Modeling the rate of HIV testing from repeated binary data amidst potential never-testers

JOHN D. RICE*

*University of Rochester Medical Center, Biostatistics and Computational Biology,
265 Crittenden Blvd., Box 630, Rochester, NY 14642, USA
john.rice@ucdenver.edu*

BRENT A. JOHNSON, ROBERT L. STRAWDERMAN

University of Rochester, Biostatistics and Computational Biology, Rochester, NY, USA

SUMMARY

Many longitudinal studies with a binary outcome measure involve a fraction of subjects with a homogeneous response profile. In our motivating data set, a study on the rate of human immunodeficiency virus (HIV) self-testing in a population of men who have sex with men (MSM), a substantial proportion of the subjects did not self-test during the follow-up study. The observed data in this context consist of a binary sequence for each subject indicating whether or not that subject experienced any events between consecutive observation time points, so subjects who never self-tested were observed to have a response vector consisting entirely of zeros. Conventional longitudinal analysis is not equipped to handle questions regarding the rate of events (as opposed to the odds, as in the classical logistic regression model). With the exception of discrete mixture models, such methods are also not equipped to handle settings in which there may exist a group of subjects for whom no events will ever occur, i.e. a so-called “never-responder” group. In this article, we model the observed data assuming that events occur according to some unobserved continuous-time stochastic process. In particular, we consider the underlying subject-specific processes to be Poisson conditional on some unobserved frailty, leading to a natural focus on modeling event rates. Specifically, we propose to use the power variance function (PVF) family of frailty distributions, which contains both the gamma and inverse Gaussian distributions as special cases and allows for the existence of a class of subjects having zero frailty. We generalize a computational algorithm developed for a log-gamma random intercept model (Conaway, 1990. A random effects model for binary data. *Biometrics* **46**, 317–328) to compute the exact marginal likelihood, which is then maximized to obtain estimates of model parameters. We conduct simulation studies, exploring the performance of the proposed method in comparison with competitors. Applying the PVF as well as a Gaussian random intercept model and a corresponding discrete mixture model to our motivating data set, we conclude that the group assigned to receive follow-up messages via SMS was self-testing at a significantly lower rate than the control group, but that there is no evidence to support the existence of a group of never-testers.

Keywords: Binary longitudinal data; Frailty models; HIV testing; Power variance function; Recurrent events.

*To whom correspondence should be addressed.

1. INTRODUCTION

Traditionally, human immunodeficiency virus (HIV) testing has taken place in a clinic or community-based organization with a doctor, nurse, or trained professional. Non-standard HIV testing, including at-home self-testing, is becoming a popular tool among HIV epidemiologists and prevention experts as a way to increase the rate of HIV testing among at-risk individuals because the tests can be performed in the convenience of their home at a time of their choosing, avoiding any stigma associated with visiting an STD clinic.

The motivation for this work comes from the *Checking In!* study (Khosropour and others, 2013), a CDC-sponsored study designed to evaluate the use of text messaging to increase retention in a cohort study of a population of HIV-negative men who have sex with men (MSM). At-home HIV test kits were distributed for free throughout the follow-up period to participants, who were then asked every 2 months throughout the 12-month follow-up period whether or not they had tested for HIV since the last survey. The study design creates statistical challenges for modeling and estimating the rate of HIV testing from repeated at-home HIV self-tests. In particular, one must account for the coarseness of the observed HIV test data, which stems from not recording the exact test dates, and also for the possibility that some fraction of the population refuses or avoids HIV testing altogether.

Let $N_i(t)$ be a counting process for the cumulative number of HIV tests by time t for the i th subject ($i = 1, \dots, m$), where $N_i(0) = 0$. Our interest lies primarily in understanding the behavior of the intensity function $\lambda_i(t)$ that corresponds to $N_i(t)$. Epidemiologically, for example, there is interest in knowing whether $\lambda_i(t)$ is constant in time because it suggests the subject is testing routinely, at least on average. If the rate of HIV testing is consistently high or low, then this information is useful for making recommendations on HIV testing policy. Statistically, we are also interested in estimating and drawing inferences on the model parameters that govern $\lambda_i(t)$ as a function of subject-level risk factors. However, in the *Checking In!* study, the HIV testing process $N_i(t)$ is unobserved; instead, the observed data is a coarsening of the unobserved testing process that occurs in continuous time. By “coarsening,” we specifically mean two distinct operations: discretization and clipping (Foufoula-Georgiou and Lettenmaier, 1986). The former is due to the grouping of data within disjoint time intervals and the latter to grouping events within intervals into the categories *zero* or *at least one*. This process generates longitudinal binary response profiles for each subject, and the goal of this article is to develop methods that allow us to estimate the same intensity model parameters in the presence of such coarsened testing data.

The presence of longitudinal binary response profiles on independent study participants suggests the use of a mixed effects modeling framework. However, a second prominent feature of the *Checking In!* study data is the substantial proportion of participants who never tested; in particular, 32.9% of subjects did not self-test during follow-up, indicating the possible existence of a group of subjects who will never self-test regardless of interventions or covariates. Although low-probability events can be accommodated by random intercept models with a large negative intercept, such models cannot properly account for the fact that some subjects may never experience the event of interest.

These features, along with a desire to preserve the interpretation of the model describing $\lambda_i(t)$ insofar as possible, motivate us to develop a novel mixed effects modeling framework for the underlying rate function when the observed data involves coarsening of a recurrent events process in the manner described earlier and where there is the possibility that an unknown proportion of subjects remain event-free. In particular, conditionally on a subject-specific frailty (i.e., random effect), we assume $N_i(t)$ follows a non-homogeneous Poisson process (NHPP) with intensity function $\lambda_i(t)$. Using the fact that only the presence or absence of any events in each time interval per subject is recorded, we obtain the corresponding exact marginal likelihood using methods proposed by Conaway (1990) for longitudinal binary responses; this approach avoids the need for numerical integration and related approximation methods provided that the indicated frailty distribution has an explicit Laplace transform. Our particular focus

will be on the use of the power variance function (PVF) frailty distribution (Aalen and others, 2008, pp. 238–242), which has such a transform, contains the gamma and inverse Gaussian distributions as special cases and has the capability of incorporating a point mass at zero (i.e., zero frailty with positive probability). Despite its flexibility, this frailty distribution has seen limited use in applied survival analysis problems and, to our knowledge, has never been used in the context of mixed effects models for binary outcomes.

