Research Article

Statistics in Medicine

Received 19 June 2012, Accepted 26 April 2013 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5857

Causal inference for bivariate longitudinal quality of life data in presence of death by using global odds ratios

Keunbaik Lee[a] and Michael J. Daniels[b*†]

In longitudinal clinical trials, if a subject drops out due to death, certain responses, such as those measuring quality of life (QoL), will not be defined after the time of death. Thus, standard missing data analyses, e.g., under ignorable dropout, are problematic because these approaches implicitly 'impute' values of the response after death. In this paper we define a new survivor average causal effect for a bivariate response in a longitudinal quality of life study that had a high dropout rate with the dropout often due to death (or tumor progression). We show how principal stratification, with a few sensitivity parameters, can be used to draw causal inferences about the joint distribution of these two ordinal quality of life measures. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords: principal stratification; ordinal data; missing data

1. Introduction

In a longitudinal clinical trial, if a subject drops out because of death, the responses will often not be defined after the dropout time. Thus, standard missing data analyses are problematic because they implicitly (or explicitly) 'impute' values of response after dropout. In this circumstance, the joint distribution of the longitudinal responses and death times can be used to conduct an analysis [1–4]. However, this approach stratifies on death times (that occur after randomization), and thus the estimated treatment effect is not a causal one. Similar problems occur by restricting analyses to only the survivors on each treatment arm.

To define the causal effect of treatment, we first introduce potential outcomes. Potential outcomes are all the outcomes that would be observed if each of the treatments had been applied to each of the subjects [5–7]. So for example, in the case of a binary treatment Z, the two potential outcomes would be $Y(Z)$ for $Z = 0, 1$. The causal effect of the treatment could be defined as the difference, $Y(1) - Y(0)$. In a randomized trial, this can be estimated as the differences in the sample mean for those randomized to each treatment. Causal inference becomes more complex when there is no randomization (e.g., in observational studies) and in settings such as when an outcomes of interest is unobserved because of death (as mentioned previously). We will explore the latter situation.

'Principal stratification' uses potential outcomes and partitions subjects into sets with respect to post-treatment variables [8]. The principal strata are not affected by treatment assignment. Causal effects are defined within these principal strata. To estimate the causal effect of treatment in our setting, one can define a 'survivor average causal effect (SACE)' [8–11]; we describe this further in Section 3, but it is

[a] *Associate Professor, Department of Statistics, Sungkyunkwan University, Seoul, 110-745, Korea*
[b] *Professor, Section of Integrative Biology, Division of Statistics & Scientific Computation, University of Texas at Austin, Austin, TX 78712, U.S.A.*
**Correspondence to: Michael J. Daniels, Section of Integrative Biology, Division of Statistics & Scientific Computation, University of Texas at Austin, Austin, TX 78712, U.S.A.*
[†] *E-mail: mjdaniels@austin.utexas.edu*

a causal treatment comparison among those who would survive on both treatment arms (which corresponds to one of the principal strata). Here, survival is the post-treatment variable. Egleston *et al.* [9] proposed assumptions to identify the SACE among survivors at a fixed time for a binary response and implemented a sensitivity analysis for some of those assumptions. Lee and Daniels [10] estimated the SACE for ordinal data in a two treatment arm setting, whereas Lee *et al.* [11] introduced a set of relevant SACEs in the setting of three treatment arms with ordinal outcomes. In a related setting, Cheng [12] defined an average causal effect for univariate ordinal outcomes.

In this paper, we define a novel SACE for *bivariate ordinal* outcomes by using the ratio of global odds ratios [13]. We also present a (simple) set of assumptions to identify the SACE. Instead of just using the data at the time point of interest (the end of the study) to estimate the SACE, we will use all available longitudinal data via models for bivariate longitudinal data which is essential because of the high rate of dropouts unrelated to death in this study. We do this by introducing a partial ignorability assumption [14] that is weaker than that used in earlier work by conditioning on baseline covariates [11].

The paper is organized as follows. In Section 2, we describe our motivating dataset. In Section 3, we formally define the causal effect of interest, provide some assumptions that are necessary to identify it, and then show how it can be estimated from these assumptions along with the observed data. In Section 4, we analyze data from the longitudinal QoL study. Finally, conclusions and extensions are provided in Section 5.

