Published in final edited form as:

Stat Med. 2010 March 15; 29(6): 627–638. doi:10.1002/sim.3825.

Modeling smoking cessation data with alternating states and a cure fraction using frailty models

Yimei Li^{a,*†}, E. Paul Wileyto^b, and Daniel F. Heitjan^a

^aDepartment of Biostatistics & Epidemiology, University of Pennsylvania, Philadelphia, U.S.A.

^bDepartment of Psychiatry, University of Pennsylvania, Philadelphia, U.S.A.

Abstract

We propose a flexible parametric model to describe alternating states recurrent-event data where there is a possibility of cure with each type of event. We begin by introducing a novel cure model in which a common frailty influences both the cure probability and the hazard function given not cured. We then extend our model to data with recurring events of two alternating types. We assume that each type of event has a gamma frailty, and we link the frailties by a Clayton copula. We illustrate the model with an analysis of data from two smoking cessation trials comparing bupropion and placebo, in which each subject potentially experienced a series of lapse and recovery events. Our analysis suggests that bupropion increases the probability of permanent cure and decreases the hazard of lapse, but does not affect the distribution of time to recovery during a lapse. The data suggest a positive but non-significant association between the lapse and recovery frailties. A simulation study suggests that the estimates have little bias and that their 95 per cent confidence intervals have nearly nominal coverage in samples of practical size.

Keywords

addiction; copula; frailty model; multiple events; recurrent events

1. Introduction

Subjects in smoking cessation trials commonly transit several times between lapse and recovery. An important feature of such data is that any quit episode may become permanent, in the sense that the subject stops smoking for good, and any lapse may also become permanent, in the sense that the subject abandons the quit attempt entirely. The notion that subjects exhibit both temporary and permanent quits is well established in smoking cessation research [1] and has found support in recent statistical analyses of smoking cessation data sets [2–6]. Another important aspect of such data is that an individual's quit pattern may reflect underlying factors that are not readily encoded in covariates. This manifests itself as correlation between repeated outcomes, such as series of quit and lapse durations within individuals. Thus, a comprehensive model for a smoking cessation data set would also need to include the possibility of correlation between events within individuals.

*Correspondence to: Yimei Li, Department of Biostatistics & Epidemiology, University of Pennsylvania, 503 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, U.S.A. [†] yimeili@mail.med.upenn.edu.

We consider these elements—recurrent events, cure modeling, and modeling of correlation—in turn below.

Recurrent events

Recurrent-event data arise when subjects can have repeated episodes of the event in question; common examples in medicine include coronary infarctions, relapses of cancer, and loss of viral control in HIV infection. Most analyses of multiple failure times have used extended versions of the Cox model, although the same extensions apply to parametric models. Unlike their single-event counterparts, there are many structural choices to make in describing the time origin and at-risk status for an individual episode [7, 8]. In smoking cessation studies, the appropriate modeling strategy is one in which the subject returns to the risk set only at the end of the preceding event [9, 10].

Cure models

In survival analysis we typically assume that all subjects are genuinely at risk and will eventually experience an event. In some applications, however, a fraction may either be or become non-susceptible. An example is time to recurrence of cancer among patients treated with surgery, where we may find that some patients never recur, presumably because the operation removed every vestige of the tumor [11]. *Cure models* were devised to describe such data. The pertinent literature begins with Farewell, who used logistic regression to model the cured fraction and Weibull regression to model the non-cured event rate [12]. Kuk and Chen proposed a Cox model for non-cured survival [13], and Peng *et al.* modeled the survival component with a generalized F [14]. Attempts to model such data semiparametrically encounter vexing identifiability problems, because non-parametric survival estimates need not decline to zero [15]. Sy and Taylor therefore proposed a proportional hazards model with a zero-tail constraint [16], and Li and Taylor described non-cured survival with a semiparametric accelerated failure time model, which handles identifiability more gracefully [17].

Modeling heterogeneity/correlation

A popular device for extending the classical survival model to account for between-unit heterogeneity (within-unit correlation) is the *frailty model*, first proposed by Vaupel *et al.* [18, 19]. A typical frailty model assumes that all subjects' survival distributions follow a common form up to a random subject-specific effect, known as a frailty. One then integrates out this latent factor to derive a new marginal distribution for the data. The introduction of frailties can describe lack of fit to classical survival models, between-subject heterogeneity, and within-subject correlation in clustered, multivariate, or recurrent outcomes. Practitioners commonly assume that frailties follow a distribution from the power variance function family, such as the gamma, inverse Gaussian, or positive stable law [20, 21].

Several authors have proposed extending the basic frailty model to incorporate the possibility of cure. Aalen included distributions with a non-susceptible subgroup [22] and developed these models further using a class of compound Poisson distributions [23]. Longini and Halloran proposed frailty mixture models [24], and Price and Manatunga elaborated the idea of applying the frailty concept to cure models [25]. Chen *et al.* and Yin

and Ibrahim described models for cancer recurrence in which a latent Poisson variable gives the number of surviving malignant clones in a treated subject, with the subject's overall time to recurrence being the minimum of the times to clinical observability generated by the clones [26–29]; a subject with zero clones is cured.

In this article we propose a model to describe data on series of lapse and recovery episodes from subjects participating in a smoking cessation trial. In Section 2, we introduce a frailty model with cure fraction, its extensions, and estimation methods. In Section 3, we apply the proposed model to a smoking cessation data set. In Section 4 we present simulation results, and we offer concluding comments in Section 5.

