Published in final edited form as:

Stat Med. 2018 November 20; 37(26): 3733–3744. doi:10.1002/sim.7838.

Exploring causality mechanism in the joint analysis of longitudinal and survival data

Lei Liu^{1,*}, Cheng Zheng², and Joseph Kang³

¹Division of Biostatistics, Washington University in St. Louis, St. Louis, MO

²Joseph. J. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI

³Centers for Disease Control and Prevention, Atlanta, GA

Abstract

In many biomedical studies, disease progress is monitored by a biomarker over time, e.g., repeated measures of CD4 in AIDS, and hemoglobin in end stage renal disease (ESRD) patients. The endpoint of interest, e.g., death or diagnosis of a specific disease, is correlated with the longitudinal biomarker. In this paper, we examine and compare different models of longitudinal and survival data to investigate causal mechanisms, specifically, related to the role of random effects. We illustrate the methods by data from two clinical trials: an AIDS study and a liver cirrhosis study.

Keywords

mediation analysis; moderator; interaction; shared random effects; repeated measures

1. INTRODUCTION

In many biomedical studies, disease progress is monitored by a biomarker over time, e.g., CD4 counts in AIDS studies, hemoglobin levels in end stage renal disease (ESRD) patients, etc. The endpoint of interest, e.g., death or diagnosis of a specific disease, is correlated with the longitudinal biomarker. The relation between the longitudinal biomarker and the time to event is of interest. For example, in the HIV disease, researchers would like to investigate whether laboratory tests, e.g., CD4 counts, CD4/CD8 cell ratios, and viral loads, are surrogate markers of disease progression. Since the true effects of anti-HIV drugs in slowing the HIV disease may not be obvious in many years, changes in these markers can be used to reveal whether treatment will have any long-term benefits in a timely manner^{1,2}. Other potential surrogate endpoints include prostate specific antigen for prostate cancer³, and intraocular pressure, visual fields, and structural measurements of the optic nerve in glaucoma clinical trials⁴.

* lei.liu@wustl.edu.

Data and software availability

The SAS codes are available at <https://github.com/joyfulstones>. Both the CPCRA and Prothrombin datasets are available in the JM package in R. For introduction, please visit <https://cran.r-project.org/web/packages/JM/JM.pdf>.

We consider two scenarios and their extensions on the relation between the longitudinal biomarker (denoted by Y) and the time to event (denoted by T): (i) Y causes T ; (ii) T and Y are both influenced by another (possible unobserved) variable U . Since the observation of the longitudinal biomarker cannot happen after the time to event (e.g., death), we would not consider the situation that T causes Y .

In practice, the first scenario: “ Y causes T ” is often described by including the longitudinal biomarker Y as a time dependent covariate in a Cox model for the time to event T . Shared latent variable models are often used to describe the second scenario when T and Y are both influenced by another unobserved variable U . Such a latent variable could be continuous, e.g., random effects^{5–11}, or discrete, e.g., latent class^{12,13}, or both¹⁴. Recently, joint models have also been incorporated in the machine learning framework through statistical boosting¹⁵ and variable selection by adaptive LASSO¹⁶. In this paper we will focus on shared random effects models, though the extension to shared latent class models is straightforward in principle.

Although several papers have been published for mediation analysis on longitudinal data^{17–20} and on survival data^{21–26}, there is limited research on the causality in the joint analysis of longitudinal and survival data^{27–30}. Joffe and Greene²⁹ considered causal frameworks for surrogate outcomes. However, they did not take into account the longitudinal nature of the biomarker. Ibrahim et al.³⁰ gave a brief introduction to the causality in the joint models of longitudinal and survival data. On the other hand, the focus of most existing causal mediation analysis methods is to estimate the natural direct effect (NDE) and natural indirect effect (NIE). Zheng and van der Laan³¹ provided flexible nonparametric estimation of the NDE and NIE for the longitudinal exposure/mediator with the survival outcome under the sequential ignorability assumption. In the survival setting with a time-varying exposure or mediator, the interventional direct effect (IDE) and interventional indirect effect (IIE)¹⁹ are used to generalize the definition of the NDE and NIE. Lin et al.³² proposed a method with a SAS Macro to estimate the IDE and IIE using g-formula. However, the estimation of these effects fails to show the essential role of the shared random effect in such joint models since the latent variable is not directly modeled. In this paper, we focus on comparing different models to explore the causal mechanism related to the shared random effect instead of quantifying the NDE and NIE.

The rest of the paper is organized as follows. In Section 2 we present four models, in the increasing order of complexity, to study the role of random effects in the causal relation between the longitudinal biomarker and survival. In Section 3 we illustrate our methods using data from two clinical trials: an AIDS study and a liver cirrhosis study. In Section 4 we summarize our findings.

2. MODELS

For subject i , we first define a linear mixed model for the repeated measures of a biomarker (e.g., CD4) at time t , denoted by $Y_i(t)$:

$$Y_i(t) = Trt_i \beta_1 + X_i \beta_2 + t \beta_3 + a_i + e_{it}, \quad (0)$$

where Trt_i is the treatment indicator with coefficient β_1 ; X_i denotes other known covariates (e.g., age, gender, race, and time of measurement t) with coefficient (vector) β_2 ; we also include a linear time trend in t with slope β_3 ; a_i is the random intercept (latent variable) to characterize heterogeneity in repeated measures; e_{it} is the measurement error which is assumed to be independent of a_i , X_i , and Trt_i . For simplicity, we only consider a random intercept throughout the paper. Other random effects, e.g., random slope, can be included in the model as well. Also, an unknown smooth function $f(t)$ can replace the linear function $t\beta_3$ to denote a nonlinear temporal trend.

