Recurrent events modelling of haemophilia bleeding events

Andrew C. Titman[1], Martin J. Wolfsegger[2], and Thomas F. Jaki[1]

[1]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
[2]Baxalta Innovations GmbH, a Takeda company, Vienna, Austria

Abstract

A pharmacokinetic-pharmacodynamic (PK-PD) approach is developed for modelling recurrent bleeding events in patients with severe haemophilia to investigate the relationship between factor VIII plasma activity level and the instantaneous risk of a bleed. The model incorporates patient-level pharmacokinetic (PK) information obtained through measurements taken prior to the study which are used to fit a non-linear mixed effects two-compartment PK model. Dosing times within the study are combined with the PK model to provide the estimated factor VIII plasma level for all patients, which is used as a time dependent covariate within the recurrent events model. Methods are developed to correct the attenuation in covariate effects that would otherwise arise due to the discrepancy between estimated and true factor VIII. In contrast to existing methods proposed for such data, such as count data regression or time-to-event analysis, the new method allows all the bleeding times to be used to investigate the relationship between current factor VIII and risk of a bleed. The performance of the proposed estimators are assessed via simulation and found to outperform the naive estimator, which treats the estimated factor VIII levels as if they were measured without error, both in terms of bias and mean squared error.

1 Introduction

1.1 Background

Patients with haemophilia suffer recurrent bleeding events due to deficiency of factor VIII (FVIII) clotting agent in their blood. An established treatment for the prevention or reduction in occurrence of bleeds is through regular administration of FVIII into the blood. It is known that patients whose natural level of FVIII is higher than 1 IU/dL have fewer spontaneous bleed events (den Uijl, Fisher et al., 2011). This has led to treatments which aim to set the dosage and dosing interval in order to keep the minimum, or trough, level of plasma FVIII above 1 IU/dL. After an intravenous infusion, FVIII activity levels peak immediately and then rapidly decrease over time. The volume of the blood which is completely cleared of FVIII per unit time by means of an elimination process (clearance) of FVIII varies considerably between individuals and hence there is potential benefit in using individual pharmacokinetic (PK) data to determine the dosing regimen (Gringeri et al., 2005; McEneny-King et al., 2016). However, there is limited evidence for whether 1 IU/dL is indeed the optimal target value.

Within clinical drug-development there are two broad approaches taken to assist decision making; exposure-response analysis, which seeks to model the relationship between a summary variable (e.g. area under the curve) and a clinical endpoint (e.g. total number of bleeds), and pharmacokinetic/pharmacodynamic (PK/PD) modelling where the concentration timecourse is linked directly to the timecourse of the response (Overgaard et al., 2015). While both approaches have their merits, in this article a PK/PD modelling approach is adopted.

1.2 Motivating dataset

Data from 71 patients with severe haemophilia A (subset of subjects with an endogenous level of FVIII < 1 IU/dL) treated in the pivotal (069901) Advate clinical study were used to illustrate the introduced method. The patients were aged between 10 and 59 years (mean 22 years). Directly prior to commencement of treatment patients received a controlled dose of FVIII. The level of FVIII in the blood at up to ten time points during the following 48 hours were taken.

Patients were subsequently followed up for a duration of around a year during which time they received multiple doses of FVIII, typically at 48 hour intervals. Among these 71 subjects considered, there were 112 spontaneous bleeding events during the study where more information on patient outcomes of this study can be found for example in Collins et al. (2009), Fischer et al. (2011), Björkman et al. (2012) and Shapiro et al. (2015). The distribution of bleeds by patient is given in Table 1. Assuming the rate of bleeds was constant and the population homogeneous, the counts would be Poisson with mean proportional to the time of follow-up. However, the counts are clearly overdispersed with a coefficient of variation of 4.0.

Table 1: Frequency of observed spontaneous bleeds per patient during the study period

Bleeds	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Frequency	31	18	7	4	4	3	2	0	0	0	0	0	1	1

Baseline explanatory factors are available for the patients including age, BMI, regimen of the patient immediately prior to the study (prophylaxis, on demand or combined).

1.3 Approaches to modelling haemophilia bleeds data

Previous analyses of data of this type have analysed the count data directly, for instance using zero-inflated Poisson or negative binomial regression (den Uijl, Fisher et al., 2011). (Wang, 2019, Chapter 5) considered the use of survival analysis techniques to model the time to first event. The effect of past doses is included as time dependent covariates which may be given a PK model form, for instance a two-compartment model. The parameters of the PK model are inferred from the pattern of events themselves rather than external PK data. In this article, a recurrent events modelling approach is proposed for the analysis of the haemophilia bleeding data. Recurrent events modelling is an extension of survival analysis to accommodate outcomes which can occur multiple times for the same patient (Lawless and Cook, 2007, Chapter 1). The intensity of bleeding events is modelled as a function of FVIII activity, taken as a time dependent covariate, and other time-fixed or time-dependent covariates. The FVIII activity level can be inferred by combining information on dosing times with PK model parameters estimated from PK measurements taken prior to a study. However as a result, the values of FVIII activity taken in the model are subject to error and estimates of the effect of FVIII activity will be subject to attenuation bias. Recently, Abrantes et al. (2019) used a recurrent events approach in the context of Bayesian forecasting of bleeding events, but did not account for measurement error. An aim of this paper is therefore to both investigate the degree of bias which results from a naive recurrent events analysis, where error in the time-dependent covariate is ignored, and also develop methods which can account for the error which arises due to the uncertainty in the predicted FVIII activity level.

The remainder of the article is organized as follows. In Section 2, the pharmacokinetic modelling used to estimate FVIII plasma levels is described. Section 3 develops the recurrent events model, including methods to account for uncertainty in the FVIII time-dependent covariate. In Section 4 the methods are applied to a study of patients with haemophilia. Section 5 presents a simulation study to investigate the extent to which naive estimates of covariate effects are biased and the performance of the proposed approach. The paper concludes with a discussion.

2 Pharmacokinetic model

A two-compartment pharmacokinetic (PK) model is assumed for the pharmacokinetic data in terms of a macro constant parametrization. Let d_i denote the dose taken, Y_{ij}^{con} denote the jth measurement taken by patient i which occurred at time t_{ij} after administration of the drug, then

$$\log Y_{ij}^{con} = \log d_i + \log\left[a_{1i}\exp\{-k_{1i}t_{ij}\} + a_{2i}\exp\{-k_{2i}t_{ij}\}\right] + \epsilon_{ij},$$

where $\epsilon_{ij} \sim N(0, \sigma^2)$ and $a_{1i}, a_{2i} \geq 0, k_{1i} > k_{2i} > 0$, $i = 1, \ldots, n$ are individual level macro constants. Note that in practice the error structure of the model could be more complicated where the method can be adapted to other forms to model the residual unexplained variability where we considered a constant proportional error for simplicity.

Potentially, estimates of a patient's parameters, $\boldsymbol{\theta}_i = (a_{1i}, a_{2i}, k_{1i}, k_{2i})'$, can be found by fitting separate non-linear fixed effects regression models, fitted separately to each patient's FVIII activity levels, using maximum likelihood. However, here a non-linear mixed effects model is considered where individual patients' PK parameters are assumed to be random effects (Davidian and Giltinan, 1995, Chapter 9). To accommodate the constraints on $\boldsymbol{\theta}_i$, it is assumed that

$$(\log(a_{1i}), \log(a_{2i}), \log(k_{1i} - k_{2i}), \log(k_{2i}))' \sim N(\boldsymbol{\nu}, \boldsymbol{\Sigma}) \tag{1}$$

where $\boldsymbol{\nu}$ is a vector of fixed effects and $\boldsymbol{\Sigma}$ is an unstructured 4×4 variance-covariance matrix. Note that a clearance parametrization is often used in PK analyses, such that the clearances

and volumes for an individual are taken to be log-normally distributed. Accommodating such a structure would be relatively straightforward via fitting a system of differential equations whereas here we use a parametrization in terms of macro constants for simplicity.

The population parameters $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}$ can be jointly estimated using restricted maximum likelihood (REML) estimation. The individual level parameter estimates, $\hat{\boldsymbol{\theta}}_i$, may then be obtained using the mode of the conditional distribution of $\boldsymbol{\theta}_i$ given the observed data. Moreover, estimates of the uncertainty in $\hat{\boldsymbol{\theta}}_i$ can also be inferred based on the second derivatives of the conditional distribution of $\boldsymbol{\theta}_i$.

The model can be fitted in **R** using the `nlme` or `lme4` packages (Pinheiro et al., 2018; Bates et al., 2015). Estimates of the individual random effects and their conditional variances may be obtained using the method in Lindstrom and Bates (1990) and implemented in the `ranef` function. These estimates are based on the mode and second derivative of the mode of the posterior log density of each random effect with $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}$ set to their REML estimates.

