

BAYESIAN APPROACHES TO JOINT LONGITUDINAL AND SURVIVAL MODELS ACCOMMODATING BOTH ZERO AND NONZERO CURE FRACTIONS

Yueh-Yun Chi and Joseph G. Ibrahim

University of Washington and University of North Carolina at Chapel Hill

Abstract: The trade-offs between survival benefits and therapeutic adverse effects on quality of life (QOL) is always an important clinical issue for cancer and AIDS patients. The International Breast Cancer Study Group (IBCSG) conducted a large clinical trial, IBCSG Trial VI, to examine the duration and timing of adjuvant therapy for advanced breast cancer patients after the initial removal surgery. We present a novel joint model for longitudinal and survival data to evaluate the relationship between QOL and breast cancer progression, and also assess issues associated with different therapeutic procedures and baseline covariates. Multidimensional longitudinal QOL measurements are modeled in a hierarchical mixed effects model to account for psychological fluctuations and measurement errors, provide estimates for time points where QOL data are not available, and to explicitly allow for direct inferences about different dependence structures in the QOL data over time and over different QOL measures (indicators). A parametric survival model is also proposed for disease-free survival (DFS) to incorporate the underlying smooth QOL trajectories and prognostic factors. This survival model is attractive and capable of accommodating *both* zero and nonzero cure fractions. With advances in modern medicine, a positive cure fraction is often tenable for breast cancer patients since many are completely cured after surgery, and are no longer susceptible to relapse. A Bayesian paradigm is adopted to facilitate the estimation process and ease the computational complexity.

Key words and phrases: Breast cancer clinical trial, Cure rate model, random effects.

1. Introduction

In cancer and AIDS clinical trials, it is becoming increasingly common to repeatedly collect one or more biologic markers during the follow-up for the primary endpoint, the time to an event. These longitudinal markers are often important indicators for disease progression, but are prone to measurement error and random fluctuations, and thus their direct use as time-varying covariates in a survival model is inappropriate (see Ibrahim, Chen and Sinha (2001, Chap.7)). Jointly modeling the longitudinal markers and survival data provides a way to account

for errors in the longitudinal measurements, and allows the presence of treatment and baseline covariates, unbalanced observations of biomarkers over time, and censoring information associated with survival time. The model building often starts from models for the longitudinal component and then characteristics of the longitudinal model are incorporated into the model for the survival component. Both Frequentist and Bayesian approaches have been examined in the literature. Tsiatis and Davidian (2004) and Ibrahim, Chen and Sinha (2001, Chap.7) provide a detailed discussion of joint modeling.

The model proposed here was primarily motivated by a clinical trial conducted by the International Breast Cancer Study Group (IBCSG). IBCSG directed a large clinical trial, IBCSG Trial VI, in premenopausal women with node-positive breast cancer to examine the efficacy of post-surgery chemotherapy procedures. The therapeutic procedure is hypothesized to affect breast cancer progression, monitored in terms of disease free survival (DFS) which corresponds to time to breast cancer relapse, through two paths, either directly or indirectly with an intermediate time-dependent factor, patients' quality of life (QOL). The toxicity of a therapeutic procedure may adversely affect a patient's QOL and QOL is typically associated with DFS. A joint modeling framework for longitudinal and survival data not only allows investigation for both paths, but accommodates all important data features and complications associated with them. Particularly in response to the characteristics of the IBCSG data, a new Bayesian joint modeling with a multivariate extension for the longitudinal component and the possible presence of cure for the survival component DFS, is therefore proposed.

In the IBCSG trial, four distinct QOL indicators were assessed repeatedly over time with a self-reported QOL questionnaire. A multivariate longitudinal model is required to incorporate treatment and baseline covariates, and account for psychological fluctuations, measurement errors and unbalanced observations. Ibrahim, Chen and Sinha (2004), Xu and Zeger (2001) and Song, Davidian and Tsiatis (2002) have proposed various types of joint models with a multivariate longitudinal component. Here, we present a hierarchical mixed effects model to explicitly allow for direct inferences about different dependence structures in the QOL data over time as well as over different indicators.

Due to advances in cancer research, particularly breast cancer, a significant proportion of patients after the initial surgery are not susceptible to cancer relapse, that is, they are "cured" of the disease. Figure 1 shows the Kaplan-Meier curves for DFS for all patients. A plateau appears to occur in the survival curve after approximately 10 years of follow-up, hence suggesting a possible cure fraction in the population. Joint longitudinal-survival-cure models have been investigated by several authors (see Brown and Ibrahim (2003), Law, Taylor and Sandler (2002) and Yu, Law, Taylor and Sandler (2004)). The joint model we propose here is quite different from what has been proposed in the literature. First

and foremost, the proposed survival model does not impose a boundary value on the parameter space, as is the case for the classic mixture (Berkson and Gage (1952)) and non-mixture cure models (Yakovlev and Tsodikov (1996)). The incorporation of longitudinal QOL assessments in the survival model helps determine the proportion of cured and non-cured patients in the trial, and hence allows the flexibility for either zero or nonzero cure fractions in the joint model. In addition, our survival model maintains a proportional hazards structure when only baseline covariates are considered in the model. The rest of the article is organized as follows. In Section 2, we describe the basic data structure from the breast cancer clinical trial. In Section 3, we review the basic setup and demonstrate some attractive properties of our model. We examine the performance and robustness of our proposed model with a set of simulation studies in Section 4, and then apply the model to the analysis of the IBCSG dataset in Section 5. A brief discussion is given in Section 6 to conclude the article.

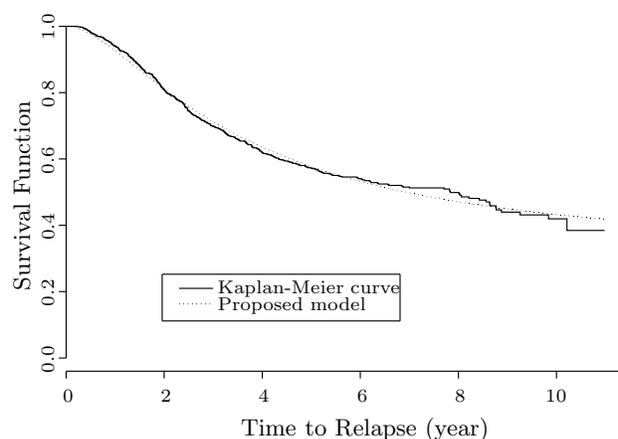


Figure 1. Superimposed survival curves.

