Title

A copula model for joint modeling of longitudinal and time-invariant mixed outcomes

Journal

Statistics in Medicine, 37(27)

ISSN

0277-6715

Authors

Kürüm, Esra
Jeske, Daniel R
Behrendt, Carolyn E
et al.

Publication Date

2018-11-30

DOI

10.1002/sim.7855

Peer reviewed

HHS Public Access

Author manuscript

Stat Med. Author manuscript; available in PMC 2021 November 12.

Published in final edited form as:

Stat Med. 2018 November 30; 37(27): 3931–3943. doi:10.1002/sim.7855.

A copula model for joint modeling of longitudinal and time-invariant mixed outcomes

Esra Kürüm[1], **Daniel R. Jeske**[1], **Carolyn E. Behrendt**[2], **Peter Lee**[3]

[1]Department of Statistics, University of California, Riverside, CA

[2]Division of Biostatistics, Department of Information Sciences, Beckman Research Institute of the City of Hope, Duarte, CA

[3]Department of Immuno-Oncology, City of Hope Comprehensive Cancer Center, Duarte, CA

Abstract

Motivated by a preclinical study in a mouse model of breast cancer, we suggest a joint modeling framework for outcomes of mixed type and measurement structures (longitudinal versus single time/time-invariant). We present an approach based on the time-varying copula models, which is used to jointly model longitudinal outcomes of mixed types via a time-varying copula, and extend the scope of these models to handle outcomes with mixed measurement structures. Our framework allows the parameters corresponding to the longitudinal outcome to be time varying and thereby enabling researchers to investigate how the response-predictor relationships change with time. We investigate the finite sample performance of this new approach via a Monte Carlo simulation study and illustrate its usefulness by an empirical analysis of the motivating preclinical study, comparing the effect of various treatments on tumor volume (longitudinal continuous response) and the number of days until tumor volume triples (time-invariant count response). Through the real-life application and the simulation study, we demonstrate that, compared with marginal modeling, the joint modeling framework offers more precision in the estimation of model parameters.

Keywords

joint model; local regression; varying coefficient model

1 | INTRODUCTION

Increasingly, studies are collecting data on multiple outcomes, usually of mixed measurement structures. That is, within a single study, some outcomes are measured at several time points (longitudinal), whereas other endpoints are time-invariant, that is, measured only at a single time point, either by definition or by lack of data collection resources. In this paper, our goal is to present a joint modeling approach that is flexible enough to model outcomes of mixed type with different measurement

Correspondence: Esra Kürüm, 1348 Olmsted Hall, Department of Statistics, University of California, Riverside, CA 92521. esra.kurum@ucr.edu.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

structures (longitudinal versus single time), while taking full advantage of the longitudinal measurements by allowing parameters corresponding to the longitudinal outcome(s) to vary over time.

Our motivating data come from a longitudinal preclinical study performed on genetically-alike mice. Preclinical studies represent the phase of drug development before human testing, in which novel compounds are studied in relevant animal models of disease. Statistical analyses of preclinical studies are the primary basis for deciding whether to advance or abandon a therapeutic idea. In preclinical trials for cancer, a common longitudinal outcome is the growth of the primary tumor. For patients with cancer, controlling the primary tumor (that is, by surgical resection) may not suffice to achieve cure: preventing metastasis of the primary tumor to sites elsewhere in the body is key to extending survival and achieving cure. However, the onset of metastasis, being microscopic, is usually hard to recognize. The growth of the primary tumor increases dramatically when cells that have left the primary tumor grow elsewhere in the body (as “metastases”) and cells released from the metastasized tumors return to the primary tumor, enhancing its growth rate.[1–5] Thus, we propose “number of days to tripling of tumor volume” as an additional outcome in preclinical cancer studies, as a potential proxy for time to metastasis, a clinically meaningful event.

The study that motivated our work was performed in a mouse model of breast cancer, and in this study, the primary tumor’s volume on each mouse was measured 2–3 times per week with the aim of comparing growth under five conditions (three active treatment arms and two control arms). The three treatment arms were EXP, an antiparasitic drug not previously used to treat cancer, STD, a widely used cytotoxic chemotherapy, and a combination of these two drugs. Typically, in literature, to study the effect of each treatment, the only longitudinal outcome (volume of the primary tumor) considered in the preclinical studies is first plotted graphically by treatment arm and growth curves under various treatments are then compared using mixed effects modeling. In these studies, the effect of treatment is likely to change over time, and if this effect has the same direction over time and is a monotone function of time, mixed effects models can be formulated to be flexible enough to capture this time-varying effect by including an interaction term between treatment and time. However, when the effect of treatment exhibits several fluctuations over time such as continuous changes in the direction and strength, it cannot be directly or efficiently estimated using mixed effects models. Moreover, as we mentioned above, we suggest a second correlated outcome with biological importance, that is, the number of days until tumor volume triples (a time-invariant count response), to be also recognized in preclinical studies. This outcome is related to the longitudinal outcome because it suggests that metastatic tumors have become established elsewhere and are contributing additional volume to the primary tumor (the longitudinal outcome). To analyze these two related outcomes that are of mixed type with different measurement structures, our new joint modeling approach, which has a higher statistical efficiency than marginal modeling and is able to recognize the dynamic nature of treatment effects, has been developed.

In the literature, a number of joint modeling approaches, primarily focusing on longitudinal binary and continuous outcomes, have been established (see, for example, other works[6–14]).

The main difficulty in joint modeling outcomes of mixed types is the lack of a natural multivariate distribution. One solution to this challenge is to introduce a continuous latent variable underlying the binary response and assume that this latent variable and the continuous outcome follow a joint normal distribution. This joint distribution is then factorized in one of two ways into two components that can be modeled separately: (i) a marginal distribution for the continuous response and a conditional distribution for the binary response given the continuous response or (ii) a marginal distribution for the binary response and a conditional distribution for the continuous response given the binary response.

Joint mixed effects models represent another solution to the lack of a natural multivariate distribution for mixed outcomes.[12,15] In these models, a random effect is assumed for each outcome and the association between the outcomes is defined through a joint distribution for these random effects. One drawback of these models, as pointed out by Verbeke et al,[16] is that the maximum-likelihood estimation is only possible when strong assumptions are made. For example, Roy and Lin[17] assumed that the random effects for various outcomes are perfectly correlated. In addition, Hodges and Reich[18] discussed that a mixed-effects model may be confounded. This may inflate the variance of fixed-effects estimators and thus prevent the discovery of important response–predictor relationships.

Copula-based approaches are another solution to the aforementioned problem. (see the work of Nelsen[19] for an introduction to copulas and other works[20–26] for information on copula-based regression models.) In this paper, we follow this solution.

In longitudinal studies, the response-predictor relationships may vary over time, and ordinary regression models cannot capture these dynamic patterns. Time-varying coefficient models overcome this limitation by allowing regression coefficients to vary over time. Kürüm et al[27] built time-varying copula models using time-varying coefficient models and allowed all parameters, including regression coefficients and dependence parameters, to vary over time. In addition, unlike previous approaches[14] that allow time-varying parameters in a joint-modeling framework but are limited to binary–continuous responses, time-varying copula models can be used to jointly model arbitrary response types and dimensions. Although Kürüm et al[27] introduced a flexible model that allows all parameters to be time varying and applied this method to outcomes that are all measured longitudinally, in practice (for example, the preclinical study that motivates this work), it may be desirable to model time-invariant outcomes jointly with longitudinal ones.