The pharmacokinetic model can be used to estimate the FVIII activity level of a patient at any time during the study follow up period. Let $x_i(t;\boldsymbol{\theta}_i) \geq 0$ represent the true level of FVIII activity at time t. On the premise of linear pharmacokinetics this level is a deterministic function of the patient's dosing history through a pharmacokinetic (PK) model. Specifically

$$x_i(t,\boldsymbol{\theta}_i) = \sum_{j:t_{ij}\leq t} d_{ij}\left[a_{1i}\exp\{-k_{1i}(t-t_{ij})\} + a_{2i}\exp\{-k_{2i}(t-t_{ij})\}\right] \tag{2}$$

where t_{ij} is the time of an infusion of dose level d_{ij}.

3 Bleeding events model

Patients are assumed to be continuously observed up to some right censoring time C_i, which is independent of the bleeding events process. It is assumed that the dosing times are non-informative of bleeding events in the sense that knowledge of the timings of doses provides no additional information on the future rate of bleeds beyond the dose's influence on $x_i(t,\boldsymbol{\theta}_i)$. In addition there may be other, possibly time-dependent, covariates $\mathbf{w}_i(t)$ which are assumed to be observed without error.

The intensity of bleeding events for a subject i at time t and past history $\mathcal{H}_t$ (encompassing the number and times of previous bleeding events) is assumed to be of the form

$$\lambda_i(t,\mathcal{H}_t) = \lambda_0(t)Z_i\exp\left[\mathbf{f}\{x_i(t;\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(t)'\boldsymbol{\gamma}\right],$$

meaning the covariates are assumed to act proportionally on the intensities, where Z_i is an individual level frailty and $\mathbf{f}(x;\boldsymbol{\psi})$ is a known function up to possible unknown parameters $\boldsymbol{\psi}$. Note that it is assumed that conditional on the frailty, the timing of previous bleeding events has no direct bearing on the future intensity of bleeds. For mathematical convenience, it is assumed $Z_i \sim \Gamma(1/\xi, 1/\xi)$. Under this formulation, the bleeding events arise as a Poisson process conditional on the time dependent covariates and the unobserved random effect, while the distribution of counts given the covariates, but marginalizing over Z_i has a negative binomial distribution.

A model of possible clinical interest is of the form

$$f(x_i(t;\boldsymbol{\theta}_i);\psi) = I(x_i(t;\boldsymbol{\theta}_i) < \psi), \tag{3}$$

such that β then represents the log-hazard ratio associated with the FVIII activity level going below a threshold, ψ. If it can be shown that the hazard of bleeds is much higher below some

particular FVIII activity level then it would indicate that the dosing regime should be devised to ensure maintenance of FVIII activity greater than or equal to ψ. While the model in (3) assumes it is the current FVIII activity that governs the bleed risk, the premise of an instantaneous effect could easily be relaxed through a lagged version of the model,

$$f(x_i(t;\boldsymbol{\theta}_i);\psi) = I(x_i(t-u;\boldsymbol{\theta}_i) < \psi),$$

for some lag u.

3.1 Analysis ignoring uncertainty in predicted FVIII activity levels

Naively, we can consider estimating the primary parameter(s) of interest, $\boldsymbol{\beta}$, by ignoring the uncertainty in our estimates, $\hat{\boldsymbol{\theta}}_i$, and replacing $x_i(t;\theta_i)$ with $x_i(t;\hat{\boldsymbol{\theta}}_i)$.

If $x_i(t;\hat{\boldsymbol{\theta}}_i)$ is assumed to be observed without measurement error then standard methods for survival and recurrent events data with time-dependent covariates can be applied (Lawless and Cook, 2007).

A semi-parametric model where $\lambda_0(t)$ is taken to be a non-parametric function of time, and incorporating the Gamma distributed random effect, can be fitted using penalized partial likelihood (PPL) (Therneau et al., 2003). Under this approach, let $\tau_{(1)},\ldots,\tau_{(N_b)}$ be the ordered set of times at which a bleed occurred and $\mathcal{R}_k = \{i : Y_i(\tau_{(k)}) = 1\}$ the set of patients at risk of a bleed at time $\tau_{(k)}$, where $Y_i(t)$ is an indicator variable of whether patient i is under observation at time t. Typically we would take $Y_i(t) = I(t \le C_i)$. Evaluation of the PPL requires evaluation of $f\{x_i(\tau_{(k)};\boldsymbol{\theta}_i),\boldsymbol{\psi}\}$ for all $i \in \mathcal{R}_k$ at $k = 1,\ldots,N_b$. To avoid time consuming and computationally wasteful repeated calculation of (2), a matrix of values of $x_i(\tau_{(k)};\boldsymbol{\theta}_i)$ can be stored upfront.

In this article we instead pursue models with a parametric form for $\lambda_0(t)$. A simple way of permitting flexibility is to let $\lambda_0(t)$ be piecewise-constant between pre-specified cut-points. If desired, the estimates will approach those of the semi-parametric model as the number of pieces increases. In practice, the cut-points would be chosen to ensure a roughly equal amount of patient-years of follow-up for each piece.

Conditional on the frailty Z_i, the likelihood for a patient with bleeding events at times $\tau_{i1},\ldots,\tau_{in_i}$, is given by

$$\begin{aligned} L_i = Z_i^{n_i} \prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \exp\left[\mathbf{f}\{x_i(\tau_{ij};\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(\tau_{ij})'\boldsymbol{\gamma}\right] \\ \times \exp\{-Z_i \int_0^{C_i} Y_i(u)\lambda_0(u)\exp\left[f\{x_i(u;\boldsymbol{\theta}_i);\boldsymbol{\psi}\} + \mathbf{w}_i(u)'\boldsymbol{\gamma}\right] du\}. \end{aligned}$$

The marginal likelihood, integrating out the frailty term is then

$$\begin{aligned} L_i = & \int_0^\infty \left(Z_i^{n_i} \textstyle\prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \exp\left[\mathbf{f}\{x_i(\tau_{ij};\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(\tau_{ij})'\boldsymbol{\gamma}\right] \exp\{-Z_i G_i(\boldsymbol{\theta}_i)\} \right) f_z(Z_i) dZ_i \\ = & \frac{\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \frac{(1/\xi)^{(1/\xi)}}{(G_i(\boldsymbol{\theta}_i)+1/\xi)^{n_i+1/\xi}} \times \textstyle\prod_{j=1}^{n_i}\{\lambda_0(\tau_{ij})\} \times B_i(\boldsymbol{\theta}_i;\boldsymbol{\tau}_i) \end{aligned}$$

where $G_i(\boldsymbol{\theta}_i) = \int_0^{C_i} Y_i(u)\lambda_0(u)\exp\left[f\{x_i(u;\boldsymbol{\theta}_i);\boldsymbol{\psi}\} + \mathbf{w}_i(u)'\boldsymbol{\gamma}\right] du$ and

$$B_i(\boldsymbol{\theta}_i;\boldsymbol{\tau}_i) = \exp\left[\sum_{j=1}^{n_i} \mathbf{f}\{x_i(\tau_{ij};\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(\tau_{ij})'\boldsymbol{\gamma}\right]. \quad (4)$$

The principal computational difficulty in evaluating the likelihood is in evaluating the terms $G_i(\boldsymbol{\theta}_i)$ since they involve analytically intractable integrals of a function of (2). Two methods for approximating these integrals are considered.

3.1.1 Grid approximation

The grid approach involves approximating the integral required for $G_i(\boldsymbol{\theta}_i)$ by using the trapezoidal rule on a fine grid of points, $g_0, g_1, \ldots, g_K$ where $g_k = k \times \delta$, for some small value δ. On this basis

$$G_i(\boldsymbol{\theta}_i) \approx \frac{\delta}{2} \sum_{k=1}^{K} Y_i(g_{k-1})\lambda_0(g_{k-1}) \exp\left[\mathbf{f}\{x_i(g_{k-1};\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(g_{k-1})'\boldsymbol{\gamma}\right] + Y_i(g_k)\lambda_0(g_k) \exp\left[\mathbf{f}\{x_i(g_k;\boldsymbol{\theta}_i);\boldsymbol{\psi}\}'\boldsymbol{\beta} + \mathbf{w}_i(g_k)'\boldsymbol{\gamma}\right].$$

As before, to avoid repeated evaluation of $x_i(t;\boldsymbol{\theta}_i)$, the levels of the time dependent covariate at the grid of points may be computed and stored in an $n \times (K+1)$ matrix $\mathbf{X}$ where $\{\mathbf{X}\}_{ij} = x_i(g_{j-1};\boldsymbol{\theta}_i)$. It may also be computationally advantageous to also store the relevant values of any time dependent elements of $\mathbf{w}_i(t)$.