The random effects model assumption generally requires that a_i is independent of treatment and other observed confounders X . We could consider a_i as the unmeasured confounder. Of note, an auxiliary model can be used to take into account the dependence between a_i and X as in Mundlak³³ if a_i and X are correlated.

The interaction between treatment and other observed confounders ($Trt * X$) can be added in the model, but omitted here for simplicity.

2.1 Cox model with a time dependent covariate: Y causes T

A classical approach to evaluating the effect of a longitudinal biomarker on the time to event is to use observed values of the biomarker as a time-varying covariate in a survival model for the event time (e.g., Cox proportional hazards model³⁴). For example, in the kidney disease setting, Robinson et al.³⁵ and Bradbury et al.³⁶ applied this approach to study the effect of time-varying hemoglobin measures on survival.

Denote the death hazard for subject i by $\lambda_i(t)$. We can write the model as

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + Y_i(t) \eta), \quad (I)$$

where the baseline hazard for death is denoted by $\lambda_0(t)$; and $Y_i(t)$ is the longitudinal marker measured at time t in Model 0. Of note, Model 0 and Model I are independent and can be fitted separately, greatly simplifying the computation.

The direct acyclic graph (DAG) is shown in Figure 1 (see also Figure 1a in Joffe and Greene²⁹), in which the longitudinal biomarker denoted by $Y_i(t)$ is an intermediate outcome (i.e., potential mediator) between treatment and the survival event. The nonzero indirect effect of treatment on survival through the longitudinal biomarker depends on the nonzero coefficients β_1 in Model 0 and η in Model I. There could also be a direct effect of treatment on survival, denoted by the arrow from treatment to survival directly (with coefficient α_1 in Model I). Note in this DAG (and hereafter) we omit the symbol for the observed covariate X , which could affect both $Y_i(t)$ and survival.

Prentice³⁷ proposed three conditions to check whether Y is the surrogate marker for the true endpoint T : (a) Y must be correlated with the true endpoint T ; (b) Y must be affected by treatment A ; and (c) treatment and the outcome should be conditionally independent given Y , i.e., $T \perp A | Y$. Condition (a) can be tested by H_{0a} : $\eta = 0$ in Model I, while Condition (b) can be tested by H_{0b} : $\beta_1 = 0$ in Model 0. Condition (c) is the same to testing the direct effect by H_{0c} : $\alpha_1 = 0$ in Model I. If H_{0c} holds, there is no direct arrow from “Treatment” to “Survival” in Figure 1. All the treatment effect on survival passes through the repeated measures of the biomarker.

We can also consider whether the longitudinal marker could be a moderator in the relation between treatment and survival, i.e., there exists interaction between treatment and $Y_i(t)$ on survival. The DAG is shown in Figure 2. The moderator effect is denoted by the vertical arrow from $Y_i(t)$ pointed to the arrow from “Treatment” to “Survival”. We add the interaction between treatment and $Y_i(t)$ in Model I:

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + Y_i(t) \eta + \omega Trt_i Y_i(t)) \quad (\text{I-B})$$

Thus, ω is used to denote the magnitude of the mediator moderator. If ω is significantly different from 0, then the biomarker is a mediator moderator for the treatment.

2.2 T and Y are both influenced by other (possible unobserved) variables

Model I has two major advantages: easy to interpret and easy to implement (estimate). However, this model requires that the biomarker should be observed continuously, while in practice, $Y_i(t)$ is often measured only at discrete time points. Although some popular *ad hoc* methods, e.g., last observation carried forward, can be used to impute $Y_i(t)$ at the event occurrence, such practice may be too crude, especially when the duration from the last biomarker measuring time to death is long. More importantly, the measurement with error could lead to bias toward the null hypothesis³⁷.

To reduce the bias, Wulfsohn and Tsiatis⁵ proposed a shared random effects model of repeated measures of CD4 counts and death. Instead of assuming that the death hazard $\lambda_i(t)$ depends on the observed biomarker $Y_i(t)$ as in Model I, they assumed that $\lambda_i(t)$ depends on the expected (underlying) value of $E(Y_i(t)|a_i)$, which is a function of the random effect after averaging over the measurement error.

An alternative formulation is to assume that $\lambda_i(t)$ depends on the random effects, e.g., random intercept and random slope separately, instead of the expected (underlying) value of $Y_i(t)$. This formulation can also remove from the survival model the redundant covariate effect present in the mixed model for the biomarker.

While Wulfsohn and Tsiatis⁵ is more appealing when the prediction of the survival outcome by the longitudinal biomarker is of interest, this alternative has been used more often in the sensitivity analysis of the impact of informative dropout on the longitudinal outcome. Liu and Liu⁹ compared these two formulations using real examples. Specifically, when only a random intercept is present, these two models are equivalent. We have

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + \gamma a_i), \quad (\text{II})$$

In Model II, the death hazard depends on the heterogeneity of repeated measures through the random effect a_i , which is shared with Model 0. The magnitude of the association between the longitudinal and survival outcomes is denoted by the coefficient γ . Of note, the linear term γa_i in Model II can be replaced by more sophisticated forms, e.g., polynomial terms to improve the approximation with the price of larger variability of the estimates.