2. Description of data

Our approach is motivated by data from a colorectal cancer clinical trial that was designed to compare survival among three treatments (IFL, FOLFOX, IROX) [15]. The main objective of this trial was to find a better treatment for colorectal cancer. The median survival for patients receiving FOLFOX was higher than IROX and IFL. IFL was the standard of care at the time of the trial. Secondarily, it was of interest to determine if the treatment differed with respect to QoL because the toxicity profiles were quite different on the treatment arms.

Previous work [4, 10, 11] focused on a single QoL measure. In this paper, we focus on two QoL measures, insomnia and outlook (measure of anxiety). Insomnia (outlook) was measured on a five-point ordinal scale: (1) it is almost impossible for me to get a decent night's sleep (I am worried and scared about things); (2) I have difficulty sleeping almost every night (I am worried and a little frightened about things); (3) I frequently have trouble getting to sleep and staying asleep (I am quite worried but unafraid); (4) I have occasional spells of sleeplessness (I am a little worried about things); (5) I sleep as well as I always have (I am not fearful or worried). The connection between these two is of interest as high insomnia is related to a more negative outlook (anxiety) [16].

Forty-one percent of those dropped out from the study did so because of tumor progression or death. A naive treatment comparison among the survivors on each treatment is not causal. We outline an approach to estimate the causal effect of the treatment difference (among survivors) next. In the following, we focus on (only) two of the treatments (IFL and FOLFOX).

3. Principal stratification approach

We first introduce some further notation for our setting. Denote Tx_i as the treatment indicator for subject i,

$$\mathrm{Tx}_i = \begin{cases} 0, & \text{if subject } i \text{ is assigned to IFL;} \\ 1, & \text{if subject } i \text{ is assigned to FOLFOX,} \end{cases}$$

and let $D_{it}(\mathrm{Tx}_i)$ be tumor-progression/death indicator for subject i at visit t,

$$D_{it}(\mathrm{Tx}_i) = \begin{cases} 0, & \text{alive;} \\ 1, & \text{tumor progressed or dead.} \end{cases}$$

Obviously, if $D_{it}(\mathrm{Tx}_i) = 1$, then $D_{is}(\mathrm{Tx}_i) = 1$ for $s > t$. To maintain clarity, we will refer to the tumor progression/death indicator $D_{it}(\cdot)$ as the death indicator in what follows.

We define four (principal) strata, $A_t(l, m) = \{i | (D_t(0), D_t(1)) = (l, m)\}$ on the basis of the pairs of death indicators (the post-treatment variable) at each time. Within each of these strata, we can define a causal effect. $A_t(0, 0)$ is the stratum for the subjects who would be alive under both arms at visit t;

$A_t(1,0)$ is for the subjects who would be alive under the experimental treatment (FOLFOX) but not alive under the standard of care treatment (IFL) at visit t; $A_t(0,1)$ is for the subjects who would be alive under IFL but not alive under the FOLFOX at visit t; $A_t(1,1)$ is for the subjects who would not be alive under either treatment at visit t.

Now, we define a more general set of potential outcomes for each subject under each treatment at each time. Let $Y_{it1}(\text{Tx}_i)$ and $Y_{it2}(\text{Tx}_i)$ be the potential outcomes for subject i at time t. The full set of potential variables for subject i at visit t is

$$\mathcal{P}_{it} = \{D_{it}(0), (Y_{it1}(0), Y_{it2}(0); D_{it}(0) = 0), D_{it}(1), (Y_{it1}(1), Y_{it2}(1); D_{it}(1) = 0)\},$$

where, for treatment tx, $Y_{it1}(\text{tx})$ and $Y_{it2}(\text{tx})$ are the potential responses if a subject i is alive at visit t ($D_{it}(\text{tx}) = 0$).

