer J. *Analysis of incomplete multivariate data*. Chapman & Hall, 1997.
32. Dempster A, Rubin D. *Incomplete data in sample surveys (Volume 2): Theory and Bibliography*, chap. :Introduction. New York: Academic Press, 1983; 3–10.
33. Little R. A class of pattern-mixture models for normal incomplete data. *Biometrika* 1994; **81**:471–483.
34. Molenberghs G, Michiels, Kenward MG, Diggle PJ. Monotone missing data and pattern-mixture models. *Statistica Neerlandica* 1998; **52**:153–161.
35. Thijs H, Molenberghs G, Michiels B, Verbeke G, Curran D. Strategies to fit pattern-mixture models. *Biostatistics* 2002; **3**:245–265.
36. Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques on dose-response studies. *Biometrics* 2005; **61**:738–748.
37. Lawless J. The analysis of recurrent events for multiple subjects. *Journal of the Royal Statistical Society, Applied Statistics* 1995; **44**:487–498.
38. Lin D, Wei J, Ying Z. Semiparametric transformation models for point processes. *Journal of the American Statistical Association- Theory and Methods* 2001; **96**:620–628.
39. Andersen P, Gill R. Cox's regression model for counting processes: A large sample study. *The Annals of Statistics* 1982; **10**:1100–1120.
40. Dabrowska D, Doksum K. Estimation and testing in in the two-sample generalized odds-rate model. *Journal of the American Statistical Association* 1988; **83**:744–749.
41. Cheng S, Wei L, Ying Z. Predicting survival probabilities with semiparametric transformation models. *Journal of the American Statistical Association* 1997; **92**:227–235.
42. Becker M, Schumacher R, Wortmann R, MacDonald P, Eustace D, Palo W, Streit J, Joseph-Ridge N. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *The New England Journal of Medicine* 2005; **353**:2450–2461.
43. Borstad G, Bryant L, Abel M, Scroggie D, Harris M, Alloway J. Colchicine for prophylaxis of acute flares when initiating allopurinol for chronic gout arthritis. *The Journal of Rheumatology* 2004; **31**:2429–2432.
44. Gustafson P. On measuring sensitivity to parametric model misspecification. *Journal of the Royal Statistical Society* 2001; **63**:81–94.

APPENDIX

The results based on the Weibull-rate for the scenarios in Table VIII and IX are given in Table X and XI, respectively.

**Simulation scenarios for a decreasing Weibull rate and
a linear dose-response relation**

Scenario	Missingness		Contrast		Dropout-Rate
	MAR	MNAR	same miss.	diff. miss.	
1	X		X		20.69%
2		X	X		19.91%
3	X		X		41.83%
4		X	X		45.03%
5	X			X	24.41/31.97/21.39%
6		X		X	23.64/30.60/20.85%
7	X			X	49.29/62.65/43.95%
8		X		X	52.42/63.22/48.09%

Table VIII. Overview of the different scenarios for the simulation study using a Weibull rate function with a linear dose-response relationship. In case of different missingness rates for the new drug and the comparator, the column 'Dropout-Rate' consists of the dropout rates for the complete data set/ only for the comparator/ only for the experimental drug.

**Simulation scenarios for a decreasing Weibull rate and
a log-linear dose-response relation**

Scenario	Missingness		Contrast		Dropout-Rate
	MAR	MNAR	same miss.	diff. miss.	
1	X		X		21.46%
2		X	X		20.69%
3	X		X		42.92%
4		X	X		46.09%
5	X			X	21.99/32.04/18.64%
6		X		X	21.49/30.66/18.44%
7	X			X	49.14/62.71/43.89%
8		X		X	52.20/63.26/47.92%

Table IX. Overview of the different scenarios for the simulation study using a Weibull rate function with a log-linear dose-response relationship. In case of different missingness rates for the new drug and the comparator, the column 'Dropout-Rate' consists of the dropout rates for the complete data set/ only for the comparator/ only for the experimental drug.