The advantage of the grid approximation approach is that it can be applied to a wide range of functions $\mathbf{f}(.)$, including the cases where $\mathbf{f}(.)$ depends on unknown parameters, $\boldsymbol{\psi}$.

3.1.2 Crossing point approximation

An alternative approach to calculating $G_i(\theta_i)$ is possible when $\mathbf{f}(x(t);\boldsymbol{\psi})$ is a vector of binary indicators. Of primary interest is the case given in (3), but the method could be extended to situations where there is a series of threshold values. Numerical pre-processing can be used to find all the times satisfying $x_i(t) = \psi,\ 0 < t \leq C_i$. Let $\boldsymbol{\eta}_i = (\eta_{i1}, \eta_{i2}, \ldots, \eta_{im_i})'$ be the set of such crossing times for patient i, then

$$\begin{aligned} G_i(\boldsymbol{\theta}_i) = &\sum_{j=0}^{d_i} \int_{t_{ij}}^{t_{ij+1} \wedge \eta_{k(j)}} Y_i(u)\lambda_0(u) \exp(\mathbf{w}_i(u)'\boldsymbol{\gamma})du \\ &+ \sum_{k=1}^{m_i} \int_{\eta_k}^{t_{j(k)}} Y_i(u)\lambda_0(u) \exp(\beta + \mathbf{w}_i(u)'\boldsymbol{\gamma})du, \end{aligned} \tag{5}$$

where $j(k) = \arg\min\{t_{ij} : t_{ij} > \eta_{ik}\}$ and $k(j) = \arg\min\{\eta_k : \eta_k > t_j\}$. The crossing times, $\boldsymbol{\eta}_i$ can also be used to establish which observed bleed times occurred when the FVIII activity level was above or below ψ. Let $u_i(\boldsymbol{\theta}_i)$ be the number of bleeding events for subject i below ψ given $\boldsymbol{\theta}_i$, then

$$B_i(\boldsymbol{\theta}_i;\boldsymbol{\tau}_i) = \exp\{\beta u_i(\boldsymbol{\theta}_i) + \sum_{i=1}^{n_i} \mathbf{w}_i(\tau_{ij})'\boldsymbol{\gamma}\}. \tag{6}$$

A drawback of this method is that the crossing times must be recalculated for each candidate value of ψ.

3.2 Accounting for uncertainty in predicted FVIII activity levels

There is a well-developed literature on methods for adjusting for measurement error in covariates for survival and event history data. For survival data, the regression-calibration approach (Carroll et al., 1995, Chapter 3) involves using an estimate of the conditional expectation of the true covariate value given the observed value, as the covariate value within the Cox partial likelihood function. Hu et al. (1998) considered a likelihood-based approach. In the context of misclassifed binary covariates, Zucker and Spiegelman (2008) developed a corrected score approach, while

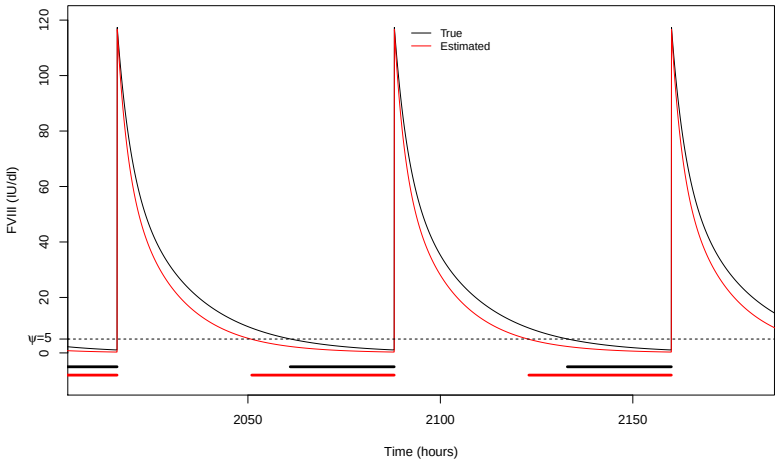

Figure 1: Example partial sequence of true and estimated FVIII activity levels for a patient. The estimated periods of time in which the patient is below the threshold $\psi = 5$ are longer than the true lengths.

Wan et al. (2019) considered an EM algorithm. Andersen and Liestol (2003) considered the attenuation bias arising from intermittent updating of time-dependent covariates.

Relatively little consideration has been given to recurrent events data with misclassification error and in all cases, normally distributed errors are assumed with time fixed covariates. Turnbull et al. (1997) and Jiang et al. (1999) obtained consistent estimators of the model parameters by adjusting the naive estimates using the asymptotic bias, in the parametric and semi-parametric cases, respectively. Yi and Lawless (2012) developed a 'corrected likelihood' function whose conditional expectation is the same as the likelihood with respect to the true covariate values. Yu et al. (2016) extended semi-parametric estimators to accommodate informative censoring in addition to covariate measurement error. In the context of clustered survival data, Li and Lin (2000) allowed the measurement error of covariates from the same cluster to be correlated through a linear mixed model.

For the haemophilia model, the attenuation arises due to the discrepancy between the true PK parameters of the patient and those estimated from the PK data. Suppose, that the risk of a bleed is elevated when the FVIII activity level is below some threshold as given in (3). Figure 1 illustrates that if $\hat{\boldsymbol{\theta}}_i$ differs from $\boldsymbol{\theta}_i$, the periods in which $f(x_i(t;\boldsymbol{\theta}_i);\psi) = 1$ are misidentified. Depending on the direction of the error, this leads to a systematic over- or under-estimation of the time at elevated risk and misclassification of the risk category for observed bleeds.

It is assumed, based on the population PK model considered in Section 2, that

$$g(\boldsymbol{\theta}_i) \sim N(\hat{\boldsymbol{\nu}}_i, \hat{\boldsymbol{\Sigma}}_i),$$

where $\hat{\boldsymbol{\nu}}_i$ is the posterior mode estimate of the random effect and $\hat{\boldsymbol{\Sigma}}_i$ is the conditional variance of the random effect, where each are obtained by plugging in the REML estimates of $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}$. While this assumption is somewhat similar to assumptions made about the the conditional distribution of the true covariates given the measured covariates in previous treatments of measurement errors in recurrent events data, the complexity of the relationship between θ_i and the intensity function

$\lambda_i(t)$ makes existing methods non-applicable.

The error in the time dependent covariate resulting from uncertainty in $\boldsymbol{\theta}_i$ is akin to the presence of an additional frailty term within the model that has a known (or assumed known) distribution. Note that $\hat{\boldsymbol{\Sigma}}_i$ depends on the number and timing of observations for patient i, the residual error variance of individual observations, σ^2 and the variance of the macro constants, $\boldsymbol{\Sigma}$. As such the method accounts for the greater uncertainty (and larger attenuation bias of the uncorrected estimate) when the residual errors, ϵ_{ij}, have a higher variance and when there is greater population heterogeneity.

Under the parametric approach the likelihood contribution for an individual then becomes

$$\begin{aligned} L_i = & \int_{\mathbf{v}\in\mathbb{R}^p} \frac{\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \frac{(1/\xi)^{(1/\xi)}}{(G_i(\mathbf{v})+1/\xi)^{n_i+1/\xi}} \times \left(\prod_{j=1}^{n_i} \lambda_0(\tau_{ij})\right) \times B_i(\mathbf{v};\boldsymbol{\tau}_i) f_{\mathbf{v}}(\mathbf{v};\hat{\boldsymbol{\theta}}_i,\hat{\boldsymbol{\Sigma}}_i) d\mathbf{v} \qquad (7) \\ L_i = & \prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \times \frac{(1/\xi)^{(1/\xi)}\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \int_{\mathbf{v}\in\mathbb{R}^p} \frac{B_i(\mathbf{v};\boldsymbol{\tau}_i) f_{\mathbf{v}}(\mathbf{v};\hat{\boldsymbol{\theta}}_i,\hat{\boldsymbol{\Sigma}}_i)}{(G_i(\mathbf{v})+1/\xi)^{n_i+1/\xi}} d\mathbf{v}, \end{aligned}$$

where p is the dimension of $\boldsymbol{\theta}_i$ and $B_i(\mathbf{v};\boldsymbol{\tau}_i)$ is the quantity as defined in (4) applied at $\boldsymbol{\theta}_i = \mathbf{v}$.

The integral over $\boldsymbol{\theta}_i$ is intractable, but can be approximated using multivariate Gaussian quadrature where the integral is approximated by a weighted sum over N_q quadrature points. Note that since the distribution of $\boldsymbol{\theta}_i$ is fixed, the quadrature points for a given subject, $\mathbf{v}_{i1},\ldots,\mathbf{v}_{iN_q}$, are also fixed. As such to avoid unnecessary computation, the matrix of values of $x_i(t,\boldsymbol{\theta}_i)$ considered in Section 3.1 can be expanded into a $n\times(K+1)\times N_q$ array where $\{\mathbf{X}\}_{ijk} = x_i(g_{j-1};\mathbf{v}_{ik})$.