2. Statistical models

2.1. A frailty model with a cure fraction

First consider univariate time-to-event data. We let T_i and $S_i(t_i)$ denote a single event time with its survival function for subject i , and assume that the subject is either cured (with probability π_i) or has a proper survival function $S_i^*(t_i)$ (with probability $1 - \pi_i$):

$$S_i(t_i) = \pi_i + (1 - \pi_i)S_i^*(t_i). \quad (1)$$

To account for heterogeneity, we assume that each subject has an unknown, latent frailty, denoted b_i , that affects both the cure fraction and the survival hazard. We further assume that, conditional on b_i , the cure probability follows a binary regression with a complementary log-log (cloglog) link, and time to event given not cured is determined by a proportional hazard model with a constant baseline hazard. Thus,

$$\pi_i(b_i) = \exp[-b_i \exp(-\eta_{\pi_i})], \quad S_i^*(t_i|b_i) = \exp[-t_i b_i \exp(\eta_{\sigma_i})], \quad (2)$$

where η_{π_i} and η_{σ_i} are linear predictors for the cure fraction and event rate, respectively. We represent the linear predictors as

$$\eta_{\pi_i} = \alpha_{\pi} + \beta_{\pi} V_i, \quad \eta_{\sigma_i} = \alpha_{\sigma} + \beta_{\sigma} W_i, \quad (3)$$

where the predictor vectors V_i and W_i may overlap. Letting h_i^* be the hazard function associated with S_i^* , we can rewrite (2) and (3) as

$$\ln(-\ln \pi_i(b_i)) = \ln b_i + \alpha_{\pi} + \beta_{\pi} V_i, \quad \ln h_i^*(t_i|b_i) = \ln b_i + \alpha_{\sigma} + \beta_{\sigma} W_i.$$

In our parameterization, a higher value of the frailty b_i is associated with both a decreased cure probability and an increased hazard of experiencing the event. This is a natural and plausible way to allow the underlying frailty to affect both cure and survival.

Following common practice, we assume that the frailty follows a gamma distribution with shape and scale both equal to $1/\theta$, ensuring that $E(b)=1$ and $\text{Var}(b)=\theta$ ($\theta>0$). Integrating out b_i , the marginal survival function is

$$\tilde{S}_i(t_i) = \int S_i(t_i | b_i) f(b_i) db_i = \int S_i(t_i | b_i) \frac{b_i^{1/\theta-1}}{e^{-b_i/\theta} \theta^{1/\theta} \Gamma(1/\theta)} db_i = \left(\frac{1}{1+\theta e^{-\eta \pi_i}} \right)^{1/\theta} + \left(\frac{1}{1+\theta t_i e^{\eta \sigma_i}} \right)^{1/\theta} - \left(\frac{1}{1+\theta e^{-\eta \pi_i} + \theta t_i e^{\eta \sigma_i}} \right)^{1/\theta}, \quad (4)$$

with corresponding density

$$\tilde{f}_i(t_i) = e^{\eta \sigma_i} \left[\left(\frac{1}{1+\theta t_i e^{\eta \sigma_i}} \right)^{1/\theta+1} - \left(\frac{1}{1+\theta e^{-\eta \pi_i} + \theta t_i e^{\eta \sigma_i}} \right)^{1/\theta+1} \right]. \quad (5)$$

Therefore, the log likelihood of the observed data is

$$l = \sum_i [d_i \ln \tilde{f}_i(t_i) + (1 - d_i) \ln \tilde{S}_i(t_i)], \quad (6)$$

where for subject i , t_i is the observed time (minimum of event and censoring times) and d_i is the indicator that an event has occurred. We can rewrite (4) as

$$\tilde{S}_i(t_i) = \tilde{\pi}_i + (1 - \tilde{\pi}_i) \tilde{S}_i^*(t_i),$$

where

$$\tilde{\pi}_i = \left(\frac{1}{1+\theta e^{-\eta \pi_i}} \right)^{1/\theta}, \quad (7)$$

and

$$\tilde{S}_i^*(t_i) = \frac{(1+\theta t_i e^{\eta \sigma_i})^{-1/\theta} - (1+\theta e^{-\eta \pi_i} + \theta t_i e^{\eta \sigma_i})^{-1/\theta}}{1 - (1+\theta e^{-\eta \pi_i})^{-1/\theta}}. \quad (8)$$

Because $\tilde{S}_i^*(t_i)$ is a proper survival function, the marginal cure fraction is $\tilde{\pi}_i$ and the marginal non-cured survival function is $\tilde{S}_i^*(t_i)$, with corresponding marginal non-cured hazard function

$$\tilde{h}_i^*(t_i) = e^{\eta\sigma_i} \frac{(1 + \theta t_i e^{\eta\sigma_i})^{-1/\theta-1} - (1 + \theta e^{-\eta\pi_i} + \theta t_i e^{\eta\sigma_i})^{-1/\theta-1}}{(1 + \theta t_i e^{\eta\sigma_i})^{-1/\theta} - (1 + \theta e^{-\eta\pi_i} + \theta t_i e^{\eta\sigma_i})^{-1/\theta}}. \quad (9)$$

We note that the cure probability and hazard given not cured are originally defined at the subject level, conditional on the frailty. Thus although we have also derived the marginal cure fraction and non-cure hazard, as a rule we are more interested in the effects of predictors at the individual level, and we will interpret the parameters in this conditional sense.

2.2. Modeling repeated events

One can extend the model to repeated events by using the underlying frailty to induce dependence among the event times. That is, we assume that for each subject, given the susceptibility b_i the event times are independent with survival function given by (2) and (3). Denoting t_{ij} as the length of the j -th episode for subject i , we have

$$S_{ij}(t_{ij} | b_i) = \pi_{ij}(b_i) + [1 - \pi_{ij}(b_i)] S_{ij}^*(t_{ij} | b_i), \quad (10)$$

$$\pi_{ij}(b_i) = \exp[-b_i \exp(-\eta\pi_{ij})], \quad S_{ij}^*(t_{ij} | b_i) = \exp[-t_{ij} b_i \exp(\eta\sigma_{ij})], \quad (11)$$

where the linear predictors can include both subject-level and episode-level predictors. The log likelihood is then

$$l = \sum_i \ln \left[\int \exp \left(\sum_j l_{ij} \right) f(b_i) db_i \right], \quad (12)$$

where

$$\begin{aligned} l_{ij} &= d_{ij} \ln f_{ij} + (1 - d_{ij}) \ln S_{ij} \\ &= d_{ij} \ln \left(-\frac{\partial S_{ij}}{\partial t_{ij}} \right) + (1 - d_{ij}) \ln S_{ij}. \end{aligned}$$

One can in principle derive closed-form expressions for the likelihood terms, but the derivation quickly becomes tedious as the number of events increases. Moreover, the analytical formulas involve numerous subtractions among quantities like those in equations (4) and (5), and therefore may lead to loss of precision when evaluated numerically. Therefore in our implementation of the model we have relied on numerical quadrature, which although only mathematically approximate is not subject to cancelation errors and therefore may be more reliable in practice.

2.3. Modeling alternating states

As we have indicated, subjects in smoking cessation trials transit between smoking and abstinence, often several times during the course of treatment. Those currently in state 1 (abstinent) are at risk for events of type 1 (lapse), whereas those currently in state 2 (smoking) are at risk for events of type 2 (recovery of abstinence). Hougaard [30] terms this situation as ‘alternating states’ (Figure 1).