2. The Data Structure

The IBCSG conducted a clinical trial, IBCSG Trial VI, in premenopausal women with node-positive breast cancer to investigate both the duration of adjuvant chemotherapy (three vs six initial cycles of oral cyclophosphamide, methotrexate, and fluorouracil (CMF)) and the reintroduction of three single courses of delayed chemotherapy. Each participant was randomly assigned in a 2×2 factorial design to receive either six initial courses of CMF at consecutive months 1 to 6 with (CMF6RE) or without (CMF6) three single courses of reintroduction CMF given on month 9, 12 and 15; or three initial courses of CMF at consecutive months 1 to 3 with (CMF3RE) or without (CMF3) three single

courses of reintroduction CMF given on month 6, 9 and 12. Randomization was stratified according to participation institution, type of surgery, and estrogen receptor (ER) status. Date of relapse is defined as the time when recurrent disease is diagnosed or first suspected, and disease-free survival (DFS) is defined as the time from randomization to any relapse, occurrence of a second primary cancer, or death. Different therapeutic procedures may directly affect DFS, as well as age, ER status (negative/positive), and number of positive nodes of the tumor. The trial is described in greater detail by the International Breast Cancer Study Group (1996).

As part of the trial, the QOL questionnaire was assigned periodically to all participants and included four indicators of health-related QOL that are especially relevant to breast cancer patients. Physical well-being (lousy—good), mood (miserable—happy), appetite (none—good) and perceived coping (“How much effort does it cost you to cope with your illness?” (a great deal—none)) were assessed with single-item linear analogue self-assessment scales. The scores were scaled between 0 to 100 with a larger number indicating a better perceived QOL. Measurement errors may occur by possible imperfect reliability of the questionnaire and are assumed to take place independently over time. Different QOL indicators may be related to each other and the correlation is most likely to be positive. For example, a happy mood mostly goes with a good appetite and physical well-being, and vice versa. These QOL measurements were important in the investigation of breast cancer progression since the toxicity of adjuvant chemotherapy may have adverse effects on patients’ QOL, and hence associate with DFS. The incorporation of longitudinal QOL measures in the model of DFS allows for the evaluation of the indirect therapeutic effects through the intermediate QOL factor. In addition, the direct therapeutic effects for optimal duration and timing of adjuvant therapy can also be examined. Both patients and physicians can make their decisions about therapies after balancing the trade-offs between survival benefits and QOL pay-offs.

We analyzed data from 831 patients from Switzerland, Sweden and New Zealand/Australia, each with more than one complete set of QOL assessments over time. The perceptions of QOL were assessed at the start of the study and at months 3 and 18 after randomization. A total of 2,152 QOL observations are included in the dataset. Figure 2 displays median QOL scores at three observational times. On average, patients perceive improvement in their QOL 18 months after the initial surgery, and the improvement is most substantial for their mood and perceived coping. However, the perceived mood and coping are relatively worse than perceived appetite and physical well-being over time. The median DFS is 7.836 years for all patients, with a censoring rate of 51.62%. For different therapeutic groups, the median DFS is 6.188 years for CMF3 therapy,

6.045 years for CMF6 therapy, 8.033 years for CMF3RE therapy and 9.257 years for CMF6RE therapy. The Kaplan-Meier curve for DFS in Figure 1 suggests a nonzero cure fraction after a long follow-up. A joint model capable of accommodating nonzero cure fractions for time-to-event data is essential for investigating the association between QOL scores and the survival outcome.

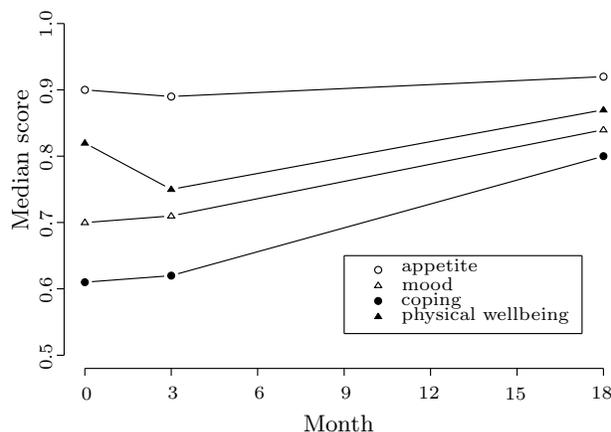


Figure 2. Median QOL scores over time.

3. A New Class of Joint Models

In this section, we review the background and setup for our model. Let Y be the possibly censored time-to-event, X be the observed vector of longitudinal QOL measures, Z be the vector of baseline covariates for time-to-event and R be the vector of covariates (possibly time-varying) for the longitudinal measures. The components of Z may be different from the components of R , which therefore allows different covariate information in the model of time-to-event and the longitudinal measures. With measurement errors and psychological fluctuations corresponding to each QOL indicator, we let a latent trajectory function X^* represent the underlying true QOL, and R affects X only through its influence on X^* . In the framework of joint modeling, we specifically assume that the clinical event time Y and vector of QOL measures X are conditionally independent given X^* . Based upon these assumptions, we have

$$P(Y, X|Z, R) = \int P(Y|Z, X^*)P(X|X^*)P(X^*|R)dX^*. \tag{3.1}$$

The joint likelihood of the complete data $\{Y, X, X^*\}$ requires the specification of three components, $P(Y|Z, X^*)$, $P(X|X^*)$ and $P(X^*|R)$. In the literature, several authors have proposed methods based on different distributional assumptions for

these three components. Our assumptions for $[X|X^*]$ and $[X^*|R]$ result in a multivariate longitudinal model with explicit dependence structures among the repeated measures, and hence allow for direct inference about each structure. We also propose a parametric survival model accommodating both zero and nonzero cure fractions for $[Y|Z, X^*]$.

3.1. The multivariate longitudinal process

With more than one QOL indicator measured over time, two levels of dependence structures are involved in the observations of a single patient. The first level involves dependence over time for each indicator, while the second level involves dependence over different indicators. A multivariate mixed effects model is proposed to explicitly model these two sources of dependence structures. Let $x_{ik}(t_{ij})$ be an assessment of the k th QOL indicator for the i th patient at time t_{ij} and $x_{ik}^*(t_{ij})$ be the corresponding trajectory function representing its underlying true value, where $k = 1, \dots, q$, $j = 1, \dots, n_i$, and $i = 1, \dots, n$. The longitudinal model for $x_{ik}(t_{ij})$ is given by

$$x_{ik}(t_{ij}) = x_{ik}^*(t_{ij}) + \epsilon_{ijk}, \quad (3.2)$$

where ϵ_{ijk} represents the residual component consisting of both the psychological fluctuation and measurement error. Let $\mathbf{R}_{ik}(t_{ij})$ be the vector of covariates (possibly time-varying) for the k th QOL indicator at time t_{ij} . The trajectory function is modeled as