Hence

$$L_i \approx \prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \times \frac{(1/\xi)^{(1/\xi)}\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \times \sum_{k=1}^{N_q} \frac{B_i(\mathbf{v}_{ik};\boldsymbol{\tau}_i) w_{ik}}{(G_i(\mathbf{v}_{ik})+1/\xi)^{n_i+1/\xi}}, \qquad (8)$$

for quadrature weights $w_{i1},\ldots,w_{iN_q}$.

3.2.1 Approximation for threshold models

Under the model in (3), and for a given $\boldsymbol{\theta}_i$ the likelihood can be characterized in terms of the crossing times where $x_i(t;\boldsymbol{\theta}_i) = \psi$. As such, the variability in $\boldsymbol{\theta}_i$ is only relevant through its effect on those crossing times. Moreover, if the dosing schedule is fairly regular and the dose fixed, the activity level curves will approach an approximate equilibrium such that $x(t+t_k;\boldsymbol{\theta}) \approx x(t+t_{k'};\boldsymbol{\theta})$ for any dose times t_k and $t_{k'}$ and $0 < t < t_{k+1} - t_k$. Specifically, under (2) and assuming a long sequence of doses of size d, spaced u apart, $x(t+t_k;\boldsymbol{\theta})$ would approach

$$D(t;\boldsymbol{\theta}) = d\left\{\frac{a_1\exp(-k_1 t)}{1-\exp(-k_1 u)} + \frac{a_2\exp(-k_2 t)}{1-\exp(-k_2 u)}\right\}, 0 \le t < t_{k+1} - t_k.$$

Let υ satisfy $D(\upsilon;\boldsymbol{\theta}) = \psi$, then $\sigma_\upsilon^2 = \mathrm{Var}(\upsilon(\hat{\theta}))$ gives an approximation to the variance of the crossing times. Hence the crossing times, η_{ik}, will all have approximately the same dependence on $\boldsymbol{\theta}$. Again using the assumption $g(\boldsymbol{\theta}_i) \sim N(g(\hat{\boldsymbol{\theta}}_i),\hat{\boldsymbol{\Sigma}}_i)$, the delta method may be used to approximate σ_υ^2. For each dose time t_j, $j = 0,1,\ldots,d_i$, let $\hat{\eta}_{j+1} > t_j$ be the time at which $x(t;\hat{\boldsymbol{\theta}}_i)$ would next cross ψ assuming no further doses after t_j. As such, $\hat{\boldsymbol{\eta}}_i = (\hat{\eta}_{1i},\hat{\eta}_{2i},\ldots,\hat{\eta}_{d_i+1,i})'$ can be thought of as the vector of *potential* crossing times for patient i.

Assuming υ itself is approximately normal, then the potential crossing times for $\boldsymbol{\theta}_i$ can be expressed as $\eta_k \approx \hat{\eta}_k + \omega$ for $k = 1,\ldots,m$ where $\omega \sim N(0,\sigma_\upsilon^2)$.

Table 2: Fixed effect estimates for the non-linear mixed effects model.

Parameter	Est	SE
ν_1	-0.081	0.070
ν_2	0.362	0.048
ν_3	-1.687	0.109
ν_4	-2.770	0.024

The expression for $G_i(\boldsymbol{\theta}_i)$ in (5) can then be adapted to account for the possibility that the patient received a new dose before the next crossing time,

$$G_i(\omega) = \sum_{j=0}^{d_i} \int_{t_j}^{\hat{\eta}_{ij+1}+\omega \wedge t_{j+1}} Y_i(u)\lambda_0(u)du + \sum_{j=1}^{d_i+1} \int_{\hat{\eta}_{ij}+\omega \wedge t_j}^{t_j} Y_i(u)\lambda_0(u)\exp(\beta)du. \tag{9}$$

In addition, the crossing times can again be used to establish the number of bleeds that occurred below τ, denoted $u_i(\omega)$, which can be used to determine $B_i(\omega) = \exp\{\beta u_i(\omega) + \sum_{i=1}^{n_i} \mathbf{w}_i(\tau_{ij})'\boldsymbol{\gamma}\}$, the analogous quantity to (6).

The overall likelihood contribution for subject i may then be be approximated by a one-dimensional integral

$$\begin{aligned} L_i \approx & \quad \prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \times \frac{(1/\xi)^{(1/\xi)}\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \times \int_{-\infty}^{\infty} \frac{B_i(\omega)}{(G_i(\omega)+1/\xi)^{n_i+1/\xi}} \frac{1}{\sqrt{2\pi\sigma^2_{v_i}}} \exp(-\omega^2/2\sigma^2_{v_i})d\omega \\ L_i \approx & \quad \prod_{j=1}^{n_i} \lambda_0(\tau_{ij}) \times \frac{(1/\xi)^{(1/\xi)}\Gamma(n_i+1/\xi)}{\Gamma(1/\xi)} \times \sum_{k=1}^{N_q} \frac{B_i(v_{ik})}{(G_i(v_{ik})+1/\xi)^{n_i+1/\xi}} w_{ik}, \end{aligned}$$

where v_{ik} and w_{ik}, $k = 1, \ldots, N_q$ are the Gauss-Hermite quadrature nodes and weights for an integral with respect to $N(0, \sigma^2_{v_i})$ variable.

The observed interval between doses or the doses themselves will not be fixed for a patient, however it may be reasonable to take the mean dose and mean dosing interval to estimate $\sigma^2_{v_i}$. In principle, more sophisticated approximations can be considered that adjust the variance associated with a given crossing time on the basis of the precise past dosing regime. For the example dataset, the enhanced approximation gave almost identical results, but this would not necessarily be the case if the dosing schedules were more variable.

4 Application

The non-linear mixed effects model specified in Section 2 is fitted to the PK data for the 71 patients. Fixed parameter estimates from the model are given in Table 2 and the variance component estimates in Table 3. Note that the estimate of $\sigma^2 = 0.017$ corresponds to a percentage coefficient of variation of 13%.

Corresponding parameter estimates in terms of a clearance parametrization obtained from the model fitted in terms of the macro-constant parametrization were 2.1 dL/h for clearance, 1.1 dL/h for inter-compartmental clearance, 23.7 dL for the central volume of distribution and 6.3 dL for the peripheral volume of distribution, which is in-line with other publications of FVIII models (Björkman et al., 2012).

The marginal cumulative intensity function estimated using the Nelson-Aalen estimator is shown in Figure 2. Pointwise 95% confidence intervals are obtained using the robust variance estimator in Lawless and Nadeau (1995). The constant gradient of the curve indicates that it

Table 3: Variance-covariance estimates of the random effects for the non-linear mixed model.

	Σ			
ν_1	0.098	-0.025	-0.033	0.003
ν_2		0.027	0.049	-0.068
ν_3			0.173	-0.120
ν_4				0.238
Residual (σ^2)	0.017			

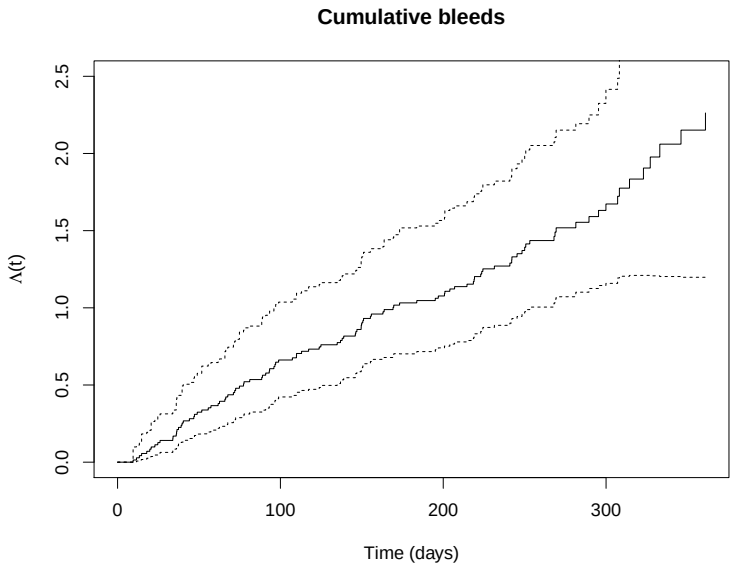

Figure 2: Nelson-Aalen estimate of the mean count of cumulative bleeds per patient over the study period. Pointwise 95% confidence intervals are indicated by the dashed lines.

may be reasonable to assume that the baseline intensity, $\lambda_0(t)$, is constant, rather than piecewise constant, over the study period.