To describe such data, we further extend the model by positing the existence of two correlated frailties $(b_i^{(1)}, b_i^{(2)})$ for each subject, where $b_i^{(k)}$ pertains to events of type k . As in (10) and (11), the survival function for episode j of type k for subject i is

$$S_{ij}^{(k)}(t_{ij}^{(k)} | b_i^{(k)}) = \pi_{ij}^{(k)}(b_i^{(k)}) + [1 - \pi_{ij}^{(k)}(b_i^{(k)})] S_{ij}^{*(k)}(t_{ij}^{(k)} | b_i^{(k)}),$$

where

$$\begin{aligned} \pi_{ij}^{(k)}(b_i^{(k)}) &= \exp[-b_i^{(k)} \exp(-\eta_{\pi_{ij}}^{(k)})], \\ S_{ij}^{*(k)}(t_{ij}^{(k)} | b_i^{(k)}) &= \exp[-t_{ij}^{(k)} b_i^{(k)} \exp(\eta_{\sigma_{ij}}^{(k)})], \quad k=1, 2, \end{aligned} \quad (13)$$

and $\eta_{ij}^{(k)}$ and $\sigma_{ij}^{(k)}$ refer to the linear predictors for cure fraction and hazard, respectively, of subject i , episode j of type k . To complete the model, we must specify the joint distribution of the frailties. A popular choice in hierarchical models is the lognormal, which can flexibly represent a range of complex dependence structures. We prefer, however, to use marginal gamma distributions, which are more commonly associated with frailty models. Smith *et al.* proposed a five-parameter bivariate gamma distribution [31], but its intractability complicates implementation. An alternative approach is to apply a copula, which allows for correlation while retaining the specified marginals. We will use the Clayton copula [32], whose joint cumulative distribution function (CDF) is:

$$F(b_i^{(1)}, b_i^{(2)}) = \begin{cases} \max[(F_1(b_i^{(1)})^{-\alpha} + F_2(b_i^{(2)})^{-\alpha} - 1)^{-1/\alpha}, 0] & \text{if } \alpha \neq 0, \\ F_1(b_i^{(1)}) F_2(b_i^{(2)}) & \text{if } \alpha = 0, \end{cases}$$

where $F_k(\cdot)$, $k=1, 2$, is the CDF of a gamma (shape= $1/\theta_k$, scale= $1/\theta_k$) distribution, and α is an association parameter defined on $(-1, +\infty)$. The copula approaches the minimum, product, and maximum, as α approaches -1 , 0 and $+\infty$, respectively. The parameter α measures the strength of association and is linked to Kendall's τ by $\tau = \alpha/(\alpha+2)$. The corresponding density is

$$f(b_i^{(1)}, b_i^{(2)}) = \begin{cases} (1+\alpha) \max[(F_1(b_i^{(1)})^{-\alpha} + F_2(b_i^{(2)})^{-\alpha} - 1)^{-1/\alpha-2}, 0] \\ \quad \times F_1(b_i^{(1)})^{-\alpha-1} F_2(b_i^{(2)})^{-\alpha-1} f_1(b_i^{(1)}) f_2(b_i^{(2)}) & \text{if } \alpha \neq 0, \\ f_1(b_i^{(1)}) f_2(b_i^{(2)}) & \text{if } \alpha = 0; \end{cases}$$

where $f_k(\cdot)$, $k=1, 2$, is the gamma density with shape and scale both equal to $1/\theta_k$. The log likelihood is

$$l = \sum_i \ln \left[\int \exp \left(\sum_j l_{ij} \right) f(b_i^{(1)}, b_i^{(2)}) db_i^{(1)} db_i^{(2)} \right], \quad (14)$$

where

$$l_{ij} = \text{Indicator}(\text{type 1 event}) \cdot l_{ij}^{(1)} + \text{Indicator}(\text{type 2 event}) \cdot l_{ij}^{(2)},$$

$$l_{ij}^{(k)} = d_{ij}^{(k)} \ln \left(-\frac{\partial S_{ij}^{(k)}}{\partial \theta_i^{(k)}} \right) + (1 - d_{ij}^{(k)}) \ln S_{ij}^{(k)}, \quad k=1, 2.$$

2.4. Estimation

As indicated above, we advocate computing the likelihood by numerical integration. A straightforward approach is to use bivariate Gaussian quadrature [33], as implemented, for example, in SAS Proc NLMIXED. Unfortunately, we cannot apply Gaussian quadrature directly in Proc NLMIXED unless the random effects are normally distributed. Nelson *et al.* described a simple computational method using the probability integral transformation (PIT) [34], and Liu and Huang applied it to various frailty proportional hazards models [35]. Recently, Liu and Yu proposed another integration method that reformulates the conditional likelihood on non-normal random effects [36]. One can implement both PIT and likelihood reformulation (LR) in Proc NLMIXED, but we favor LR because it offers reduced computing times and a potentially broader range of application.

Following the notation in the previous section, the likelihood from subject i is

$$L_i = \int \exp \left(\sum_j l_{ij} \right) f(b_i^{(1)}, b_i^{(2)}) db_i^{(1)} db_i^{(2)}. \quad (15)$$

Multiplying and dividing the integrand by the standard bivariate normal density $\phi(\cdot, \cdot)$ gives

$$L_i = \int \exp \left(\sum_j l_{ij} \right) \frac{f(e^{a_i^{(1)}}, e^{a_i^{(2)}}) e^{a_i^{(1)} + a_i^{(2)}}}{\phi(a_i^{(1)}, a_i^{(2)})} \phi(a_i^{(1)}, a_i^{(2)}) da_i^{(1)} da_i^{(2)}, \quad (16)$$

where $a_i^{(k)} = \ln b_i^{(k)}$, $k=1, 2$; the extra $e^{a_i^{(1)} + a_i^{(2)}}$ arises from the Jacobian. We can then apply Gaussian quadrature in SAS Proc NLMIXED to integrate

$$\exp \left(\sum_j l_{ij} \right) \frac{f(e^{a_i^{(1)}}, e^{a_i^{(2)}}) e^{a_i^{(1)} + a_i^{(2)}}}{\phi(a_i^{(1)}, a_i^{(2)})}.$$