$$x_{ik}^*(t_{ij}) = \mathbf{R}_{ik}(t_{ij})\boldsymbol{\eta}_k + \mathbf{W}_{ik}(t_{ij})\mathbf{b}_{ik}, \quad (3.3)$$

where $\mathbf{W}_{ik}(t_{ij})$ is the random effects design matrix and may be a subset of the fixed effects design matrix $\mathbf{R}_{ik}(t_{ij})$, and $\boldsymbol{\eta}_k$ and \mathbf{b}_{ik} are vectors of the corresponding fixed and random effects parameters of length m_k and v_k , respectively. Let $\boldsymbol{\epsilon}_{ij} = [\epsilon_{ij1}, \dots, \epsilon_{ijq}]^T$. We assume that $\boldsymbol{\epsilon}_{ij} \stackrel{i.i.d.}{\sim} N_q(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathbf{b}_{ik} \stackrel{ind}{\sim} N_{v_k}(\mathbf{0}, \boldsymbol{\Psi}_k)$, and $\boldsymbol{\epsilon}_{ij}$ is independent of \mathbf{b}_{ik} . The residuals of the longitudinal QOL indicators observed at the same time may be correlated with each other due to psychological fluctuation, but are assumed to be independent among observations assessed at different time points. Due to the significant separation in time between observations, correlation (induced by psychological fluctuation) among residuals over time is assumed to be negligible. Measurement errors are assumed to occur independently over time and over different QOL indicators. After some algebraic derivations, the structure of $\boldsymbol{\Sigma}$ characterizes the association between QOL indicators measured at the same time, and is assumed to be common across time and patients. A common correlation structure among QOL indicators over time is psychologically and biologically plausible. For instance, a better mood may

relate to a better appetite, and the correlation is likely to be constant over time. On the other hand, the introduction of the random effects \mathbf{b}_{ik} in the trajectory function induces a common factor for repeated QOL measures of the i th patient at the k th indicator. The independence between \mathbf{b}_{ik} 's over different QOL indicators is tenable, since we assumed a distinct underlying mechanism in driving each QOL dimension. The structure of Ψ_k describes the association between repeated observations of the k th QOL indicator and is allowed to be distinct for different indicators. In addition, QOL indicators measured at different time points are set to be uncorrelated. An extensive set of simulations was conducted in Section 4 to assess the robustness of our joint model for these longitudinal assumptions, such as independence and normal distribution of the random effects b_{ik} 's.

Xu and Zeger (2001) proposed a joint model to evaluate multiple surrogate endpoints in a schizophrenia clinical trial. They assumed $\epsilon_{ijk} \sim N(0, \sigma_k^2)$ and $\mathbf{b}_i \sim N_v(\mathbf{0}, \Psi)$ independently, where $\mathbf{b}_i = [\mathbf{b}_{i1}^T, \dots, \mathbf{b}_{iq}^T]^T$ and $v = v_1 + \dots + v_q$. The joint distributional assumption on the random effects \mathbf{b}_i simultaneously models the two dependence structures. The specification of Ψ alone carries all of the information about dependence between the repeated measures, including correlations between markers measured at different time points. With Xu and Zeger's model, making separate inferences about the different dependence structures becomes less straightforward. The difference between the number of covariance parameters in their model and our model is $(1/2) \sum_{k \neq k'=1}^q v_k v_{k'} - q(q-1)/2$, which increases with the dimension of the random effects and the number of markers measured repeatedly over time. The two models are equivalent in the case of a univariate longitudinal process.

3.2. The time-to-event model

Joint models with a survival component incorporating a nonzero cure fraction have been established either by a two-component mixture (Law et al. (2002) and Yu et al. (2004)) or a non-mixture (Brown and Ibrahim (2003)) approach. The survival model we propose here is quite different from what has been presented in the literature and, particularly, does not impose a boundary value on the parameter space to require a nonzero cure fraction. Motivated by the promotion time model (Yakovlev and Tsodikov (1996)), we propose a novel generalization that allows for a zero as well as a nonzero cure fraction. We do this by specifying the population survival function as

$$S(y) = \exp \left\{ - \int_0^y \lambda(t) \tilde{F}(y-t) dt \right\}, \tag{3.4}$$

where $\lambda(t)$ is a non-negative function over time, and $\tilde{F}(t)$ is the distribution function of some non-negative random variable with $\tilde{F}(0) = 0$. Aside from the statistical properties of this survival function, described later in the section, this model

may be derived from biological considerations characterizing tumor growth. In the progression of cancer relapse, there exist clonogenic tumor cells (clonogens) that are capable of producing a detectable tumor mass. These clonogens may occur in clumps over time and, once they appear, are assumed to independently develop into a detectable tumor mass. If any of the clonogens occur and fully develop during the follow-up period, the patient would be observed to experience a relapse, otherwise would be considered as a censored case. The statistical link of this disease process to our survival function in (3.4) is through the assumptions that the number of clonogens over time, $N(t)$, follows a nonhomogeneous Poisson process with mean $\lambda(t)$, and the time of a clonogen to become a detectable tumor comes from the distribution function $\tilde{F}(t)$. This mechanism helps facilitate our estimation process and computational development. For instance, the introduction of N^* , the total number of clonogens, makes it straightforward for the specification of the joint likelihood in Section 3.3 and the construction of the MCMC algorithm.

Based on its definition, the cure fraction at (3.4) is given by

$$S(\infty) = \exp \left\{ - \lim_{y \rightarrow \infty} \int_0^y \lambda(t) \tilde{F}(y-t) dt \right\}. \quad (3.5)$$

When $\int_0^y \lambda(t) F(y-t) dt$ is bounded as $y \rightarrow \infty$, the survival function has a nonzero cure fraction, otherwise the survival function in (3.4) leads to a proper survival function (i.e. $S(\infty) = 0$). Using the properties of a distribution function $\tilde{F}(t)$ and the fact that $\lambda(t)$ is non-negative, we have

$$\tilde{F}\left(\frac{y}{2}\right) \int_0^{\frac{y}{2}} \lambda(t) dt \leq \int_0^y \lambda(t) \tilde{F}(y-t) dt \leq \tilde{F}(y) \int_0^y \lambda(t) dt.$$