A common approach to analyzing longitudinal binary data is mixed effects logistic regression. In this model, the regression coefficients do not have the desired interpretation as log hazard ratios for self-testing events occurring over time. With mixed effects logistic regression, the use of any absolutely continuous random effects distribution for modeling within-subject correlation also does not allow for the existence of a group of non-responders. Indeed, as pointed out in Carlin and others (2001), the only way in which such a model can accommodate a large proportion of subjects who never have the event of interest is with a large negative intercept and high random effect variance. To overcome this limitation, they instead propose using a discrete mixture logistic model, where the mixture components correspond to whether a subject is “susceptible” or “immune” to the event under study and inference for the logistic model parameters is carried out in a Bayesian framework. In related work, Hall (2000) discusses an EM method for estimation in zero-inflated Poisson and binomial mixed models for longitudinal data, also assuming Gaussian random intercepts. There has been comparatively little investigation of more flexible random effects distributions for binary longitudinal data with a substantial proportion of non-responders.

The remainder of the article is organized as follows: in Section 2, we introduce the modeling framework, explain how we use the method of Conaway (1990) to compute exact marginal likelihoods under general frailty distributions having an explicit Laplace transform, and review the PVF frailty distribution. A detailed comparison between the proposed model and an appropriate version of the discrete mixture modeling approach proposed in Carlin and others (2001) is also provided. In Section 3, we apply the proposed approach to our motivating data set, contrast results across competing methods, and show that our method provides a better overall fit to the *Checking In!* data compared with a discrete mixture modeling approach as well as added precision for estimating the effects of interest. Section 4 presents simulation results demonstrating the improvements in estimation efficiency under the PVF model when there is a fraction of subjects who never experience the event of interest. We discuss our conclusions in Section 5 and provide some possible directions for future research.

2. METHODS

2.1. The continuous-time Poisson process model

For a given subject i , we initially conceive of the testing process $\{N_i(t), t \geq 0\}$ as a NHPP with intensity $\lambda_i^\circ(t) = \tilde{g}_0(t) \exp(\mathbf{z}_i' \boldsymbol{\beta})$, where $\mathbf{z}_i$ is a p -vector of subject-specific baseline risk factors, $\boldsymbol{\beta}$ is a corresponding p -vector of unknown regression coefficients, and $\tilde{g}_0(t) > 0$ for $t > 0$ but is otherwise unspecified. This semiparametric multiplicative intensity model is widely used in the analysis of recurrent events (Kalbfleisch and Prentice, 2002) and includes the homogeneous Poisson process (HPP) as a special case.

The independent increments property of the NHPP implies that the probabilistic behavior of the testing process is independent across any set of disjoint time intervals, between or within subjects. This is an untenable assumption in most applications and we would like to be able to account for the potential dependence of the event sequence within a subject. To be specific, multiplying the Poisson intensity $\lambda_i^\circ(t)$ by a non-negative subject-specific frailty V_i to account for this within-subject correlation, we obtain the intensity model

$$\lambda_i(t) = \tilde{g}_0(t) \exp(\mathbf{z}_i' \boldsymbol{\beta}) V_i; \quad (2.1)$$

that is, the underlying testing process is assumed to follow a NHPP conditionally on both V_i and $\mathbf{z}_i$. In applications, V_i is often taken to have a gamma distribution parameterized to have mean 1 and variance σ^2 to be estimated from the data (Kalbfleisch and Prentice, 2002, Chapter 10). However, this assumption is made primarily for mathematical convenience (Lawless, 1987) and is not required. Until Section 2.4, we assume only that V_i is a nonnegative random variable with distribution $F(\cdot; \sigma^2, \xi)$ where $\mathbb{E} V_i = 1$, $\text{Var} V_i = \sigma^2 \geq 0$ and ξ is a (possibly null) vector of parameters that controls other aspects of the frailty distribution.

2.2. The coarsened Poisson process model

Without loss of generality, we suppose that the process $N_i(t)$ is observed over the unit interval, which we partition into k disjoint intervals, $0 \equiv t_0 < t_1 < \dots < t_k \equiv 1$. Let the observed data be $Y_{ij} = \mathbb{1}(N_{ij} > 0)$, where the actual interval counts $N_{ij} = N_i(t_j) - N_i(t_{j-1})$, $j = 1, \dots, k$ are not observed. For convenience, we first reparameterize (2.1) such that $\lambda_i(t) = g_0(t) \exp(\beta_0 + \mathbf{z}_i' \boldsymbol{\beta}) V_i$, where $\beta_0 = \log \int_0^1 \tilde{g}_0(t) dt \in (-\infty, \infty)$ and $\int_0^1 g_0(t) dt = 1$. Then, it is easy to show that

$$\mathbb{E}(Y_{ij} | \mathbf{z}_i, V_i) = P(N_{ij} > 0 | \mathbf{z}_i, V_i) = 1 - e^{-G_j \exp(\beta_0 + \mathbf{z}_i' \boldsymbol{\beta}) V_i}. \quad (2.2)$$

where $G_j = \int_{t_{j-1}}^{t_j} g_0(t) dt$, $j = 1, \dots, k$. Under our assumptions on $g_0(\cdot)$, it follows that $G_j > 0$ for each j and $\sum_{j=1}^k G_j = 1$. Next, as in Brown (1975), we introduce a set of $k - 1$ dummy variables for the k observation intervals, with associated coefficients $\boldsymbol{\gamma} = (\gamma_2, \dots, \gamma_k)$. It may be shown that this implies that $e^{\gamma_j} = G_j / G_1$, $j = 2, \dots, k$. Defining

$$\eta_{ij} = \beta_0^* + \gamma_j + \mathbf{z}_i' \boldsymbol{\beta}, \quad j = 1, \dots, k, \quad (2.3)$$

where $\beta_0^* = \beta_0 + \log G_1$ and $\gamma_1 \equiv 0$, it follows that $\mathbb{E}(Y_{ij} | \mathbf{z}_i, V_i) = 1 - \exp\{-\exp(\eta_{ij} + \log V_i)\}$. The corresponding marginal expectation is

$$\mathbb{E}(Y_{ij} | \mathbf{z}_i) = \int_0^\infty (1 - e^{-e^{\eta_{ij}} v_i}) dF(v_i; \sigma^2, \xi) = 1 - \mathcal{L}(e^{\eta_{ij}}; \sigma^2, \xi), \quad (2.4)$$

where $\mathcal{L}(\cdot; \sigma^2, \xi)$ is the Laplace transform of the frailty distribution. It can further be shown that the marginal covariance between Y_{ij} and $Y_{ij'}$ is

$$\text{Cov}(Y_{ij}, Y_{ij'} | \mathbf{z}_i) = \mathcal{L}(e^{\eta_{ij}} + e^{\eta_{ij'}}; \sigma^2, \xi) - \mathcal{L}(e^{\eta_{ij}}; \sigma^2, \xi) \mathcal{L}(e^{\eta_{ij'}}; \sigma^2, \xi), \quad (2.5)$$

a quantity that is in general not equal to zero. The correlation structure induced by the model is exchangeable when G_j (and therefore η_{ij}) is constant with respect to j (i.e., for a HPP); otherwise, the dependence structure is more complex.