Specifically, the new conditional log likelihood for subject i consists of three elements:

$$l_i^A = \sum_j l_{ij},$$

$$l_i^B = \ln f(e^{a_i^{(1)}}, e^{a_i^{(2)}}) + a_i^{(1)} + a_i^{(2)},$$

$$l_i^C = -0.5(a_i^{(1)})^2 - 0.5(a_i^{(2)})^2 - \ln(2\pi),$$

leading to

$$L_i = \int \exp(l_i^A + l_i^B - l_i^C) \phi(a_i^{(1)}, a_i^{(2)}) da_i^{(1)} da_i^{(2)}.$$

We have obtained maximum likelihood estimates (MLEs) using quasi-Newton optimization in SAS Proc NLMIXED, with covariance estimates from inversion of a finite-differences approximation to the Hessian [37]. As a check on the validity of standard errors (SEs), we have also computed estimated sample variances using a parametric bootstrap [38]. Specifically, denoting ξ as the parameter vector, we estimated the model from the actual data to obtain the MLE $\hat{\xi}$. We then generated a large number Q of data sets simulated under the assumed model with ξ fixed at $\hat{\xi}$, obtaining from data set q the new estimate $\hat{\xi}^{*(q)}$. The estimated variance-covariance matrix is then the sample variance of the bootstrapped estimates $\hat{\xi}^{*(q)}$ ($q=1, \dots, Q$). In simulations, confidence intervals (CIs) based on bootstrap SEs reliably gave near-nominal coverage probabilities.

2.5. A marginal approach

An alternative to likelihood-based modeling is a *marginal* approach [5], in which one estimates the model as though the correlated events are independent, then constructs a covariance matrix that adjusts for potential correlation [39]. For our model, we assume the survival function for the length $T_{ij}^{(k)}$ of episode j of type k on subject i to be

$$\tilde{S}_{ij}^{(k)}(t_{ij}^{(k)}) = \left(\frac{1}{1 + \theta^{(k)} e^{-\eta_{\pi_{ij}}^{(k)}}} \right)^{1/\theta^{(k)}} + \left(\frac{1}{1 + \theta^{(k)} t_{ij}^{(k)} e^{\eta_{\sigma_{ij}}^{(k)}}} \right)^{1/\theta^{(k)}} - \left(\frac{1}{1 + \theta^{(k)} e^{-\eta_{\pi_{ij}}^{(k)}} + \theta^{(k)} t_{ij}^{(k)} e^{\eta_{\sigma_{ij}}^{(k)}}} \right)^{1/\theta^{(k)}}, \quad (17)$$

for $i=1, \dots, m$, $j=1, \dots, n_i^{(k)}$, $k=1, 2$, where m is the total number of subjects and $n_i^{(k)}$ is the number of episodes of type k on subject i . Here $\eta_{\pi_{ij}}^{(k)}$ and $\eta_{\sigma_{ij}}^{(k)}$ refer to linear predictors for

subject i , event j of type k , for cure fraction and event rate given not cured, respectively.

Note that the marginal distribution of an event time $T_{ij}^{(k)}$ follows (17) in all of the models in Sections 2.1, 2.2, and 2.3. But in Sections 2.2 and 2.3, the likelihood is calculated based on the joint distribution of $\{T_{ij}^{(k)}; j=1, 2, \dots, n_i^{(k)}\}$, instead of their marginal distributions.

We denote $U(\hat{\xi})$ as the estimated score function and $I(\hat{\xi})$ as the estimated information matrix. Then we estimate the variance–covariance matrix of $\hat{\xi}$ by the Williams robust estimate $I^{-1}(\hat{\xi})\widehat{\text{Var}}_w(U(\hat{\xi}))I^{-1}(\hat{\xi})$, where

$$\widehat{\text{Var}}_w(U(\xi)) = \widehat{\text{Var}}\left(\sum_{i=1}^m \sum_{k=1}^2 \sum_{j=1}^{n_i^k} U_{ij}^{(k)}(\xi)\right) = \frac{m}{m-1} \sum_{i=1}^m (U_i(\xi) - \bar{U}(\xi))(U_i(\xi) - \bar{U}(\xi))^T, \quad (18)$$

$U_{ij}^{(k)}(\xi)$ is the score for episode j of type k on subject i , $U_i(\xi) = \sum_{k=1}^2 \sum_{j=1}^{n_i^k} U_{ij}^{(k)}(\xi)$ is the score function for subject i , and $\bar{U}(\xi) = \sum_{i=1}^m U_i(\xi)/m$ is the mean of the individual scores.

Note that observing a set of recurrent event times over a fixed length of observation C induces dependent censoring in that the censoring time of T_{ij} is $C - T_{i1} - \dots - T_{i,j-1}$ (for $j \geq 2$) [40, 41]. Specifically, the excess of shorter events observed for later episodes (except the first episode) can bias frequentist data summaries such as marginal model fits and simple Kaplan–Meier (KM) estimates. But because the data are coarsened at random, in the sense that censoring times depend only on fully observed data, likelihood-based inferences that ignore the randomness of the coarsening mechanism are nevertheless valid [42–44]. We will revisit this issue in the simulations.

Code is available from the first author.

3. Application

The data are from two clinical trials of bupropion for smoking cessation—one ($n=555$) conducted at Georgetown University [45] and the other ($n=559$) at Brown University [46]. Because the designs were nearly identical, the two studies have been published in combined analyses [47]. Briefly, eligible participants were randomized to receive 10 weeks of treatment with either sustained-release bupropion (Zyban, 300mg/day) or placebo, plus seven sessions of in-person behavioral group counseling. Subjects began taking assigned medication on the day of the first counseling session, and were instructed to smoke as normal until the target quit date (TQD), scheduled for 14 days after the initiation of drug treatment. Telephone follow-ups were conducted at the end of treatment (EOT) and 6 and 12 months post-TQD. Before each counseling session and during follow-up assessments, participants reported the number of cigarettes smoked each day since the previous assessment. From the daily cigarette consumption records we constructed the outcomes, time to lapse (a day that includes smoking) and time to recovery from a lapse (an entire day without smoking).

We restrict our analysis to the treatment phase of the study, an 8-week period between TQD and EOT. Among the 1114 randomized subjects, 757 succeeded in quitting on the TQD and 357 failed. Experience has shown that subjects who are unable to quit on the TQD have sharply different outcome patterns, and therefore we excluded them from this analysis.