Therefore, in this paper, we introduce a new statistical framework to joint modeling. Our main contribution is twofold. First, our method brings arbitrary response type with different measurement structures (longitudinal versus single time/time-invariant) to joint modeling. For instance, for the aforementioned preclinical study, our method will be used to jointly model the tumor volume, longitudinal continuous outcome, and the number of days until the tumor volume triples, the time-invariant count outcome. Second, for the longitudinal outcome, our methodology includes time-varying parameters, which will allow researchers to uncover complex dynamic patterns of response-predictor relationships. This method will lead to more efficient estimators for both time-varying and time-invariant response-predictor

relationships because, in the presence of dependence among outcomes, joint modeling leads to more precise estimators than marginal modeling.[15,27,28] This gain in efficiency increases as the strength of dependence increases, especially for smaller samples. To achieve our goal, we employ time-varying copula models and broaden the applicability of this framework to joint modeling of responses with different measurement structures in addition to joint modeling of arbitrary response types and dimensions, and we also provide a novel application of our approach.

The remainder of this paper is organized as follows. Section 2 describes our joint modeling framework for outcomes of mixed type with different measurement structures. In this section, we also describe the estimation procedure and address issues that might arise in practice. Section 3 illustrates the proposed approach by analyzing data from motivating preclinical study. Section 4 presents our simulation study, which shows the finite sample behavior of our approach. Section 5 includes our concluding remarks.

2 | TIME-VARYING COPULA MODELS

In this section, we introduce the statistical methodology to jointly model responses of mixed type and measurement structures. To achieve this goal, we employ the time-varying copula models[27] and extend the applicability of these models to joint modeling of longitudinal and time-invariant outcomes. These models were built on varying coefficient models.[29,30]

In longitudinal studies, the relationship between response and predictor(s) may change with time. Ordinary regression models cannot capture these dynamic features that might exist in the data. To address the inability of ordinary regression models to capture these dynamic relationships, time-varying coefficient models were developed to allow the regression coefficients to vary over time. A linear time-varying coefficient model takes the form

$$Y(t) = \boldsymbol{x}^{\mathrm{T}}(t)\boldsymbol{\beta}(t) + \varepsilon(t),$$

where $Y(t)$ is the response variable and $\boldsymbol{x}(t) = \{x_1(t), \ldots, x_p(t)\}^{\mathrm{T}}$ is a vector of predictors collected at time t, $\boldsymbol{\beta}(t) = \{\beta_1(t), \ldots, \beta_p(t)\}^{\mathrm{T}}$ are unknown coefficient functions, and $\varepsilon(t)$ is an error such that $\mathbb{E}(\varepsilon \mid \boldsymbol{x}, t) = 0$. A brief summary of the utilization of varying coefficient models in time-varying copula model framework is given below.

For subject i, $i = 1, \ldots, n$, the outcomes measured at time point t_{ij} are denoted as a d-variate process $\boldsymbol{Y}_i(t_{ij}) = \{Y_{i1}(t_{ij}), \ldots, Y_{id}(t_{ij})\}^{\mathrm{T}}$, some of which may be time-invariant outcomes and $j = 1, \ldots, n_i$. In this framework, longitudinal data could be collected at irregular time points; in other words, the number of observations and the observation times may vary from subject to subject. The marginal distribution and density/mass function of the kth outcome are assumed to be F_{ik} and f_{ik}, respectively, both of which may depend on time-varying parameters $\boldsymbol{\theta}_k(t)$ and some of which may be regression coefficients $\boldsymbol{\beta}_k(t)$. For instance, in the aforementioned mouse cancer study, for the bivariate continuous–count outcomes, we might use normal linear and poisson models

$$Y_{i1}(t) \sim \mathcal{N}\left\{\boldsymbol{x}_{i1}(t)^{\mathrm{T}}\boldsymbol{\beta}_1(t), \sigma^2(t)\right\} \quad \text{and} \quad Y_{i2}(t) \sim \mathcal{P}\left[\exp\left\{-\boldsymbol{x}_{i2}(t)^{\mathrm{T}}\boldsymbol{\beta}_2(t)\right\}\right],$$

where $\mathcal{P}(\lambda)$ denotes a Poisson random variable with mean λ, and $\boldsymbol{x}_{i1}(t)$ and $\boldsymbol{x}_{i2}(t)$ are vectors of predictors for subject i, measured at time t.

In time-varying copula models, the dependence between the outcomes is modeled using a time-varying Gaussian d-copula $C_{\boldsymbol{\eta}(t)}\{u_1(t), \ldots, u_d(t)\}$,[31,32] with $\boldsymbol{\eta}(t)$ as the copula parameters,[19]

$$\boldsymbol{\Phi}_{\mathbf{Q}(t)}\left[\boldsymbol{\Phi}^{-1}\{u_1(t)\}, \ldots, \boldsymbol{\Phi}^{-1}\{u_d(t)\}\right],$$

where $u_i(t) = F_i\{y_i(t)\}$, $\Phi_{\mathbf{Q}(t)}$ is the CDF of a d-variate multinormal random variable with mean vector $\mathbf{0}$ and correlation matrix $\mathbf{Q}(t)$, and Φ^{-1} is the univariate standard normal quantile function.

The estimation of the parameters $\theta(t) = \left\{\theta_1^{\mathrm{T}}(t), \ldots, \theta_d^{\mathrm{T}}(t), \boldsymbol{\eta}^{\mathrm{T}}(t)\right\}^{\mathrm{T}}$ at time t_0 are performed via using local constant fitting techniques, that is, $\boldsymbol{\theta}(t_0)$ is assumed to be constant on a neighborhood of t_0. It is straightforward to adopt the estimation to employ a higher-order polynomial approximation to $\boldsymbol{\theta}(t_0)$, but even a linear approximation increases the computational burden quite a bit while reducing bias only slightly.

2.1 | Estimation in time-invariant Gaussian copula models

Before we proceed to the estimation in the proposed copula-based models, in this subsection, we will briefly describe the estimation procedure in time-invariant Gaussian copula models. This will allow us to introduce the notation and motivate our approach in the next section more clearly. The density for the Gaussian d-copula is given as

$$c_{\mathbf{Q}}(\boldsymbol{u}) = \frac{\phi_{\mathbf{Q}}\left\{\boldsymbol{\Phi}^{-1}(u_1), \ldots, \boldsymbol{\Phi}^{-1}(u_d)\right\}}{\prod_{i=1}^{d}\phi\left\{\boldsymbol{\Phi}^{-1}(u_i)\right\}} \propto |\mathbf{Q}|^{-1/2}\exp\left\{-\frac{1}{2}\boldsymbol{w}'\left(\mathbf{Q}^{-1} - \mathbf{I}\right)\boldsymbol{w}\right\},$$

where $\boldsymbol{u} = (u_1, \ldots, u_d)^{\mathrm{T}}$, $\boldsymbol{w} = (w_1, \ldots, w_d)^{\mathrm{T}} = \{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_{\mathrm{d}})\}^{\mathrm{T}}$, $\mathbf{Q}$ is the correlation matrix and $\mathbf{I}$ is the $d \times d$ identity matrix. Let f_i be the density function corresponding to the marginal distribution F_i with $i = 1, \ldots, d$. If the marginal distributions $F_1, \ldots, F_d$ are all continuous, the likelihood of the parameters $\boldsymbol{\theta}$ is

$$L(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto c_{\mathbf{Q}}\{F_1(y_1), \ldots, F_d(y_d)\}\prod_{i=1}^{d} f_i(y_i).$$

This leads to the following the log likelihood:

$$\ell(\boldsymbol{\theta} \mid \mathbf{y}) = -\frac{1}{2}\log|\mathbf{Q}| - \frac{1}{2}\boldsymbol{w}'\left(\mathbf{Q}^{-1} - \mathbf{I}\right)\boldsymbol{w} + \sum_{i=1}^{d}\log f_i(y_i), \quad (1)$$

where $w_i = \Phi^{-1}\{F_i(y_i)\}$. The optimization of this log likelihood will give us the maximum likelihood estimate of $\boldsymbol{\theta}$.