Model (2.2) may be viewed as a generalization of a model for grouped survival data (Prentice and Gloeckler, 1978) with an added frailty for correlated time-to-event data; see Lawless and Zhan (1998) for related developments in the case where the N_{ij} s are all observed. Model (2.2) under parameterization (2.3) can also be put in notational correspondence with a generalized linear mixed model with complementary log–log link and random intercept equal to $\log V_i$ (Diggle and others, 2002, Chapter 9). When there is no frailty variable, the rate ratios in the continuous-time model can be estimated as the (exponentiated) regression coefficients in a binary regression model with complementary log–log link. However, there is also an important distinction between the indicated frailty formulation and the more classical formulation

of mixed models for longitudinal binary data (e.g., [Fitzmaurice and others, 2004](#), Chapter 12), where the conditional mean of the response given covariates is modeled on the scale of the link function. In particular, a typical random-intercept model is formulated as $h\{\mathbb{E}(Y_{ij}|U_i)\} = \eta_{ij} + U_i$, where $h(\cdot)$ is a known link function and U_i is a random effect with mean zero and finite variance (e.g., Gaussian). The use of model (2.2), which is focused on modeling the conditional rate of the recurrent event given covariates, evidently corresponds to choosing $h(x) = \log\{-\log(1-x)\}$ and replaces U_i with $\log V_i$. However, the constraint $\mathbb{E} V_i = 1$ and Jensen's inequality imply that $\log \mathbb{E} V_i = 0 > \mathbb{E}(\log V_i)$ whenever V_i has positive variance, showing that the corresponding linear predictor is not quite the same as the conditional mean response on the scale of the link function (i.e., the intercept is different).

These differences can be highlighted somewhat more explicitly when V_i is assumed to be lognormal with parameters m and s^2 . In order for $\mathbb{E} V_i = 1$, we must set $m = -s^2/2$, leading to $\text{Var } V_i = e^{s^2} - 1$; this corresponds to $\log V_i$ being normally distributed with mean $m \neq 0$ and variance s^2 , showing that the marginal mean function depends on the frailty variance (on the link scale). It is also not possible to directly compare the frailty variance with that from the mixed model approach since these are on different scales.

2.3. Maximum likelihood estimation for general frailties

Consider model (2.2) under parameterization (2.3). Assuming that all subjects are independent, the relevant marginal log-likelihood is

$$\ell(\beta_0^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \xi) = \sum_{i=1}^m \log \left\{ \int_0^\infty \prod_{j=1}^k P(Y_{ij}|\mathbf{z}_i, V_i = v_i) dF(v_i; \sigma^2, \xi) \right\}, \quad (2.6)$$

where $P(Y_{ij}|\mathbf{z}_i, V_i = v_i) = (1 - e^{-\eta_{ij} v_i})^{Y_{ij}} e^{-\eta_{ij} v_i (1 - Y_{ij})}$. Maximum likelihood estimation is straightforward in the absence of frailty (i.e. $V_i = 1$ with probability one). In the presence of frailty, the marginal probabilities $\pi_{i, \mathbf{y}_i} = \int_0^\infty \prod_{j=1}^k P(Y_{ij}|\mathbf{z}_i, V_i = v_i) dF(v_i; \sigma^2, \xi)$ are not generally available in closed form. However, note that

$$\pi_{i, T}^* = \int_0^\infty e^{-\sum_{j \in T} \eta_{ij} v_i} dF(v_i; \sigma^2, \xi) = \mathcal{L} \left(\sum_{j \in T} e^{\eta_{ij}}; \sigma^2, \xi \right),$$

where T is a subset of the indices $\{1, 2, \dots, k\}$, are easily computed when $\mathcal{L}(\cdot; \sigma^2, \xi)$ has a closed form. It follows that one may use the methods of [Conaway \(1990\)](#) to calculate the exact value of $\pi_{i, \mathbf{y}_i}$ and hence (2.6). Essentially, this involves specific linear combinations of the elements of $\boldsymbol{\pi}_i^*$, where $\boldsymbol{\pi}_i^*$ is the vector containing $\pi_{i, T}^*$ for T in all 2^k possible subsets of the index set $\{1, 2, \dots, k\}$ (see [Conaway, 1990](#), p. 320, for details). Parameter estimates are then obtained by maximizing (2.6). In practice, the effective use of this methodology requires that $\mathcal{L}(\cdot; \sigma^2, \xi)$ be easily computed, and preferably available in closed form. We consider estimation for a specific choice of frailty distribution in the next subsection.

2.4. The PVF frailty distribution: modeling and estimation

An interesting and useful choice for the distribution of V_i in models (2.1) and (2.2), hence in (2.4), is the PVF distribution. The PVF distribution has the closed-form Laplace transform

$$\mathcal{L}(x; \sigma^2, \xi) = \exp \left[-\frac{\xi + 1}{\xi \sigma^2} \left\{ 1 - \left(1 + \frac{\sigma^2 x}{\xi + 1} \right)^{-\xi} \right\} \right]. \quad (2.7)$$

As parameterized, $\mathbb{E} V_i = 1$, $\text{Var} V_i = \sigma^2$, and $\xi > -1$ controls aspects of the shape of the frailty distribution. This flexible family of distributions includes the gamma ($\xi = 0$) and inverse Gaussian distributions ($\xi = -1/2$) as special cases.

In addition to its flexibility, a particularly relevant feature of the family of PVF distributions in the context of the application that motivated this work is its ability to mimic a cure model (Aalen and others, 2008, p. 239). Price and Manatunga (2001) considered the merits of both standard frailty models and frailty mixture models (i.e., where there is a point mass at zero) for modeling correlated survival data with a cured fraction. The PVF frailty model unifies the treatment of several frailty and frailty mixture models in a natural way; in particular, when $\xi > 0$, the PVF frailty distribution reduces to a compound Poisson distribution (Wienke and others, 2009) that has a point mass of $e^{-(\xi+1)/(\xi\sigma^2)}$ at $v = 0$. The presence of a point mass at $v = 0$ implies that some subjects will never experience the event of interest; in the context of our motivating example, this would correspond to a subgroup of subjects that never self-test, regardless of intervention or any other covariates. Importantly, the decision to include a point mass at $v = 0$ need not be made *a priori*; the existence of such a point mass is instead captured by the estimable parameter ξ .

