Neuhaus, Augustin, Heumann, Daumer:

A Review on Joint Models in Biometrical Research

Sonderforschungsbereich 386, Paper 506 (2006)

Online unter: http://epub.ub.uni-muenchen.de/

University of Munich
Discussion Paper 506 - SFB 386

A Review on Joint Models in Biometrical Research

A. Neuhaus[†1], T. Augustin[‡], C. Heumann[‡], M. Daumer[†]

[†] Sylvia Lawry Centre for MS Research, Hohenlindenerstr. 1, D-81677 Munich
[‡] Department of Statistics, Ludwigstr. 33, D-80539 Munich

Abstract

In some fields of biometrical research joint modelling of longitudinal measures and event time data has become very popular. This article reviews the work in that area of recent fruitful research by classifying approaches on joint models in three categories: approaches with focus on serial trends, approaches with focus on event time data and approaches with equal focus on both outcomes. Typically longitudinal measures and event time data are modelled jointly by introducing shared random effects or by considering conditional distributions together with marginal distributions. We present the approaches in an uniform nomenclature, comment on sub-models applied to longitudinal measures and event time data outcomes individually and exemplify applications in biometrical research.

Key words: Joint model, shared random effects, mixed effects model, survival analysis.

1 Introduction

Clinical trials and epidemiological studies often collect more than one outcome for each subject. In addition to the outcome for which the study was primarily initiated, secondary and tertiary outcomes are collected during an investigation. These data are often time to event data or repeated measurements. Various approaches have been described in the past decades to handle repeated measurements and survival data separately, but in the situation that both outcomes were selected on one subject, classical modelling does not consider dependencies between the two types of responses. A powerful method to overcome this problem is a joint modelling of survival and repeated measurements. Well known examples in which repeated measurements and event time data are generated are studies in the field of the acquired immunodeficiency syndrome (e.g. Tsiatis et al., 1995). In these studies disease

[1] neuhaus@slcmsr.org

markers (e.g. viral load or CD4 counts) are measured repeatedly and disease specific events (e.g. seroconversion or death) are documented at the same time. A joint examination of the longitudinal process of such disease markers and the time to event is possible using joint model approaches. This example also indicates that the issue of surrogate markers is a natural area for the application of joint models, since the course of the longitudinal disease marker process might serve as surrogate for the event.

This article classifies approaches on joint modelling, spread around the literature, in the categories 'focus on serial trend', 'focus on event times' and 'equal focus on both outcomes'. We concentrate on methodological aspects of joint model approaches, details on forming estimates are only noted marginally.

Throughout the paper, we assume that k subjects are observed, each with possibly different visit schedules, i.e. at different time points, $t_{i1} := 0, t_{i2}, \ldots, t_{in_i}$. Thus altogether n_i individual observations are collected for the ith subject ($i = 1, \ldots, k$). In addition to the outcome measured longitudinally, the time to a specific event is recorded.

We use the following notation for the ith subject to harmonise diverse approaches:

$\boldsymbol{y}_i = (y_{i1}, y_{i2}, ..., y_{in_i})'$	$(n_i \times 1)$ vector containing longitudinal observations
$\boldsymbol{t}_i = (t_{i1}, t_{i2}, ..., t_{in_i})'$	$(n_i \times 1)$ vector with corresponding observation time points, with $t_{i1} := 0$
$\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, ..., \epsilon_{in_i})'$	$(n_i \times 1)$ vector of errors independent from $\boldsymbol{y}_i$
$\boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}'_{i1} \\ \vdots \\ \boldsymbol{x}'_{in_i} \end{pmatrix}$	$(n_i \times p)$ matrix of possibly time-varying covariates with $\boldsymbol{X}_i[t]$ the corresponding step function which is $\boldsymbol{x}_{ij}$ if $t_{ij} \leq t < t_{i(j+1)}$ for $j = 1, \ldots, n_i - 1$ and $\boldsymbol{x}_{in_i}$ if $t_{in_i} \leq t$
$\boldsymbol{\beta}$	$(p \times 1)$ fixed effects corresponding to $\boldsymbol{X} = (\boldsymbol{X}'_1, \boldsymbol{X}'_2, ..., \boldsymbol{X}'_k)'$
$\boldsymbol{Z}_i$	$(n_i \times q)$ matrix of covariates with $\boldsymbol{Z}_i[t]$ defined in the same way as $\boldsymbol{X}_i[t]$
$\boldsymbol{b}_i$	$(q \times 1)$ random effects corresponding to $\boldsymbol{Z}_i$
τ_i	event time of survival outcome
c_i	censoring time
τ_i^*	$\min(\tau_i, c_i)$
δ_i	censoring indicator with $\delta_i = I(\tau_i \leq c_i)$

Ideally, the complete longitudinal process $\boldsymbol{y}_i$ is known and measurement times are non-informative. The latter means that the t_{ij}s are not affected by the trend or values of $\boldsymbol{y}_i$. Therefore the $\boldsymbol{t}_i$s might differ in length and schedules for different subjects. If the $\boldsymbol{t}_i$s are identical for all subjects panel data are available and corresponding models can be applied.