Hence as $y \rightarrow \infty$, the population survival function in (3.4) reduces to

$$S(\infty) = \exp \left\{ - \int_0^\infty \lambda(t) dt \right\} = \exp \left\{ - \lim_{y \rightarrow \infty} \Lambda(y) \right\}, \quad (3.6)$$

where $\Lambda(y) = \int_0^y \lambda(t) dt$. In other words, a cure rate model is characterized by a bounded cumulative risk, $\Lambda(t)$, while a proper survival model is characterized by an unbounded cumulative risk. Specific demonstrations about how a cure fraction is governed by the longitudinal trajectories and their relationship with the survival time will be given later in the section. The hazard function of (3.4) is given by

$$h(y) = \int_0^y \lambda(t) \tilde{f}(y-t) dt, \quad (3.7)$$

where $\tilde{f}(y) = \frac{d}{dy}\tilde{F}(y)$. Equation (3.7) has the proportional hazards structure when the baseline covariates are modeled through $\lambda(t)$, and no time-varying covariates are considered. The proposed survival function in (3.4) can also be mathematically linked to the two-component mixture model. We can write (3.4) as

$$S(y) = \exp\{-\Lambda(y)\} + \{1 - \exp(-\Lambda(y))\}S_1(y), \tag{3.8}$$

where

$$S_1(y) = \frac{\exp\{-\int_0^y \lambda(t)\tilde{F}(y-t)dt\} - \exp\{-\Lambda(y)\}}{1 - \exp\{-\Lambda(y)\}}.$$

When the cure fraction is positive, $\exp\{-\Lambda(y)\}$ indicates the probability of cure, and $S_1(y)$ represents the survival function for the “non-cured” group. We note that $S_1^*(0) = 1$ and $S_1^*(\infty) = 0$, so that $S_1^*(y)$ is a proper survival function.

To incorporate information from both the longitudinal trajectories $x_k^*(t)$, $k = 1, \dots, q$, and baseline covariates, Z , in our survival model, we let all covariates depend on $\lambda(t)$ through the relationship

$$\lambda(t) = \exp \left\{ \sum_{k=1}^q \gamma_k x_k^*(t) + \mathbf{Z}\boldsymbol{\delta} \right\}. \tag{3.9}$$

Entering the covariates in this fashion corresponds to a canonical link in a Poisson generalized linear model, with $N(t)$ being the Poisson count in the disease process. All covariates may be assumed to affect survival biologically through their impact on the mean number of clonogens over time. Specifically, a negative regression coefficient leads to a smaller hazard, whereas a positive coefficient leads to a larger hazard, when the corresponding covariate takes a positive value. As mentioned earlier, the limiting behavior of $\Lambda(t)$ determines the property of our survival function, and hence the trajectory functions $x_k^*(t)$, and the relationship between the longitudinal trajectories and survival outcome, that is the γ_k 's, together account for the presence or absence of a cure fraction. For example, with linear trajectories $x_k^*(t) = \beta_{k0} + \beta_{k1}t$ for $k = 1, \dots, q$, the sign of $\sum_{k=1}^q \gamma_k \beta_{k1}$ determines the limiting behavior of $\Lambda(t)$ as well as the presence of a cure fraction: negative $\sum_{k=1}^q \gamma_k \beta_{k1}$ leads to a bounded $\Lambda(t)$ and results in a positive cure fraction, whereas a non-negative $\sum_{k=1}^q \gamma_k \beta_{k1}$ leads to an unbounded $\Lambda(t)$ and results in a zero cure fraction. For positive β_{k1} 's, when the coefficients of all the γ_k 's are negative, $\Lambda(t)$ is bounded and a positive cure fraction is obtained in the model, with smaller coefficients corresponding to larger cure fractions. On the other hand, when the coefficients of all the γ_k 's are positive, the model results in an unbounded $\Lambda(t)$ and a zero cure fraction. Furthermore, with quadratic

trajectories $x_k^*(t) = \beta_{k0} + \beta_{k1}t + \beta_{k2}t^2$ for $k = 1, \dots, q$, the sign of $\sum_{k=1}^q \gamma_k \beta_{k2}$ determines the limiting behavior of $\Lambda(t)$ and the presence of a cure fraction.

3.3 Joint likelihood and priors

In this subsection, we construct the joint likelihood with a specific choice of the trajectory functions and distribution function $\tilde{F}(t)$. This setup is later used in the analysis of the IBCSG data. Let $\mathbf{X} = \{x_{ik}(t_{ij}), i = 1, \dots, n, j = 1, \dots, n_i, k = 1, \dots, q\}$, $\mathbf{X}^* = \{x_{ik}^*(t_{ij}), i = 1, \dots, n, j = 1, \dots, n_i, k = 1, \dots, q\}$ and $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_n)$, where $x_{ik}(t_{ij})$ and $x_{ik}^*(t_{ij})$ are the respective observed and true k th QOL indicator for the i th patient at time t_{ij} , and \mathbf{R}_i is the corresponding vector of baseline covariates for longitudinal QOL indicators. For each QOL indicator, we consider a linear trajectory function over time as

$$x_{ik}^*(t_{ij}) = \mathbf{R}_i \boldsymbol{\eta}_k + b_{ik0} + b_{ik1}t_{ij} = \tau_{ik0} + \tau_{ik1}t_{ij},$$

where we hierarchically center (Gelfand, Sahu and Carlin (1996)) $\tau_{ik0} = \mathbf{R}_i \boldsymbol{\eta}_k + b_{ik0}$ and $\tau_{ik1} = b_{ik1}$ to facilitate convergence of the MCMC algorithms. Let $\boldsymbol{\tau}_0 = (\tau_{110}, \dots, \tau_{nq0})$ and $\boldsymbol{\tau}_1 = (\tau_{111}, \dots, \tau_{nq1})$. Further, let $\mathbf{Y} = (Y_1, \dots, Y_n)$, $\boldsymbol{\nu} = (\nu_1, \dots, \nu_n)$, $\mathbf{N}^* = (N_1^*, \dots, N_n^*)$ and $\mathbf{Z} = (Z_1, \dots, Z_n)$, where Y_i is the observed DFS for the i th patient, ν_i is the corresponding censoring indicator which equals to 1 if Y_i is a failure time and 0 if it is right censored, N_i^* indicates the total number of clonogens by the observed failure time for the i th patient as described in Section 3.2, and Z_i represents the corresponding vector of baseline covariates for DFS. The observed data is then given by $\mathbf{D}_{obs} = \{n, \tilde{\mathbf{n}}, \mathbf{X}, \mathbf{Y}, \boldsymbol{\nu}, \mathbf{R}, \mathbf{Z}\}$ and the complete data is given by $\mathbf{D} = \{n, \tilde{\mathbf{n}}, \mathbf{X}, \mathbf{Y}, \boldsymbol{\nu}, \mathbf{R}, \mathbf{Z}, \boldsymbol{\tau}_0, \boldsymbol{\tau}_1, \mathbf{N}^*\}$, where $\tilde{\mathbf{n}} = (n_1, \dots, n_n)$. As discussed in (3.1), the joint likelihood of the complete data requires the specification of three components, $P(\mathbf{X} | \mathbf{X}^*)$, $P(\mathbf{X}^* | \mathbf{R})$ and $P(\mathbf{Y} | \mathbf{X}^*, \mathbf{Z}, \mathbf{N}^*)$. The three components can be equivalently expressed as $P(\mathbf{X} | \mathbf{X}^*) = P(\mathbf{X} | \boldsymbol{\tau}_0, \boldsymbol{\tau}_1)$, $P(\mathbf{X}^* | \mathbf{R}) = P(\boldsymbol{\tau}_0, \boldsymbol{\tau}_1 | \mathbf{R})$ and $P(\mathbf{Y} | \mathbf{X}^*, \mathbf{Z}, \mathbf{N}^*) = P(\mathbf{Y} | \boldsymbol{\tau}_0, \boldsymbol{\tau}_1, \mathbf{Z}, \mathbf{N}^*)$, since the information carried by \mathbf{X}^* can be fully recovered by $\boldsymbol{\tau}_0$ and $\boldsymbol{\tau}_1$. The introduction of the latent vector \mathbf{N}^* in the third component can facilitate the development of the MCMC algorithm, and hence ease the estimation process.