The likelihood has the simple form given above only when all the marginal distributions are continuous; however, when some of the marginal distributions are discrete, the true likelihood has a more complicated form[25,33] because $w_i = \Phi^{-1}\{F_i(y_i)\}$ is not standard normal (since $F_i(y_i)$ is not standard uniform if F_i has jumps). In addition, it becomes cumbersome to use the true likelihood as the number of discrete responses increases and the distributional transform can be used to approximate the true likelihood in these cases (see other works[34–36] for detailed information on this transformation).

2.2 | Estimation in time-varying copula models

In this section, we present the estimation in time-varying copula models. Although, these models have been introduced in a completely time-varying manner, these models are flexible enough to accommodate time-varying and time-invariant parameters simultaneously. More specifically, when one or more of the outcomes are time invariant, in this approach, parameters corresponding to these outcomes in $\boldsymbol{\theta}(t)$ can be assumed to be time invariant and estimated accordingly.

The parameters $\theta(t_0) = \left\{\theta_1^{\mathrm{T}}(t_0), \ldots, \theta_d^{\mathrm{T}}(t_0), \boldsymbol{\eta}^{\mathrm{T}}(t_0)\right\}^{\mathrm{T}}$ are estimated via maximizing the local kernel-weighted log likelihood

$$\ell\{\theta(t_0) \mid \mathbf{T}\} = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ell\left\{\theta(t_0) \mid \mathbf{y}_{ij}\right\}K_h(t_0 - t_{ij}),$$

where $\mathbf{T} = (\boldsymbol{t}_1 \cdots \boldsymbol{t}_m)$ with $\boldsymbol{t}_i = \left(t_{i1}, \ldots, t_{in_i}\right)^{\mathrm{T}}$, $\ell\{\boldsymbol{\theta}(t_0) \mid \mathbf{y}_{ij}\}$ is the log likelihood of $\boldsymbol{\theta}(t_0)$ given the outcomes for subject i at time point t_{ij}, $K_h(t) = h^{-1}K(t/h)$ is the scaled kernel function $K(\cdot)$ with bandwidth h. The response vector is partitioned so that the first d_1 coordinates are continuous and the remaining coordinates are discrete,

$$\begin{aligned}\ell\left\{\boldsymbol{\theta}(t_0) \mid \mathbf{y}_{ij}\right\} = &-\frac{1}{2}\log|\mathbf{Q}| \\ &-\frac{1}{2}\left(\boldsymbol{w}_{ij}^{\mathrm{T}}, \boldsymbol{w}_{ij}^{*\mathrm{T}}\right)\left\{\mathbf{Q}^{-1} - \mathbf{I}\right\}\left(\boldsymbol{w}_{ij}^{\mathrm{T}}, \boldsymbol{w}_{ij}^{*\mathrm{T}}\right)^{\mathrm{T}} \\ &+\sum_{k=1}^{d}\log f_{ik}\left\{y_{ik}(t_{ij})\right\},\end{aligned}$$

where

$$\boldsymbol{w}_{ij} = \left(\boldsymbol{\Phi}^{-1}\left[F_{i1}\left\{y_{i1}(t_{ij})\right\}\right], \ldots, \boldsymbol{\Phi}^{-1}\left[F_{id_1}\left\{y_{id_1}(t_{ij})\right\}\right]\right)^{\mathrm{T}},$$

$$\boldsymbol{w}_{ij}^{*} = \left(\Phi^{-1}\left\{u_{i(d_1+1)}(t_{ij})\right\}, \ldots, \Phi^{-1}\left\{u_{id}(t_{ij})\right\}\right)^{\mathrm{T}}.$$

The computation of $u_{ik}(t_{ij})(k = d_1 + 1, \ldots, d)$ is performed through the following distributional transform approximation[34–36]:

$$u_{ik}(t_{ij}) = \frac{F_{ik}\left\{y_{ik}^{-}(t_{ij})\right\} + F_{ik}\left\{y_{ik}(t_{ij})\right\}}{2}.$$

The quasi-Newton method[37] is employed to obtain $\boldsymbol{\theta}(t_0)$ to ensure that estimated dependence and scale parameters can be appropriately constrained.

In order to construct pointwise confidence intervals, the variance of $\hat{\theta}(t_0)$ is estimated using the results obtained by Fan et al.[38] The conditional variance has a sandwich form

$$\begin{aligned}\boldsymbol{\Sigma}(t_0) &= \mathbb{V}\left\{\hat{\boldsymbol{\theta}}(t_0) \mid \mathbf{T}\right\} \\ &\approx \kappa(t_0)\mathscr{H}^{-1}(t_0)\mathscr{J}(t_0)\mathscr{H}^{-1}(t_0) \\ &= \kappa(t_0)[\,\ell''\{\boldsymbol{\theta}(t_0) \mid \mathbf{T}\}]^{-1}\mathbb{V}[\,\ell'\{\boldsymbol{\theta}(t_0) \mid \mathbf{T}\}][\,\ell''\{\boldsymbol{\theta}(t_0) \mid \mathbf{T}\}]^{-1},\end{aligned}$$

where $\kappa(t_0) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} K^2\left\{(t_0 - t_{ij})/h\right\}/h^2$, $\mathscr{H}(t_0) = \ell''\{\boldsymbol{\theta}(t_0) \mid \mathbf{T}\}$ is the Hessian matrix, which can be estimated by $\ell''\left\{\hat{\boldsymbol{\theta}}(t_0) \mid \mathbf{T}\right\}$, and the variance of the score, $\mathscr{J}(t_0)$, can be estimated by

$$\frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} \nabla\nabla^{\mathrm{T}}\,\ell\left\{\hat{\boldsymbol{\theta}}(t_0) \mid \mathbf{y}_{ij}\right\}K_h(t_0 - t_{ij})}{\sum_{i=1}^{m}\sum_{j=1}^{n_i} K_h(t_0 - t_{ij})},$$

where ∇ denotes the gradient.

Based on the asymptotic normality results shown in the works of Hall and Tajvidi[39] and De Melo and Mendes,[40] the asymptotic pointwise $(1 - \alpha)100\%$ confidence interval for the kth element of $\boldsymbol{\theta}(t_0)$ can be given as follows:

$$\hat{\theta}_k(t_0) \pm \Phi^{-1}(1 - \alpha/2)\sqrt{\hat{\boldsymbol{\Sigma}}_k(t_0)},$$

where $\hat{\boldsymbol{\Sigma}}_k(t_0)$ is the kth diagonal element of $\hat{\boldsymbol{\Sigma}}(t_0)$.

Although the procedure described above does not account for intrasubject dependence, theory suggests that this type of dependence can safely be ignored (for estimation of $\boldsymbol{\theta}(t)$). More specifically, it is shown that regardless of the working correlation structure, the method of kernel generalized estimation equations (kernel GEE) yields a root-n consistent estimator.[41] Moreover, kernel GEE with working independence correlation matrix leads to the most efficient estimator for the nonparametric regression function in a longitudinal

setting. As the procedure used in this article shares the spirit of kernel GEE, we suspect that our estimator for $\boldsymbol{\theta}(t)$ is root-n consistent and is likely the most efficient. Theoretical justification is beyond the scope of this paper and is left to a future investigation.

In practical application of methods based on kernel smoothing, selection of the kernel function and the bandwidth are important issues. In terms of kernel functions, we suggest employing a bimodal kernel[42] as using this kernel results in more accurate estimates in the presence of intra-subject dependence. We recommend using a member of the so called ε-optimal class of bimodal kernels.[42] Specifically,

$$K_\varepsilon(t) = \frac{4}{4 - 3\varepsilon - \varepsilon^2} \begin{cases} \frac{3}{4}\left(1 - t^2\right)\mathbf{1}\{|t| \le 1\} & \text{if } |t| \ge \varepsilon \\ \frac{3}{4}\frac{1 - \varepsilon^2}{\varepsilon}|t| & \text{if } |t| < \varepsilon \end{cases}$$

with $\varepsilon = 0.1$, where $\mathbf{1}\{\cdot\}$ denotes the indicator function.