Two approaches are available to arrive at a joint distribution for repeated measures $\boldsymbol{y}_i$ and survival outcome τ_i: (1) the introduction of *shared random effects* and (2) the use of *mixture and selection models*. In the first approach random effects $\boldsymbol{b}_i$ are used to connect $\boldsymbol{y}_i$ and τ_i.

Conditioning on these random effects provides assumed independence of $\boldsymbol{y}_i$ and τ_i. That is

$$f(\boldsymbol{y}_i, \tau_i|\boldsymbol{b}_i) = f(\boldsymbol{y}_i|\boldsymbol{b}_i)f(\tau_i|\boldsymbol{b}_i). \tag{1}$$

Assuming a certain distribution $f(\boldsymbol{b}_i)$ for the random effects gives the joint distribution $f(\boldsymbol{y}, \boldsymbol{\tau})$ for k independent subjects:

$$f(\boldsymbol{y}, \boldsymbol{\tau}) = \prod_{i=1}^{k} \int f(\boldsymbol{y}_i, \tau_i|\boldsymbol{b}_i)f(\boldsymbol{b}_i)\, d\boldsymbol{b}_i = \prod_{i=1}^{k} \int f(\boldsymbol{y}_i|\boldsymbol{b}_i)f(\tau_i|\boldsymbol{b}_i)f(\boldsymbol{b}_i)\, d\boldsymbol{b}_i. \tag{2}$$

Unless explicitly stated we rely on the common assumption $\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_b)$.
The second approach to arrive at the joint distribution is based on a factorisation of $\boldsymbol{y}$ and $\boldsymbol{\tau}$ using conditional and marginal distributions. That is

$$f(\boldsymbol{y}, \boldsymbol{\tau}) = f(\boldsymbol{y}|\boldsymbol{\tau})f(\boldsymbol{\tau}) \qquad \text{or} \tag{3}$$

$$f(\boldsymbol{y}, \boldsymbol{\tau}) = f(\boldsymbol{\tau}|\boldsymbol{y})f(\boldsymbol{y}). \tag{4}$$

These models are known as mixture and selection models (Little, 1993).
In both approaches joint distributions are constructed on basis of sub-models for the longitudinal process and the survival outcome. A variety of models can be fitted to both outcomes. Longitudinal measurements are easiest described by a linear mixed effects model $\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$. Possible extensions allow for more complex relationships, for example polynomial specifications of $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ or any functions $f(\boldsymbol{X}_i, \boldsymbol{\beta})$ and $f(\boldsymbol{Z}_i, \boldsymbol{b}_i)$.
In general survival models are constructed within the class of multiplicative hazards models and are built without random effects. The hazard rate λ_i, conditioned on covariates at time t, has the form:

$$\lambda_i(t|\boldsymbol{X}_i[t]) = \lambda_0(t)c(\boldsymbol{X}_i[t]'\boldsymbol{\beta}). \tag{5}$$

Mostly known representatives are the Cox proportional hazards model and the Weibull model. In both models $c(.)$ is specified as $\exp(.)$. Whereas the baseline hazard rate $\lambda_0(t)$ is left completely unspecified in the Cox model it is taken as $\alpha\mu t^{\alpha-1}$ in the Weibull model (e.g Klein & Moeschberger, 2003). The latter model is also a representative of another model class used for specification of the survival outcome within joint models, the accelerated failure time (AFT) models in which covariates are assumed to have linear influence on $\log(\tau_i)$, that is $\log(\tau_i) = \boldsymbol{X}_i[t]'\boldsymbol{\beta} + e_i$, with e_i the error. The distribution of the error specifies the model (e.g. extreme value distribution which leads to the Weibull regression model). The conditional hazard rate of an AFT model has the form

$$\lambda_i(t|\boldsymbol{X}_i[t]) = \lambda_0(t\exp(\boldsymbol{X}_i[t]'\boldsymbol{\beta}))\exp(\boldsymbol{X}_i[t]'\boldsymbol{\beta}) \tag{6}$$

where λ_0 is specified by the error e as mentioned above. Extensions to semiparametric approaches are possible by leaving λ_0 unspecified (Lin & Zhiliang, 1995).