Assuming a threshold model with $\psi = 1$ and without covariates, the ratio of rates of bleeds when FVIII activity is below 1% compared to above is estimated to be 4.00 95 % CI: $(2.64 - 6.05)$ ignoring the attenuation bias, 4.31 (95 % CI: $2.83 - 6.56$) if the bias is adjusted for using the approximate approach in Section 3.2.1 based on 10 point unidimensional quadrature and 4.31 (95 % CI: $2.82 - 6.58$) if the bias is adjusted for using four dimensional Gaussian quadrature with 256 node points. Hence under the assumption of normally distributed PK parameters and independent normal errors on the log-scale, ignoring uncertainty in the estimated FVIII activity levels has a small, but discernible effect in this case. In addition, the approximate method agrees quite well with the more computationally intensive full quadrature method, including giving a similar value for the deviance ($-2 \log L$). The model parameters for each of the analyses is given in Table 4.

Table 5 presents the estimated effect of explanatory variables on the intensity of bleeding events based on a multi-variable model including all variables together. The estimates presented are based on the unidimensional approximation for correction, but the estimates using the other

Table 4: Comparison of parameter estimates for a threshold model with $\psi = 1$ % based on a naive analysis and an analysis correcting for uncertainty in the individual PK parameter estimates. 'Uncorrected' = Estimates ignoring measurement uncertainty, 'Approx correct' = Estimates using the approximation in Section 3.2.1, 'Full correct' = method using 4-d Gaussian quadrature.

	Uncorrected			Approx corrected			Full corrected		
Parameter	Est	SE	95% CI	Est	SE	95% CI	Est	SE	95% CI
$\log \lambda_0$	-8.639	0.188	(-9.01, -8.27)	-8.665	0.190	(-9.04, -8.29)	-8.665	0.190	(-9.04,-8.29)
β	1.386	0.213	(0.97, 1.80)	1.461	0.215	(1.04, 1.88)	1.460	0.216	(1.04, 1.88)
$\log \xi$	0.404	0.287	(-0.16, 0.97)	0.400	0.288	(-0.17, 0.96)	0.398	0.288	(-0.17,0.96)
$-2 \log L$	1968.02			1964.82			1965.22		

Table 5: Estimated effects of explanatory variables using the unidimensional approximation approach.

Covariate	log HR	SE	95% CI	p
Current FVIII $< 1\%$	1.479	0.215	(1.057, 1.901)	< 0.001
Age (yrs)	0.018	0.018	(-0.017, 0.052)	0.313
BMI	-0.087	0.048	(-0.180, 0.007)	0.069
Previous: Prophylaxis	0	-	-	-
Previous: On Demand	0.350	0.437	(-0.506, 1.206)	0.423
Previous: Combined	0.747	0.399	(-0.035, 1.529)	0.061

methods are broadly similar. The inclusion of other covariates has little effect on the estimated effect of a current FVIII activity below 1%, with $\hat{\beta}$ increasing from 1.461 to 1.478. There is some indication of a lower incidence of spontaneous bleeds in patients with higher BMI. In addition, patients who were on a combined therapy before the start of the study had a higher rate of bleeds than those who were on prophlyaxis. Calculation of the gradient function (Verbeke and Molenberghs, 2013) suggested the assumption of a Gamma distributed patient-level random effect for the bleeds is reasonable, although the relative small sample size makes it difficult to diagnose the random effects distribution.

4.1 Spline effect model

While the threshold model considered in (3) is clinically interesting, it does not reflect the observed pattern of bleeds within the data. An alternative model assumes that the rate of bleeds is a smooth function of the FVIII activity level. For instance, $f(x_i(t; \boldsymbol{\theta}_i))$ may be taken to be a set of natural spline basis functions, such that the log-hazard is a piecewise polynomial function over the range of values of $x_i(t; \boldsymbol{\theta}_i)$. Since it is reasonable to assume that the incidence of bleeds will not increase with increasing FVIII activity level, a monotone spline is applied by using I-splines with non-positive coefficients (Ramsay, 1988). A cubic I-spline is chosen with internal knot points at FVIII activity levels 1, 5 and 25. The grid approximation approach is used to compute the likelihood and four dimensional Gaussian quadrature with 256 nodes used to integrate over the distributions of $\boldsymbol{\theta}_i$ for the corrected method.

Figure 3 gives the estimated intensity curves as a function of FVIII activity for the corrected analyses. The spline model suggests the rate of bleeds continues to increase for decreasing FVIII activity below 1%. Qualitatively the estimated relationship is similar to the estimated relationship between bleed rate and baseline FVIII activity in den Uijl, Mauser Bunschoten et al. (2011), who considered patients with mild, moderate and severe haemophilia but did not adjust for the effect of treatments.

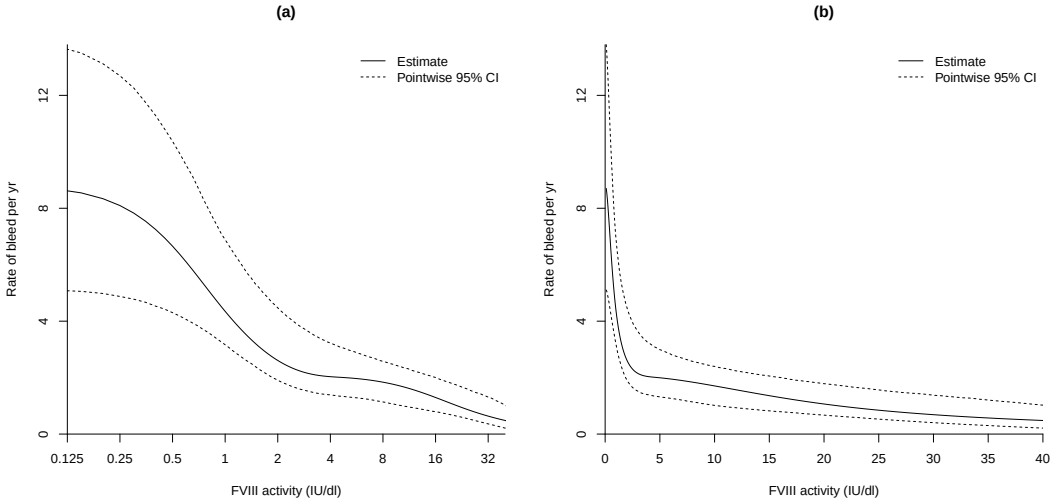

Figure 3: Estimated rate of bleeds by FVIII activity based on a monotonic I-spline using the corrected method based on Gaussian quadrature. Panel (a) presents the estimates with a log axis for FVIII activity, and panel (b) is on the natural scale.

For the spline effect model, the naive estimates almost entirely coincided with the corrected estimates and are therefore not presented. The similarity of estimates suggests that neglecting the uncertainty in estimated FVIII activity has considerably less effect for smooth models compared to threshold models.

5 Simulation study

5.1 Performance of the threshold model estimators

The procedures proposed in Section 3.2 increase the complexity of modelling. It is therefore of interest to determine under which circumstances failure to account for uncertainty in $\boldsymbol{\theta}_i$ will lead to a practical level of bias and also to investigate the effectiveness of the proposed corrected likelihood.

We again focus on the case where the influence of FVIII activity on bleeding events is through a known threshold value $\psi > 0$ such that $f(x; \boldsymbol{\beta}) = I(x < \psi)\beta$.

The PK model fitted to the example dataset in Section 4 is used to devise potential degrees of estimation uncertainty in the parameters $\boldsymbol{\theta}$. It is assumed that the distribution of true PK parameters is the same as those estimated in the population PK model. The level of uncertainty is varied by increasing the variance in the observation errors for the PK observations, or alternatively by decreasing the number of PK observations per patient. Three cases are considered,

1. Ten PK observations and the observation error has standard deviation equal to that estimated in the actual dataset.

2. Ten PK observations and the observation error has double the observed standard deviation estimated in the actual dataset.

3. Six PK observations and the observation error has standard deviation equal to that estimated in the actual dataset.

In each case constant proportional residual error is assumed, as was also assumed in Section 2. In addition, two values are considered for the critical threshold, $\psi = 1$ and $\psi = 5$. In all cases there is assumed to be a hazard of ratio of 20/3 between low and high FVIII activity levels, meaning the true $\beta = 1.897$. In each simulation, a non-linear mixed effects model is fitted to the simulated PK data to gain the estimated PK parameters and estimates of the associated uncertainty. In some samples the full model with four random effects does not converge, in which case the best fitting model with three random effects, and the fourth treated as a fixed effect, is used instead. This aims to automatically emulate what an investigator might do when faced with non-convergence of the proposed PK model. However, in practice, one should try different starting values and ensure any simplified model is still biologically plausible.