Joint Models 0 and II provides an approach of sensitivity analysis to describe the impact of unknown confounders. We can use the following DAG to describe the shared random effects model for the joint analysis of longitudinal and survival data:

The conditional independence is often assumed: $T_i \perp Y_i | a_i$. In the causality term, conditioning on a_i blocks the backdoor path between the longitudinal biomarker and survival. In other words, after conditioning on all known confounders, T_i and Y_i are correlated only through shared a_i . If a_i is a known confounder, we can treat it as a covariate (as X) in Models 0 and II. However, in reality, a_i is unknown, therefore, the relation between T_i and Y_i cannot be identified. To make it identifiable, we can assume a known distribution for a_i . As a result, we can integrate it out in the marginal likelihood. Normal random effects are often assumed due to the ease of implementation, although other distributions can be taken as well in different models of sensitivity analysis^{38–40}.

Denote by $\pi_\theta(a_i)$ the density function for random effect a_i , where θ is the parameter, e.g., variance of the random effect. The likelihood for joint Models 0 and II can be written as:

$$\prod_{i=1}^n \int L_i(T_i | a_i) L_i(Y_i | a_i) \pi_\theta(a_i) da_i,$$

where $\underline{Y}_i$ represents the whole profile of $Y_i(t)$. For example, assume a normal random error $e_{it} \sim^{iid} N(0, \sigma_e^2)$, we have

$$L_i(\underline{Y}_i | a_i) = \frac{1}{(\sqrt{2\pi}\sigma_e)^{n_i}} \exp\left[-\frac{1}{2\sigma_e^2} \sum_{j=1}^{n_i} e_{ij}^2\right],$$

and

$$L_i(T_i | a_i) = [\lambda_0(t_i) \exp(Trt_i \alpha_1 + X_i \alpha_2 + \gamma a_i)]^{\Delta_i} \exp[-\Lambda_0(t_i) \exp(Trt_i \alpha_1 + X_i \alpha_2 + \gamma a_i)],$$

where t_i is the realization of the death time T_i ; $\Lambda_0(t_i)$ is the cumulative baseline hazard; Δ_i is the event indicator ($\Delta_i = 1$ for death and $\Delta_i = 0$ for censoring); n_i is number of observed repeated measures; $e_{it} = Y_i(t) - Trt_i \beta_1 - X_i \beta_2 - t \beta_3 - a_i$ is the measurement error.

As a summary, the shared random effects model for the longitudinal and survival data accommodates the correlation between longitudinal and survival outcomes through the shared random effect a_i . To block the back door path through a_i , we make a distributional assumption on a_i . If we also know the parametric relation between the confounder and the survival outcome, e.g., γa_i in Model II, we can integrate a_i out, and estimate the parameters in the resulting marginal likelihood.

2.3 Arrow from Y to T in the joint model

In Figure 3, there is no direct effect from Y to T . In this section we consider the addition of an arrow from Y to T in the joint model. Basically, we relax the conditional independence assumption $(T_i \perp Y_i | a_i)$ in Model II. The DAG is given below:

Figure 4 can be also considered as the addition of a_i in Figure 1, by adding the path from a_i to Y and from a_i to T . Of note, Figure 4 is the same as Figure 1b in Joffe and Greene⁹. However, they did not consider a_i as a random effect.

Model III is given as

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + Y_i(t) \eta + \gamma a_i). \quad (\text{III})$$

Thus, both Models I and II are special cases of Model III: Model I is obtained if $\gamma = 0$ in Model III, while Model II is obtained if $\eta = 0$. Model III can be used to investigate whether survival depends on the observed biomarker or unknown heterogeneity, or both.

2.4 Interaction between random effects and biomarker

In the joint model literature, few considered the interaction between random effects and repeated measures of the biomarker, which can be denoted by the vertical arrow from a_i to the arrow from $Y_i(t)$ to survival as shown in Figure 5. We can add the interaction between a_i and $Y_i(t)$ to describe such a moderating effect:

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + Y_i(t) \eta + \gamma a_i + \delta a_i Y_i(t)), \quad (\text{IV})$$

where δ is used to denote the magnitude of the moderating effect. If δ is significantly higher (lower) from 0, then there is a synergistic (antagonistic) effect between a_i and $Y_i(t)$ on survival.

2.5 Estimation

Model I - the Cox model with time dependent covariates of biomarkers can be fitted by SAS Proc PHREG, using the counting process notation⁴¹.

For Models II-IV, the estimation can be carried out by approximating the integration with respect to the random effect using, e.g., adaptive Gaussian quadrature method or Monte Carlo EM algorithm. A challenge is to tackle the high-dimensional nature of the baseline hazards parameter, which invalidates the standard asymptotic arguments for MLEs⁴². On the

other hand, a profile likelihood approach for joint models encounters the difficulty of reliable estimates for standard errors, which has to be tackled by computationally intensive bootstrap procedures⁴¹. Thus, a parametric baseline hazard, e.g., Weibull or piecewise constant, is often used in joint models (JM package in R⁴³; Liu et al⁴⁴). In this paper we assume a piecewise constant baseline hazard for death^{45,46}. For example, for the CPCRA study, we construct the piecewise constant baseline hazard by dividing the follow-up time into 2 month intervals, i.e., months 2, 4, 6..., so the baseline hazard and $Y(t)$ change at the same time. The implementation can be conveniently carried out in SAS Proc NLMIXED. In Liu and Huang⁴⁴, we have shown by extensive simulation studies that the resulted estimated cumulative baseline hazards are very close to the true values, demonstrating the consistency of our estimation procedure.