Subsequent sections of this paper are organised as follows: In Section 2, we describe joint model approaches with focus on serial trends in which the pattern of repeated measurements given a survival outcome are of main concern. Models with focus on event times are described in Section 3. These models specify how the longitudinal measurements affect survival outcomes. Section 4 reviews approaches that jointly focus on serial trends and event times. Within these approaches covariate effects on longitudinal measurements and survival outcome are jointly estimated. In Section 5, we give examples in which joint models have been applied in practice. A brief look at further approaches and extensions is presented in Section 6.

2 Joint models with focus on serial trends

Models with primary focus on serial trends are applied when the description of repeated measurements is of main concern. Informative dropouts or events are considered within these models to avoid biased estimates for the longitudinal process. We describe two approaches, one based on a shared random effects model and the other one is based on a mixture model, to handle such data.

The approach introduced by Vonesh et al. (2006) is based on shared random effects; the joint density is factorised as in (2). Within this factorisation the class of generalised non-linear mixed-effects models is assumed for the sub-model $\boldsymbol{y}_i|\boldsymbol{b}_i$.

The conditional distribution of $\tau_i|\boldsymbol{b}_i$ is modelled using multiplicative hazards models that include subject-specific intercept and time trends via a function $g_i(.)$ that may also depend on fixed effects,

$$\lambda_i(t|\boldsymbol{b}_i) = \lambda_0(t)\exp[g_i(\boldsymbol{\beta}, \boldsymbol{b}_i, \boldsymbol{X}_i[t], \boldsymbol{Z}_i[t])]. \tag{7}$$

Vonesh et al. specify $\lambda_0(t)$ in two ways, similar to a Weibull or a piecewise exponential model. In the latter case the time scale is partitioned in disjoint exogenously given intervals. The baseline hazard rate is assumed to be constant within each interval but may vary from interval to interval. That is for p disjoint intervals $\lambda_0(t) = \sum_{h=1}^{p} \lambda_{0h} I(t \in (t_{h-1}, t_h])$.

Replacing $g_i(.)$ by $g_{ih}(.)$ in (7) additionally allows the inclusion of time-dependent covariates in the model.

The estimation of the unknown parameters is done via the likelihood

$$L = \prod_{i=1}^{k} \int f(\boldsymbol{y}_i|\boldsymbol{b}_i) f(\tau_i^*|\boldsymbol{b}_i)^{\delta_i} S_i(\tau_i^*|\boldsymbol{b}_i)^{1-\delta_i} f(\boldsymbol{b}_i)\, d\boldsymbol{b}_i$$

where $S(\tau_i^*|\boldsymbol{b}_i)$ is the conditional survivor function.

Since the integral has no closed form solution, numerical integration or alternatively the Laplace approximation is needed.

Hogan & Laird (1997) assume a mixture model for longitudinal measurements given time to event. The joint density function for the mutually independent pairs $(\boldsymbol{y}_i, \tau_i)$ is obtained from the factorisation given by (3). Thereby no parametric form is assumed for the cumulative function $F(\boldsymbol{\tau})$, the Kaplan-Meier product-limit estimator replaces $F(\boldsymbol{\tau})$, which implicitly means that survival times are homogenous.
$f(\boldsymbol{y}|\boldsymbol{\tau})$ is described in two stages and the corresponding model contains fixed effects, random effects and information on the event time. To implement these components we assume that the first and second column of $\boldsymbol{Z}_i$ contains 1 and t_i respectively. This allows the construction of a subject-specific intercept and time trend. Furthermore, a $(q \times p)$ matrix $\boldsymbol{W}_i(\tau_i)$ is introduced to cover information on the event time.
We choose the following structure for $\boldsymbol{W}_i(\tau_i)$ so that $\boldsymbol{X}_i = \boldsymbol{Z}_i\boldsymbol{W}_i(\tau_i)$: let m be the number of covariate effects that are modelled as both random and fixed effects. Thus

$$\boldsymbol{W}_i(\tau_i) = (\boldsymbol{E}_m, \tilde{\boldsymbol{\tau}}_i, \tilde{\boldsymbol{X}}_i)$$

with $\boldsymbol{E}_m$ a $(q \times m)$ matrix containing canonical unit vectors $\boldsymbol{e}_l$, $l \in \{1, 2, ..., q\}$, $\tilde{\boldsymbol{\tau}}_i = (g(\tau_i), 0, ..., 0)'$ with $g(.)$ allowing for transformations of τ_i and $\tilde{\boldsymbol{X}}_i$ a $(q \times (p-m-1))$ matrix containing covariates and interaction terms corresponding to fixed effects only. Thereby the jth column is $(\tilde{x}_{ij1}, 0, ..., 0)$ in constant covariates and $(0, \tilde{x}_{ij2}, 0, ..., 0)$ in covariates with time-interactions.
Let $\boldsymbol{y}_{i\tau_i}$ denote the longitudinal observations for a known τ_i. In the first step a linear mixed effects model with subject-specific random intercept and time trends ($\boldsymbol{\alpha}_{i\tau_i} = (\alpha_{1i\tau_i}, ..., \alpha_{qi\tau_i})'$) is constructed for $\boldsymbol{y}_{i\tau_i}$. In the second step the $\boldsymbol{\alpha}_{i\tau_i}$s are described via $\boldsymbol{W}_i(\tau_i)$. That is