For the bandwidth selection, we suggest using the following leave-one-subject-out cross validation score, which is a form of the cross-validation approach proposed by Fan and Zhang,[43]

$$\mathrm{CV}(h) = -\sum_{i=1}^{m}\sum_{j=1}^{n_i} \ell\left\{\hat{\theta}^{\setminus i}(t_{ij}) \mid \mathbf{y}_{ij}, h\right\},$$

where $\hat{\theta}^{\setminus i}(t_{ij})$ is the leave-i-out estimate for time t_{ij}. We compute this cross-validation score for a range of bandwidths and select the bandwidth that minimizes the score.

3 | A NOVEL APPLICATION TO PRECLINICAL STUDY

In this section, we apply our joint modeling method to the preclinical study mentioned in the introduction. To induce a primary breast tumor per subject, breast cancer cells were injected into the mammary fat pad of 68 mice on day zero. Tumors became measurable on day four, from which point tumor volume (($Length * Width^2$)/2) was measured 2–3 times per week, yielding 14 measurements per mouse over the course of 34 days. Starting on day four, groups of 12 mice received either: standard drug alone (STD), experimental drug alone (EXP), both drugs together (EXP+STD), or only the vehicle in which the experimental drug was administered (Vehicle). A further 20 mice served as controls, receiving no treatment. In our analysis, we evaluate the four treatments relative to control using (1) tumor volume over time and (2) number of days until tumor volume triples. Traditionally, only tumor volume is recognized as an endpoint to be analyzed using longitudinal mixed effects modeling. Our recognition of a second, correlated outcome (the number of days until the tumor volume triples), creates the need for a joint modeling framework capable of accommodating outcomes of mixed type and measurement structure. Figure S1 presents the density plot of this second outcome, the number of days until the tumor volume triples, for each group. Although this figure indicates that the three drug treatment groups (STD, EXP, and

EXP+STD) exhibit multimodal distributions with long tails, all groups have a concentration of mass around day six.

In our analysis, the natural logarithm of the tumor volume, $Y_{i1}(t)$, is assumed to follow a normal distribution with mean $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_1(t)$ and variance $\sigma^2(t)$, and number of days until tumor volume triples, Y_{i2}, is assumed to follow a Poisson distribution with mean $\exp\left(-\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}_2\right)$, where $\boldsymbol{x}_{i1}^{\mathrm{T}} = (1, x_{i1}, x_{i2}, x_{i3}, x_{i4})$, $\boldsymbol{\beta}_1(t) = \{\beta_{10}(t), \beta_{11}(t), \beta_{12}(t), \beta_{13}(t), \beta_{14}(t)\}^{\mathrm{T}}$, and $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21}, \beta_{22}, \beta_{23}, \beta_{24})^{\mathrm{T}}$ with

$$x_{i1} = \begin{cases} 1, & \text{if subject } i \text{ is in the vehicle alone group} \\ 0, & \text{otherwise} \end{cases}$$

$$x_{i2} = \begin{cases} 1, & \text{if subject } i \text{ is in the STD group} \\ 0, & \text{otherwise} \end{cases}$$

$$x_{i3} = \begin{cases} 1, & \text{if subject } i \text{ is in the EXP group} \\ 0, & \text{otherwise} \end{cases}$$

$$x_{i3} = \begin{cases} 1, & \text{if subject } i \text{ is in the EXP+STD group} \\ 0, & \text{otherwise.} \end{cases}$$

The reason we only include the treatment indicators as predictors in our model is that the mice included in the study were bred to be genetically identical, so the only difference between each mouse was the treatment they received. In addition, as we created dummy variables corresponding to vehicle and treatment groups, the intercept term corresponding to the longitudinal outcome describes the natural logarithm of the tumor volume for subjects in the no treatment group.

In this application, we chose a bandwidth of $h = 3$ by using the leave-one-out cross-validation method described in Section 2, and we used a bimodal kernel. The estimated time-varying regression coefficient functions along with their confidence intervals for the natural logarithm of the longitudinal continuous outcome (tumor volume) are presented in Figure 1.

- The plot in panel A shows that the intercept function is time varying and increases with time, describing how tumor volume grows among subjects in the control group.
- From panel B, we observe that the coefficient for vehicle alone group may be time invariant. Moreover, the effect of Vehicle is significantly different from no treatment group and positive between days 10 and 20, indicating that the Vehicle is able to promote tumor growth when administered alone.

- The confidence bands in panels C-E suggest that the three treatments (EXP, STD, and EXP+STD) have effects almost always significantly different from no treatment group but that only combination treatment (EXP+STD) has a time-varying effect. In other words, we can draw a straight line through all confidence band, indicating the coefficient could be a constant, except for the confidence band for the EXP+STD group. In addition, the regression coefficients in these panels are always negative, which suggests that throughout the study, these drug treatments have greater efficacy than Vehicle alone and no treatment.
- According to the plot in panel F, we see that the standard deviation of the longitudinal outcome is time varying and increases after day 13.

Although observing the behavior of regression coefficient functions over time is informative, observing differences between regression coefficients would be more useful in interpreting and comparing treatment effects. Figure 2 demonstrates the comparison of regression coefficients.

- In panel A, we see that the difference between the regression coefficients for STD and Vehicle alone $\left(\hat{\beta}_{12}(t) - \hat{\beta}_{11}(t)\right)$ is negative and is statistically significant after day seven. The negative difference indicates that the STD drug controls growth in tumor volume better than Vehicle alone does.
- Panel B demonstrates the difference between the regression coefficients for EXP and STD treatments $\left(\hat{\beta}_{13}(t) - \hat{\beta}_{12}(t)\right)$. The confidence band in this plot shows that there is no difference between these treatments in terms of their effects on the tumor volume.
- The plot in panel C compares the EXP+STD and EXP treatments. The difference between the regression coefficients for these groups $\left(\hat{\beta}_{14}(t) - \hat{\beta}_{13}(t)\right)$ is statistically significant and negative, indicating greater efficacy for EXP+STD than for EXP alone.

Figure S2 presents the fitted response values for the tumor volume along with their corresponding confidence bands. The results we conclude from this Figure are similar to the ones we obtained from the contrasts presented in Figure 2. More specifically, although the tumor volume increases in each group, EXP+STD treatment achieves the greatest efficacy compared to other treatments and the control group.

The results of the analysis for the time-invariant outcome (number of days until tumor volume triples) are depicted in Table 1. According to this Table, there is no difference between the four treatments compared to no treatment group in their ability to delay the tripling of tumor volume. This result agrees with the distributions presented in Figure S1.

In the introduction, we stated that, in the presence of dependence among outcomes, joint modeling methods may lead to more precise estimators of marginal parameters than marginal modeling approaches, especially for smaller sample sizes. In order to investigate this statement using the preclinical study in mice, we fitted a univariate time-varying model to the natural logarithm of the longitudinal outcome, which is assumed to follow a normal

distribution, and a traditional generalized linear model to the time-invariant outcome. Figure 3 and Table 1 depict a comparison of standard errors obtained from the joint model to the marginal models. In Figure 3, we observe that, for selected parameters, joint modeling has a considerable gain in efficiency compared to the univariate time-varying model for all days except at the beginning of the study. Table 1 shows a similar efficiency gain in joint modeling when compared with the generalized linear model except for a slight loss in the coefficient of EXP treatment. In summary, our joint modeling method exhibits efficiency gain compared to marginal modeling for parameters corresponding to both longitudinal and time-invariant outcomes.