For estimation purpose, it is desirable to fit Models I-IV sequentially. For example, the estimates from Models I and II can be taken as initial values in estimating Model III. On the other hand, since the four models are nested (Models I and II are nested in Model III which is nested in Model IV), we can use likelihood ratio test to compare these models.

3. DATA ANALYSIS

3.1 CPCRA study

We illustrate our method first using data from the Terry Bein Community Programs for Clinical Research on AIDS (CPCRA)^{47,48}, which enrolled 467 patients. About half (230) were randomized to receive didanosine (ddI) and the others received zalcitabine (ddC). The primary endpoint is survival. During the follow-up (median: 13 months, range: 1–21 months), there are 100 and 88 deaths in the didanosine and zalcitabine groups, respectively.

The intermediate outcome is repeated measures of CD4 counts, which were taken at baseline and every two months up to 20 months for each patient. CD4 measures are skewed to the right and heteroscedastic (mean: 74; median: 30), and a small portion (2.3%) are zero. We take $\log(\text{CD4}+1)$ (hereafter denoted by $\log \text{CD4}$) to tackle these issues.

Five covariates are incorporated in both the repeated measures and survival models: treatment (Trt; 1 = *ddC*, 0 = *ddI*), gender (1 = *female*, 0 = *male*), previous opportunistic infection (PrevOI; 1 = *AIDS diagnosis at baseline*, 0 = *no AIDS diagnosis*), stratum (1 = *AZT intolerance*, 0 = *AZT failure*), and baseline hemoglobin (Base Hb; centered at mean = 12). Time since study entry (in years) is included as a covariate in Model 0 for repeated CD4 measures. The results from Model I to Model IV are given in Table 1.

From the log likelihood and AIC shown in Table 1, we find that Model II yields the best fit statistics (smallest AIC). Thus Model II should be preferred for the CD4 data. In this model, there is no significant difference between ddC and ddI in CD4 values. However, the treatment effect is marginally significant for death hazard, as ddC reduces the mortality rate by 25% ($p = 0.07$). For the relation between CD4 values and survival, the estimate for γ is -0.50 ($p < 0.0001$), suggesting the existence of an unobserved confounder which is correlated with both CD4 values and death hazard.

Results from other models give us deeper understanding of the causal mechanism, in particular with respect to the role of the random effects. For example, in Model IV, the estimate of δ is -0.07 ($p=0.19$), suggesting no interaction between random effect a_i and $Y_i(t)$. Thus, the random effect is not a moderator of the CD4 repeated measures. In Model III, the direct effect from treatment to survival is again marginally significant ($OR=0.77$, $p=0.10$). The estimate of η is not significant ($p=0.40$), while the estimate of γ is highly significant ($OR=0.53$, $p=0.0005$). Therefore, there is no evidence for the arrow from CD4 to survival, while the effect of random effect on survival (the arrow from a_i to T) is significant. The coefficients for the confounding (a_i) effect is -0.64 while the CD4 effect is 0.11 on survival. Therefore, unmeasured confounders, e.g., patient characteristics related to underlying health (comorbidity, clinical stage, and treatment history)^{49,50}, may be correlated with both CD4 and survival, which leads to the correlation between CD4 and survival. In contrast, CD4 itself has no direct significant effect on survival. Compared to the estimate of η ($\hat{\eta}=-0.32$, $p<0.0001$) which is highly significant in Model I, we can see that the confounding effect is so strong that it even changes the sign of $\hat{\eta}$.

Since the treatment effect on CD4 is not significant, there is no evidence for the indirect path from treatment to survival through CD4. Consequently, CD4 is not a mediator between treatment and survival. Furthermore, the second condition of the definition of surrogate marker¹⁵ (Y must be affected by treatment A) does not hold. Therefore, CD4 is not a surrogate marker for survival in this trial.

We conduct a sensitivity analysis by adding a quadratic term a_i^2 in Model III, i.e.,

$$\lambda_i(t) = \lambda_0(t) \exp(Trt_i \alpha_1 + X_i \alpha_2 + Y_i(t) \eta + \gamma' a_i + \gamma'_2 a_i^2).$$

We find that the estimate for γ'_2 is -0.10 and not significant ($p=0.13$). Thus, the quadratic term did not improve fit.

A large portion of CD4 measures are missing intermittently. We investigate the heterogeneity in the binary outcome of intermittent missingness by a generalized linear mixed model. We include the same covariates as in the joint models. We find that the variance of the random effect in this model is very close to 0 ($=6 \times 10^{-8}$), suggesting very little heterogeneity in the missingness process. Thus, the missing at random⁵¹ assumption is reasonable for the intermittent missingness of CD4 repeated measures.