$$\begin{aligned} (I) \quad & \boldsymbol{y}_{i\tau_i} = \boldsymbol{Z}_i\boldsymbol{\alpha}_{i\tau_i} + \boldsymbol{\epsilon}_i \\ (II) \quad & \boldsymbol{\alpha}_{i\tau_i} = \boldsymbol{W}_i(\tau_i)\boldsymbol{\beta} + \boldsymbol{b}_i. \end{aligned}$$

The combination of (I) and (II) can be formulated as a standard linear mixed effects model that specifies $f(\boldsymbol{y}|\boldsymbol{\tau})$

$$\boldsymbol{y}_{i\tau_i} = \boldsymbol{Z}_i\boldsymbol{\alpha}_{i\tau_i} + \boldsymbol{\epsilon}_i \; = \boldsymbol{Z}_i(\boldsymbol{W}_i(\tau_i)\boldsymbol{\beta} + \boldsymbol{b}_i) + \boldsymbol{\epsilon}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i. \tag{8}$$

The EM algorithm is used to obtain ML estimates based on $f(\boldsymbol{y}, \boldsymbol{\tau})$ (see Hogan & Laird, 1997).

3 Joint models with focus on event times

The aim of joint models with focus on event time data is the description on how the longitudinal measurements affect the survival outcome. Thereby the longitudinal process is mostly considered as time-dependent covariate within a survival model which is specified as Cox proportional hazards model or accelerated failure time model.

Tsiatis et al. (1995) assume the $\boldsymbol{y}_i$ to be measured with error and use a two-stage approach in which information on the longitudinal process is incorporated in a Cox proportional hazards model. First, the $\boldsymbol{y}_i$ are modelled with a linear mixed effects model. The resulting model is used to estimate the status of the longitudinal process at time t. In the second stage, these so-called empirical Bayes estimates replace the observations measured with error and serve as time-dependent covariate in the Cox model.

Wulfson & Tsiatis (1997) improved this approach by modelling the process of the time-dependent covariate and the survival data simultaneously, assuming linear growth for $\boldsymbol{y}_i$ which might be measured with an error. Using the assumption $b_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_b)$, y_{ij} can be written as

$$y_{ij} = \boldsymbol{X}_i[t_{ij}]'\boldsymbol{\beta} + b_{0i} + b_{1i}t_{ij} + \epsilon_{ij} \tag{9}$$

with mutually independent error $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$. Furthermore ϵ_{ij} is taken as independent of the random intercept b_{0i} and slope b_{1i}. The Cox proportional hazards model with a covariate described by the random effects model given above is written as

$$\lambda(t|\boldsymbol{b}_i, \boldsymbol{y}_i) = \lambda(t|\boldsymbol{b}_i) = \lambda_0(t)\exp\{(\boldsymbol{X}_i[t]'\boldsymbol{\beta} + b_{0i} + b_{1i}t)\gamma\}.$$

Estimates for this semiparametric approach are obtained by the EM algorithm (Wulfson & Tsiatis, 1997). Hsieh et al. (2006) made a simulation study to examine the robustness and efficiency of these estimates. They concluded that as long as there is sufficient information on the longitudinal outcome, the estimates are both, robust and efficient.

Tsiatis & Davidian (2001) and Song et al. (2002b) generalised this model by setting the normality assumption of the random effects aside. A further generalisation to multiple time-dependent covariates was introduced by Song et al. (2002a). Dang et al. (2007) give a full likelihood approach in which a bivariate growth curve from two longitudinal measures of an individuum is used to predict recurrences of an event with a Cox model.

Another approach introduced by Tseng et al. (2005) is based on an accelerated failure time model combined with a linear mixed effects model for the longitudinal observations. The advantage of the class of accelerated failure time models is that they are applicable to situations in which the proportional hazards assumption on the event time data fails. The longitudinal process $y_i(t)$ is considered as a time-varying covariate and replaces $\boldsymbol{X}_i[t]$ in (6). Tseng et al. (2005) assume a step function, but no parametric form for $\lambda_0(.)$, and takes $\lambda_0(.)$ as constant between two consecutive event times. The restriction to a step function is necessary to enable the application of the EM-algorithm.