After hierarchical centering, $P(\mathbf{X} | \boldsymbol{\tau}_0, \boldsymbol{\tau}_1)$ is given by

$$L_1(\boldsymbol{\theta}_1) \propto |\boldsymbol{\Sigma}|^{-\frac{M}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{n_i} (\mathbf{X}_{ij\cdot} - \boldsymbol{\tau}_{i\cdot 0} - \boldsymbol{\tau}_{i\cdot 1} t_{ij})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X}_{ij\cdot} - \boldsymbol{\tau}_{i\cdot 0} - \boldsymbol{\tau}_{i\cdot 1} t_{ij}) \right\}, \quad (3.10)$$

where $M = \sum_{i=1}^n n_i$, $\boldsymbol{\theta}_1 = \boldsymbol{\Sigma}$, $\mathbf{X}_{ij\cdot} = [x_{i1}(t_{ij}) \cdots x_{iq}(t_{ij})]^T$, $\boldsymbol{\tau}_{i\cdot 0} = [\tau_{i10} \cdots \tau_{iq0}]^T$ and $\boldsymbol{\tau}_{i\cdot 1} = [\tau_{i11} \cdots \tau_{iq1}]^T$. Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_q)$, $\boldsymbol{\Psi} = (\boldsymbol{\Psi}_1, \dots, \boldsymbol{\Psi}_q)$ and $\boldsymbol{\eta} =$

$(\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_q)$, where $\boldsymbol{\beta}_k = [\beta_{k0} \ \beta_{k1}]^T$. With the assumption $\mathbf{b}_{ik} = [b_{ik0} \ b_{ik1}]^T \sim N_2(\boldsymbol{\beta}_k, \boldsymbol{\Psi}_k)$, the second component, $P(\boldsymbol{\tau}_0, \boldsymbol{\tau}_1 \mid \mathbf{R})$, is given by

$$L_2(\boldsymbol{\theta}_2) \propto \left(\prod_{k=1}^q |\boldsymbol{\Psi}_k|^{-\frac{n}{2}} \right) \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \sum_{k=1}^q (\boldsymbol{\tau}_{ik} - \boldsymbol{\mu}_{ik})^T \boldsymbol{\Psi}_k^{-1} (\boldsymbol{\tau}_{ik} - \boldsymbol{\mu}_{ik}) \right\}, \quad (3.11)$$

where $\boldsymbol{\theta}_2 = \{\boldsymbol{\beta}, \boldsymbol{\Psi}, \boldsymbol{\eta}\}$, $\boldsymbol{\tau}_{ik} = [\tau_{ik0} \ \tau_{ik1}]^T$, and $\boldsymbol{\mu}_{ik} = [\beta_{k0} + \mathbf{R}_i \boldsymbol{\eta}_k \ \beta_{k1}]^T$. To complete the last piece of the joint likelihood, we further assume that $\tilde{F}(t) = 1 - e^{-\alpha t}$, and thus

$$\lambda_i(t) = \exp \left\{ \sum_{k=1}^q \gamma_k x_{ik}^*(t) + \mathbf{Z}_i \boldsymbol{\delta} \right\} = \exp \left\{ \sum_{k=1}^q \gamma_k \tau_{ik0} + \gamma_k \tau_{ik1} t + \mathbf{Z}_i \boldsymbol{\delta} \right\}.$$

The incorporation of patient-specific trajectory functions in the model of $\lambda_i(t)$ allows patients to have different cure rate structures. Given $\boldsymbol{\theta}_3 = \{\alpha, \boldsymbol{\gamma}, \boldsymbol{\delta}\}$ where $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)$, and assuming independent censoring, the third component of the complete data joint likelihood is given by

$$L_3(\boldsymbol{\theta}_3) = \left\{ \prod_{i=1}^n \tilde{S}_i(y_i)^{N_i^* - \nu_i} (N_i^* \tilde{f}_i(y_i))^{\nu_i} \right\} \times \exp \left\{ \sum_{i=1}^n N_i^* \log(\Lambda_i(y_i)) - \log(N_i^*) - \Lambda_i(y_i) \right\}, \quad (3.12)$$

where

$$\Lambda_i(y_i) = \frac{\exp(\sum_{k=1}^q \gamma_k \tau_{ik0} + \mathbf{Z}_i \boldsymbol{\delta})}{\sum_{k=1}^q \gamma_k \tau_{ik1}} \{ \exp(\sum_{k=1}^q \gamma_k \tau_{ik1} y_i) - 1 \}$$

$$\tilde{S}_i(y_i) = \frac{\sum_{k=1}^q \gamma_k \tau_{ik1}}{\alpha + \sum_{k=1}^q \gamma_k \tau_{ik1}} \times \frac{\exp(\sum_{k=1}^q \gamma_k \tau_{ik1} y_i) - \exp(-\alpha y_i)}{\exp(\sum_{k=1}^q \gamma_k \tau_{ik1} y_i) - 1},$$

and $\tilde{f}_i(y_i) = \alpha \tilde{S}_i(y_i)$. Let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$ denote the set of all the parameters. The joint likelihood of the complete data is given by

$$L(\boldsymbol{\theta} \mid \mathbf{D}) = L_1(\boldsymbol{\theta}_1) L_2(\boldsymbol{\theta}_2) L_3(\boldsymbol{\theta}_3). \quad (3.13)$$