Using the methods of Conaway (1990) and the Laplace transform (2.7) to compute the terms that make up the marginal log-likelihood (2.6), we are able to obtain maximum likelihood estimates of all model parameters without having to rely on an approximation of the marginal likelihood. For convenience, we reparameterize the PVF model such that σ^2 is modeled on the log scale and $\log(\xi + 1)$ is modeled instead of ξ ; the latter is used because of the boundary condition $\xi > -1$. The BLUPs for the fitted models may also be computed if desired; see Appendix C of the [supplementary material](#) available at *Biostatistics* online for details.

2.5. Discrete mixture modeling: an alternative approach

Collectively, Sections 2.1–2.4 develop an approach to maximum likelihood estimation of the parameters of the rate function (2.1) in the case where observation of the recurrent event process is subject to extreme coarsening and V_i follows the PVF frailty distribution. This model is interesting in the context of the *Checking In!* study because it allows for, but does not *a priori* require, the existence of a never-tester group.

This approach may be contrasted with the discrete mixture model as formulated by Carlin and others (2001), who present a Bayesian approach to estimation in a model involving a latent binary group indicator S_i where (i) the probability of testing among those considered “susceptible” (i.e., $S_i = 1$, or not being in the never-tester group) follows a mixed effects logistic model with a Gaussian random intercept and (ii) the probability of testing among those that are “immune” (i.e., $S_i = 0$, or being in the never-tester group) is zero. Although developed in a different context, substituting the complementary log–log for the logistic link function results in a model that can be written in the form of (2.2), but where the frailty V_i has a distribution that is described by a mixture of a lognormal density and a point mass at zero; see Appendix A of the [supplementary material](#) available at *Biostatistics* online for details. This approach separates the model for the probability of susceptibility from the unobserved heterogeneity model used for the population of susceptibles. Such separation may not be desirable in the context of frailty modeling for time-to-event processes, where it has been suggested that the probability of susceptibility should be related to the frailty distribution (Aalen and Hjort, 2002). The example given in Aalen and others (2008, p. 245) pertains to the contrast between rare diseases, which have a high-risk for few individuals, and common diseases, which exhibit less variation in risk between individuals. In the context of HIV testing, consider a vulnerable population like MSM: high-risk behaviors may be correlated with demographics that do not perceive themselves as high risk, or perhaps do not fully comprehend the value of HIV testing; the probability that these individuals will be never-testers is high. On the other hand, individuals who are not never-testers may accurately perceive their elevated risk status and therefore test at a higher rate. Such

cases support the use of a frailty model that links the probability of susceptibility with overall heterogeneity among subjects. The PVF model formulation retains such a link.

There are other notable disadvantages associated with using the Gaussian discrete mixture model within the context of this article. First, this model requires the inclusion of a never-tester group unless the intercept parameter in the susceptibility model lies on the boundary of the parameter space (i.e., at $-\infty$), whereas the PVF frailty model allows the data to determine whether its existence is likely. Second, this model does not allow the shape of the random effects distribution to vary, while the parameter ξ in the PVF frailty model allows for additional flexibility, whether or not a point mass is present (see Appendix B of the [supplementary material](#) available at *Biostatistics* online for more details and examples of PVF density functions). Finally, estimation using a discrete mixture model generally requires the use of numerical integration to approximate the relevant marginal likelihood and its derivatives, while use of the method of [Conaway \(1990\)](#) allows us to calculate the marginal likelihood for the PVF model exactly.

Although not the focus of this article, it is possible to allow ξ in the PVF frailty model to depend on covariates, as can the probability of susceptibility in [Carlin and others \(2001\)](#). If one estimates the PVF model under the constraint that $\xi > 0$, then the model can be reparameterized in a way that corresponds very closely with the discrete mixture model. In particular, we can reparameterize the PVF model such that $(1 + e^\varphi)^{-1} = e^{-(\xi+1)/(\xi\sigma^2)}$, where φ has the same interpretation as it does in the discrete mixture model.

3. EFFECT OF SMS REMINDERS ON HIV TESTING BEHAVIOR

3.1. Description of the data

We applied the proposed methods to data from the *Checking In!* study ([Khosropour and others, 2013](#)). This study randomized eligible MSM to either receive text message (SMS) or online follow-up, and collected data at 2-month intervals for a period of 12 months. Variables on which data was collected included baseline demographics, such as age, education level, and race; as well as whether or not a subject had self-tested for HIV during each of the 2-month intervals for which that subject was followed. In our notation, Y_{ij} is the variable indicating that the i th subject has self-tested at least once for HIV during the j th interval, $j = 1, \dots, k = 6$.

A sample of size $m_1 = 329$ subjects was available for the online group, while the SMS group comprised $m_2 = 317$ subjects. We note one difference in testing patterns between the two treatment groups in particular: the online group had a large spike in tests in the first interval, which thereafter declined to roughly the same level as for the SMS group. The SMS group, by contrast, appears to have tested less in the first interval than in subsequent intervals (see Table D.1 in Appendix D of the [supplementary material](#) available at *Biostatistics* online).

We are interested in several scientific questions regarding this data. First, we would like to assess the effect of treatment (online-only versus SMS follow-up messages) on patterns of self-testing among the study subjects. Second, we would like to determine which, if any, baseline covariates might affect the rate of self-testing. Third, we would like to determine whether there might be a group of never-testers, and if so, to estimate the probability that any given subject belongs to it.

3.2. Modeling the rate of self-testing

In addition to the indicator for treatment group and the dummy variables for time, three baseline covariates were included in the model: age in years (centered at the overall mean and divided by 10), race (white, black, or Hispanic), and education (at least a college degree, some college, high school or GED, and less than a high school degree). The parameter estimates obtained using the proposed PVF model and a Gaussian random intercept model with and without a discrete mixture component are presented in

Table 1. Coefficient estimates and standard errors from the PVF frailty model and two Gaussian random intercept models (with and without a discrete mixture component). The intercept is excluded from this table because it is not meaningful to compare this between the Gaussian intercept models and the PVF models as their random effects are formulated on different scales. Reference categories are shown in parentheses. Asymptotic standard errors are given in the SE column, while bootstrap standard errors (based on 1000 replicates) are given in the BSE column.