Assuming noninformative censoring and a measurement schedule t_{ij}, that is independent of the random effects and covariate history, the joint likelihood $L(\boldsymbol{\theta}) = L(\boldsymbol{\beta}, \boldsymbol{b}, \boldsymbol{\Sigma}_b, \sigma_\epsilon^2, \lambda_0)$ for event time data and longitudinal measurements (as described in (9)) is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{k}\left[\int\left\{\prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{t}_i, \sigma_\epsilon^2)\right\} f(\tau_i^*, \delta_i|\boldsymbol{b}_i, \boldsymbol{t}_i, \lambda_0, \boldsymbol{\beta}) f(\boldsymbol{b}_i|\boldsymbol{\Sigma}_b)\, d\boldsymbol{b}_i\right]$$

where $f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{t}_i, \sigma_\epsilon^2)$ and $f(\boldsymbol{b}_i|\boldsymbol{\Sigma}_b)$ are the densities of y_{ij} and $N(\mathbf{0}, \boldsymbol{\Sigma}_b)$ respectively and

$$f(\tau_i^*, \delta_i|\boldsymbol{b}_i, \boldsymbol{t}_i, \lambda_0, \beta) = \{\lambda_0(\tilde{\tau}_i^*(\beta, y_i)) \exp[y_i(\tau_i^*)\beta]\}^{\delta_i} \exp\left\{-\int_0^{\tilde{\tau}_i^*(\beta,y_i)} \lambda_0(t)\, dt\right\}.$$

with $\tilde{\tau}_i^*(\beta, y_i) = \int_0^{\tau_i^*} \exp[y_i(s)\beta]\, ds$. The specific steps of the EM-algorithm to arrive at the parameter estimates are described in detail in Tseng et al. (2005).

4 Joint models with equal focus on both outcomes

The models described in the previous sections consider one of the jointly measured outcomes as primary and explain this with the other outcome and covariates. In the following we describe models that quantify covariates effects on longitudinal measurements and survival outcome jointly. For solving this task different approaches have been proposed, for instance a common latent stochastic process that is shared by both outcomes (Henderson et al., 2000), shared random effects that affect outcomes differently (Zeng & Cai, 2005), and joint models within the framework of hierarchical generalised linear models (Ha et al., 2003):

Henderson et al. (2000) introduce a model class for joint modelling of repeated measurements and event time data, including recurrent events. The idea of this model is similar to shared random effects models. Here, the random effect turns into a dynamic bivariate Gaussian process $\boldsymbol{B}(t) = \{B^{(1)}(t), B^{(2)}(t)\}$. It is assumed that the longitudinal observations and the event process are conditionally independent given $\boldsymbol{B}(t)$ and covariates. The association between both outcomes is described by the cross-correlation between $B^{(1)}(t)$ and $B^{(2)}(t)$. Thus, the joint distribution of both outcomes for a subject i is modelled via a latent zero-mean Gaussian process $\boldsymbol{B}_i(t) = \{B_i^{(1)}(t), B_i^{(2)}(t)\}$ that is realised for each i independently. The sub-model for the repeated measurements of a subject i is assumed to be of the form

$$\boldsymbol{y}_i = \boldsymbol{X}_i^{(1)}\boldsymbol{\beta}^{(1)} + \boldsymbol{B}_i^{(1)}(\boldsymbol{t}_i) + \boldsymbol{\epsilon}_i$$

with $\boldsymbol{X}_i^{(1)}$ (time-varying) covariates and $\boldsymbol{\beta}^{(1)}$ the corresponding coefficients associated to the repeated measurements. The number of events is specified by a counting process $N_i(t)$ that allows the modelling of recurrent events. The semi-parametric model for the event process and corresponding hazard function is

$$\lambda_i(t) = H_i(t)\lambda_0(t) \exp\{\boldsymbol{X}_i^{(2)}[t]'\boldsymbol{\beta}^{(2)} + B_i^{(2)}(t)\}.$$