The prior specifications for $\boldsymbol{\theta}$ are given as follows. We specify the joint prior as

$$\pi(\boldsymbol{\theta}) = \left\{ \prod_{k=1}^q \pi(\boldsymbol{\beta}_k) \pi(\boldsymbol{\eta}_k) \pi(\boldsymbol{\Psi}_k) \pi(\gamma_k) \right\} \pi(\boldsymbol{\Sigma}) \pi(\alpha) \pi(\boldsymbol{\delta}). \quad (3.14)$$

Priors for the $\boldsymbol{\beta}_k$'s, $\boldsymbol{\eta}_k$'s, $\boldsymbol{\Psi}_k$'s, γ_k 's, $\boldsymbol{\Sigma}$, α and $\boldsymbol{\delta}$ are assumed to be independent a priori. For $k = 1, \dots, q$, we take normal priors for $\boldsymbol{\beta}_k$ and $\boldsymbol{\eta}_k$, and Wishart priors

for Ψ_k^{-1} and Σ^{-1} . Priors for β_k , η_k , Ψ_k and Σ are motivated by their conjugacy. We take a normal prior for all the γ_k 's as well as the vector of survival regression coefficients δ . Finally, we specify an inverse gamma prior for α . In the analysis of the IBCSG data, only non-informative priors are considered.

Since we use conjugate priors for β_k , η_k , Ψ_k , $k = 1, \dots, q$, and Σ , the full conditionals of these parameters have a closed form and are thus easy to sample. The conditional posterior distribution for an element of the latent vector \mathbf{N}^* also has a closed form, and samples are easily obtained from the Poisson density. For parameters without closed form posteriors, the adaptive rejection algorithm, proposed by Gilks and Wild (1992), is used to get samples of δ from its log-concave conditional posterior. An extra Metropolis step is incorporated in the algorithm by Gilk, Best and Tan (1995), to obtain samples of γ , α and the latent τ , whose conditional posteriors are not log-concave.

4. Simulation Study

We first conducted a set of simulation studies to evaluate the performance and computational feasibility of our joint model. This simulation is based on two longitudinal markers monitored over time, and a survival event in the presence of cure. Using a total of $n = 800$ patients, each observed longitudinal marker $x_{ik}(t_{ij})$ was simulated as the sum of the trajectory function $x_{ik}^*(t_{ij})$ and the error term ϵ_{ijk} , for $i = 1, \dots, 800$, $k = 1, 2$ and $j = 1, \dots, 6$. The trajectory function was taken as $x_{ik}^*(t_{ij}) = \eta_{k0} + \eta_{k1}t_{ij} + \eta_{k2}r_i + b_{ik0} + b_{ik1}t_{ij}$, where the r_i 's were generated from a standard normal distribution to represent a baseline covariate in the longitudinal model. Let $\eta_k = [\eta_{k0} \ \eta_{k1} \ \eta_{k2}]^T$, $\mathbf{b}_{ik} = [b_{ik0} \ b_{ik1}]^T$ and $\epsilon_{ij.} = [\epsilon_{ij1} \ \epsilon_{ij2}]^T$. We took $\eta_1 = [1.0 \ 0.5 \ 1.0]^T$, $\eta_2 = [1.0 \ 1.0 \ 2.0]^T$,

$$\Psi_1 = \begin{bmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}, \quad \Psi_2 = \begin{bmatrix} 1.0 & -0.3 \\ -0.3 & 1.0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2.0 & 1.0 \\ 1.0 & 2.0 \end{bmatrix},$$

in which $\mathbf{b}_{ik} \sim N_2(\mathbf{0}, \Psi_k)$ and $\epsilon_{ij.} \sim N_2(\mathbf{0}, \Sigma)$. Out of a total of 4,800 longitudinal observations, 10% of them were randomly chosen to be missing. For the survival time, we took $\lambda_i(t)$ in (3.9) to be $\lambda_i(t) = \exp\{\gamma_1 x_{i1}^*(t) + \gamma_2 x_{i2}^*(t) + \delta z_i\}$, for $i = 1, \dots, 800$, where z_i is a binary baseline covariate with half of the patients having a 0 and the other half having a 1. We also chose $\gamma_1 = -0.3$, $\gamma_2 = -0.5$, and $\delta = 1.0$. The promotion time was assumed to be an exponential distribution with a common rate $\alpha = 1.0$. This set-up leads to a cure rate structure for the survival time.

We fit two different models to the simulated data. One was a joint model of longitudinal and survival outcomes (model I) and the other was not (model II). Model (I) is the model proposed in this paper that accounts for the association between the longitudinal and survival data, while model II is the model that

assumes longitudinal and survival data are independent. In other words, we fit model II by specifying separate models for the longitudinal and survival components, which is equivalent to assuming all the γ_k 's ($k = 1, 2$) are zero in (3.9). A total of 100 replications were conducted for this simulation, with 3,000 Gibbs samples after 500 burn-in for each replication. The convergence of each MCMC chain was checked by a combination of trace plots and autocorrelations. The program was run on a PC cluster configured with 352 CPUs, and the approximate execution time was 70 hours for 100 replications.

Both models were evaluated in terms of estimation of the longitudinal coefficients η_{12} and η_{22} , and the survival regression coefficient δ for the baseline covariate. The Deviance Information Criterion (DIC), proposed by Spiegelhalter, Best, Carlin and Van der Linde (2002) was also computed for each model as a Bayesian measure of fit and complexity for model selection. The smaller the DIC, the better the fit of the model. Table 1 summarizes the results of the parameter estimates (posterior means and standard deviations) and DIC statistics, averaged over replications. The proposed joint model (Model I) appears advantageous against the non-joint model (Model II), in terms of both the parameter estimates and DIC statistics. When there is a nonzero association between the longitudinal and survival data, ignoring this association would lead to biased estimates for important parameters, and thus result in a lack of fit for the data.

Table 1. Posterior summaries and DIC statistics in the simulation.