4 | SIMULATED APPLICATION

In this section, we demonstrate the finite sample performance of our proposed methodology via a Monte Carlo simulation study that resembles the preclinical study presented in Section 3. In this simulation study, we used the $K_{0.1}$ bimodal kernel and a set of equidistant grid points $\{t_k, k = 1, \ldots, n_{\text{grid}}\}$ between 0 and 1 with $n_{\text{grid}} = 200$. Since each subject in the preclinical study is measured at 14 time points, for the ith subject, we chose 14 equidistant measurement times between [0, 1]. Note that we used this interval for the ease of construction of this simulation study, other intervals or discrete time points could also be chosen. We used a sample size of $n = 68$ and generated 500 data sets. Figure S3 displays a few of the generated data sets.

The outcomes in our simulation study are longitudinal continuous and time-invariant count variables. Therefore, we let $Y_{i1}(t)$ be a Gaussian variable with mean $\boldsymbol{x}_{i1}^{\mathrm{T}}\boldsymbol{\beta}_1(t)$ and variance $\sigma^2(t) = 0.25\cos(2.3\pi t) + 0.5$ and Y_{i2} be a Poisson variable with mean $\exp\left(-\boldsymbol{x}_{i2}^{\mathrm{T}}\boldsymbol{\beta}_2\right)$, where $\boldsymbol{\beta}_1(t) = \{\beta_{10}(t), \beta_{11}(t)\}^{\mathrm{T}} = \{1 + \sin(2.5\pi t), \sin(0.75\pi t)\}^{\mathrm{T}}$ and $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21})^{\mathrm{T}} = (0.9, 0.4)^{\mathrm{T}}$, with $\beta_{10}(t)$ and β_{20} as the intercepts. In the preclinical study, there were treatment groups, and the dummy variables corresponding to these treatment groups were the only predictors. In order to create a resemblance of that in this simulation, we assumed that there is one treatment group and simulated the corresponding predictor from a binomial distribution with success probability 0.6. In other words, each subject has a 60% chance to be in the treatment group. Therefore, $\boldsymbol{x}_{i1}^{\mathrm{T}} = \boldsymbol{x}_{i2}^{\mathrm{T}} = (1, x_i)^{\mathrm{T}}$ with x_i as the predictor that is generated from a binomial distribution with success probability 0.6. The correlation between continuous outcomes measured at different time points is defined as $0.2^{|t_j - t_{j'}|}$. A brief description of our simulation procedure is presented below. For subject i with $n_i = 14$ and $t_i = (t_{i1}, \ldots, t_{i14})^{\mathrm{T}}$,

1. construct the 2 × 2 correlation matrix with off-diagonal entries that correspond to the correlation between the responses and apply the Cholesky decomposition to impose this correlation on $\left\{Z_{i1}(t_{ij}), Z_{i2}\right\}^{\mathrm{T}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ with $j = 1, \ldots, 14$ and $\mathbf{I}$ as the identity matrix;
2. similarly, construct the correlation matrix for the longitudinal outcome using the correlation structure defined above and use the Cholesky decomposition to impose the correlation structure on $\{Z_{i1}(t_{i1}), \ldots, Z_{i1}(t_{i14})\}^{\mathrm{T}}$;

3. apply the probability integral transformation to each $Z_{i\cdot}\boldsymbol{U} = \{U_{i1}(t_{i1}), \ldots, U_{i2}\}^{\mathrm{T}} = [\Phi^{-1}\{Z_{i1}(t_{i1})\}, \ldots, \Phi^{-1}(Z_{i2})]^{\mathrm{T}}$;
4. finally, apply the inverse probability integral transformation and obtain the responses $\{Y_{i1}(t_{ij}), Y_{i2}\}^{\mathrm{T}} = \left[F_{i1}^{-1}\{U_{i1}(t_{ij})\}, F_{i2}^{-1}(U_{i2})\right]^{\mathrm{T}}$,

where F_{i1} and F_{i2} are the cdfs corresponding to the Gaussian and Poisson distributions described above.

We used the leave-one-out cross-validation bandwidth selector defined in Section 2 to obtain the optimal bandwidth. Figure 4 shows the results for $h = 0.1$: the bandwidth that minimized the cross-validation score. According to this Figure, our estimation procedure performed well with respect to bias. We see from the plots in Figure 4 that the empirical pointwise confidence bands for all time-varying functions were close to their corresponding mean theoretical confidence band, and they all cover the true function. Therefore, the standard errors obtained using our procedure were accurate for all time-varying functions. Figure S4 demonstrates that the coverage rates, that is, the percentage of confidence bands that covered the true value of the regression coefficient at time t based on 500 Monte Carlo simulation runs, were close to the desired 95%.

The results for the time-invariant outcome are displayed in Table 2. In this Table, we present the bias and mean-squared errors for $\hat{\beta}_{2j}(\mathrm{j} = 0, 1)$ obtained using our joint modeling approach. According to this Table, we estimated time-invariant components $\beta_{2j}(\mathrm{j} = 0, 1)$ with small biases and mean-squared errors.

We present the comparison of standard errors obtained from the joint model to the marginal model for the longitudinal outcome, that is, a univariate time-varying coefficient model in Figure 5. Note that although the estimated standard errors for $\hat{\beta}_{10}(t)$ and $\hat{\beta}_{11}(t)$ exhibit a similar pattern over time, the magnitudes are different. According to this Figure, for both regression coefficients, joint modeling leads to an apparent efficiency gain. The efficiency gain is slightly less pronounced between time points 0.6 and 0.8. In Table 2, we present the 2.5% and 97.5% percentiles of the estimated standard errors for $\hat{\beta}_{2\cdot}(\mathrm{j} = 0, 1)$ obtained from the joint model and the marginal model, that is, a generalized linear model, for the time-invariant outcome. Table 2 shows that the 2.5% and 97.5% percentiles of the estimated standard errors for $\hat{\beta}_{20}$ in the marginal model were (0.176, 0.236), whereas these percentiles for the joint models were calculated as (0.113, 0.170). Similarly, the 2.5% and 97.5% percentiles of the estimated standard errors for $\hat{\beta}_{21}$ were (0.065, 0.086) and (0.044, 0.060) for the marginal and joint models, respectively. These values clearly indicate a considerable efficiency gain in joint modeling as the estimated standard errors obtained from joint modeling are less than the ones calculated from marginal modeling.

5 | CONCLUSION

In this article, we employed the time-varying copula procedure to jointly model outcomes of mixed type with different measurement structures (longitudinal versus single time/time-invariant). In this approach, we use local constant fitting techniques to obtain estimators of

the time-varying and time-invariant parameters. We performed a novel application of our method to a preclinical study in a mouse model of breast cancer to illustrate joint analysis of two associated outcomes: tumor volume, a longitudinal continuous outcome and number of days until tumor volume triples, a time-invariant count response. Our analysis demonstrated not only that combined treatment with experimental and standard drug achieves the largest reduction in tumor growth but also that all four treatments produce similar delay in number of days until tumor volume triples, suggesting that the observed gain in efficacy is not achieved through delay in onset of metastatic growth. The analysis of preclinical study data showed that the joint modeling framework presents more precision in estimation of model parameters compared to marginal modeling. Then, through a simulation study, we demonstrated that our procedure works well on estimating both the time-varying and time-invariant parameters for the longitudinal and time-invariant outcomes, respectively. Note that although we demonstrated the applicability of our approach on balanced longitudinal data, our method can also be applied to longitudinal outcomes that are measured irregularly, that is, when subjects have unequal numbers of measurements and different measurement times. The estimation in the data application and simulation study has been performed using the R statistical software.