Thereby, $H_i(t)$ is the zero-one process that indicates whether individuum i is at risk or not. $\lambda_0(t)$ is the baseline hazard with unspecified form.
Henderson et al. propose $B^{(1)}(t)$ and $B^{(2)}(t)$ to be of the form

$$B^{(1)}(t) = b_1 + b_2 t \qquad \text{and} \qquad B^{(2)}(t) = \gamma_1 b_1 + \gamma_2 b_2 + \gamma_3 B^{(1)}(t) + b_3$$

with (b_1, b_2) bivariate normal with mean zero and b_3 independent of (b_1, b_2) normally distributed with mean zero. The shape of $B^{(2)}$ is an extension of the usual proportionality assumption between $B^{(1)}$ and $B^{(2)}$.
A general expression of the components $B_i^{(l)}(t), l = 1, 2$ of the bivariate Gaussian process is

$$B_i^{(l)}(t) = \boldsymbol{Z}_i^{(l)}[t]'\boldsymbol{b}_{li} + V_i^{(l)}(t)$$

where $V_i^{(l)}(t)$ is a stationary Gaussian process with mean 0, variance σ_{vl}^2 and correlation function

$$r_l(u) = cov\{V_i^{(l)}(t), V_i^{(l)}(t-u)\}/\sigma_{vl}^2.$$

Parameters of for the longitudinal and the counting process can be estimated using the EM-algorithm.

Similar to Henderson et al. (2000), Zeng & Cai (2005) introduce shared random effects to obtain a joint model. Random effects are designed in a way that they affect both outcomes differently. The longitudinal measurements are described by a linear mixed effects model $\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and the survival outcome by a multiplicative hazards model that contains random effects which partly affect both outcomes and partly the survival outcome only. That is $\tilde{\boldsymbol{b}_i} = \phi\boldsymbol{b}_i + \boldsymbol{c}_i$ with $\boldsymbol{c}_i$ the subject-specific effect which only affects the survival outcome and ϕ a scale parameter corresponding to the random effect that affects both outcomes, but with different intensities. The hazard rate for the survival outcome is $\lambda(t) = \lambda_0(t)\exp(\tilde{\boldsymbol{X}}_i[t]'\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{Z}}_i[t]'\tilde{\boldsymbol{b}_i})$ where $\tilde{\boldsymbol{X}}_i[t]$ and $\tilde{\boldsymbol{Z}}_i[t]$ might differ from $\boldsymbol{X}_i[t]$ and $\boldsymbol{Z}_i[t]$ in the linear mixed effects model. Zeng & Cai (2005) set $\boldsymbol{c}_i$ to zero since their application allows the assumption that any unobserved factor that affects the longitudinal measurements also affects the survival outcome. The joint likelihood is built as in (2) by integrating over the random effects and parameters $\sigma_\epsilon^2, \boldsymbol{\Sigma}_b, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \phi$ and $\Lambda(t) = \int_0^t \lambda(s)\,ds$ are estimated using the EM-algorithm to maximise the likelihood.
A similar approach has been proposed by Lin et al. (2002a). The authors use a frailty model for the survival outcome and a mixed effects model to explain the longitudinal measurements. The joint likelihood contains covariates having an effect on both outcomes. Likewise Ratcliffe et al. (2004) analyse survival and longitudinal data when data clustering is present. They use a Cox frailty model for the survival outcome and a mixed effects model for the longitudinal one. A common cluster-level random effect links both outcomes and allows the simultaneous analysis of survival and longitudinal measurements.

Ha et al. (2003) introduce a joint model approach within the framework of hierarchical generalised linear models (HGLM). Shared random effects link the linear mixed model proposed for the longitudinal measurements and the Weibull frailty model proposed for the survival outcome. As in Zeng & Cai (2005) both outcomes are affected differently by the random

effects. Since frailty models can be presented in form of Poisson HGLM (Ha & Lee, 2003) two HGLMs together with a random effect provide the basis for this joint model. The vector of longitudinal measurements of subject i is assumed to follow

$$\boldsymbol{y}_i|b_i \sim N(\boldsymbol{X}_i\boldsymbol{\beta} + (\gamma_1^* b_i)\mathbf{1}, \rho^2\boldsymbol{I})$$

with $b_i \sim N(0,1)$ the random effects with scale parameter $\gamma_1^* = \sqrt{\gamma_1}$ ($\gamma_1 > 0$), $\mathbf{1}$ a vector with all elements set to 1.
The hazard function of $\tau_i|b_i$ with Weibull baseline hazard is given by

$$\lambda_i(t|b_i) = \alpha t^{\alpha-1} \exp(\tilde{\boldsymbol{X}}_i[t]'\tilde{\boldsymbol{\beta}} + \gamma_2 b_i)$$

with $\gamma_2 \in (-\infty, \infty)$ the scale parameter for the random effect and α the shape parameter of the Weibull distribution.
The estimation of the unknown parameters $\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \alpha, \rho^2, \gamma_1$ and γ_2 is done via a hierarchical likelihood approach using

$$h = h(\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \alpha, \rho^2, \gamma_1, \gamma_2) = \sum_i l_{1i} + \sum_i l_{2i} + \sum_i l_{3i}$$

with l_{1i} the logarithm of the conditional density function for y_i given b_i, l_{2i} is that for τ_i^* and δ_i given b_i and l_{3i} is the logarithm of the density for b_i.