To examine the model fit, we draw the residual plots for both the repeated measures of CD4 and the martingale residuals for the death hazard, along the lines of Liu et al.⁵² We show the results for Model IV in Figure 6. It can be seen that the residuals are close to 0, and no clear pattern can be found, suggesting a good fit of our model.

3.2 Prothrombin

Our second example comes from a clinical trial to compare prednisone versus placebo in liver cirrhosis patients⁵³. A total of 488 liver cirrhosis patients were followed longitudinally with an average of 6.1 repeated measures of prothrombin time. Prothrombin is a protein

made by the liver, which helps blood to make normal clots. The “prothrombin time” (PT) is used to measure how long it takes blood to form a clot. When the PT is high, it takes longer for the blood to clot, which usually happens because the liver is not making enough prothrombin. Therefore, a high PT is associated with a high likelihood of serious liver damage or cirrhosis.

The primary outcome is death: 292 subjects died, other were censored. We are interested in the causal relation of PT and the time to death, and fit the data by Models I-IV. The results are present in Table 2.

From the log likelihood and AIC shown in Table 2, we find that Model IV yields the best fit statistics (smallest AIC), which should be preferred for the prothrombin data.

In Model IV, the interaction between random effect and $Y(t)$ is positive and highly significant ($p < .0001$). Therefore, the random effect (unobserved confounder) moderates the effect of PT on survival in a synergistic way. Both the main effects of a_i and $Y(t)$ are negative and highly significant. Due to the small magnitude of interaction compared to the main effects, we expect that a higher level of random effect and observed PT are associated with a lower death rate.

In Model IV, we find the treatment effect on survival is not significant ($\hat{\alpha}_1 = -0.15$, $p = 0.20$), but the effect of PT is highly significant ($\hat{\eta} = -0.038$, $p < 0.0001$) on survival. Coupling with the highly significant effect of treatment on PT (slope = -6.79 , $p < 0.0001$), PT is a significant mediator: there exists an indirect path from treatment to survival through PT. Also, it can be seen that all three conditions on the surrogate marker are fulfilled, demonstrating PT’s surrogacy for survival. To test if PT is a mediator moderator, we include the interaction term of PT and treatment as in Model I-B. It is not significant ($p = 0.22$), indicating the absence of the mediator moderator.

Similar to the last subsection, we conduct a sensitivity analysis by adding a quadratic term a_i^2 in Model III. The estimate for γ_2' is close to 0 ($= 0.00023$) and not significant ($p = 0.21$), indicating that the random effect contributes to the survival in a linear way. Thus, the relation between PT and survival outcome is not affected by an unknown confounder in either linear or quadratic way.

We draw the residual plots for both the repeated measures of PT and the martingale residuals for the death hazard in Figure 7 for Model IV. We can see that the overall goodness-of-fit is acceptable, even though in the presence of a single potential outlier in the martingale residual plot.

3.3 Summary of the Application Studies

From Model I (Cox model with time dependent covariates), we find no indirect path from CD4 to survival, and CD4 is not a surrogate marker for death. PT has an indirect effect on survival, and is a surrogate marker for death. In Model II, we found a significant association (denoted by γ in Tables 1 and 2) between the random effect and the survival outcome in both CD4 and PT studies. From Model III, the association between CD4 repeated measures

and survival is due to unmeasured confounding (denoted by random effects), but no causality between CD4 itself and survival is found. In comparison, both the main effects of random effects and observed PT repeated measures are significant on survival. Finally, there could also exist interaction between random effect and biomarker, e.g., the unknown heterogeneity may be a moderator of the biomarker (as shown in Model IV for PT), which was not noted before in literature.

In conclusion, though Model II yields the same conclusion for both the CD4 and PT data, other 3 models clearly show differences in interpreting causality from the two data. Therefore, the commonly used joint models in the form of Model II may not tell the full story of the causal relation between the longitudinal biomarker and the survival outcome. Other models should be used to give a more thorough picture of the causality mechanism.

4. DISCUSSION

In this paper we used four models to explore causality in the relation between a longitudinal biomarker and the time to event. In practice, Model I (Cox model with time dependent covariates) is the most often used tool to study the association between treatment, repeated measures and survival due to its simplicity in implementation. Model II has gained much popularity during the past decades due to the availability of software for joint model with shared random effects, e.g., JM package⁴³ in R and Proc NLMIXED in SAS^{52,54}. Model III combines features of Models I and II, taking both as special cases. Finally, Model IV augments Model III by including an interaction between the observed biomarker values and the unobserved confounder denoted by random effects. The two data examples illustrate the methods well.

Of note, these models are not entirely new. Rather, they are extensions of current literature on joint models of longitudinal and survival data. However, our paper is among the first to consider the causality mechanism for parametric regression coefficients in such joint models. In particular, we illustrate the role of the random effects in the causality relation, which has been less studied before.

Identifiability issue may occur when the extended joint models fail to capture the true latent variable. No model can perfectly disclose the unseen fact, but at least the goodness-of-fit statistics of this paper suggest that our natural extensions of the traditional joint models help avoid any serious identifiability issues. Furthermore, the sensitivity analyses with an additional quadratic random effect in Model III for both CD4 and prothrombin data also suggest a reasonably good fit of the methods.