	η_{12}		η_{22}		δ_1		DIC
	Mean	SD	Mean	SD	Mean	SD	
True	1.000		2.000		1.000		
Model I	1.007	0.005	2.005	0.005	0.999	0.012	12632.2
Model II	1.019	0.007	2.012	0.005	0.002	0.011	14195.1

We then conducted a series of simulation studies to examine the robustness of our joint model. Several longitudinal assumptions, such as the independence and normality of the random effects \mathbf{b}_{ik} 's, were relaxed. We first generated $\mathbf{b}_i \sim N_4(\mathbf{0}, \Psi)$, where $\mathbf{b}_i = [\mathbf{b}_{i1}^T \mathbf{b}_{i2}^T]^T$,

$$\Psi = \begin{bmatrix} \Psi_1 & \rho \mathbf{I}_2 \\ \rho \mathbf{I}_2 & \Psi_2 \end{bmatrix},$$

and \mathbf{I}_2 is a 2×2 identity matrix. The proposed joint model was fit. When $\rho = 0.1$, the posterior means (with standard deviations in the parenthesis) for η_{12} , η_{22} and δ , averaged over 100 replications, are 0.994 (0.008), 2.004 (0.009), and 0.995 (0.017), respectively. As the correlation between \mathbf{b}_{i1} and \mathbf{b}_{i2} increases to $\rho = 0.2$, the posterior summaries for η_{12} , η_{22} and δ , averaged over 100 replications, are

1.007 (0.009), 2.000 (0.008), and 0.997 (0.018), respectively. Given the simulation setup for $\eta_{12} = 1.0$, $\eta_{22} = 2.0$ and $\delta = 1.0$, our joint model performs quite well in terms of the estimation of parameters of interest even with moderately correlated random effects. We then relaxed the normality assumption and simulated the random effects \mathbf{b}_i from a multivariate t distribution with mean zero, $a = 10$ degrees of freedom, and scale matrix $(a - 2)\Psi/a$ to ensure $Var(\mathbf{b}_i) = \Psi$. When $\rho = 0.1$, the posterior summaries for η_{12} , η_{22} and δ are 1.009 (0.010), 2.001 (0.008), and 0.966 (0.016), respectively. The proposed joint model is robust with respect to the independence and normality assumptions for all random effects.

5. Application to the IBCSG data

We applied the methodology to data from the breast cancer clinical trial. Before incorporating the patients' QOL information in the model of DFS, we first examined the performance of our proposed time-to-event model. Figure 1 shows two superimposed plots for time to breast cancer relapse. The solid line corresponds to the Kaplan-Meier estimate of survival, while the dashed line corresponds to the maximum likelihood estimate of the marginal survival function based on our model. No covariates were used in constructing the plots. We see that the two curves are nearly identical and appear to plateau after approximately 10 years of follow-up. In the proposed survival model, $\lambda(t)$ decreases over time, and thus suggests a positive cure fraction in the population.

We then applied our joint model to the data to investigate the relationship between QOL and DFS. To satisfy the normality assumption for each longitudinal QOL indicator, we transformed the observed QOL to $\sqrt{(100 - \text{QOL})}$ (see Hurny et al. (2000)). With this transformation, the transformed QOL decreased over time and was scaled between 0 and 10, with smaller values reflecting better QOL. For each transformed indicator, an individual linear trajectory was considered to account for patient-specific perception of QOL over time. After visually inspecting the individual changes of the transformed QOL over time, a linear trajectory function for each indicator appears to be a reasonable choice, especially based on a maximum of three observations over time. The same baseline covariates were incorporated (time after randomization, adjuvant therapy, age and residency) in the prediction of each QOL assessment. Residency is thought of as a proxy for "culture", and thus may affect the patients' perception of QOL. For the model of DFS, QOL measures are included as time-varying covariates, and adjuvant therapy, age, number of positive nodes and estrogen receptor (ER) status are included as baseline covariates. The derivation of the joint likelihood as well as the assumptions are described in Section 3.3. The selections of the baseline covariates in the longitudinal and survival model were essentially science-driven.

We use noninformative priors for all the parameters. With G standing for gamma and W standing for Wishart distribution, the prior specifications are $\beta_k \sim N_2(\mathbf{0}, 1000\mathbf{I}_2)$, $\eta_k \sim N_6(\mathbf{0}, 1000\mathbf{I}_6)$, $\Psi_k^{-1} \sim W_2(5, 1000\mathbf{I}_2)$, $\Sigma^{-1} \sim W_4(7, 1000\mathbf{I}_4)$, $\alpha^{-1} \sim G(0.01, 0.01)$, $\gamma_k \sim N(0, 1000)$ and $\delta \sim N_6(\mathbf{0}, 1000\mathbf{I}_6)$, for $k = 1, \dots, 4$. Four parallel MCMC chains of 25,000 iterations with overdispersed starting points and with 5,000 iterations as burn-in were run. We visually inspected these chains by overlaying their sampled values on a common graph for each parameter. Each parameter is annotated with the shrinkage factor proposed by Gelman and Rubin (1992). In order to eliminate autocorrelation among samples within a sequence, we selected every 10th iteration in each chain. The results are then presented from the combined chain with 10,000 iterations. All Highest Posterior Density (HPD) intervals were computed using a Monte Carlo method proposed by Chen and Shao (1999).

Table 2. Posterior means (with 95% HPD intervals in parentheses) for parameters in the longitudinal model.

	Appetite	Coping	Mood	Physical
Intercept	0.354 (0.324, 0.370)	0.535 (0.503, 0.550)	0.434 (0.405, 0.448)	0.403 (0.374, 0.417)
Time (in year)	-0.044 (-0.060, -0.037)	-0.083 (-0.097, -0.076)	-0.048 (-0.063, -0.040)	-0.012 (-0.025, -0.006)
# Initial cycle	-0.005 (-0.033, 0.009)	0.030 (0.000, 0.044)	0.037 (0.010, 0.050)	0.013 (-0.013, 0.025)
Reintroduction	-0.013 (-0.041, 0.001)	0.007 (-0.023, 0.021)	0.008 (-0.020, 0.021)	-0.021 (-0.047, -0.009)
# Initial cycle × Reintroduction	0.006 (-0.034, 0.025)	-0.020 (-0.061, 0.000)	-0.028 (-0.065, -0.001)	0.010 (-0.025, 0.027)
AGE > 40	0.042 (0.017, 0.054)	0.022 (-0.004, 0.033)	0.046 (0.023, 0.057)	0.054 (0.031, 0.065)
Residency: Swiss	0.010 (-0.012, 0.021)	0.037 (0.013, 0.049)	0.029 (0.007, 0.039)	0.005 (-0.016, 0.014)
Residency: Sweden	0.0167 (-0.008, 0.029)	0.046 (0.020, 0.058)	0.094 (0.070, 0.105)	0.059 (0.036, 0.070)

Table 2 displays posterior summary statistics for parameters related to the QOL assessments. All transformed indicators for health-related quality of life decrease over time with HPD intervals excluding 0, suggesting improvement of QOL after initial surgery. A longer duration of the initial therapy adversely affects a patient’s mood and coping score, but the reintroduction of a delayed therapy helps improve a patient’s physical well-being. Younger patients (under 40) have a better quality of life than older patients, and the differences reach

significance for assessments of appetite, mood and physical well-being. Residency also plays a role in influencing a patient's perception of QOL. In general, patients living in Australia and New Zealand have a better quality of life than patients living in Switzerland or Sweden. Finally, given the positive estimates for all off-diagonal elements of Σ (not displayed in Table 2), there is evidence that all QOL indicators are positively correlated to each other. Better physical well-being, for instance, leads to a better appetite, mood and coping.