As an additional benchmark, we also consider the case where the true individual PK parameters are known.

To investigate the performance of the proposed estimators, the approach proposed in Section 3.2.1 are applied to each of the scenarios. Total sample sizes of $N = 100$ and $N = 200$ patients are considered. For each scenario, 5000 replicate datasets are simulated.

Table 6 gives a summary of the results for the six scenarios. Since the true log-hazard ratio is positive, the attenuation bias is negative in this case. In all scenarios the adjusted estimator is less negatively biased than the naive unadjusted estimate and also has lower mean squared error. There is a higher degree of attenuation bias in the unadjusted estimates when $\psi = 1$ compared to when $\psi = 5$, reflecting a greater degree of uncertainty in the true FVIII activity level crossing times. The adjusted estimator also retains some bias in the case where $\psi = 1$, which is perhaps due to a violation of the assumption that the crossing times are normally distributed. Increasing the standard deviation of individual PK measurements appears to have a greater impact than reducing the number of PK measurements in terms of the degree of attenuation bias. It is notable that for the models with higher residual error or fewer PK measurements the full four dimensional mixed effects model fails to converge a high proportion of the time and a simplified model is fitted instead. This is most likely due to the occurrence of boundary estimates (i.e. where the estimated variance matrix of the random effects is singular). The good performance of the adjusted estimates despite this indicates some robustness to misspecification of the precise PK model.

Table 7 gives the corresponding standard deviation of the estimates, average standard errors and empirical coverage of nominal 95% confidence intervals for each method for the six scenarios. In most cases the adjusted estimator results in intervals with close to 95% coverage. The only exceptions are when $\psi = 1$ and the residual standard deviation of the PK measurements is higher.

5.2 Performance under model misspecification

Since the proposed method makes various assumptions about the PK model, a further set of simulations are carried out to assess the robustness of the method to model misspecification.

Some key assumptions of the PK model are;

1. An underlying multivariate normal random effects distribution for the patient's individual PK parameters based as defined in (1).

2. Normally distributed errors with homogeneous variance with respect to the measured logged FVIII activity levels.

Table 6: Simulation results for impact on the bias and mean squared error (MSE) for six scenarios with varying levels of measurement error in the PK model. Prop converge displays the percentage of samples for which the full four dimensional random effects model did not converge and a simplified model was used instead.

			Bias			MSE			Prop
N	ψ	Case	Unadj	Adj	No error	Unadj	Adj	No error	converge
100	5	(a)	-0.0855	-0.0076	0.0055	0.0424	0.0388	0.0371	94.6%
200	5	(a)	-0.0842	-0.0077	0.0065	0.0238	0.0180	0.0172	96.3%
100	5	(b)	-0.1435	-0.0329	0.0118	0.0543	0.0390	0.0363	39.1%
200	5	(b)	-0.1467	-0.0370	0.0061	0.0382	0.0202	0.0180	47.0%
100	5	(c)	-0.0883	-0.0022	0.0189	0.0421	0.0382	0.0371	65.4%
200	5	(c)	-0.1008	-0.0162	0.0054	0.0271	0.0190	0.0181	72.4%
100	1	(a)	-0.1603	-0.0578	-0.0133	0.0907	0.0702	0.0615	95.3%
200	1	(a)	-0.1540	-0.0518	-0.0097	0.0537	0.0334	0.0285	97.0%
100	1	(b)	-0.2516	-0.1057	-0.0106	0.1280	0.0794	0.0612	61.2%
200	1	(b)	-0.2411	-0.0986	-0.0086	0.0889	0.0419	0.0284	51.6%
100	1	(c)	-0.1774	-0.0523	-0.0087	0.0972	0.0714	0.0605	35.4%
200	1	(c)	-0.1681	-0.0497	-0.0051	0.0607	0.0353	0.0301	28.2%

Table 7: Simulation results for impact on uncertainty estimates for six scenarios with varying levels of measurement error in the PK model

			SD			$\widehat{\text{SE}}$			Coverage		
N	ψ	Case	Unadj	Adj	No error	Unadj	Adj	No error	Unadj	Adj	No error
100	5	(a)	0.1873	0.1968	0.1926	0.1841	0.1924	0.1890	0.9122	0.9492	0.9544
200	5	(a)	0.1295	0.1339	0.1311	0.1295	0.1353	0.1328	0.8870	0.9466	0.9496
100	5	(b)	0.1835	0.1946	0.1901	0.1813	0.1930	0.1890	0.8518	0.9436	0.9534
200	5	(b)	0.1291	0.1371	0.1340	0.1274	0.1355	0.1328	0.7739	0.9356	0.9558
100	5	(c)	0.1852	0.1954	0.1918	0.1834	0.1924	0.1888	0.9024	0.9472	0.9504
200	5	(c)	0.1302	0.1370	0.1343	0.1289	0.1352	0.1327	0.8574	0.9466	0.9490
100	1	(a)	0.2549	0.2585	0.2477	0.2462	0.2506	0.2431	0.9146	0.9454	0.9458
200	1	(a)	0.1731	0.1753	0.1686	0.1721	0.1751	0.1701	0.8731	0.9455	0.9534
100	1	(b)	0.2543	0.2611	0.2472	0.2505	0.2560	0.2438	0.8380	0.9306	0.9514
200	1	(b)	0.1753	0.1794	0.1683	0.1740	0.1780	0.1703	0.7251	0.9152	0.9511
100	1	(c)	0.2564	0.2621	0.2459	0.2506	0.2551	0.2432	0.9078	0.9486	0.9538
200	1	(c)	0.1800	0.1813	0.1735	0.1742	0.1774	0.1703	0.8521	0.9417	0.9495

3. No inter-occasional variability (IOV), meaning an individual's PK parameters are not different at different dosing occasions (Karlsson et al., 1993).

To investigate the sensitivity to each of these assumptions, the first case of the simulations in the main simulation ($\psi = 5$, 10 PK observations and residual standard error as that in the illustrative dataset) is adopted but with varying forms of misspecification with respect to each of the assumptions in turn.

To create a skewed, non-normal distribution, data are generated from a two-point mixture distribution that matches the mean and variance. The degree of skewness is varied by altering the mixture probability, we consider $\pi = 0.5$, $\pi = 0.35$ and $\pi = 0.2$, corresponding to symmetric, moderately skewed and heavily skewed distributions.

To investigate the assumption of homogeneous error variance, we generate PK data from a model where the measurement error depend on the log-mean response. Specifically, let $\mu_{ij} = \mathbb{E}(\log Y_{ij}^{con})$ then

$$\text{Var}(\epsilon_{ij}) = \frac{\sigma_0^2}{(1 + \xi\mu_{ij})^{2\zeta}}$$

where ξ and ζ controls the strength of the mean-variance relationship. We consider cases where $\xi = 1$ and $\xi = 2$ and $\zeta = 1$, corresponding to mild and moderate dependence, and a further case where $\xi = 0.2$ and $\zeta = 2$, giving a more severe scenario. In each case σ_0 is chosen to ensure the average error variance is similar to in the original study.

To investigate IOV, the data for the PK measurements and also the subsequent bleeding events are generated based on a model that assumes a three-level structure for the PK model. The PK parameters governing the clearance of the dose d_{ij} for patient i at occasion j are assumed to be governed by

$$\mathbf{g}_{ij} = \mathbf{g}_0 + \mathbf{b}_i + \mathbf{u}_{ij}$$

where $\mathbf{g}_{ij} = (\log(a_{1ij}), \log(a_{2ij}), \log(k_{1ij} - k_{2ij}), \log(k_{2i}))'$, $\mathbf{b}_i \sim N(\mathbf{0}, \Sigma^*)$ and $\mathbf{u}_{ij} \sim N(\mathbf{0}, \Omega)$ are independent random effects. For simplicity we assume $\Sigma^* = \rho\Sigma$ and $\Omega = (1-\rho)\Sigma$, where Σ is the random effects variance used in the other simulations. As such ρ controls the proportion of the total variance which is common to all dose occasions. We consider cases where $\rho = 0.95, \rho = 0.9$ and $\rho = 0.8$.

The results, shown in Table 8, indicate that non-normality of the distribution of PK parameters has the least impact on the proposed method, with the results in each of the three scenarios being comparable to the first scenario in Table 6. The presence of heterogeneity of variance with respect to the logged FVIII activity levels in the PK data has a greater impact. While the adjusted method reduces the bias and mean squared error compared to the unadjusted method, the estimates are somewhat more biased than the comparable results when the residual error model is correctly specified. The presence of inter-occasional variability (IOV) has the greatest impact on the performance of the model. Both the unadjusted and adjusted estimates are more biased, with the bias increasing when the IOV accounts for a higher proportion of the total variance.