There are many sophisticated extensions of joint models. For example, the joint models with time-dependent coefficients have been studied by Song and Wang^{55,56}. Second, the measuring times of repeated measures may be informative, i.e., correlated to the values of the biomarkers and death, resulting in the so called “recurrent marker” data.⁵² Third, multiple longitudinal biomarkers, e.g., cholesterol and blood pressure, may be correlated with the event of interest, e.g., cardiovascular event. Causality analysis for these more sophisticated joint models merits further research interest.

In this manuscript, we focus on explaining the causal mechanism instead of quantifying the causal effect in each pathway. It will be of interest in our future study to estimate the pathway effect such as the NIE and NDE in the presence of unmeasured confounders as in the joint random effects model of longitudinal and survival data.

Acknowledgements

We are thankful to Dr. Xinhong Lin for helpful discussion.

Funding

This research is partly supported by AHRQ R01 HS020263.

References

1. Tsoukas CM, Bernard NF. Markers predicting progression of human immunodeficiency virus-related disease. *Clinical Microbiology Reviews*. 1994; 7:14–28. [PubMed: 8118788]
2. Hughes MD, Daniels MJ, Fischl MA, Kim S, Schooley RT. CD4 cell count as a surrogate endpoint in HIV clinical trials: a meta-analysis of studies of the AIDS Clinical Trials Group. *AIDS* 1998; 12:1823–32. [PubMed: 9792383]
3. Colloca G Prostate-specific antigen kinetics as a surrogate endpoint in clinical trials of metastatic castration-resistant prostate cancer: a review. *Cancer treatment reviews*. 2012; 38:1020–1026. [PubMed: 22503300]
4. Medeiros FA. Biomarkers and surrogate endpoints in glaucoma clinical trials. *British Journal of Ophthalmology* 2015; 99:599–603. [PubMed: 25034049]
5. Wulfsohn MS and Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997; 53: 330–339. [PubMed: 9147598]
6. Liu L, Wolfe RA, and Kalbfleisch JD. A shared random effects model for censored medical costs and mortality. *Statistics in Medicine* 2007; 26: 139–155. [PubMed: 16526006]
7. Liu L, Ma JZ, and O’Quigley J. Joint analysis of multi-level repeated measures data and survival: an application to the end stage renal disease (ESRD) data. *Statistics in Medicine* 2008; 27: 5679–5691. [PubMed: 18693300]
8. Liu L and Huang X. Joint analysis of correlated repeated measures and recurrent events processes in the presence of a dependent terminal event. *Journal of Royal Statistical Society Series C: Applied Statistics* 2009; 58: 65–81.
9. Liu Y and Liu L. Joint models for longitudinal and time-to-event occurrence Book chapter in the *Routledge International Handbook of Advanced Quantitative Methods in Nursing Research*, edited by Henly Susan. Routledge, 2016.
10. Tsiatis A and Davidian M. Joint modelling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; 14: 809–834.
11. Rizopoulos D Joint models for longitudinal and time-to-event data: with applications in R *Chapman & Hall/CRC Biostatistics Series*. Taylor & Francis, 2012.
12. Lin H, Turnbull BW, McCulloch CE, Slate EH. 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* 2002; 97: 53–65.
13. Proust-Lima C, Taylor JM. Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach. *Biostatistics* 2009; 10: 535–549. [PubMed: 19369642]
14. Liu Y, Liu L, Zhou J. A joint latent class model of survival and longitudinal data. *Computational Statistics & Data Analysis* 2015; 91: 40–50.
15. Waldmann E, Taylor-Robinson D, Klein N, Kneib T, Pressler T, Schmid M, Mayr A. Boosting joint models for longitudinal and time-to-event data. *Biometrical Journal* 2017; 59: 1104–121. [PubMed: 28321912]