Covariate	Value	Gaussian intercept			Gaussian intercept (discrete mixture)			PVF frailty		
		Est.	SE	BSE	Est.	SE	BSE	Est.	SE	BSE
Treatment	(Online)									
	SMS	-2.78	0.21	0.26	-2.74	0.14	0.30	-2.74	0.21	0.23
Time	(int. 1)									
	int. 2	-1.93	0.16	0.21	-1.91	0.15	0.20	-1.87	0.15	0.17
	int. 3	-2.25	0.17	0.23	-2.23	0.16	0.20	-2.18	0.17	0.18
	int. 4	-2.27	0.17	0.24	-2.25	0.16	0.21	-2.21	0.17	0.19
	int. 5	-2.17	0.17	0.24	-2.16	0.16	0.21	-2.13	0.16	0.19
	int. 6	-2.24	0.17	0.23	-2.22	0.16	0.20	-2.17	0.17	0.19
Time by treatment	(SMS * int. 1)									
	SMS * int. 2	2.66	0.27	0.31	2.63	0.21	0.30	2.60	0.27	0.27
	SMS * int. 3	2.54	0.29	0.37	2.49	0.23	0.40	2.46	0.29	0.32
	SMS * int. 4	2.92	0.28	0.34	2.87	0.22	0.30	2.85	0.28	0.30
	SMS * int. 5	2.63	0.28	0.36	2.60	0.22	0.34	2.59	0.28	0.31
	SMS * int. 6	2.62	0.29	0.36	2.58	0.22	0.42	2.56	0.28	0.31
Age		0.11	0.05	0.05	0.10	0.05	0.06	0.08	0.05	0.05
Race	(White)									
	Black	0.37	0.15	0.14	0.32	0.15	0.16	0.33	0.14	0.14
	Hispanic	0.11	0.14	0.12	0.07	0.14	0.15	0.05	0.13	0.14
Education	(≥ college)									
	Some college	-0.15	0.12	0.10	-0.16	0.12	0.13	-0.17	0.11	0.11
	High school	-0.17	0.17	0.15	-0.19	0.17	0.21	-0.17	0.16	0.15
	< high school	-0.62	0.37	0.30	-0.74	0.38	0.31	-0.65	0.36	0.33

Table 1. The Gaussian discrete mixture model with a complementary log-log link is a likelihood-based implementation of the model in [Carlin and others \(2001\)](#), and involves maximizing the likelihood

$$\prod_{i=1}^m \left\{ \frac{1}{1 + e^\varphi} \prod_{j=1}^k (1 - Y_{ij}) + \frac{e^\varphi}{1 + e^\varphi} \int_0^\infty \prod_{j=1}^k P(Y_{ij} | \mathbf{z}_i, W_i = w) \phi \left(\frac{\log w}{\tau} \right) \frac{dw}{w\tau} \right\}, \quad (3.8)$$

where $\phi(\cdot)$ is the standard normal density function; see Appendix A of the [supplementary material](#) available at *Biostatistics* online for further details. With an intercept-only model for probability of being susceptible, having $\varphi < \infty$ or $\xi > 0$ both give rise to a positive proportion of subjects who are not susceptible to the event.

The PVF model was fit by maximizing the likelihood of Section 2.3 using the R function `optim()` ([R Core Team, 2015](#)) with the quasi-Newton optimization method BFGS. The discrete mixture model was also fit to the data using `optim()`, with direct numerical integration being used to compute the marginal likelihood. The corresponding Gaussian intercept model, which excludes the point mass at zero, was fit using the `lme4` package ([Bates and others, 2016](#)). Asymptotic standard errors were

obtained by inverting the information matrix, which we estimated using numerical differentiation. We also used bootstrap resampling as an alternative method for obtaining standard error estimates for model parameters (Efron and Tibshirani, 1993). We generated 1000 bootstrap samples by resampling individual subjects.

The estimated regression coefficients for these three models are summarized in Table 1, along with their corresponding asymptotic and bootstrap standard errors. The sign and magnitude for each coefficient is similar across all three models. Each model shows that there is an effect of treatment on the rate of testing, with the intervention (SMS) group self-testing at a significantly lower rate than the control (online only) group. In addition, black race is seen to be associated with significantly higher testing rates.

These conclusions hold regardless of which standard error estimate is used. The asymptotic standard errors of the PVF model are somewhat smaller than for the Gaussian intercept without discrete mixture component. The Gaussian discrete mixture model also exhibits smaller asymptotic standard errors than the PVF for the main effects of treatment and time as well as the time-by-treatment interactions; the standard errors for the PVF model are slightly smaller than the Gaussian discrete mixture model for other covariates.

However, there are some substantial discrepancies between the asymptotic and bootstrap standard errors for both the Gaussian intercept and discrete mixture models, whereas differences are comparatively minor in the case of the PVF frailty model. For example, the treatment effect is estimated to have an asymptotic standard error of 0.14 under the discrete mixture model, whereas the estimated bootstrap standard error is 0.30. Good agreement between asymptotic and bootstrap standard errors is expected in cases where the parametric model is a good-fit, with greater divergence expected as model misspecification increases (Efron and Tibshirani, 1993, Section 21.8). These results point to the PVF model providing a better fit to these data, which we investigate further in Section 3.5. In most cases, the bootstrap standard errors under the PVF model are also smaller than those for the discrete mixture and Gaussian intercept models, suggesting that it provides improved precision for estimating the effects of interest, including the treatment effect.

3.3. Assessing treatment effect on patterns of self-testing

Our modeling approach is based on the assumption that, conditional on a frailty variable, the events occur according to a Poisson process. We would like to address the question of whether these groups differ from one another in their patterns of self-testing. To do this, we test the interaction between time and treatment.

The results of the likelihood ratio chi-square tests for similarity of self-testing patterns between the two treatment groups revealed significant differences associated with treatment. The models are each adjusted for the baseline covariates race, education, and age and the results show highly significant evidence ($P < 0.0001$ regardless of assumed frailty model) against the null hypothesis of identical testing patterns in the two groups. These results appear to be due primarily to the testing behavior of subjects in the first interval. As noted above, the Online subjects tested at a much greater rate in this interval than in subsequent intervals, while those in the SMS group tested at a somewhat lower rate in the first interval compared with later intervals (see Figure D.2 in Appendix D of the [supplementary material](#) available at *Biostatistics* online).