The 757 subjects had varying number of episodes, from 1 to 42. Among them, 314 subjects (42 per cent) had only one episode, implying that they never lapsed during the entire 8 weeks. Ninety-five subjects (13 per cent) had only two episodes, with an average censoring time 42, implying that they lapsed shortly after the initial quit and then continued smoking to the end of the study. Most of the rest of the subjects experienced many episodes with short durations, on average 8 episodes of length 7 days.

Figure 2 depicts the KM curves in the two groups for the two types of events, disregarding episode number and clustering. This is a rough way to understand the pattern of the data, because as we have indicated the censoring induces a potential bias in event times. Nevertheless we see that all curves exhibit an initial rapid decline followed by a plateau. The recovery events decline steeply to a nearly horizontal plateau, signaling the existence of a cured fraction. For lapse events the decline is more gradual and the plateau less perfectly flat, making the hypothesis of a cure model less certain but still plausible. For lapse events the curve for the intervention group appears to decline more slowly and to a higher plateau compared to the placebo group, suggesting possible drug effects on both cure fraction and hazard of lapse given not cured.

We analyzed the data under three models. First we fit the *independent frailty by episode and type (IFET)* model, which assumes separate, independent frailties for each event, adjusting for potential correlation using the cluster-correlated robust variance estimate of Williams [39] (Section 2.5). The second model *independent frailty by type (IFT)* assumes separate, independent frailties by event type (Section 2.2). That is, we again model the two types of events separately, but let the recurrent episodes of each type on one subject share a common underlying frailty, as shown in equations (10) and (11). The third model *correlated frailty (CF)* extends IFT by allowing the possibility of correlation between the two frailties through the gamma Clayton copula (Section 2.3). In all the three models we assume that the linear predictor includes only an intercept and a drug treatment effect. Results appear in Table I.

In the IFET model, robust SEs exceed naïve SEs by 9–48 per cent for all parameters pertaining to recovery events, whereas for lapse events, robust SEs associated with the variance and the intercept in the cure component are deflated slightly (15–19 per cent). The differences do not affect statistical significance, however. The Akaike information criterion (AIC) suggests a substantially better fit with IFT. Applying the parametric bootstrap with $Q=20$ changes the SEs only slightly and has no effect on statistical significance. Examination of the CF results reveals that incorporating the association parameter α has little effect on the parameter estimates or the quality of the fit; the log likelihood increases by only 0.2 and the AIC by 1.6, and the 95 per cent Wald CI for α includes 0 ($[-0.334, 0.417]$). Thus we henceforth restrict attention to the IFT model.

To check the model fit, we calculated the marginal survival function based on the MLEs (equation (4)), for the two types of events and two treatment arms separately. In Figure 3, we compare these fitted marginal survival curves to KM estimates from the first episode only. We limited our analysis to the first episode because KM estimates from later episodes ($j \geq 2$) are potentially biased due to dependent censoring, as indicated in Section 2.5. The plots suggest that the model fits the data well. Based on equation (7), the estimated marginal cure probabilities for lapse are 0.39 for bupropion and 0.27 for placebo, and for recovery are 0.21 for bupropion and 0.23 for placebo, respectively. For recovery events, these estimates are very close to the plateau levels in the KM plots in Figure 3. For lapse events, these estimates are slightly lower than the graphical estimates, presumably because the curves are not completely flat by the end of the observation period.

Based on the IFT model results, bupropion increases the probability of permanent quit, with cure probability 0.11 for drug and 0.01 for placebo (for a standard individual with frailty 1). For non-cured events bupropion reduces the hazard of lapse, with hazard ratio (HR) 0.56. Bupropion has no significant effect on the probability of abandoning quit attempts (0.15 for drug vs. 0.17 for placebo for a standard individual), nor does it accelerate recovery for those who persist in trying to quit (HR=1.15 for drug vs. placebo). Note that the individual-level cure probabilities are quite different from the marginal cure probabilities, especially for lapse events. The individual and marginal probabilities agree when the frailty variance is zero but can diverge substantially when the frailty variance increases. For our data the estimated frailty variances are 1.63 for lapse and 0.26 for recovery, whence the disparity between marginal and individual quit probabilities.

In the CF model we expected that shorter times to recovery would be correlated with longer times to relapse, reflecting stronger addiction and implying a negative association parameter. The estimated α is slightly positive, however, although not significant. We postulate that individuals with a heavier addiction may have a stronger impulse to quit, which induces a higher frailty to recovery, whereas at the same time the heavy addiction makes abstinence harder and thus corresponds to a higher frailty to lapse.

4. Simulation

We conducted a series of simulations to evaluate the performance of our method in repeated samples. Each simulation consisted of 100 replications of 800 subjects. We fixed the length of follow-up at $C=60$ and assigned treatment randomly with probability 1/2, as in the trial. We chose the parameters to be similar to estimates from the data: For the lapse parameters, we set the intercept at -1.5 and the drug coefficient at 0.7 , corresponding to a cure probability of 0.01 for placebo and 0.11 for drug. We set the intercept and drug coefficients in the survival part at -1 and -0.6 , giving a baseline hazard of 0.37 and HR of 0.55 for drug vs. placebo. For recovery events, we set the intercept and drug coefficients in the cure component to be 0.5 and -0.1 , and those in the survival part to be -1 and 0.1 . We fixed the frailty variances at 1 and 0.3 and the association parameter at 1 .

For each simulated subject, we first generated two frailties from the Clayton copula using function `mvdC()` from the package `copula` in R 2.6.2 [48]. This gave an average correlation

of 0.35 in the 100 replicates, with range from 0.21 to 0.41. Next, we generated lapse times $T_j^{(1)}$ and recovery times $T_j^{(2)}$. To incorporate the cure fraction, we drew random variates $R_j^{(k)}$, $k=1,2$ from a Uniform (0,1); if $R_j^{(k)} < \pi_j^{(k)}$, $T_j^{(k)}$ was changed to a number beyond the end of study ($T_j^{(k)} = C+10$), representing a ‘cure’. The recurrent event times were then lined up as: $T_1^{(1)}, T_1^{(2)}, T_2^{(1)}, T_2^{(2)}, \dots$, and the cumulative time was used to determine censoring. That is, if the sum Ψ_l of times to the l -th event was less than C , and the sum Ψ_{l+1} of times to the $l+1$ -st event was greater than C , then the $l+1$ -st event was censored at $C-\Psi_l$. We fitted the CF model in Proc NLMIXED.