5 Applications in biometrical research

A large number of studies generate repeated measurements and event time data, which often depend on each other. Joint model approaches allow a modelling of both outcomes by taking those dependencies into account. This technique has been applied in the following settings: In acquired immunodeficiency syndrome (AIDS) studies viral load and CD4 counts are important disease markers. The longitudinal process of these markers together with time to seroconversion or death has been investigated using models with focus on event time data (Tsiatis et al., 1995; Wang & Taylor, 2001). Guo & Carlin (2004) transferred the idea of Henderson et al. (2000) into a Bayesian framework and modelled CD4 counts and time to death simultaneously via a latent bivariate Gaussian process.
The issue of surrogacy of a disease marker for an endpoint can also be addressed with joint models (Taylor & Wang, 2002). This has been done for example in prostate cancer where the prostate-specific antigen (PSA) level was evaluated as surrogate for survival (Renard et al., 2003). Furthermore Lin et al. (2002b) found subpopulations that differ in the longitudinal PSA trajectories among prostate cancer patients using latent class models for joint modelling.
In patients with cystic fibrosis the pattern of pulmonary function was explored by Schluchter et al. (2002). Since patients with poorest lung function are oftentimes censored by death,

a joint model with focus on serial trend regarding informative dropout was used to avoid biased estimates.
Additional to disease-specific markers many studies also measure quality of life or depression measures together with survival data. Bowman & Manatunga (2005) focus on the serial trend of a depression score that includes the risk of study discontinuation to determine a treatment effect. In data from a pancreatic cancer study the treatment effect was measured in terms of quality and quantity of life (Billingham & Abrams, 2002). The focus of the analysis was survival which then was adjusted for the quality of life.
Joint model approaches with equal focus on repeated measures and survival outcome were for example applied to data that arose from a drug therapy of schizophrenia patients. The course of scores describing psychiatric disorder together with the time of withdraw were analysed jointly (Henderson et al., 2000). Furthermore Ha et al. (2003) investigated the serum creatinine level in patients with renal transplants together with the failure of kidney graft. The quality of life together with the time to the first episode of severe hypoglycaemia was analysed in patients that participated in the Diabetes Control and Complications Trial (Rochon & Gillespie, 2001).

6 A brief look at further models and some extensions

Besides the presented models Bayesian approaches that address the simultaneous analysis of longitudinal measurements and event time data are reasonable. For details on these approaches see Ibrahim et al. (2004) and Brown & Ibrahim (2003). Chi & Ibrahim (2006) give further extensions to the likelihood approach for models with focus on event times based on Bayesian inference. Their model allows both outcomes, longitudinal and survival, to be multidimensional. Multivariate longitudinal processes are modelled using a multivariate mixed effects model which captures both, the dependence among the longitudinal measures over time and the dependence between different longitudinal processes. Multivariate event time data are described by a survival model with proportional hazards structure that additionally includes a frailty to account for correlations between event times.
Tsiatis & Davidian (2004) give additional insight into the structure of likelihoods used to estimate model parameters within joint models. Extensions are particularly necessary when approaches are applied to different types of data. In case of non-continuous event times Rochon & Gillespie (2001) propose to combine a generalised estimating equation (GEE) model for the longitudinal outcome with a GEE for the survival endpoint. But this approach restricts the longitudinal process to equally-spaced intervals.
Many approaches use a linear mixed effects model to describe the longitudinal process. However, this linearity assumption is not always appropriate in practice. Therefore it would be important to extent joint model approaches to more general trajectories for the longitudinal outcome. Splines with high-dimensional basis functions might be a starting point at this

place. Further extensions would also be necessary if one outcome is discrete. That is either a discrete longitudinal outcome, a discrete time scale or both. An example is the disease multiple sclerosis, which motivated this review. There disability status is rated according to an ordinal scale (Kurtzke, 1983). In further research we aim to expand joint model approaches to data from multiple sclerosis patients to increase the understanding of interactions between exacerbations, permanent disability and demographic factors.
Predominantly joint model approaches focus on single event times. In Henderson et al. (2000) an embedded counting process allows to include recurrent events in the model. To examine to which extent recurrent events have an effect on the longitudinal disease process analogous ways have to be followed within the model class focusing on longitudinal processes. As a starting point Hogan et al. (2004) use varying coefficients random effects models to allow the coefficients that describe the trajectories to depend on a single event time.

Acknowledgement
Financial support from the German Research Foundation (DFG), Collaborative Research Center (SFB) 386 - Statistical Analysis of Discrete Structures (Projects B2 and C2), is gratefully acknowledged.

References

Billingham, L. J. & Abrams, K. R. (2002). Simultaneous analysis of quality of life and survival data. *Statistical Methods in Medical Research* **11**, 25–48.