3.4. Existence of a group of never-testers

Because so many subjects in the data set did not self-test during the study period, we wanted to assess the possible existence of a group of “never-testers.” This would comprise subjects who would not self-test regardless of intervention or other covariates. To answer this question in our modeling framework

Table 2. *Estimates of ξ for each fitted model. This table shows the MLE and associated 95% confidence intervals, as well as the median of the bootstrap distribution and its associated 0.025 and 0.975 quantiles. LCL = Lower confidence limit; UCL = Upper confidence limit.*

	Likelihood			Bootstrap		
	$\hat{\xi}$	LCL	UCL	Median	LCL	UCL
Constant hazard	−0.882	−0.934	−0.787	−0.882	−0.929	−0.795
Same hazard for both groups	−0.887	−0.939	−0.788	−0.883	−0.927	−0.770
Unrestricted hazard	−0.825	−0.899	−0.697	−0.827	−0.901	−0.679

requires us to interpret estimates of the parameter ξ , which governs aspects of the PVF distribution beyond variance. Care should be taken, however, when interpreting these results, because the relatively short time horizon of the study does not allow us to conclude definitively that someone identified by the model as a never-tester truly belongs to this group (see Section 5 for further discussion of this point).

Table 2 shows both the MLE of ξ under several models, along with associated 95% confidence intervals based on asymptotic theory (calculated using the Hessian matrix at the solution); as well as the median of the bootstrap estimates and 95% bootstrap confidence intervals (based on the 0.025 and 0.975 quantiles of the bootstrap distribution) under each model. With the exception of the model including only the SMS treatment group, we see remarkably close agreement between the asymptotic and bootstrap estimates and confidence intervals. This suggests that in at least some cases there may be utility in using standard likelihood theory to conduct inference on variance parameters, at least when they are parameterized on a log scale.

The estimates of ξ are quite consistent between these models, suggesting that ξ lies in the interval $(-0.9, -0.8)$. Hence, there is no evidence of the existence of a group of so-called “never-testers,” a finding that is also in agreement with the results of the discrete mixture model with Gaussian intercept (3.8): with $\hat{\phi} \approx 14$, this model estimated the probability of being a never tester to be essentially zero. This conclusion has important implications for policy and the planning of interventions to improve rates of HIV testing in at-risk populations. Had we found that ξ was significantly greater than zero for this sample, it would suggest that no intervention would be effective for some proportion of the population. However, our finding that ξ is significantly less than zero suggests that there is not a fraction of the population recalcitrant to efforts to increase HIV self-testing rates.

3.5. Goodness-of-fit and model comparison

To evaluate goodness-of-fit and compare alternative distributions for the random effect, we examined the gradient function diagnostic plots proposed by Verbeke and Molenberghs (2013). These plots for the full model including both treatment groups (i.e., assuming common frailty parameters but allowing all other model parameters to differ between the online and SMS groups) are shown in Figure 1. The gradient function for a given frailty distribution F is

$$\Delta(F, v) = \frac{1}{m} \sum_{i=1}^m \frac{\prod_j P(Y_{ij}; \eta_{ij}|v)}{\int \prod_j P(Y_{ij}; \eta_{ij}|s) dF(s)}$$

and measures the improvement in likelihood that would be obtained by a frailty distribution with additional mass at the point v (relative to the mass placed there by F). A good fit is implied by a gradient function that is close to but does not exceed 1 within its support, where the support bounds are given by

Fig. 1. Gradient functions (Verbeke and Molenberghs, 2013) for the HIV self-testing study. To create these plots, the PVF frailty model was fit to the full data set (i.e. including both the online and SMS groups) under the alternative hypothesis that hazards in the two groups may differ. A gradient function that is close to 1 within the support bounds is indicative of good fit for that frailty distribution, while a bad fit is reflected by a gradient function exceeding 1. The support bounds, or the region about which the data provide information on the frailty distribution, are given by equation (3.9).

Verbeke and Molenberghs (2013) as

$$\min_{i:0 < \sum_j Y_{ij} < k} \left\{ \sup_v \prod_j P(Y_{ij}; \eta_{ij} | v) \right\} \leq v \leq \max_{i:0 < \sum_j Y_{ij} < k} \left\{ \sup_v \prod_j P(Y_{ij}; \eta_{ij} | v) \right\}. \quad (3.9)$$

Figure 1 shows clearly that some non-degenerate frailty is necessary, since the gradient function for the model with no frailty substantially exceeds 1. The lognormal and discrete mixture frailties are almost indistinguishable from one another, which is to be expected since the discrete mixture component was estimated to be zero in the latter. The PVF model shows substantial improvement over both of these competing choices, with gradient function almost exactly 1 up to approximately $v = 2$, and then much closer to 1 for larger v compared with the competing models.

Since the graphical diagnostic methods of Verbeke and Molenberghs (2013) do not account for differences in dimension between the models under consideration, we also examined AIC values (see Table D.2 in Appendix D of the [supplementary material](#) available at *Biostatistics* online). AIC is also useful in

this setting since the Gaussian intercept models are not nested within either the gamma or PVF models. In particular, the lower AIC values observed for the PVF frailty model as fitted to the full data set are consistent with the implications of Figure 1. A comparison on the basis of AIC between the PVF model and the discrete mixture model with Gaussian random intercept reveals substantially better fit for the PVF model (3409.9 for PVF versus 3421.3 for discrete mixture model). This provides evidence of the increased flexibility offered by the PVF model as compared with a discrete mixture model.

4. SIMULATION STUDIES

In order to evaluate the performance of the proposed modeling approach based on the PVF frailty, we compared it in simulation studies with three competing methods: a naive analysis ignoring the correlation within subjects, a gamma frailty model, and a Gaussian intercept model with a complementary log–log link. We remind the reader that the Gaussian intercept model is similar, but not identical, to a lognormal frailty model due to the latter having a different mean and variance on the scale of the frailty. While we do model the pattern of events over time (in the form of the γ_j 's), McCulloch and Neuhaus (2011) have shown that the estimates of these “within-subject” parameters are apt to be quite robust with respect to misspecification of the random effects distribution. This was investigated and confirmed, and hence we do not report these results. We also do not include results from marginal model fits (e.g., generalized estimating equations) in our simulations: due to our use of a nonlinear link function (the complementary log–log), estimates derived from marginal models are not directly comparable with those obtained using the mixed models that are the focus of this article (Fitzmaurice and others, 2004, Section 12.3).