5.3 Continuous effect models

In Section 4.1 it was notable that the naive and adjusted estimates of the spline model were almost identical, which indicates that bias in the naive method may be less of a problem for models that assume a continuous effect of FVIII activity level on the intensity of events. To investigate this further, PK data are generated from scenario (a) from Section 5.1, using $n = 100$ but with the bleeding events simulated using a monotone spline effect (as in Section 4.1). The averaged estimates over 500 simulations under the naive method, compared to the true effect

Table 8: Simulation results for nine scenarios with varying levels of model misspecification of the PK or PD model. 'No error' estimates for the IOV sceanrios are absent due to no clear method of estimation.

			Bias			SD			MSE		
N	Case	Par	Unadj	Adj	No error	Unadj	Adj	No error	Unadj	Adj	No error
100	(a)	$\pi = 0.5$	-0.0827	-0.0020	0.0087	0.1882	0.1966	0.1930	0.0422	0.0386	0.0373
200	(a)	$\pi = 0.5$	-0.0845	-0.0053	0.0056	0.1299	0.1353	0.1329	0.0240	0.0183	0.0177
100	(a)	$\pi = 0.35$	-0.0815	-0.0043	0.0084	0.1845	0.1926	0.1892	0.0407	0.0371	0.0359
200	(a)	$\pi = 0.35$	-0.0835	-0.0062	0.0059	0.1295	0.1348	0.1328	0.0238	0.0182	0.0177
100	(a)	$\pi = 0.2$	-0.0824	0.0028	0.0151	0.1882	0.1968	0.1919	0.0422	0.0387	0.0370
200	(a)	$\pi = 0.2$	-0.0922	-0.0068	0.0056	0.1292	0.1359	0.1339	0.0252	0.0185	0.0180
100	(b)	$\xi = 1, \zeta = 1$	-0.1272	-0.0335	0.0093	0.1835	0.1902	0.1879	0.0499	0.0373	0.0354
200	(b)	$\xi = 1, \zeta = 1$	-0.1334	-0.0379	0.0054	0.1317	0.1375	0.1354	0.0351	0.0203	0.0184
100	(b)	$\xi = 2, \zeta = 1$	-0.1388	-0.0421	0.0109	0.1842	0.1937	0.1920	0.0532	0.0393	0.0370
200	(b)	$\xi = 2, \zeta = 1$	-0.1431	-0.0475	0.0036	0.1289	0.1349	0.1330	0.0371	0.0204	0.0177
100	(b)	$\xi = 0.2, \zeta = 2$	-0.1025	-0.0151	0.0099	0.1848	0.1921	0.1889	0.0447	0.0371	0.0358
200	(b)	$\xi = 0.2, \zeta = 2$	-0.1072	-0.0215	0.0047	0.1263	0.1330	0.1312	0.0275	0.0181	0.0172
100	(c)	$\rho = 0.95$	-0.1732	-0.0792	-	0.1818	0.1919	-	0.0630	0.0431	-
200	(c)	$\rho = 0.95$	-0.1777	-0.0844	-	0.1276	0.1347	-	0.0479	0.0253	-
100	(c)	$\rho = 0.9$	-0.2218	-0.1276	-	0.1806	0.1888	-	0.0818	0.0519	-
200	(c)	$\rho = 0.9$	-0.2283	-0.1349	-	0.1271	0.1327	-	0.0683	0.0358	-
100	(c)	$\rho = 0.8$	-0.3052	-0.2114	-	0.1786	0.1872	-	0.1250	0.0801	-
200	(c)	$\rho = 0.8$	-0.3057	-0.2149	-	0.1244	0.1300	-	0.1089	0.0631	-

curve and compared to the estimator using the true individual PK parameter values is given in Figure 4.

There is no discernible difference in the mean estimate for the naive estimates, the corrected estimates and the estimates assuming the true PK parameters. All estimates display some negative bias in the estimated rate for low FVIII activity levels for both estimators. This is presumably due to small sample bias and the limited amount of person-time spent at low FVIII activity levels.

Additional simulations were performed for the other scenarios of PK data generation from Section 5.1, without calculating the corrected estimate, but giving very similar results with respect to the naive and true estimates. The naive estimator was observed to produce close to unbiased estimates of the log-intensity function in a corresponding set of simulations which instead assumed a log-linear effect of FVIII activity on bleeds.

6 Discussion

This paper has presented an approach to incorporating pharmacokinetic information about patients into a recurrent events model of bleeding events and proposed methods for adjusting the analysis to avoid attenuation bias due to uncertainty in the PK parameter estimates. A limitation of the example analysis in Section 4 is the size of the dataset. In particular, the degree of patient heterogeneity potentially makes the results quite sensitive to the inclusion or exclusion of patients who had a higher propensity for bleeds.

While the paper has concentrated upon a specific application, the methods have broad application to any drug where the drug level is expected to be a covariate for the occurrence of events. In addition, this approach might be also useful for the purpose of assessing a potential impact of concentrations on the occurrence of adverse events.

The proposed method is able to account for uncertainty in the estimated FVIII activity levels arising from the residual unexplained variation and the population heterogeneity. The methods proposed make an assumption that the conditional distribution, $\boldsymbol{\theta}_i$ is multivariate normal, at least on a transformed scale. This may be unreasonable in some situations due to the limited

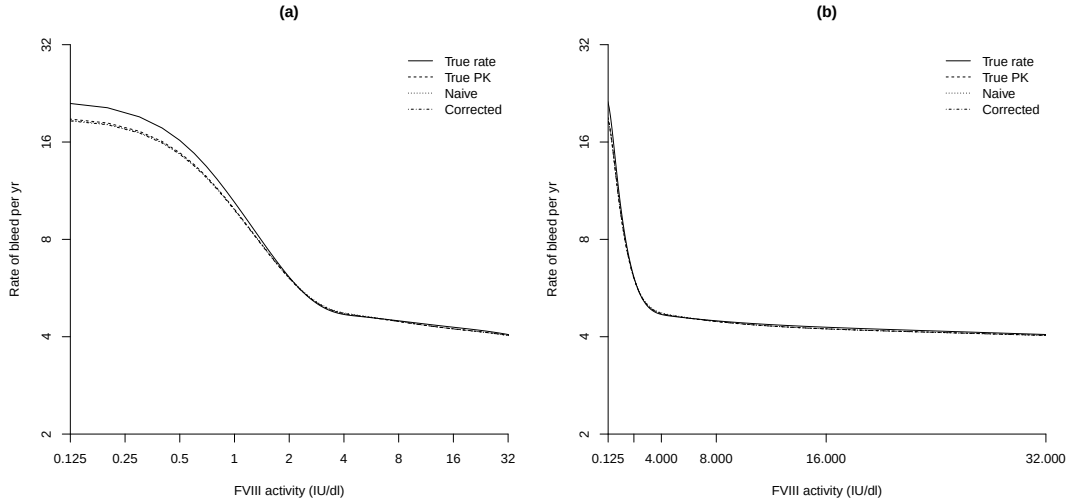

Figure 4: Average estimated rate of bleeds by FVIII activity based on 500 simulated datasets assuming a monotone I-spline. (a) Displayed with FVIII activity on the log-scale, (b) FVIII activity on the natural scale.

information about the parameters. An alternative approach would be to use MCMC techniques to simulate from the posterior distributions of $\boldsymbol{\theta}_i$. In this case the form of the likelihood would still be as in (8), but $\mathbf{v}_{ik}$ would represent the kth sample from the posterior for patient i, N_q the number of posterior samples and the weights would be taken to be $w_{ik} \equiv 1/N_q$. A greater number of posterior samples may be required to get a similar degree of accuracy to the quadrature points. The estimate of conditional variance of $\boldsymbol{\theta}_i$ assumed in our approach does not incorporate the uncertainty about the fixed model parameters. In the motivating data example, the PK model parameters were obtained from extensive sampling based on up to 8 measurement time points per subject and as a result the intra-subject variability is the main driving factor of the overall uncertainty. However, in other circumstances the PK parameters might be estimated from sparse sampling or obtained from a population PK model built using external data and, as such, PK parameter uncertainty may have a greater influence. Further accuracy could potentially be gained through adopting a fully Bayesian approach to the PK modelling which could incorporate the additional uncertainty from the fixed model parameters. Moreover, there could be additional uncertainty in the specification of the PK model, which could in principle be accommodated through model averaging approaches.