In the analysis of the preclinical mouse study, we treat the number of days until tumor volume triples as a time-invariant count outcome because of no censoring in the data (tumor volume of all mice tripled during the study), whereas it could also be considered as a survival outcome. Although as noted by Laird and Olivier[44], a survival outcome could also be considered to follow a Poisson distribution and analyzed accordingly, one could also be interested in analyzing the data from this study via a joint modeling framework for longitudinal and survival outcomes. Currently, there is no such joint modeling framework that includes time-varying coefficient models for the longitudinal outcome. Developing such a joint modeling framework would also be of interest and requires further research, for which the joint modeling approaches described in the work of Rizopoulos[45] may prove useful.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The preclinical study reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award number P30CA033572. The content of this report is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding information

National Cancer Institute of the National Institutes of Health, Grant/Award Number: P30CA033572

REFERENCES

1. Norton L, Massagué J. Is cancer a disease of self-seeding? Nat Med. 2006;12(8):875–878. [PubMed: 16892025]

2. Kim M-Y, Oskarsson T, Acharyya S, et al. Tumor self-seeding by circulating cancer cells. Cell. 2009;139(7):1315–1326. [PubMed: 20064377]
3. Comen E, Norton L. Self-seeding in cancer. Minimal Residual Disease and Circulating Tumor Cells in Breast Cancer. Berlin, Germany: Springer-Verlag Berlin Heidelberg; 2012;13–23.
4. Newton PK, Mason J, Bethel K, et al. Spreaders and sponges define metastasis in lung cancer: a Markov chain Monte Carlo mathematical model. Cancer Res. 2013;73(9):2760–2769. [PubMed: 23447576]
5. Newton PK, Mason J, Venkatappa N, et al. Spatiotemporal progression of metastatic breast cancer: a Markov chain model highlighting the role of early metastatic sites. NPJ Breast Cancer. 2015;1.
6. Catalano PJ, Ryan LM. Bivariate latent variable models for clustered discrete and continuous outcomes. J Am Stat Assoc. 1992;87(419):651–658.
7. Cox DR, Wermuth N. Response models for mixed binary and quantitative variables. Biometrika. 1992;79(3):441–461.
8. Fitzmaurice GM, Laird NM. Regression models for a bivariate discrete and continuous outcome with clustering. J Am Stat Assoc. 1995;90(431):845–852.
9. Sammel MD, Ryan LM, Legler JM. Latent variable models for mixed discrete and continuous outcomes. J R Stat Soc Ser B. 1997;59(3):667–678.
10. Regan MM, Catalano PJ. Likelihood models for clustered binary and continuous outcomes: application to developmental toxicology. Biometrics. 1999;55(3):760–768. [PubMed: 11315004]
11. Dunson DB. Bayesian latent variable models for clustered mixed outcomes. J R Stat Soc Ser B. 2000;62(2):355–366.
12. Gueorguieva RV, Agresti A. A correlated probit model for joint modeling of clustered binary and continuous responses. J Am Stat Assoc. 2001;96(455):1102–1112.
13. Liu X, Daniels MJ, Marcus B. Joint models for the association of longitudinal binary and continuous processes with application to a smoking cessation trial. J Am Stat Assoc. 2009;104(486):429–438. [PubMed: 20161053]
14. Kürüm E, Li R, Shiffman S, Yao W. Time-varying coefficient models for joint modeling binary and continuous outcomes in longitudinal data. Stat Sin. 2016;26(3):979–1000. [PubMed: 27667908]
15. Gueorguieva R. A multivariate generalized linear mixed model for joint modelling of clustered outcomes in the exponential family. Stat Model. 2001;1(3):177–193.
16. Verbeke G, Molenberghs G, Rizopoulos D. Random effects models for longitudinal data. In: van Montfort K, Oud J, Satorra A, eds. Longitudinal Research with Latent Variables. Berlin, Germany: Springer-Verlag Berlin Heidelberg; 2010:37–96.
17. Roy J, Lin X. Latent variable models for longitudinal data with multiple continuous outcomes. Biometrics. 2000;56(4):1047–1054. [PubMed: 11129460]
18. Hodges JS, Reich BJ. Adding spatially-correlated errors can mess up the fixed effect you love. Am Stat. 2010;64(4):325–334.
19. Nelsen RB. An Introduction to Copulas. New York, NY: Springer Science+Business Media; 2006.
20. De Leon AR, Chough KC. Analysis of Mixed Data: Methods & Applications. Boca Raton, FL: Chapman and Hall/CRC Press; 2013.
21. Heinen A, Rengifo EW. Multivariate reduced rank regression in non-Gaussian contexts, using copulas. Comput Stat Data Anal. 2008;52(6):2931–2944.
22. Madsen L, Fang Y. Joint regression analysis for discrete longitudinal data. Biometrics. 2011;67(3):1171–1175. [PubMed: 21039391]
23. Masarotto G, Varin C. Gaussian copula marginal regression. Electron J Stat. 2012;6:1517–1549.
24. Smith M, Min A, Almeida C, Czado C. Modeling longitudinal data using a pair-copula decomposition of serial dependence. J Am Stat Assoc. 2010;105(492):1467–1479.
25. Song PX-K, Li M, Yuan Y. Joint regression analysis of correlated data using Gaussian copulas. Biometrics. 2009;65(1):60–68. [PubMed: 18510653]
26. Wu B, de Leon AR. Gaussian copula mixed models for clustered mixed outcomes, with application in developmental toxicology. J Agric Biol Environ Stat. 2014;19(1):39–56.
27. Kürüm E, Hughes J, Li R, Shiffman S. Time-varying copula models for longitudinal data. Stat Interface. 2018;11(2):203–221. [PubMed: 29686744]

28. Song PX-K. Correlated Data Analysis: Modeling, Analytics, and Applications. Springer Series in Statistics. New York, NY: Springer Science+Business Media; 2007.
29. Cleveland WS, Grosse E, Shyu WM. Local regression models. Statistical Models in S. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software; 1992.
30. Hastie T, Tibshirani R. Varying-coefficient models. J R Stat Soc Ser B. 1993;55(4):757–796.
31. Joe H. Multivariate Models and Dependence Concepts. Boca Raton, FL: Chapman & Hall/CRC; 1997.
32. Song PX-K. Multivariate dispersion models generated from Gaussian copula. Scand J Stat. 2000;27(2):305–320.
33. Li M, Boehnke M, Abecasis GR, Song PX-K. Quantitative trait linkage analysis using Gaussian copulas. Genetics. 2006;173(4):2317–2327. [PubMed: 16751671]
34. Rüschendorf L. Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. Ser Stat. 1981;12(3):327–338.
35. Rüschendorf L. On the distributional transform, Sklar's theorem, and the empirical copula process. J Stat Plan Infer. 2009;139(11):3921–3927.
36. Burgert C, Rüschendorf L. On the optimal risk allocation problem. Stat Decis. 2006;24(1):153–171.
37. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput. 1995;16(5):1190–1208.
38. Fan J, Farmen M, Gijbels I. Local maximum likelihood estimation and inference. J R Stat Soc Ser B. 1998;60(3):591–608.
39. Hall P, Tajvidi N. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. Stat Sci. 2000;15(2):153–167.
40. De Melo EFL, Mendes BVM. Local estimation of copula based value-at-risk. Braz Rev Financ. 2009;7(1):19–50.
41. Lin X, Carroll RJ. Semiparametric regression for clustered data using generalized estimating equations. J Am Stat Assoc. 2001;96(455):1045–1056.
42. De Brabanter K, De Brabanter J, Suykens JAK, De Moor B. Kernel regression in the presence of correlated errors. J Mach Learn Res. 2011;12:1955–1976.
43. Fan J, Zhang W. Statistical methods with varying coefficient models. Stat Interface. 2008;1:179–195. [PubMed: 18978950]
44. Laird N, Olivier D. Covariance analysis of censored survival data using log-linear analysis techniques. J Am Stat Assoc. 1981;76(374):231–240.
45. Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. Boca Raton, FL: CRC Press; 2012.