Simulation results appear in Table II. Bias for the CF model is modest, ranging from -0.010 to 0.014 , and all of the estimates except those for α are tightly distributed around their means. Hessian-based 95 per cent CIs achieve roughly nominal coverage, with values ranging from 89 to 98 per cent. Thus, the overall performance of the estimates appears to be satisfactory. Moreover, 94 per cent of the α estimates are significant at level 0.05, suggesting that the CF model has substantial power to detect a moderate underlying correlation.

We also fit the IFT model (fixing $\alpha=0$) to the simulated data (Table II). Estimated biases under the incorrect IFT model are 3 to 30 times larger than under the correct CF model, and the coverage of 95 per cent CIs is much poorer, ranging as low as 39 per cent. Thus, the CF model can provide a substantially better fit than the IFT model when the frailty correlation is moderate.

To better understand the power in testing the significance of α , we conducted two other simulations (results not shown) where the true α was set to be 0.4 or 0. With $\alpha=0.4$, the numerical correlation is around 0.17, and 29 per cent of the α estimates are significant. With $\alpha=0$, 10 per cent of the estimates are significant, so the type I error is somewhat inflated. We also investigated the likelihood ratio test (LRT), which gives power 95, 40, and 6 per cent at $\alpha=1, 0.4$, and 0, respectively, suggesting its superiority over the Wald test.

To investigate potential bias in the IFET (marginal) model (as suggested in Section 2.5), we estimated both the IFT and IFET models when α was set to 0 (Table III). As expected, biases in the parameter estimates under the IFT model are small (absolute value 0 to 0.019), whereas biases under the IFET model range are as high as 0.405. Accordingly, 95 per cent CIs have good coverage under the IFT model (91–97 per cent) and poor coverage under the IFET model (0–78 per cent).

Lastly, we estimated the parametric bootstrap SE for the CF model in each scenario. In all cases it turns out to be close to the Hessian-based SE (± 0.005), with 95 per cent CIs based on both approaches achieving near-nominal coverage probability. So the reliability of the simple Hessian-based SE seems adequate in these settings. For some other parameter values that we investigated, however, the Hessian-based SEs give reduced coverage probabilities; by contrast, the bootstrap method is well calibrated in every case.

5. Discussion

We have proposed a novel cure-mixture frailty model that allows for correlated events of alternating types, as one encounters in smoking cessation studies. Our model is unique in positing that a common frailty influences both the cure fraction and the survival rate. The use of a cloglog link for the cure fraction and a constant hazard for the non-cured survival offers analytic convenience, but because we use numerical integration and optimization even with small numbers of events, one could easily extend the model by substituting other, less tractable link functions and survival distributions. Moreover, although the models we estimated include only a randomization indicator in the linear predictor, it is straightforward to incorporate baseline and episode-specific predictors. In particular, this would allow for estimating trends across episodes.

Our method offers an alternative to mover–stayer Markov models with alternating states and a possibility of cure [2, 3] by incorporating time-to-event models for correlated outcomes. Use of the Clayton copula maintains the simplicity of marginally gamma frailties while avoiding the complexity of the bivariate gamma. Moreover, as with the link and hazard specification, one can readily substitute different and more complex frailty models. A potential disadvantage is that with dimension three or higher the Clayton copula allows only positive association.

As evaluation of the likelihood requires numerical integration in all but the simplest cases, parameter estimation is computationally intensive. We have employed the LR method of Liu and Yu [36], implemented in SAS Proc NLMIXED with the recommended 30 quadrature points and default quasi-Newton optimization. Simulation results suggest that the procedure works well, and moreover we confirmed results for fitting the IFET model in an independent implementation in R.

In simulations with smaller sample sizes (e.g. $m=200$), estimates sometimes failed to converge. One can avoid this problem by identifying better starting values, reparameterizing to avoid boundary constraints, or employing a more robust optimization technique such as Nelder–Mead. Changing optimization algorithms should be a last resort, however, as the default quasi-Newton method typically attains convergence in several minutes compared to possibly several hours for Nelder–Mead.

With this numerically intensive approach, the largest feasible number of simulation replicates was 100, and thus estimates of frequentist operating characteristics have substantial uncertainty. Nevertheless, we were able to establish that one can fit the model reliably in samples of practical size, that departures from underlying assumptions can cause severe bias, and that SEs derived from a finite-differences estimate of the Hessian can be accurate. Because for some parameter settings the Hessian-based SEs underestimated sampling variability, we recommend that in key analyses one employ the parametric bootstrap, which gave properly calibrated CIs in every simulation we attempted.

We applied our model to smoking data collected during the relatively brief 8-week treatment period, where time-to-lapse curves had not reached a plateau and consequently the cure fraction might be poorly estimated. Nevertheless we are convinced that there is a cured

fraction, as evidence from other trials suggests strongly that this is a real feature of smoking cessation data. It is well known that when cure is a possibility, failure to incorporate it in the model can cause severe bias in survival hazard parameters [16, 25]. The consequences of incorrectly assuming a cured fraction are perhaps less well known but amenable to investigation in any specific case. In fact we find that the cure model is most helpful as an analytic tool when there is enough censoring to create some ambiguity about it; otherwise, one can easily classify subjects as cured or not and then estimate cure and survival models directly. Nevertheless, we recommend restricting the use of the model to problems where the possibility of cure has a strong basis in the science.

There is a growing literature on the analysis of smoking cessation data. Banerjee and Carlin extended cure models to allow for spatial correlation using a Bayesian approach [4], and Yu and Peng applied a univariate cure model, adjusting for clustering by a one-step jackknife [5]. Both of these models focused on lapse events only. Yu and Peng's model is similar to our IFET model, but substitutes a jackknife for the sandwich variance estimate. A marginal approach can work well with clustered survival data but, as our simulations show, is subject to bias when applied to recurrent events. Another recent article by Luo *et al.* used a discrete-time stochastic model to describe the transitions among smoking, transient quitting, and permanent quitting [6]. A major difference with our work is that they model visit-to-visit transitions of smoking status at lengthy intervals, whereas we model transitions in more nearly continuous time. Moreover, their model fails to account for the probability of abandoning the quit attempt.

A potential further application of our model is to estimate individual frailties as a means of characterizing subjects with respect to amenability to treatment. This is important because nicotine addiction is most helpfully thought of as a chronic disease that may take years to successfully treat. Estimating patients' underlying chances of success may ultimately lead to better prediction of outcomes and personalization of therapies.