16. He Z, Tu W, Wang S, Fu H, Yu Z. Simultaneous variable selection for joint models of longitudinal and survival outcomes. *Biometrics* 2015; 71: 178–187. [PubMed: 25223432]
17. Kenny DA, Korchmaros JD, Bolger N. Lower level mediation in multilevel models. *Psychological Methods* 2003; 8:115–128. [PubMed: 12924810]
18. Bind M-AC, Vanderweele TJ, Coull BA, et al. Causal mediation analysis for longitudinal data with exogenous exposure. *Biostatistics* 2016; 17: 122–134. [PubMed: 26272993]
19. VanderWeele TJ and Tchetgen EJ. Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society, Series B* 2017; 79: 917–938.
20. Bind M-AC, Vanderweele TJ, Schwartz JD, Coull BA. Quantile causal mediation analysis allowing longitudinal data. *Statistics in Medicine*. 2017; 36:4182–4195 [PubMed: 28786129]
21. Lange T and Hansen JV. Direct and indirect effects in a survival context. *Epidemiology* 2011; 22: 575–581 [PubMed: 21552129]
22. Martinussen T, Vansteelandt S, Gerster M, et al. Estimation of direct effects for survival data by using the Aalen additive hazards model. *Journal of the Royal Statistical Society, Series B* 2011; 73:773–788.
23. Zheng C, and Zhou XA. Causal mediation analysis on failure time outcome without sequential ignorability. *Lifetime data analysis* 2017; 23, 533–559. [PubMed: 27464958]
24. Huang YT, Yang HI Causal Mediation Analysis of Survival Outcome with Multiple Mediators. *Epidemiology*. 2017;28(3):370–378. [PubMed: 28296661]
25. Luo P, Geng Z Causal mediation analysis for survival outcome with unobserved mediator–outcome confounders. *Computational Statistics & Data Analysis* 2016; 93:336–347.
26. Gelfand Lois A., MacKinnon David P., DeRubeis Robert J. and Baraldi Amanda N.. Mediation Analysis with Survival Outcomes: Accelerated Failure Time vs. Proportional Hazards Models. *Front. Psychol*, 30 3 2016
27. Zheng W and van der Laan M. Longitudinal mediation analysis with time-varying mediators and exposures, with application to survival outcomes. *Journal of Causal Inference* 2017; 5: Issue 2, 20160006. [PubMed: 29387520]
28. Lin SH, Young JG, Logan R, VanderWeele TJ, Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders. *Statistics in Medicine*. 2017;36:4153–4166. [PubMed: 28809051]
29. Joffe MM and Greene T. Related causal frameworks for surrogate outcomes. *Biometrics* 2009; 65: 530–538. [PubMed: 18759836]
30. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *Journal of Clinical Oncology* 2010; 28: 2796–2801. [PubMed: 20439643]
31. Zheng W and van der Laan M. Longitudinal mediation analysis with time-varying mediators and exposures, with application to survival outcomes. *Journal of Causal Inference* 2017; 5: Issue 2, 20160006 [PubMed: 29387520]
32. Lin SH, Young JG, Logan R, VanderWeele TJ, Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders. *Statistics in Medicine*. 2017; 36:4153–4166. [PubMed: 28809051]
33. Mundlak Y On the Pooling of Time Series and Cross Section Data. *Econometrica* 1978; 46:69–85
34. Cox DR. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* 1972; 34, 187–220.
35. Robinson BM, Joffe MM, Berns JS, et al. Anemia and mortality in hemodialysis patients: Accounting for morbidity and treatment variables updated over time. *Kidney International* 2005; 68: 2323–2330. [PubMed: 16221236]
36. Bradbury BD, Wang O, Critchlow CW, et al. Exploring relative mortality and epoetin alfa dose among hemodialysis patients. *American Journal of Kidney Disease* 2008; 51: 62–70.
37. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in Medicine* 1989; 8: 431–40. [PubMed: 2727467]
38. Liu L and Yu Z. A likelihood reformulation method in non-normal random effects models. *Statistics in Medicine* 2008; 27: 3105–3124. [PubMed: 18038445]

39. Rizopoulos D, Verbeke G, Lesaffre E, et al. A two-part joint model for the analysis of survival and longitudinal binary data with excess zeros. *Biometrics* 2008; 64: 611–619. [PubMed: 17725808]
40. Tsonaka R, Verbeke G, Lesaffre E. A semi-parametric shared parameter model to handle nonmonotone nonignorable missingness. *Biometrics* 2009; 65: 81–87. [PubMed: 18373713]
41. SAS documentation. available at https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_phreg_sect050.htm
42. Hsieh F, Tseng Y-K, Wang J-L. Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics* 2006; 62: 1037–1043. [PubMed: 17156277]
43. Rizopoulos D The JM package. Available at <http://cran.rproject.org/web/packages/JM/JM.pdf>
44. Liu L, Huang X. The use of Gaussian quadrature in frailty proportional hazards models. *Statistics in Medicine* 2008; 27: 2665–2683. [PubMed: 17910008]
45. Liu L Joint modeling longitudinal semi-continuous data and survival, with application to longitudinal medical cost data. *Statistics in Medicine* 2009; 28: 972–986. [PubMed: 19040210]
46. Liu L, Huang X, Yaroshinsky A, Cormier J. Joint frailty models for zero-inflated recurrent events in the presence of a terminal event. *Biometrics* 2016; 72: 204–214. [PubMed: 26295794]
47. Abrams DI, Goldman AI, Launer C, et al. A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine* 1994; 330: 657–662. [PubMed: 7906384]
48. Neaton JD, Wentworth DN, Rhame F, et al. Considerations in choice of a clinical endpoint for AIDS clinical trials. *Statistics in Medicine* 1994; 13: 2107–2125. [PubMed: 7846414]
49. Hansen AB, Gerstoft J, Kirk O, et al. Unmeasured confounding caused slightly better response to HAART within than outside a randomized controlled trial. *J Clin Epidemiol.* 2008; 61: 87–94. [PubMed: 18083465]
50. Shen Z, Zhu Q, Tang Z, et al. Effects of CD4 cell counts and viral load testing on mortality rates in patients with HIV infection receiving antiretroviral treatment: an observational cohort study in rural southwest China. *Clin Infect Dis.* 2016; 63: 108–114. [PubMed: 27001800]
51. Little RJA, Rubin DB. *Statistical analysis with missing data* (2nd ed.). New York, NY: John Wiley & Sons, 2002.
52. Liu L, Huang X, O’Quigley J. Analysis of longitudinal data in the presence of informative observational times and a dependent terminal event, with application to medical cost data. *Biometrics* 2008; 64: 950–958. [PubMed: 18162110]
53. Andersen PK, Borgan O, Gill RD, et al. *Statistical Models Based on Counting Processes*. New York: Springer, 1993.
54. Vonesh EF, Greene T, Schluchter MD. Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine* 2006; 25: 143–163. [PubMed: 16025541]
55. Song X, Wang CY. Semiparametric Approaches for Joint Modeling of Longitudinal and Survival Data with Time-Varying Coefficients. *Biometrics* 2008, 64: 557–566 [PubMed: 17725812]
56. Song X, Wang L. Partially time-varying coefficient proportional hazards models with error-prone time-dependent covariates—an application to the AIDS Clinical Trial Group 175 data. *Ann. Appl. Stat* 2017, 11:274–296.