Let $Y_{it} = (Y_{it1}, Y_{it2})$. If the patient is alive ($D_{it}(\text{Tx}_i) = 0$), define R_{it} to be the indicator that $Y_{it} = Y_{it}(\text{Tx}_i)$ is observed. In the following, we assume monotone dropout ($R_{it} = 1 \Rightarrow R_{it-1} = 1$). So, the observed data for individual i is

$$\mathcal{O}_{it} = \{\text{Tx}_i, D_{it}, (R_{it}; D_{it} = 0), (Y_{it}; D_{it} = 0, R_{it} = 1)\}.$$

Next, we define the causal effect of interest that will be of interest in the strata corresponding to surviving on both treatment arms. We then discuss assumptions to identify it given the observed data.

Causal effect. We suppress the subscript i here for clarity. We define a causal effect that captures the treatment difference on the joint distribution of insomnia and outlook. More specifically, the causal effect of interest will be the ratio of generalized odds ratios (GOR) [17] under each treatment in the strata corresponding to surviving at time T on both treatments,

$$SACE_{k_1,k_2} = \frac{GOR(k_1, k_2, \text{Tx} = 1; A_T(0,0))}{GOR(k_1, k_2, \text{Tx} = 0; A_T(0,0))}. \tag{1}$$

The GOR for two ordinal responses is defined as

$$GOR(k_1, k_2, \text{Tx} = j; A_T(0,0)) = \frac{P(Y_{T1}(j) \leq k_1, Y_{T2}(j) \leq k_2 | A_T(0,0)) P(Y_{T1}(j) > k_1, Y_{T2}(j) > k_2 | A_T(0,0))}{P(Y_{T1}(j) \leq k_1, Y_{T2}(j) > k_2 | A_T(0,0)) P(Y_{T1}(j) > k_1, Y_{T2}(j) \leq k_2 | A_T(0,0))}. \tag{2}$$

Generalized odds ratios are convenient summaries of the relationship between ordered categorical responses and are an extension of the simple odds ratio for a 2×2 contingency table. As can be seen in (2), the GOR corresponds to collapsing the contingency table into a 2×2 table [18] via cumulative probabilities; in particular, collapsing (or dichotomizing) the ordinal variable $Y_{T1}(j)$ as $\leq k_1$ vs $> k_1$ and similarly for $Y_{T1}(j)$ with k_2. In our case, each GOR is the odds ratio of probabilities that both QoL responses at time T are smaller than k_1 and k_2 (in our example, high insomnia, and outlook) for subjects who would be alive under both treatment arms. $SACE_{k_1,k_2}$ is the ratio of GORs for the two treatment arms. k_1 and k_2 would be chosen by the investigator.

The ratio of odds ratios was first proposed as a measure of the interaction effects of two dichotomous exposure factors on a binary response in [19]. Here, we have extended this idea to accommodate bivariate ordinal outcomes in the setting of causal inference.

The GOR in (2) consists of four joint probabilities. By using the following relationship from [17], it is clear we will only need to estimate $F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0)) = P(Y_{T1}(j) \leq k_1, Y_{T2}(j) \leq k_2 | A_T(0,0))$ to estimate (2),

$$GOR(k_1, k_2, \text{Tx} = j; A_T(0,0)) = \frac{F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0)) \{1 - F_{iT1}(k_1 | \text{Tx} = j, A_T(0,0)) - F_{iT2}(k_2 | \text{Tx} = j, A_T(0,0)) + F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0))\}}{\{F_{iT1}(k_1 | \text{Tx} = j, A_T(0,0)) - F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0))\} \{F_{iT2}(k_2 | \text{Tx} = j, A_T(0,0)) - F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0))\}}, \tag{3}$$

where

$$F_{iT12}(k_1, k_2 | \text{Tx} = j, A_T(0,0)) = P(Y_{T1}(j) \leq k_1, Y_{T2}(j) \leq k_2 | A_T(0,0)),$$
$$F_{iT1}(k_1 | \text{Tx} = j, A_T(0,0)) = P(Y_{T1}(j) \leq k_1 | A_T(0,0)) = P(Y_{T1}(j) \leq k_1, Y_{T2}(j) \leq K_2 | A_T(0,0)),$$
$$F_{iT2}(k_2 | \text{Tx} = j, A_T(0,0)) = P(Y_{T2}(j) \leq k_2 | A_T(0,0)) = P(Y_{T1}(j) \leq K_1, Y_{T2}(j) \leq k_2 | A_T(0,0)),$$

for $j = 0, 1$. Next, we outline assumptions necessary to identify the SACE.