The simulations indicated that the threshold model can be well estimated by using a one dimensional approximation to the dependence of the random effects. It may therefore be reasonable to use this approximation rather than the full quadrature method, provided dosing schedules are reasonably regular. The simulations also suggest the proposed method is reasonably robust to violations of the model assumptions relating to the random effects distribution of the PK parameters and to a lesser extent assumptions about the residual variance of the FVIII activity levels in the PK data. However, if inter-occasional variability is present then the adjusted estimate only removes a portion of the attenuation bias. Provided repeated PK data were available to establish the IOV, the approximate method for threshold models in Section 3.2.1 could be extended to accommodate the additional variability arising through IOV. Extending the general

approach, for instance for spline effect models, is more challenging since the IOV would induce a sequence of random effects for each patient, which in principle would require integrating out to obtain the marginal likelihood.

The dataset considered in the paper considered only patients with severe haemophilia A, for whom it might be plausible to assume their FVIII activity levels could go down to zero. To include non-severe haemophilia A patients (i.e. those with an endogenous FVIII level greater than 1 IU/dL) it would be necessary to adjust the PK model for FVIII activity to account for each patients' baseline endogenous level (Bauer and Wolfsegger, 2014).

Determining the optimal target FVIII activity threshold, which could be interpreted as the level of ψ associated as the most significant increase in the hazard of a bleed, is of particular clinical interest. In principle, ψ may be treated like any other parameter within the model and jointly maximized. However, although there are two parameters to be optimized, the likelihood ratio between the full model and a null model where FVIII activity is assumed to have no effect such that $\beta = 0$ does not have a standard asymptotic χ^2_2 distribution under the null, due to ψ not being identifiable when $\beta = 0$ (Davies, 1987). There may be scope to extend results for the simpler situation of an optimal cut-point for the dichtomization of a continuous time fixed covariate (Lausen and Schumacher, 1992), to provide a test for an overall effect of FVIII activity via a threshold. A further difficulty is in establishing standard error estimates for the parameters which incorporate selection of ψ, for which bootstrap may be a possible, if time-consuming, solution.

Acknowledgements

We thank Kristof Haraszti for assistance in data management and thank Alexander Bauer and Werner Engl for very helpful discussions on this topic.

This report is independent research arising in part from Prof Jaki's Senior Research Fellowship (NIHR-SRF-2015-08-001) supported by the National Institute for Health Research. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care (DHCS).

This study is also partly funded by Baxalta Innovations GmbH, a Takeda company, Vienna, Austria. Martin Wolfsegger is an employee of Takeda and owns Takeda stock/shares.

References

Abrantes JA, Solms A, Garmann D, Nielsen EI, Jönsson S, Karlsson MO. (2019) Bayesian Forecasting Utilizing Bleeding Information to Support Dose Individualization of Factor VIII. *CPT: Pharmacometrics and Systems Pharmacology.* **8**:894–903.

Andersen PK, Liestol K. (2003) Attenuation caused by infrequently updated covariates in survival analysis. *Biostatistics* **4**: 633–649.

Bates D, Mächler M, Bolker B, Walker S. (2015) Fitting Linear Mixed-Effects Models Using R. *Journal of Statistical Software* **67**: 1–48.

Bauer A, Wolfsegger MJ. (2014) Adjustment of endogenous concentrations in pharmacokinetic modeling. *European Journal of Clinical Pharmacology* **70**: 1465–1470.

Björkman S, Oh M, Spotts G, Schroth P, Fritsch S, Ewenstein BM, Casey K, Fischer K, Blanchette VS, Collins PW. (2012). Population pharmacokinetics of recombinant factor VIII: the relationships of pharmacokinetics to age and body weight. *Blood* **119**: 612–618.

Carroll RJ, Ruppert D, Stefanski LA. (1995) *Measurement Error in Nonlinear Models.* Chapman and Hall, London, UK.

Collins PW, Blanchette VS, Fischer K, Björkman S, Oh M, Fritsch S, Schroth P, Spotts G, Astermark J, Ewenstein B; rAHF-PFM Study Group. (2009). Break-through bleeding in relation to predicted factor VIII levels in patients receiving prophylactic treatment for severe hemophilia A. *Journal of Thrombosis and Haemostasis* **7**: 413–420.

Davidian M, Giltinan DM. (1995) *Nonlinear models for repeated measurement data.* Chapman and Hall, London, UK.

Davies RB. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **74**: 33-43.

den Uijl IE, Fischer K, Van Der Bom JG, Grobbee DE, Rosendaal FR, Plug I. (2011). Analysis of low frequency bleeding data: the association of joint bleeds according to baseline FVIII activity levels. *Haemophilia* **17**: 41-44.

den Uijl IEM, Mauser Bunschoten EP, Roosendaal G, Schutgens REG, Biesma DH, Grobbee DE, Fischer. (2011). Clinical severity of haemophilia A: does the classification of the 1950s still stand? *Haemophilia* **17**: 849-853.

Fischer K, Collins P, Björkman S, Blanchette V, Oh M, Fritsch S, Schroth P, Spotts G, Ewenstein B. (2011). Trends in bleeding patterns during prophylaxis for severe haemophilia: observations from a series of prospective clinical trials. *Hemophilia*, **17**: 433–438

Gringeri A, Wolfsegger M, Steinitz KN, Reininger AJ. (2005) Recombinant full-length factor VIII (FVIII) and extended half-life FVIII products in prophylaxis – new insight provided by pharmacokinetic modelling. *Haemophilia* **21**: 300–306.

Hu P, Tsiatis AA, Davidian M. (1998) Estimating the Parameters in the Cox Model When Covariate Variables are Measured with Error. *Biometrics* **54**: 1407–1419.

Jiang W, Turnbull BW, Clark LC. (1999) Semiparametric regression models for repeated events with random effects and measurement error. *Journal of the American Statistical Association.* **94**: 111–124.

Karlsson MO, Sheiner LB. (1993). The importance of modeling interoccasion variability in population pharmacokinetic analyses. *Journal of Pharmacokinetics and Biopharmaceutics* **21**: 735–750.

Lausen B, Schumacher M, (1992) Maximally selected rank statistics. *Biometrics.* **48**: 73–85.

Lawless JF, Cook RJ. (2007) *The Statistical Analysis of Recurrent Events* Springer, New York.

Lawless JF, Nadeau RJ. (1995) Some simple robust methods for the analysis of recurrent events. *Technometrics.* **37**: 158–168.

Li Y, Lin X. (2000) Covariate measurement errors in frailty models for clustered survival data. *Biometrika.* **87**: 849–866.

Lindstrom MJ, Bates DM. (1990) Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics.* **46**: 673-687.

McEneny-King A, Iorio A, Foster G, Edginton AN. (2016) The use of pharmacokinetics in dose individualization of factor VIII in the treatment of hemophilia A. *Expert Opinion on Drug Metabolism & Toxicology.* **12**: 1313–1321.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. (2018). *nlme: Linear and Nonlinear Mixed Effects Models.* R package, version 3.1-137.

Overgaard RV, Ingwersen SH, Tornoe CW. (2015). Establishing good practices for exposure-response analysis of clinical endpoints in drug development. *CPT: Pharmacometrics and Systems Pharmacology*, **4**: 565–575.

Ramsay JO. (1988). Monotone Regression Splines in Action. *Statistical Science*, **4**: 425–441.

Shapiro D, Schoenig-Diesing C, Silvati-Fidell L, Wong WY, Romanov V. (2015). Integrated analysis of safety data from 12 clinical interventional studies of plasma- and albumin-free recombinant factor VIII (rAHF-PFM) in haemophilia A. *Hemophilia* **21**:791–798.

Therneau TM, Grambsch PM, Pankratz VS. (2003) Penalized survival models and frailty. *Journal of Computational and Graphical Statistics* **12**: 156–175.

Turnbull BW, Jiang W, Clark LC. (1997). Regression models for recurrent event data: Parametric random effects models with measurement error. *Statistics in Medicine*, **16**:853–864.

Verbeke G, Molenberghs G. (2013) The gradient function as an exploratory goodness-of-fit assessment of the random-effects distribution in mixed models. *Biostatistics* **14**: 477-490.

Wan F, Titman AC, Jaki TF. (2019). Subgroup analysis of treatment effects for misclassified biomarkers with time-to-event data. *Journal of the Royal Statistical Society, Series C,* **68**:1447–1463.

Wang J. (2019) *Exposure-Response Modeling. Methods and Practical Implementation*, Chapman and Hall.

Yi GY, Lawless JF. (2012) Likelihood-based and marginal inference methods for recurrent event data with covariate measurement error. *Canadian Journal of Statistics*, **40**:530–549.

Yu H, Cheng Y-J, Wang C-Y. (2016). Semiparametric Regression Estimation for Recurrent Event Data with Errors in Covariates under Informative Censoring. *International Journal of Biostatistics* **12** (2).

Zucker DM, Spiegelman D. (2008). Corrected score estimation in the proportional hazards model with misclassified discrete covariates. *Statistics in Medicine* **27**:1911–1933.