The data was generated according to an NHPP, with baseline hazard function given by

$$g_0(t) = e^1 \frac{\Gamma(2.5)}{\Gamma(1)\Gamma(1.5)} t^{1-1} (1-t)^{1.5-1} \approx 4.077\sqrt{1-t}, \quad 0 \leq t \leq 1.$$

This is a decreasing, concave function based on the beta density, and was chosen to reflect an event process where subjects are likely to experience an event early on in the study, but increasingly less likely later in the study period. Two covariates were included in the true model, both with coefficients equal to 1. One of these was standard normal, the other Bernoulli with mean 0.5 (intended to reflect a randomized treatment variable, for example). The sample size for all simulations was $m = 500$.

We are interested in examining the potential ramifications of fitting a misspecified model when the true model contains a random intercept following the PVF distribution. We therefore varied ξ and σ^2 in order to evaluate a range of possible shapes for the true frailty distribution, including the inverse Gaussian ($\xi = -1/2$), gamma ($\xi = 0$), and compound Poisson ($\xi > 0$) models as special cases.

Table 3 gives the bias and variance of the regression coefficient estimates for a selection of true values of ξ and σ^2 . It is clear that for no frailty ($\sigma^2 = 0$), all methods perform similarly, with slight underestimation of the true variance of the parameters for the PVF method. This is likely due to assuming the presence of, and estimating, an additional parameter that has no meaning when σ^2 is truly zero. As σ^2 increases, the advantages of the PVF model become more apparent. Although the bias of the estimates under the Gaussian intercept model is not substantial except for large values of both ξ and σ^2 , its precision is uniformly lower than for the PVF model. This may help explain the lower MSE for the PVF method seen in Figure 2. Additionally, as σ^2 increases, the accuracy of the asymptotic standard error estimates decreases for the gamma frailty and Gaussian intercept models. In the case of the gamma frailty model, this tends to result in overestimation, whereas variability under the Gaussian intercept model tends to be slightly underestimated.

As noted above, one of the features of the proposed random effects distribution is its ability to model a positive probability that subjects will never experience the event of interest. Figure 2 shows the mean-

Table 3. *Simulation results: bias and variance of parameter estimates for fixed covariates. The label "Gauss. Int." denotes the Gaussian intercept model.*

σ^2	ξ	Method	Standard normal			Binary (mean = 0.5)		
			Bias	ESD	ASE	Bias	ESD	ASE
0	0.0	Gamma	0.009	0.040	0.041	0.008	0.063	0.064
		Gauss. Int.	0.009	0.040	0.041	0.008	0.063	0.064
		None	0.004	0.040	0.040	0.002	0.062	0.062
	1.0	PVF	0.009	0.040	0.037	0.008	0.063	0.061
		Gamma	0.010	0.040	0.041	0.010	0.065	0.064
		Gauss. Int.	0.010	0.040	0.041	0.010	0.065	0.064
	4.0	None	0.005	0.039	0.040	0.005	0.064	0.062
		PVF	0.011	0.040	0.036	0.011	0.065	0.061
		Gamma	0.009	0.040	0.041	0.013	0.065	0.064
		Gauss. Int.	0.008	0.040	0.041	0.013	0.065	0.064
		None	0.003	0.040	0.040	0.008	0.065	0.062
		PVF	0.009	0.041	0.036	0.013	0.065	0.061
2	0.0	Gamma	0.003	0.095	0.094	−0.004	0.165	0.167
		Gauss. Int.	0.021	0.103	0.103	0.019	0.185	0.187
		None	−0.416	0.059	0.033	−0.408	0.110	0.064
	1.0	PVF	0.002	0.097	0.096	−0.004	0.167	0.168
		Gamma	0.170	0.131	0.132	0.173	0.216	0.235
		Gauss. Int.	0.049	0.163	0.131	0.055	0.284	0.245
	4.0	None	−0.528	0.065	0.033	−0.513	0.130	0.066
		PVF	0.014	0.111	0.103	0.012	0.173	0.170
		Gamma	0.386	0.161	0.178	0.404	0.276	0.317
		Gauss. Int.	−0.005	0.222	0.193	0.002	0.402	0.370
		None	−0.613	0.069	0.034	−0.600	0.141	0.068
		PVF	0.014	0.100	0.097	0.017	0.151	0.146
4	0.0	Gamma	0.012	0.133	0.130	0.011	0.236	0.235
		Gauss. Int.	0.022	0.148	0.134	0.024	0.275	0.253
		None	−0.510	0.067	0.036	−0.500	0.132	0.070
	1.0	PVF	0.010	0.136	0.133	0.011	0.239	0.237
		Gamma	0.359	0.213	0.225	0.357	0.351	0.409
		Gauss. Int.	−0.073	0.251	0.231	−0.092	0.482	0.455
	4.0	None	−0.646	0.079	0.038	−0.646	0.160	0.076
		PVF	0.031	0.151	0.149	0.032	0.237	0.239
		Gamma	0.841	0.339	0.355	0.916	0.572	0.664
		Gauss. Int.	−0.470	0.389	0.296	−0.416	0.679	0.583
		None	−0.743	0.085	0.041	−0.730	0.182	0.082
		PVF	0.019	0.142	0.140	0.027	0.212	0.202

square error of the parameter estimates for each of the assumed frailty distributions as a function of the true proportion of never-responders. For a low proportion of never-responders, the model with no frailty performs worst, with minimal differences in MSE between the three random-effects models. By contrast, for a high proportion of never-responders, the model assuming a gamma-distributed frailty

Fig. 2. Simulation results, MSE of fixed covariate effects. This figure displays the mean-square error of the parameter estimates for each of the two covariates in the model and each of the four estimation methods we consider.

variable performs worst, the model assuming no frailty performs similarly to the Gaussian random intercept model and the PVF model performs the best.

It is also of interest to examine how the model performs with respect to estimation of ξ . Boxplots of the MLE of ξ for each of the simulations appear in Figure E.4 of Appendix E of the [supplementary material](#) available at *Biostatistics* online. These boxplots suggest that the sampling distribution of $\hat{\xi}$ is quite skewed. However, the median of the distribution tends to be close to the true value of ξ , especially for smaller values of σ^2 . In particular, it is encouraging that there does not appear to be a wide spread in the distribution of $\hat{\xi}$ under the condition $\sigma^2 = 0$. This suggests good behavior of the estimators despite a non-standard parameter space.