Although we developed our model to describe cigarette consumption histories in a randomized smoking cessation trial, it is potentially applicable in observational studies and in research on cessation of other substances of abuse. The modeling of alternating states will also be applicable in the study of relapsing–remitting diseases such as multiple sclerosis and depression, and diseases such as heart failure that involve repeated hospitalizations.

Acknowledgments

The bupropion trials were supported by the U.S. National Cancer Institute and National Institute on Drug Addiction under grants P50 CA/DA 84718 and R01 CA 63562 (Caryn Lerman, PI); National Heart, Lung, and Blood Institute under grant HL32318 (Ray Niaura, PI); and National Institute on Drug Abuse under grant DA08511 (Richard Brown, PI). The National Cancer Institute supported the research of Drs Heitjan and Wileyto under grant R01 CA 116723 (Daniel F. Heitjan, PI). We are grateful to Dr Lerman for permission to use the data, and to Drs Mary Putt and Robert Schnoll for helpful comments.

References

1. U.S. Department of Health and Human Services. DHHS Publication No. (CDC) 90-8416. USDHHS, PHS, CDC, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1990. The health benefits of smoking cessation.

2. Cook RJ, Kalbfleisch JD, Yi GY. A generalized mover-stayer model for panel data. *Biostatistics*. 2002; 3:407–420. [PubMed: 12933606]
3. Mannan HR, Koval JJ. Latent mixed Markov modelling of smoking transition using Monte Carlo bootstrapping. *Statistical Methods in Medical Research*. 2003; 12:125–246. [PubMed: 12665207]
4. Banerjee S, Carlin BP. Parametric spatial cure rate models for interval-censored time-to-relapse data. *Biometrics*. 2004; 60:268–275. [PubMed: 15032798]
5. Yu B, Peng Y. Mixture cure models for multivariate survival data. *Computational Statistics and Data Analysis*. 2008; 52:1524–1532.
6. Luo S, Crainiceanu CM, Louis TA, Chatterjee N. Analysis of smoking cessation patterns using a stochastic mixed-effects model with a latent cured state. *Journal of the American Statistical Association*. 2008; 103:1002–1013. [PubMed: 19305513]
7. Hosmer DW., Lemeshow, S. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. Wiley; New York: 1999.
8. Kelly PJ, Lim LL. Survival analysis for recurrent event data: an application to childhood infectious diseases. *Statistics in Medicine*. 2000; 19:13–33. [PubMed: 10623910]
9. Andersen, PK., Borgan, O., Grill, RD., Keilding. *Statistical Models Based on Counting Process*. Springer; New York: 1993.
10. Therneau TM, Hamilton SA. rhDNase as an example of recurrent event analysis. *Statistics in Medicine*. 1997; 16:2029–2047. [PubMed: 9308130]
11. Withers HR, Peters LJ, Taylor JM, Owen JB, Morrison WH, Schultheiss TE, Keane T, O'Sullivan B, van Dyk J, Gupta N. Local control of carcinoma of the tonsil by radiation therapy: an analysis of patterns of fractionation in nine institutions. *International Journal of Radiation Oncology, Biology, Physics*. 1995; 33:548–562.
12. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*. 1983; 38:1041–1046.
13. Kuk AYC, Chen CH. A mixture model combining logistic regression with proportional hazards regression. *Biometrika*. 1992; 79:531–541.
14. Peng Y, Dear KBG, Denham JW. A generalized F mixture model for cure rate estimation. *Statistics in Medicine*. 1998; 17:813–830. [PubMed: 9595613]
15. Peng Y, Carriere KC. An empirical comparison of parametric and semiparametric cure models. *Biometrical Journal*. 2002; 44:1002–1014.
16. Sy JP, Taylor JMG. Estimation in a Cox proportional hazards cure model. *Biometrics*. 2000; 56:227–236. [PubMed: 10783800]
17. Li CS, Taylor JMG. A semi-parametric accelerated failure time cure model. *Statistics in Medicine*. 2002; 21:3235–3247. [PubMed: 12375301]
18. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty and the dynamics of mortality. *Demography*. 1979; 16:439–454. [PubMed: 510638]
19. Duchateau, L., Janssen, P. *The Frailty Model*. Springer; New York: 2008.
20. Hougaard P. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*. 1986; 71:75–83.
21. Hougaard P. Heterogeneity models of disease susceptibility, with application to diabetic nephropathy. *Biometrics*. 1994; 50:1178–1188. [PubMed: 7787000]
22. Aalen OO. Heterogeneity in survival analysis. *Statistics in Medicine*. 1988; 7:1121–1137. [PubMed: 3201038]
23. Aalen OO. Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Annals of Applied Probability*. 1992; 2:951–972.
24. Longini IM, Halloran ME. A frailty mixture model for estimating vaccine efficacy. *Applied Statistics*. 1996; 45:165–173.
25. Price DL, Manatunga AK. Modelling survival data with a cured fraction using frailty models. *Statistics in Medicine*. 2001; 20:1515–1527. [PubMed: 11343371]
26. Chen MH, Ibrahim JG, Sinha D. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*. 1999; 94:909–919.