3.1. Assumptions for identification of a causal effect

We first make the stable unit treatment value assumption (SUTVA, [6]) which states that a subject's potential outcomes are unrelated to the treatment status of other subjects. Next, we list the other assumptions necessary to identify the causal effect of interest. However, before doing that, we need to define a few more quantities.

Let $g_T(\text{tx}, w) = P(D_T(\text{tx}) = 0|w)$ for $\text{tx} = 0, 1$ where w represent baseline covariates. g_T is the marginal probability of being alive at time T given baseline covariates. We assume that

$$\text{logit}[g_T(\text{tx}, w)] = w^T \psi_{\text{tx}}. \tag{4}$$

Note that we can easily be more flexible than the simple logistic regression model above (e.g., by using regression trees [20]). Also, define

$$p_T^*(w) = P(D_T(1) = 0|D_T(0) = 0, w).$$

$p_T^*(w)$ is the probability of being alive on FOLFOX at time T, given that also alive on IFL at time T for a patient with baseline covariates, w.

Assumption 1 (Ignorability)
$\text{Tx} \perp \bar{\mathcal{P}}_T$.
where $\bar{\mathcal{P}}_T = (\mathcal{P}_1, \cdots, \mathcal{P}_T)$ is the history of potential outcomes up to and including the outcomes at visit T.

Assumption 1 states that the treatment arm is unrelated to the set of potential outcomes. This holds since the treatments were randomized.

Assumption 2 (Stochastic monotonicity)
We assume $p_T^*(w)$ is given by

$$p_T^*(w) = g_T(1, w) + \rho\{U(w) - g_T(1, w)\}, \quad 0 \leqslant \rho \leqslant 1,$$

where $U(w) = \min\left\{1, \frac{g_T(1,w)}{g_T(0,w)}\right\}$. The term, $U(w)$, corresponds to the bounds on the conditional distribution given $g_T(1, w)$. This modeling assumption has been used previously in Roy *et al.* [21] and Lee *et al.* [11]. Note that if $\rho = 0$, then $p_T^*(w) = g_T(1, w)$, which implies that $D_T(1)$ is independent of $D_T(0)$. If $\rho = 1$, then $p_T^*(w) = U(w)$, which is the largest possible probability that is compatible with the marginal distributions. This assumption implies that the probability of being alive on the experimental arm (FOLFOX) is more likely conditional that one would also be alive on the standard of care arm (IFL).

Assumption 3 (Proportional odds reduction of sensitivity parameters)

$$\begin{aligned}
\frac{\text{odds}(Y_{T1}(1) \leqslant k_1, Y_{T2}(1) \leqslant k_2|A_T(1,0))}{\text{odds}(Y_{T1}(1) \leqslant k_1, Y_{T2}(1) \leqslant k_2|A_T(0,0))} &= \tau, \\
\frac{\text{odds}(Y_{T1}(0) \leqslant k_1, Y_{T2}(0) \leqslant k_2|A_T(0,1))}{\text{odds}(Y_{T1}(0) \leqslant k_1, Y_{T2}(0) \leqslant k_2|A_T(0,0))} &= \lambda,
\end{aligned} \tag{5}$$

where $0 < \tau, \lambda < \infty$. This assumptions characterizes the relationship between potential outcomes of interest in different strata.

From (5), the joint probability of insomnia and outlook under the FOLFOX arm among the group that is alive under the FOLFOX arm but not alive under the standard of care treatment at visit T is given by,

$$P(Y_{T1}(1) \leqslant k_1, Y_{T2}(1) \leqslant k_2|A_T(1,0)) = \frac{\tau P(Y_{T1}(1) \leqslant k_1, Y_{T2}(1) \leqslant k_2|A_T(0,0))}{1 - (1-\tau)P(Y_{T1}(1) \leqslant k_1, Y_{T2}(1) \leqslant k_2|A_T(0,0))}. \tag{6}$$

The functional form (6) was introduced in Marshall and Olkin [22] for a family of distributions from a survival function and Sankaran and Jayakumar [23] interpreted the family by using the odds function. The sensitivity parameters, τ and λ, are the odds ratios of the joint probability of insomnia and outlook under the FOLFOX arm among the group that is alive under the FOLFOX arm but not alive under the standard of care (IFL) treatment at visit T as compared with the group that is alive under both arms.