**Simulation results for a decreasing Weibull rate and
a linear dose-response relation**

Scenario	Parameter	CD	CC	LORCF	LCCF	DL	PMM
1	mean(η_{td})	2.002	1.984	-3.188	1.991	2.002	1.999
1	median($t\eta_{td}$)	2.002	1.984	2.006	1.990	2.002	1.999
1	$\hat{\sigma}_{\eta_{td}}$	0.06	0.08	164.91	0.07	0.07	0.07
1	$\hat{\gamma}$	[1.90,2.11]	[1.87,2.10]	[1.59,2.55]	[1.88,2.11]	[1.89,2.12]	[1.89,2.12]
2	mean(η_{td})	2.002	1.987	2.031	1.992	2.002	2.001
2	median(η_{td})	2.002	1.985	2.014	1.991	2.001	2.000
2	$\hat{\sigma}_{\eta_{td}}$	0.06	0.07	0.31	0.07	0.07	0.07
2	$\hat{\gamma}$	[1.90,2.11]	[1.87,2.10]	[1.56,2.58]	[1.88,2.11]	[1.89,2.12]	[1.89,2.12]
3	mean(η_{td})	2.001	1.980	-3.081	1.955	2.002	2.009
3	median(η_{td})	1.998	1.979	1.984	1.957	2.001	2.008
3	$\hat{\sigma}_{\eta_{td}}$	0.07	0.09	160.35	0.10	0.07	0.08
3	$\hat{\gamma}$	[1.90,2.10]	[1.85,2.12]	[1.65, 2.41]	[1.82,2.10]	[1.89,2.12]	[1.88,2.14]
4	mean(η_{td})	2.001	1.935	1.998	1.894	1.990	1.995
4	median(η_{td})	1.998	1.953	1.994	1.893	1.993	1.996
4	$\hat{\sigma}_{\eta_{td}}$	0.07	0.17	0.35	0.10	0.08	0.10
4	$\hat{\gamma}$	[1.90,2.10]	[1.77,2.11]	[1.65, 2.44]	[1.75,2.05]	[1.87,2.11]	[1.84,2.15]
5	mean(η_{td})	1.999	2.316	1.581	2.320	1.997	2.317
5	median(η_{td})	1.998	2.316	1.571	2.325	1.994	2.320
5	$\hat{\sigma}_{\eta_{td}}$	0.07	0.09	0.40	0.08	0.07	0.08
5	$\hat{\gamma}$	[1.90,2.10]	[2.19,2.45]	[1.27,1.97]	[2.19,2.45]	[1.88,2.11]	[2.18,2.45]
6	mean(η_{td})	1.999	2.365	0.934	2.400	2.047	2.381
6	median(η_{td})	1.998	2.372	0.943	2.412	2.047	2.382
6	$\hat{\sigma}_{\eta_{td}}$	0.07	0.13	0.45	0.16	0.07	0.09
6	$\hat{\gamma}$	[1.90,2.10]	[2.23,2.52]	[0.62,1.31]	[2.27,2.55]	[1.93,2.16]	[2.24,2.53]
7	mean(η_{td})	2.003	3.032	1.703	2.680	1.994	2.977
7	median(η_{td})	2.002	3.028	1.716	2.685	1.990	2.983
7	$\hat{\sigma}_{\eta_{td}}$	0.07	0.15	0.50	0.15	0.08	0.16
7	$\hat{\gamma}$	[1.90,2.11]	[2.82,3.27]	[1.40,2.11]	[2.53,2.85]	[1.87,2.12]	[2.77,3.19]
8	mean(η_{td})	2.003	3.207	1.600	2.938	2.075	3.236
8	median(η_{td})	2.002	3.460	1.610	2.963	2.077	3.240
8	$\hat{\sigma}_{\eta_{td}}$	0.07	0.31	0.54	0.28	0.09	0.18
8	$\hat{\gamma}$	[1.90,2.11]	[2.97,3.52]	[1.30,1.99]	[2.76,3.12]	[1.93,2.22]	[2.98,3.50]

Table X. Simulation results for a Weibull rate and a linear dose-response relationship. The amount of missingness and the missingness process vary according to Table VIII. The estimate $\hat{\sigma}_{\eta_{td}}$ denotes the standard error of $\hat{\eta}_{td}$ and $\hat{\gamma}$ the 90% range of the η_{td} -estimates.