Figure 1.
DAG for mediation analysis.

Figure 2.
Mediator moderator.

Figure 3.
DAG for the shared random effects model

Figure 4.
DAG for Model III with an arrow from Y to T .

Figure 5.
Interaction of random effect and biomarker.

Figure 6: Residual plots for the Community Programs for Clinical Research on AIDS data. A, CD4 residuals vs treatment; B, Martingale residuals vs treatment; C, CD4 residuals vs time; D, Martingale residuals vs baseline hemoglobin

Figure 7:
Residual plots for the prothrombin data. A, Prothrombin index (PI) residuals vs treatment;
B, martingale residuals vs treatment; C, PI residuals vs time

Table I:

Results from the CPCRA data

	Model I			Model II			Model III			Model IV		
Longitudinal												
Intercept	4.21	0.18	<.0001	4.09	0.18	<.0001	4.09	0.18	<.0001	4.10	0.18	<.0001
Trt	-0.11	0.11	0.31	-0.12	0.11	0.29	-0.11	0.11	0.29	-0.11	0.11	0.31
Year	-0.64	0.05	<.0001	-0.68	0.04	<.0001	-0.69	0.04	<.0001	-0.69	0.04	<.0001
Gender	0.26	0.19	0.18	0.26	0.19	0.18	0.26	0.19	0.18	0.26	0.19	0.17
PrevOI	-0.96	0.15	<.0001	-0.96	0.14	<.0001	-0.96	0.14	<.0001	-0.97	0.14	<.0001
Stratum	0.10	0.14	0.45	0.11	0.13	0.43	0.11	0.14	0.42	0.10	0.13	0.45
HgB	0.24	0.033	<.0001	0.24	0.03	<.0001	0.24	0.03	<.0001	0.24	0.03	<.0001
Survival												
Trt	-0.34	0.15	0.02	-0.29	0.16	0.07	-0.27	0.16	0.10	-0.29	0.16	0.08
Gender	0.023	0.25	0.93	-0.13	0.26	0.63	-0.17	0.28	0.55	-0.14	0.28	0.61
PrevOI	0.71	0.23	0.0022	1.09	0.24	<.0001	1.20	0.28	<.0001	1.14	0.29	<.0001
Stratum	-0.085	0.16	0.60	-0.11	0.17	0.52	-0.13	0.18	0.49	-0.12	0.18	0.52
HgB	-0.31	0.05	<.0001	-0.4	0.05	<.0001	-0.43	0.07	<.0001	-0.43	0.07	<.0001
Association												
γ				-0.50	0.08	<.0001	-0.64	0.18	0.0005	-0.41	0.25	0.10
η	-0.32	0.06	<.0001				0.11	0.14	0.40	0.07	0.14	0.60
δ										-0.07	0.06	0.19
Variance Components												
σ_a^2	1.24	0.09	<.0001	1.22	0.09	<.0001	1.23	0.09	<.0001	1.22	0.09	<.0001
σ_e^2	0.41	0.02	<.0001	0.41	0.02	<.0001	0.41	0.02	<.0001	0.41	0.02	<.0001
Fit Statistics												
-2 logL		27583			27453			27453			2743.4	
ATC		28083			27953			27973			2797.4	

γ is the coefficient for random effect a in the survival model

η is the coefficient for $Y(t)$ in the survival model

g is the coefficient for interaction of $Y(t)$ and z in the survival model

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

Results from the Prothrombin data

	Model I			Model II			Model III			Model IV		
Longitudinal												
Intercept	76.2	1.4	<.0001	76.0	1.4	<.0001	76.0	1.4	<.0001	76.0	1.3	<.0001
Trt	-6.50	1.90	0.0006	-6.51	1.92	0.0007	-6.79	1.90	0.0004	-6.79	1.87	0.0003
Time	1.74	0.17	<.0001	1.45	0.17	<.0001	1.77	0.18	<.0001	1.81	0.18	<.0001
Survival												
Trt	-0.14	0.12	0.25	0.07	0.14	0.51	-0.07	0.12	0.55	-0.15	0.12	0.20
Association												
γ				-0.039	0.004	<.0001	0.002	0.005	0.70	-0.028	0.009	0.0013
η	-0.035	0.002	<.0001				-0.036	0.004	<.0001	-0.038	0.004	<.0001
δ										0.0005	0.0001	<.0001
Variance Components												
σ_a^2	359	28	<.0001	359	27	<.0001	359	29	<.0001	359	28	<.0001
σ_e^2	345	10	<.0001	345	10	<.0001	344	10	<.0001	344	10	<.0001
Fit Statistics												
-2 logL		22602			22694			22601			22582	
AIC		22636			22728			22637			22620	

γ is the coefficient for random effect a in the survival model
 η is the coefficient for $Y(t)$ in the survival model
 δ is the coefficient for interaction of $Y(t)$ and a in the survival model