Bowman, F. D. & Manatunga, A. K. (2005). A joint model for longitudinal data profiles and associated event risks with application to a depression study. *Applied Statistics* **54**, 301–316.

Brown, E. R. & Ibrahim, J. G. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics* **59**, 221–228.

Chi, Y.-Y. & Ibrahim, J. G. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics* **62**, 432–445.

Dang, Q., Mazumdar, S., Anderson, S. J., Houck, P. R. & Reynolds, C. F. (2007). Using trajectories from a bivariate growth curve as predictors in a Cox regression model. *Statistics in Medicine* **26**, 800–811.

Guo, X. & Carlin, B. P. (2004). Separate and joint modelling of longitudinal and event time data using standard computer packages. *The American Statistician* **58**, 16–24.

Ha, I. D. & Lee, Y. (2003). Estimating frailty models via poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics* **12**, 663–681.

Ha, I. D., Park, T. & Lee, Y. (2003). Joint modelling of repeated measures and survival time data. *Biometrical Journal* **45**, 647–658.

Henderson, R., Diggle, P. & Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**, 465–480.

Hogan, J. W. & Laird, N. M. (1997). Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine* **16**, 239–257.

Hogan, J. W., Lin, X. & Herman, B. (2004). Mixtures for varying coefficient models for longitudinal data with discrete or continuous nonignorable dropout. *Biometrics* **60**, 854–864.

Hsieh, F., Tseng, Y.-K. & Wand, J.-L. (2006). Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics* **62**, 1037–1043.

Ibrahim, J. G., Chen, M.-H. & Sinha, D. (2004). Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine trials. *Statistica Sinica* **14**, 863–883.

Klein, J. P. & Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, New York, 2nd ed.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* **33**, 1444–1452.

Lin, D. Y. & Zhiliang, Y. (1995). Semiparametric inference for the accelerated life model with time-dependent covariates. *Journal of Statistical Planning and Inference* **44**, 47–63.

Lin, H., McCulloch, C. E. & Mayne, S. E. (2002a). Maximum likelihood estimation in the joint analysis of time-to-event and multiple longitudinal variables. *Statistics in Medicine* **21**, 2369–2382.

Lin, H., Turnbull, B. W., McCulloch, C. E. & Slate, E. H. (2002b). Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* **97**, 53–65.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association* **88**, 125–134.

Ratcliffe, S. J., Guo, W. & Ten Have, T. R. (2004). Joint modeling of longitudinal and survival data via a common frailty. *Biometrics* **60**, 892–899.

Renard, D., Geys, H., Molenberghs, G., Burzykowski, T., Buyse, M., Vangeneugden, T. & Bijnens, L. (2003). Validation of a longitudinally measured surrogate marker for a time-to-event endpoint. *Journal of Applied Statistics* **30**, 235–247.

Rochon, J. & Gillespie, B. W. (2001). A methodology for analysing a repeated measures and survival outcome simultaneously. *Statistics in Medicine* **20**, 1173–1184.

Schluchter, M. D., Konstan, M. W. & Davis, P. B. (2002). Joint modelling the relationship between survival and pulmonary function in cystic fibrosis patients. *Statistics in Medicine* **21**, 1271–1287.

Song, X., Davidian, M. & Tsiatis, A. A. (2002a). An estimator for the proportional hazards model with multiple longitudinal covariates measured with error. *Biostatistics* **3**, 511–528.

Song, X., Davidian, M. & Tsiatis, A. A. (2002b). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics* **58**, 742–753.

Taylor, J. M. G. & Wang, Y. (2002). Surrogate markers and joint models for longitudinal and survival data. *Controlled Clinical Trials* **23**, 626–634.

Tseng, Y.-K., Hsieh, F. & Wang, J.-L. (2005). Joint modelling of accelerated failure time and longitudinal data. *Biometrika* **92**, 587–603.

Tsiatis, A. A. & Davidian, M. (2001). A semiparametric estimator for proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88**, 447–458.

Tsiatis, A. A. & Davidian, M. (2004). Joint modelling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14**, 809–834.

Tsiatis, A. A., DeGruttola, V. & Wulfson, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association* **90**, 27–37.

Vonesh, E. F., Greene, T. & Schluchter, M. D. (2006). Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine* **25**, 143–163.

Wang, Y. & Taylor, J. M. G. (2001). Joint modelling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* **96**, 895–905.

Wulfson, M. S. & Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53**, 330–339.

Zeng, D. & Cai, J. (2005). Simultaenous modelling of survival and longitudinal data with an application to repeated quality of life measures. *Lifetime Data Analysis* **11**, 151–174.