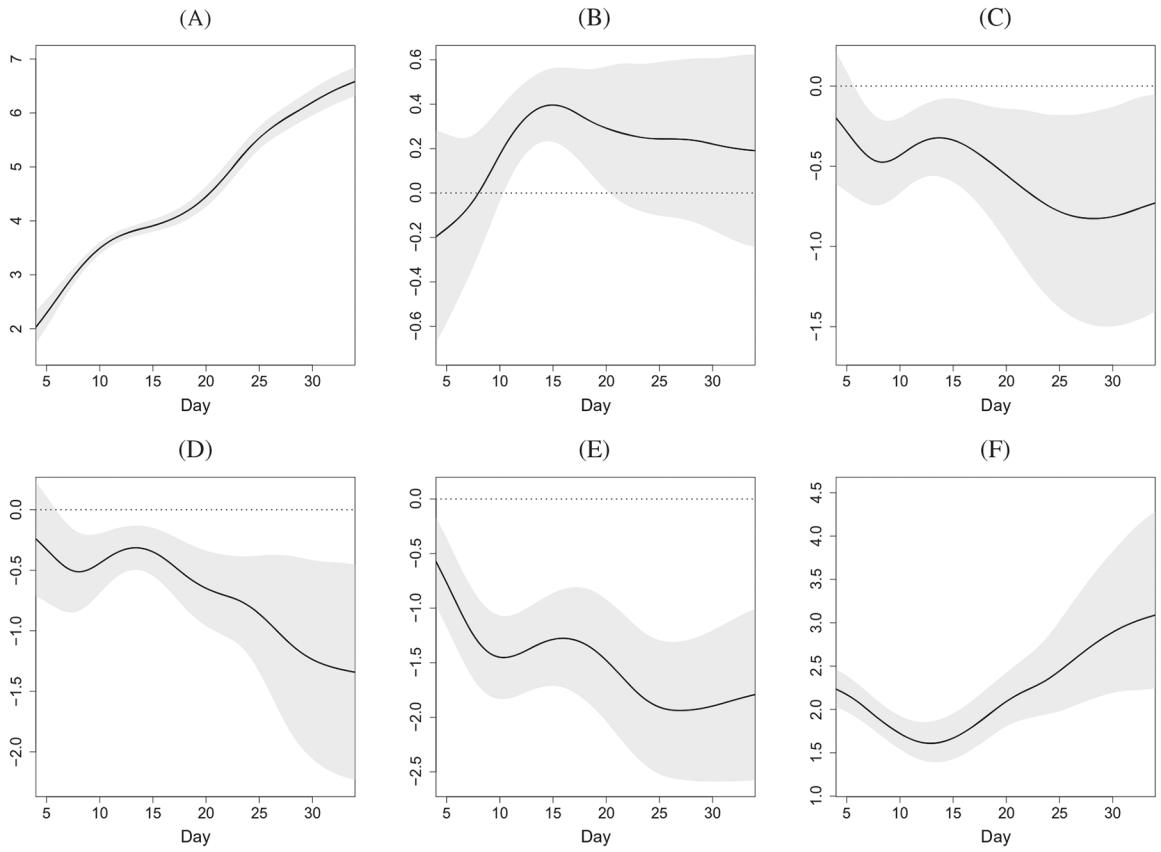

FIGURE 1.

Estimated time-varying coefficient functions (solid) for the analysis of natural logarithm of the longitudinal continuous response (tumor volume) along with the estimated 95% pointwise confidence band. For panels A-E and F, the y-axis shows the estimated value of the corresponding regression coefficient in terms of natural logarithm of mm^3 (tumor volume was measured in mm^3) and the estimated standard error, respectively. For each panel, the dotted line marks zero. A, Intercept; B, Vehicle; C, EXP; D, STD; E, EXP+STD; F, $\hat{\sigma}(t)$

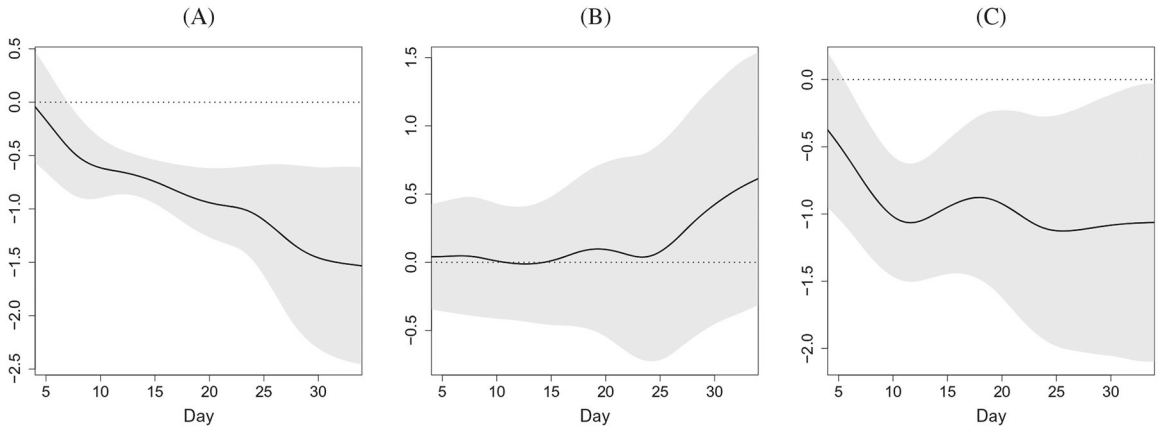

FIGURE 2.

The comparison of regression coefficients for the natural logarithm of the longitudinal continuous response (tumor volume). For each panel, the y-axis shows the difference between estimated values of the corresponding regression coefficients in terms of natural logarithm of mm^3 (tumor volume was measured in mm^3). In this Figure, panels A, B, and C show the differences between estimated regression coefficients of STD and vehicle groups $\left(\hat{\beta}_{12}(t) - \hat{\beta}_{11}(t)\right)$, EXP and STD groups $\left(\hat{\beta}_{13}(t) - \hat{\beta}_{12}(t)\right)$, and EXP+STD and EXP groups $\left(\hat{\beta}_{14}(t) - \hat{\beta}_{13}(t)\right)$, respectively

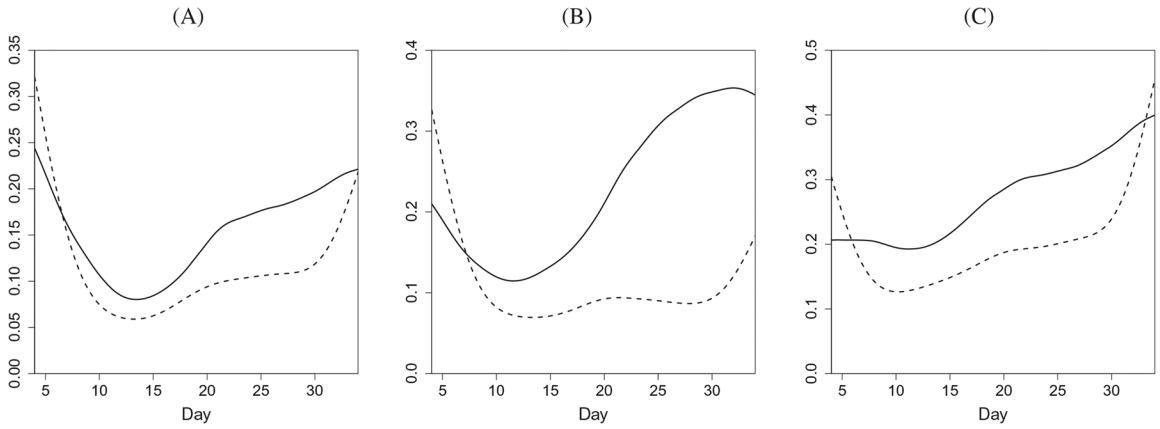

FIGURE 3.