5. DISCUSSION

We have proposed a model for recurrent event data subject to extreme coarsening based on the class of NHPPs assuming that only the presence or absence of at least one event per equally spaced observation interval is available. Consideration of the PVF frailty distribution on the unobserved continuous time scale ([Aalen and others, 2008](#)), which encompasses the gamma and inverse Gaussian distributions as special cases, leads to a new and flexible class of models for analyzing repeated binary outcomes. In combination with methods introduced in [Conaway \(1990\)](#), which can be used with any frailty distribution having an easily computed Laplace transform, the proposed approach eliminates the need for numerical integration, Monte Carlo simulation or approximations in obtaining maximum likelihood estimates. The flexibility of the PVF frailty distribution, combined with its allowance for the presence of subgroups that never experience the event of interest, presents an interesting and useful alternative to the Gaussian random effects model commonly used with repeated measures data.

Our data analysis shows that the increased flexibility provided by the PVF frailty model results in a better fit when compared with competitors. In more detail, both the PVF model and the Gaussian intercept with discrete mixture component have two parameters describing their respective frailty distribution: a variance parameter and a parameter that controls the probability of being a non-responder. In the case of the PVF model, this latter parameter also controls other aspects of the distribution; see Figure B.1 in Appendix B of the [supplementary material](#) available at *Biostatistics* online for examples of the variety of shapes its density can take. The parameter φ in the discrete mixture model, by contrast, is limited

to determining the fraction of nonresponders in the population; the random effects distribution for the responders will have the same shape regardless of φ . This limits the ability of discrete mixture models to describe data with random effects distributions that deviate from normality.

Like other methods for grouped forms of survival and recurrent event data, our approach provides consistent estimates of the relative risk parameters for the underlying proportional hazards intensity model (e.g., Prentice and Gloeckler, 1978). Despite the restricted nature of the observations available, we are still able to construct model-based estimates of the mean number of events in each of the study intervals for each subject. It is anticipated that the PVF frailty model can be extended in a useful way to grouped recurrent event data with exact event counts observed in each interval.

A referee has noted the practical difficulty of differentiating between zero- and low-propensity testers in finite samples. It is acknowledged that this is a technical difficulty in settings involving finite samples as well as in those having finite follow-up. From an equally practical point of view, making this distinction may not be so important compared to separating the low-or-never group from those inclined to test. In this regard, models that permit the possibility of a subject “never” testing continue to be advantageous compared to those that do not, even in settings when zero- and low-propensity testers are not well-differentiated. In particular, in a random intercept model, the presence of a substantial proportion of low-propensity testers would shrink the overall propensity for all subjects towards zero; by contrast, the PVF frailty model may assign these subjects to a “never tester” group without shrinking the overall propensities for other subjects.

SUPPLEMENTARY MATERIAL

Supplementary material is available at <http://biostatistics.oxfordjournals.org>. A GitHub repository (<https://github.com/jdrice8/pvf-frailty-binary-regression>) contains the code used for the analyses presented here, along with a simulated data set that mimics characteristics of the real data analyzed in the paper.

ACKNOWLEDGMENTS

We gratefully acknowledge the comments and suggestions of the Associate Editor and anonymous referee that led to stronger presentation of ideas. We also thank Patrick Sullivan for permission to use the *Checking In!* data set. *Conflict of Interest*: None declared.

FUNDING

We further acknowledge the partial support from the University of Rochester CTSA award UL1TR000042 from the National Center for Advancing Translational Sciences of the NIH and from the University of Rochester Center for AIDS Research grant P30AI078498 (NIH/NIAID) and the University of Rochester School of Medicine and Dentistry.

REFERENCES

- AALEN, O. O., BORGAN, Ø. AND GJESSING, H. (2008). *Survival and Event History Analysis: A Process Point of View*. New York, NY: Springer.
- AALEN, O. O. AND HJORT, N. (2002). Frailty models that yield proportional hazards. *Statistics and Probability Letters* **58**, 335–342.
- BATES, D., MAECHLER, M., BOLKER, B., WALKER, S., CHRISTENSEN, R. H. B., SINGMANN, H., DAI, B., GROTHENDIECK, G. AND GREEN, P. (2016). *Package ‘lme4’: Linear Mixed-Effects Models using ‘Eigen’ and S4*.
- BROWN, C. C. (1975). On the use of indicator variables for studying the time-dependence of parameters in a response-time model. *Biometrics* **31**, 863–872.

- CARLIN, J. B., WOLFE, R., BROWN, C. H. AND GELMAN, A. (2001). A case study on the choice, interpretation and checking of multilevel models for longitudinal binary outcomes. *Biostatistics* **2**, 397–416.
- CONAWAY, M. R. (1990). A random effects model for binary data. *Biometrics* **46**, 317–328.
- DIGGLE, P. J., HEAGERTY, P., LIANG, K.-Y. AND ZEGER, S. L. (2002). *Analysis of Longitudinal Data*, 2nd edition. Oxford, UK: Oxford.
- EFRON, B. AND TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall.
- FITZMAURICE, G. M., LAIRD, N. M. AND WARE, J. H. (2004). *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons.
- FOUFOULA-GEORGIOU, E. AND LETTENMAIER, D. P. (1986). Continuous-time versus discrete-time point process models for rainfall occurrence series. *Water Resources Research* **22**, 531–542.
- HALL, D. B. (2000). Zero-inflated poisson and binomial regression with random effects: a case study. *Biometrics* **56**, 1030–1039.
- KALBFLEISCH, J. D. AND PRENTICE, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edition. Hoboken, NJ: John Wiley & Sons.
- KHOSROPOUR, C. M., JOHNSON, B. A., RICCA, A. V. AND SULLIVAN, P. S. (2013). Enhancing retention of an internet-based cohort study of men who have sex with men (MSM) via text messaging: randomized controlled trial. *Journal of Medical Internet Research* **15**, e194.
- LAWLESS, J. F. (1987). Regression methods for Poisson process data. *Journal of the American Statistical Association* **82**, 808–815.
- LAWLESS, J. F. AND ZHAN, M. (1998). Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *The Canadian Journal of Statistics* **26**, 549–565.
- MCCULLOCH, C. E. AND NEUHAUS, J. M. (2011). Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. *Statistical Science* **26**, 388–402.
- PRENTICE, R. L. AND GLOECKLER, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* **34**, 57–67.
- PRICE, D. L. AND MANATUNGA, A. K. (2001). Modelling survival data with a cured fraction using frailty models. *Statistics in Medicine* **20**, 1515–1527.
- R CORE TEAM. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- VERBEKE, G. AND 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.
- WIENKE, A., RIPATTI, S., PALMGREN, J. AND YASHIN, A. (2009). A bivariate survival model with compound Poisson frailty. *Statistics in Medicine* **29**, 275–283.

[Received March 9, 2017; revised November 27, 2017; accepted for publication December 10, 2017]