We assume the two sensitivity parameters, τ and λ, do not depend on k_1 and k_2 (proportional odds type assumption). A large value of τ means that the joint probability of insomnia and outlook under the FOLFOX arm among patients in stratum $A_T(1,0)$ is higher than that among patients in stratum $A_T(0,0)$. When $\tau = 1$, the two probabilities are same. We also have similar interpretation of λ for stratum $A_T(0,1)$ vs $A_T(0,0)$.

Similar assumptions in related contexts were used in Egleston *et al.* [9], Lee and Daniels [10], and Lee *et al.* [11] in the context of a *univariate* response. Here, we have provided a related extension in the more complex setting of a bivariate response. Assumptions with just a few sensitivity parameters, such as Assumption 3, are essential to conducting a feasible sensitivity analysis [24].

3.2. *Identification of the survivor average causal effect*

The following theorem confirms that the observed data, $\mathcal{O}_t$, along with Assumptions 1–3 are sufficient to identify the *SACE*.

Theorem 1
Under assumptions 1, 2, and 3, $SACE_{k_1,k_2}$ is identified.

The proof is given in the Supplementary Materials/Web Appendix[‡].

3.3. *Estimation of survivor average causal effect*

Recall that a lot of missing data in this study was due to dropout unrelated to death. To use all available data on survivors to estimate the SACE, we need one more assumption.

Assumption 4 (Conditional MAR)
$R_t \perp \{Y_t, \ldots, Y_T\} | D_T = 0, \bar{Y}_{t-1}, \mathrm{Tx}, w$ for $t < T$.

This assumption states that missingness of outcome at visit t is independent of the value of the outcome given the previous history of outcomes ($\bar{Y}_{t-1}$), treatment (Tx), and baseline covariates (w) for all subjects with progression/death time larger than T ($D_T = 0$). It is MAR conditional on progression/death time; assumptions such as these have been termed 'partially ignorable' [14]. This assumption will allow us to use *all* the available data on survivors for the analysis. The addition of baseline covariates is a simple weakening of the assumption proposed in [11].

Models for the observed data: To use all the available longitudinal data, we can use any joint model for bivariate longitudinal ordinal data. In this paper, we consider the following marginalized model that was introduced in Lee *et al.* [25],

$$\begin{aligned}
&\mathrm{logit}\{P(Y_{it1} \leqslant k | x_{it}, D_T = 0, \mathrm{tx})\} = \beta_{10k} + x_{it}^T \beta_1 + \alpha_1 \mathrm{tx}, && (7)\\
&\mathrm{logit}\{P(Y_{it2} \leqslant k | x_{it}, D_T = 0, \mathrm{tx})\} = \beta_{20k} + x_{it}^T \beta_2 + \alpha_2 \mathrm{tx}, && (8)\\
&\mathrm{logit}\{P(Y_{it1} \leqslant k | b_{it1}, x_{it}, D_T = 0, \mathrm{tx})\} = \triangle_{it1k} + b_{it1}, && (9)\\
&\mathrm{logit}\{P(Y_{it2} \leqslant k | b_{it2}, x_{it}, D_T = 0, \mathrm{tx})\} = \triangle_{it2k} + b_{it2}, && (10)\\
&b_i^T \sim \text{i.i.d. } N(0, \Sigma_i),
\end{aligned}$$

where $b_i = (b_{i11}, b_{i12}, \cdots, b_{iT1}, b_{iT2})^T$, x_{it} is a vector of covariates including w_i, and tx is an indicator for treatment (tx = 0(IFL) and 1 (FOLFOX)). In marginal models (7) and (8), β_{j0k} are intercepts for the models for each ordinal response that satisfy the monotonicity property,

$$\beta_{j01} < \cdots < \beta_{j0K-1},$$

β_j are the regression coefficients of the covariates x_{it}, and α_j are the coefficients of tx for $j = 1, 2$. The intercepts in the two random effects models (9) and (10), Δ_{itjk} are determined based on the relationship between the marginal, (7) and (8), and the conditional, (9) and (10), cumulative probabilities; they can be calculated by a Newton-Raphson algorithm. These types of marginalized models were first introduced in Heagerty [27]. For further details on the computations, see the Supplementary Materials/Web Appendix.