The comparison of standard errors for joint and univariate analyses of the natural logarithm of the longitudinal outcome (tumor volume). In this univariate analysis, the other outcome was included as an additional predictor. For each panel, the y-axis shows the value of the estimated standard error, and the solid and dashed curves show the standard errors for the univariate and joint models, respectively. A, Vehicle; B, EXP; C, EXP+STD

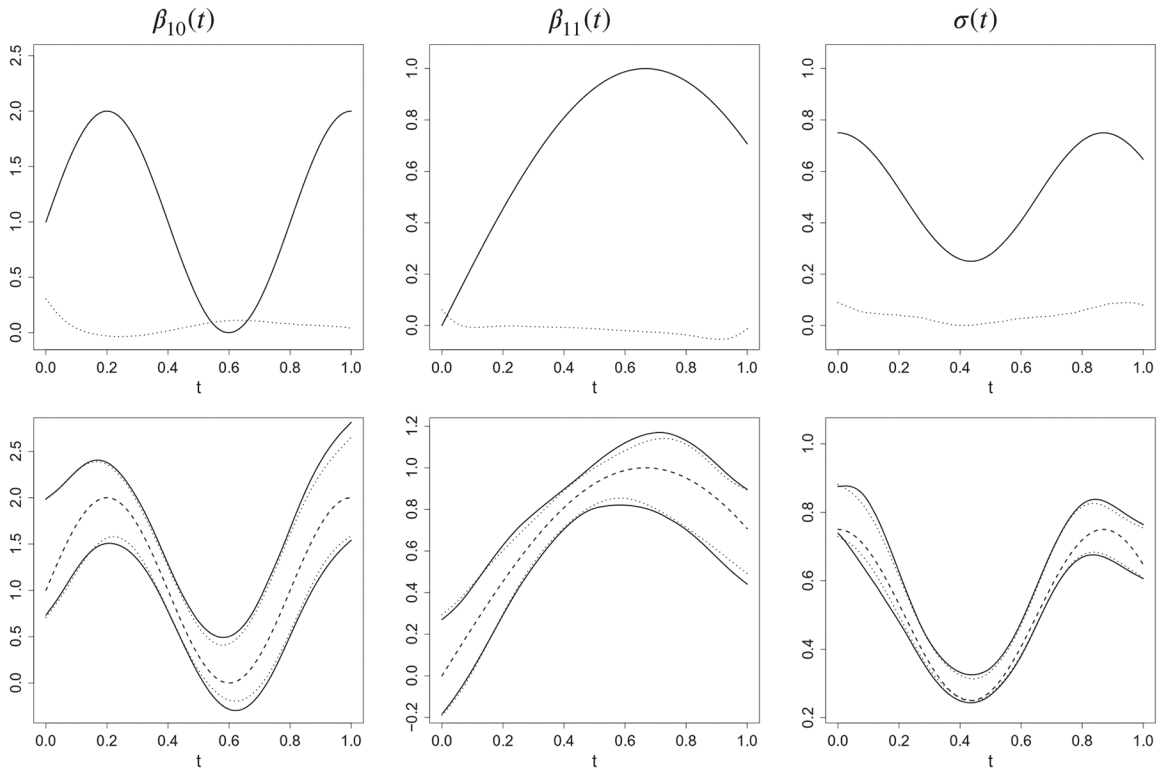

FIGURE 4.

Bias and estimated pointwise 95% confidence band for each time-varying coefficient function based on 500 Monte Carlo simulation runs. The first row shows the true function (solid) along with the empirical bias of our estimator (dotted). The second plot shows the empirical pointwise 95% confidence band (solid) and the mean theoretical pointwise 95% confidence band (dotted) along with the true function (dashed)

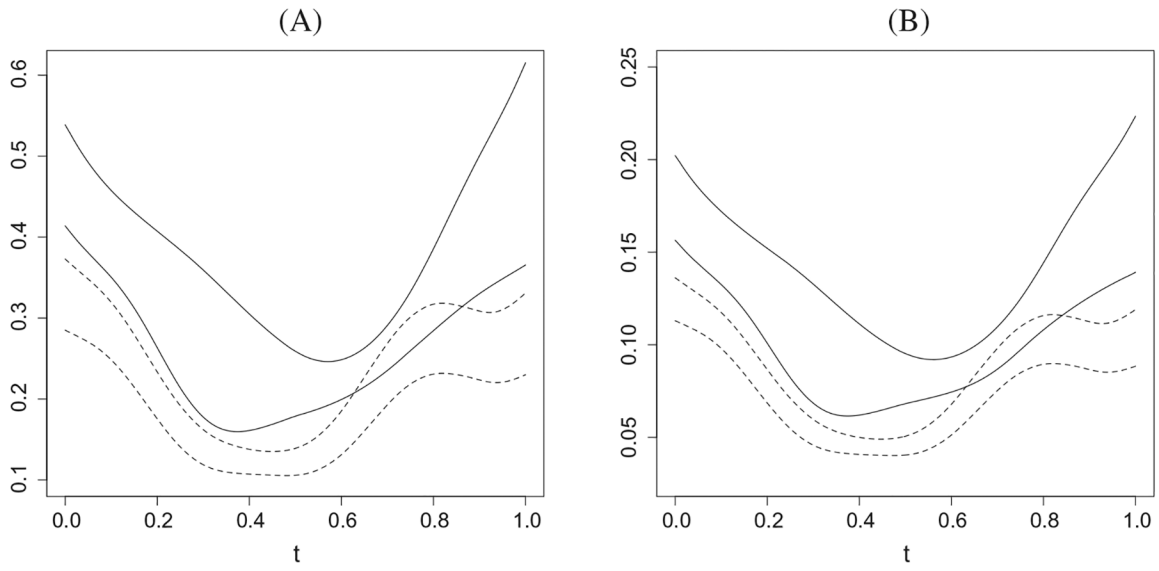

FIGURE 5.

The comparison of standard errors for joint and univariate analyses of the longitudinal outcome for each regression coefficient. In this univariate analysis, the other outcome was included as an additional predictor. For each plot, the y-axis shows the value of the estimated standard error, the solid and dashed curves show the 2.5% and 97.5% percentiles of the estimated standard errors for the univariate and joint models, respectively. A, $\hat{\beta}_{10}(t)$; B, $\hat{\beta}_{11}(t)$

TABLE 1

Estimated regression coefficient for each time-invariant outcome along with 95% confidence interval and a comparison of standard errors obtained in joint and marginal modeling

Variable	$\hat{\beta}_{2.}$ (Joint Modeling)	95% CI (Joint Modeling)	$\hat{\sigma}_{\hat{\beta}_{2.}}$ (Joint Modeling)	$\hat{\sigma}_{\hat{\beta}_{2.}}$ (Marginal Modeling)
Intercept	1.917	(1.762, 2.072)	0.079	0.145
Vehicle	−0.125	(−0.280, 0.029)	0.079	0.247
EXP	0.393	(−0.038, 0.825)	0.220	0.212
STD	0.212	(−0.071, 0.496)	0.145	0.222
EXP+STD	0.206	(−0.081, 0.492)	0.146	0.223

TABLE 2

Bias and mean square error (MSE) for the regression coefficients of the time-invariant outcome along with a comparison of standard errors obtained in joint and marginal modeling based on 500 Monte Carlo simulation runs

Regression Coefficient	Bias (Joint Modeling)	MSE (Joint Modeling)	2.5% and 97.5% Percentiles $\hat{\sigma}_{\hat{\beta}_{2.}}$ (Joint Modeling)	2.5% and 97.5% Percentiles $\hat{\sigma}_{\hat{\beta}_{2.}}$ (Marginal Modeling)
$\hat{\beta}_{20}$	0.008	0.001	(0.113, 0.170)	(0.176, 0.236)
$\hat{\beta}_{21}$	0.007	0.002	(0.044, 0.060)	(0.065, 0.086)