Simulation results for a decreasing Weibull rate and a log-linear dose-response relation

Scenario	Parameter	CD	CC	LORCF	LCCF	DL	PMM
1	mean(η_{td})	2.006	1.980	1.350	2.01	2.009	1.982
1	median(η_{td})	1.998	1.969	1.616	1.999	1.994	1.972
1	$\hat{\sigma}_{\eta_{td}}$	0.13	0.14	1.09	0.14	0.14	0.13
1	$\hat{\gamma}$	[1.82,2.21]	[1.77,2.20]	[1.2·10 ⁻²¹ ,2.99]	[1.79,2.26]	[1.81,2.24]	[1.77,2.22]
2	mean(η_{td})	2.006	1.979	1.433	2.011	2.010	1.981
2	median(η_{td})	1.998	1.972	1.685	2.003	2.001	1.976
2	$\hat{\sigma}_{\eta_{td}}$	0.13	0.14	1.12	0.14	0.14	0.14
2	$\hat{\gamma}$	[1.82,2.21]	[1.77,2.22]	[1.5·10 ⁻²¹ ,3.15]	[1.80,2.25]	[1.80,2.24]	[1.77,2.22]
3	mean(η_{td})	2.005	1.975	1.874	2.031	2.005	1.983
3	median(η_{td})	1.999	1.965	1.951	2.017	1.999	1.975
3	$\hat{\sigma}_{\eta_{td}}$	0.12	0.17	0.81	0.19	0.14	0.15
3	$\hat{\gamma}$	[1.82,2.22]	[1.72,2.27]	[1.6·10 ⁻²⁰ ,3.00]	[1.74,2.36]	[1.77,2.25]	[1.76,2.24]
4	mean(η_{td})	2.005	1.955	1.870	2.039	2.004	1.972
4	median(η_{td})	1.999	1.943	1.959	2.023	1.998	1.968
4	$\hat{\sigma}_{\eta_{td}}$	0.12	0.19	0.83	0.27	0.15	0.16
4	$\hat{\gamma}$	[1.82,2.22]	[1.67,2.28]	[1.1·10 ⁻²⁰ ,3.00]	[1.66,2.49]	[1.77,2.26]	[1.73,2.24]
5	mean(η_{td})	2.001	2.765	1.057	2.882	1.996	2.720
5	median(η_{td})	1.992	2.741	1.215	2.863	1.983	2.703
5	$\hat{\sigma}_{\eta_{td}}$	0.13	0.24	0.62	0.24	0.15	0.23
5	$\hat{\gamma}$	[1.80,2.22]	[2.41,3.21]	[1.3·10 ⁻²¹ ,1.83]	[2.51,3.30]	[1.78,2.26]	[2.37,3.14]
6	mean(η_{td})	2.001	2.990	0.700	3.344	2.116	2.966
6	median(η_{td})	1.992	2.978	0.696	3.331	2.104	2.948
6	$\hat{\sigma}_{\eta_{td}}$	0.13	0.26	0.20	0.32	0.17	0.27
6	$\hat{\gamma}$	[1.80,2.22]	[2.59,3.44]	[0.45,1.00]	[2.88,3.85]	[1.88,2.40]	[2.56,3.42]
7	mean(η_{td})	2.006	5.726	1.401	4.490	1.991	5.147
7	median(η_{td})	2.003	5.653	1.494	4.452	1.983	5.112
7	$\hat{\sigma}_{\eta_{td}}$	0.13	0.83	0.60	0.49	0.16	0.64
7	$\hat{\gamma}$	[1.80,2.23]	[4.58,7.13]	[4.1·10 ⁻²¹ ,2.12]	[3.78,5.37]	[1.74, 2.25]	[4.20,6.24]
8	mean(η_{td})	2.006	7.394	1.263	7.036	2.176	6.721
8	median(η_{td})	2.003	7.223	1.342	7.117	2.168	6.642
8	$\hat{\sigma}_{\eta_{td}}$	0.13	1.24	0.52	1.77	0.18	1.04
8	$\hat{\gamma}$	[1.80,2.23]	[5.74,9.77]	[5.6·10 ⁻²¹ ,1.90]	[3.95,9.61]	[1.89,2.52]	[5.24,8.52]

Table XI. Simulation results for a Weibull rate and a log-linear dose-response relationship. The amount of missingness and the missingness process vary according to Table VIII. The estimate $\hat{\sigma}_{\eta_{td}}$ denotes the standard error of $\hat{\eta}_{td}$ and $\hat{\gamma}$ the 90% range of the η_{td} -estimates.