27. Chen MH, Ibrahim JG, Sinha D. Bayesian inference for multivariate survival data with a cure fracture. *Journal of Multivariate Analysis*. 2002; 80:101–126.
28. Yin G. Bayesian cure rate frailty modes with application to a root canal therapy study. *Biometrics*. 2005; 61:552–558. [PubMed: 16011704]
29. Yin G, Ibrahim JG. A class of Bayesian shared gamma frailty models with multivariate failure time data. *Biometrics*. 2005; 61:208–216. [PubMed: 15737095]
30. Hougaard, P. *Analysis of Multivariate Survival Data*. Springer; New York: 2000.
31. Smith, OE., Adelfang, SI., Tubbs, JD. NASA Technical Memorandum 82483. National Aeronautics and Space Administration; 1982. A bivariate Gamma probability distribution with application to gust model.
32. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65:141–152.
33. Abramowitz, M., Stegun, I. *Handbook of Mathematical Functions*. Dover; New York: 1972.
34. Nelson KP, Lipsitz SR, Fitzmaurice GM, Ibrahim J, Parzen M, Strawderman R. Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics*. 2006; 15:39–57.
35. Liu L, Huang X. The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine*. 2007; 27:2665–2683.
36. Liu L, Yu Z. A likelihood reformulation method in non-normal random effects models. *Statistics in Medicine*. 2008; 27:3105–3124. [PubMed: 18038445]
37. Dennis, JE., Schnabel, RB. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall; Englewood Cliffs, NJ: 1983.
38. Efron, B., Tibshirani, RJ. *An Introduction to the Bootstrap*. Chapman & Hall; London: 1993.
39. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics*. 2000; 56:645–646. [PubMed: 10877330]
40. Huang Y, Chen YQ. Marginal regression of gaps between recurrent events. *Lifetime Data Analysis*. 2003; 9:293–303. [PubMed: 14649847]
41. Chen YQ, Wang MC, Huang Y. Semiparametric regression analysis on longitudinal pattern of recurrent gap times. *Biostatistics*. 2004; 5:277–290. [PubMed: 15054031]
42. Heitjan DF, Rubin DB. Ignorability and coarse data. *The Annals of Statistics*. 1991; 19:2244–2253.
43. Heitjan DF. Ignorability and coarse data: some biomedical examples. *Biometrics*. 1993; 49:1099–1109. [PubMed: 8117903]
44. Heitjan DF. Ignorability in general incomplete-data models. *Biometrika*. 1994; 81:701–708.
45. Lerman C, Roth D, Kaufmann V, Audrain J, Hawk L, Liu A, Niaura R, Epstein L. Mediating mechanisms for the impact of bupropion in smoking cessation treatment. *Drug and Alcohol Dependence*. 2002; 67:219–223. [PubMed: 12095672]
46. Brown RA, Niaura R, Lloyd-Richardson EE, Strong DR, Kahler CW, Abrantes AM, Abrams D, Miller IW. Bupropion and cognitive-behavioral treatment for depression in smoking cessation. *Nicotine and Tobacco Research*. 2007; 9:721–730. [PubMed: 17577801]
47. Wileyto EP, Patterson F, Niaura R, Epstein LH, Brown RA, Audrain-McGovern J, Hawk LW, Lerman C. Recurrent event analysis of laps and recovery in a smoking cessation clinical trial using bupropion. *Nicotine and Tobacco Research*. 2005; 7(2):257–268. [PubMed: 16036283]
48. Yan J. Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software*. 2007; 21(4)

Figure 1.
Alternating states data.

Figure 2.
Kaplan–Meier estimates for the two event types, by treatment arms.

Figure 3.

Fitted marginal survival curves and Kaplan–Meier estimates for the first episode.

Table 1

Parameter estimates from the bupropion data.

	IPET		IFT		CF	
	Estimate	SE(adjusted)	Estimate	SE	Estimate	SE
<i>Lapse</i>						
Cure fraction						
Intercept	-1.627	0.210 (0.171)	-1.479	0.203	-1.475	0.207
Drug	0.669	0.141 (0.150)	0.680	0.221	0.681	0.220
Survival model						
Intercept	-1.581	0.059 (0.072)	-2.096	0.090	-2.096	0.090
Drug	-0.447	0.082 (0.126)	-0.586	0.122	-0.587	0.122
Variance	0.808	0.125 (0.106)	1.634	0.144	1.629	0.157
<i>Recovery</i>						
Cure fraction						
Intercept	-0.734	0.062 (0.071)	-0.578	0.062	-0.570	0.074
Drug	-0.019	0.081 (0.088)	-0.059	0.091	-0.057	0.091
Survival model						
Intercept	-0.815	0.048 (0.055)	-0.932	0.060	-0.942	0.075
Drug	0.117	0.069 (0.082)	0.141	0.087	0.140	0.087
Variance	0.201	0.029 (0.044)	0.255	0.043	0.261	0.044
Association	—	—	—	—	0.041	0.191
Log likelihood	-8024.1		-7716.2		-7716.0	
AIC	16068.2		15452.4		15454.0	

Table II

Results of simulations with $\alpha = 1$.

Parameter	True value	CF			IFT		
		Bias	$\sqrt{\text{MSE}}$	95 per cent CI, CP (per cent)	Bias	$\sqrt{\text{MSE}}$	95 per cent CI, CP (per cent)
<i>Lapse</i>							
Cure fraction							
Intercept	−1.5	0.010	0.113	93	−0.052	0.127	88
Drug	0.7	0.000	0.119	93	0.025	0.123	95
Survival model							
Intercept	−1.0	−0.010	0.078	89	0.026	0.082	86
Drug	−0.6	−0.005	0.095	93	−0.022	0.097	94
Variance	1.0	0.001	0.079	96	0.030	0.084	96
<i>Recovery</i>							
Cure fraction							
Intercept	−0.5	0.000	0.049	97	−0.099	0.108	39
Drug	−0.1	−0.003	0.072	96	−0.038	0.083	90
Survival model							
Intercept	−1.0	0.002	0.056	94	0.099	0.110	41
Drug	0.1	−0.006	0.066	98	0.022	0.070	93
Variance	0.3	0.004	0.040	95	−0.056	0.063	61
Association	1.0	0.014	0.343	96	—	—	—

Note: 'CP', Coverage probability.

Table III

Results of simulations with $\alpha = 0$.

Parameter	True value	IF-T			IFET		
		Bias	$\sqrt{\text{MSE}}$	95 per cent CI, CP (per cent)	Bias	$\sqrt{\text{MSE}}$	95 per cent CI, CP (per cent)
<i>Lapse</i>							
Cure fraction							
Intercept	-1.5	0.019	0.093	97	0.241	0.271	49
Drug	0.7	-0.014	0.116	97	-0.117	0.166	72
Survival model							
Intercept	-1.0	0.001	0.064	91	0.332	0.340	0
Drug	-0.6	0.003	0.094	95	0.112	0.148	54
Variance	1.0	-0.016	0.099	91	-0.405	0.413	0
<i>Recovery</i>							
Cure fraction							
Intercept	-0.5	-0.004	0.054	93	-0.252	0.266	2
Drug	-0.1	-0.008	0.076	95	-0.055	0.100	78
Survival model							
Intercept	-1.0	0.001	0.051	95	0.278	0.285	1
Drug	0.1	0.000	0.070	95	-0.059	0.101	70
Variance	0.3	-0.004	0.039	93	-0.079	0.104	44
Association	0	—	—	—	—	—	—

Note: 'CP', Coverage probability.