Table 3. Posterior summaries for parameters in DFS model.

	Mean	HPD interval
Appetite	1.163	(-0.227, 1.825)
Coping score	0.833	(0.044, 1.220)
Mood	0.114	(-0.456, 1.298)
Physical well-being	1.087	(0.431, 2.501)
# Initial cycle	-0.141	(-0.427, -0.002)
Reintroduction	-0.528	(-0.816, -0.390)
# Initial cycle \times Reintroduction	0.194	(-0.233, 0.397)
AGE > 40	-0.503	(-0.759, -0.378)
# positive nodes > 4	0.930	(0.725, 1.027)
ER (1=Positive)	-0.423	(-0.637, -0.321)

Table 3 summarizes posterior distributions for parameters related to DFS. With positive regression coefficients of the transformed QOL measures, patients having a better quality of life are less likely to have cancer relapse, and the effects reach significance for coping and physical well-being. All posterior distributions of the γ_k 's appear quite symmetric with positive modes. Increased duration of initial adjuvant chemotherapy and reintroduction of a single course of delayed chemotherapy are able to delay time to cancer relapse with 95% HPD intervals excluding zero. Younger patients are more likely to have a relapse than older patients. Different mechanisms of cancer progression may work for patients under 40 and patients over 40. A positive ER status may reduce the risk of cancer relapse; however, a greater number of positive nodes may increase the risk of relapse. Both the effect of ER status and the number of positive nodes are important effects as indicated by the exclusion of zero in their 95% HPD intervals.

The positive regression coefficients of the transformed QOL indicators (which decrease over time) in the model of DFS imply a bounded cumulative hazard, and suggest a nonzero cure fraction in the population. We note that when covariates are included in the model, each patient has an individual cure fraction. Figure 3 shows the distribution of posterior means of cure rate fractions for all patients. The mean cure fraction for time to relapse is 0.316 and the standard deviation

is 0.134. On average, 31.6% of patients are cured and thus are not susceptible to cancer relapse after surgery.

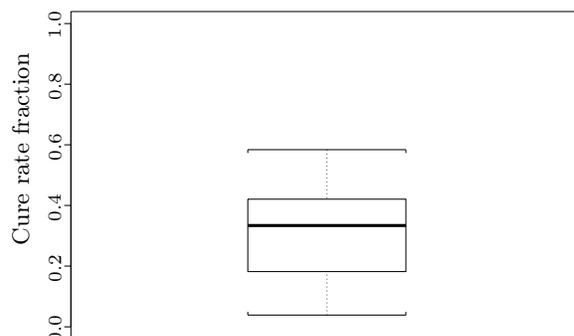


Figure 3. Boxplot of the posterior means of the cure rates for all patients.

6. Discussion

With more studies being conducted that repeatedly take measures over time in an effort to evaluate a patient's health or risk to some event, a joint modeling approach is essential. We have presented a latent process model for multivariate repeated measures and a flexible survival model to incorporate the characteristics of the longitudinal model. Our longitudinal model explicitly acknowledges two sources of dependence among multidimensional repeated measures and allows direct inferences on both association over time and over different markers. A parametric survival model is proposed and is able to accommodate both a zero and nonzero cure fraction in the population. Aside from the biological motivation, our time-to-event model is suitable for any type of event time data as long as the data can be thought of as being generated by a process of latent potential risks. Thus the model can be useful for analyzing various types of survival data, including time to relapse, time to death, time to first infection, and so forth. Future work with this model includes developing methods for hypothesis testing, model selection and model adequacy assessment.

References

- Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *J. Amer. Statist. Assoc.* **47**, 501-515.
- Brown E. R. and Ibrahim, J. G. (2003). Bayesian approaches to joint cure-rate and longitudinal models with applications to cancer vaccine trials. *Biometrics* **59**, 686-693.
- Chen, M. H. and Shao, Q. M. (1999). Monte carlo estimation of Bayesian credible and HPD intervals. *J. Comput. Graph. Statist.* **8**, 69-92.

- Gelfand, A. E., Sahu, S. K. and Carlin, B. P. (1996). Efficient parametrizations for generalized linear mixed models. *Bayesian Statist.* **5**, 165-180.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7**, 457-551.
- Gilks, W. R., Best, N. G. and Tan, K. C. C. (1995). Adaptive rejection Metropolis sampling within Gibbs sampling. *Appl. Statist.* **44**, 455-472.
- Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Appl. Statist.* **41**, 337-348.
- Hurny, C., Bernhard, J., Coates, A. S., Castiglione M., Peterson, H. F., Gelber, R. D., Forbes, J. F., Rudenstam, C. M., Simoncini, E., Crivellari, D., Goldhirsch, A. and Senn, H. J. (2000). Impact of adjuvant therapy on quality of life in women with node-positive operable breast cancer. *Lancet* **347**, 1279-1284.
- Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer-Verlag, New York.
- Ibrahim, J. G., Chen, M. H. and Sinha, D. (2004). Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine studies. *Statist. Sinica* **14**, 863-883.
- Law, N. J., Taylor, J. M. G. and Sandler, H. M. (2002). The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure. *Biostatistics* **3**, 547-63.
- Song, X., Davidian, M. and Tsiatis, A. A. (2002). An estimator for the proportional hazards model with multiple longitudinal covariates measured with errors. *Biostatistics* **3**, 511-28.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc. Ser. B* **64**, 583-616.
- The International Breast Cancer Study Group (1996). Duration and reintroduction of adjuvant chemotherapy for node-positive premenopausal breast cancer patients. *J. Clinical Oncology* **14**, 1885-1894.
- Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statist. Sinica* **14**, 793-818.
- Xu, J. and Zeger, S. L. (2001). The evaluation of multiple surrogate endpoints. *Biometrics* **57**, 81-87.
- Yakovlev, A. Y. and Tsodikov, A. D. (1996). *Stochastic Models for Tumor Latency and their Biostatistical Applications*. World Scientific, New Jersey.
- Yu, M., Law, N. J., Taylor, J. M. G. and Sandler, H. M. (2004). Joint longitudinal-survival-cure models and their applications to prostate cancer. *Statist. Sinica* **14**, 835-862.

Department of Biostatistics, University of Washington, Seattle, WA 98195, U.S.A.

E-mail: yychi@u.washington.edu

Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 27599, U.S.A.

E-mail: ibrahim@bios.unc.edu

(Received June 2005; accepted November 2005)