[‡]*Supporting information may be found in the online version of this article.*

We specify $\Sigma_i = R_i \otimes \Sigma_2$ where R_i is a correlation matrix, Σ_2 is a covariance matrix between responses at the same time, and $\otimes$ denotes the Kronecker product. We consider the following forms for R_i and Σ_2,

$$R_i = \begin{pmatrix} 1 & \rho_{i12} & \cdots & \rho_{i,1,n_i} \\ \rho_{i12} & 1 & \cdots & \rho_{i,2,n_i} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{i,1,n_i} & \rho_{i,1,n_i} & \cdots & 1 \end{pmatrix} \text{ and } \Sigma_2 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}. \tag{11}$$

R_i can be a relatively simple structure such as AR(1) or more complex by using partial autocorrelations ([28, 29]). The covariance matrix Σ_i captures the correlation of the observed bivariate responses, both longitudinally and inter-response. For further details on this model, see Lee *et al.* [25] and for details on the form of R_i in our application, see Section 4. As is typical in non-Bayesian approaches, we implicitly assume conditional independence between the potential outcomes. In principle, we could do sensitivity to this assumption, but the only impact would be minor changes (if any) to the standard errors.

To estimate the SACE, we need to estimate the joint probability of insomnia and outlook for those who did not progress/die before the end of the study, $P(Y_{iT1} \leqslant k_1, Y_{iT2} \leqslant k_2 | D_{iT} = 0, \text{tx})$. As stated earlier, this can be estimated using any model for the observed data. For the observed data model specified previously, this probability can be expressed as

$$\begin{aligned} &P\left(Y_{iT1} \leqslant k_1, Y_{iT2} \leqslant k_2 | D_{iT} = 0, \text{tx}\right) \\ &\quad = E_w \left\{ \int P(Y_{iT1} \leqslant k_1 | b_{iT1}, D_{iT} = 0, \text{tx}, w) P(Y_{iT2} \leqslant k_2 | b_{iT2}, D_{iT} = 0, \text{tx}, w) \phi(b_{iT}) \mathrm{d}b_{iT} \right\} \\ &\quad \stackrel{\text{let}}{=} E_w \left\{ h_{T,\text{tx}}(k_1, k_2, w) \right\}. \end{aligned} \tag{12}$$

The inner part of the integral in (12) can be estimated using the maximum likelihood estimates of the parameters (7)–(10). We prefer to make inference here unconditional on baseline covariates. As a result, we take expectations over the distribution of the baseline covariates, w. The expectations in (12) can be estimated using their empirical expectations. That is, for an arbitrary function $f(w; \text{tx})$, $E_w\{f(w; \text{tx})\}$, can be estimated as

$$\hat{E}_w \left\{ f(w; \text{tx}) \right\} = \frac{1}{N} \sum_{i=1}^{N} \hat{f}(w_i; \text{tx}).$$

Related work (Sjolander *et al.* [26]) defines the causal effect conditional on baseline covariates.

Overall algorithm to estimate the SACE: Now we present the complete algorithm to estimate the SACE. For a fixed ρ, τ, and λ,

(1) Estimate $g_T(\text{tx}, w)$ in (4) for tx = 0, 1. Then use these to estimate p_T^* in Assumption 2.
(2) Estimate $h_{T,\text{tx}}(k_1, k_2; w)$ for tx = 0, 1 in (12) by using the maximum likelihood estimation of the parameters in (7)–(10).
(3) Estimate $P(Y_{T_1}(\text{tx}) \leqslant k_1, Y_{T_2}(\text{tx}) \leqslant k_2 | A_T(0,0))$ for tx = 0, 1 by using the estimates in steps 1 and 2 and the expressions (A.5) and (A.8) given in the Supplementary Materials/Web Appendix.
(4) Estimate $GOR(k_1, k_2, \text{Tx} = \text{tx}; A_T(0,0))$ for tx = 0, 1 by plugging in the estimates from step 3 into the expression for the GOR in (3). Then, plug these estimates into (1) to estimate $SACE_{k_1,k_2}$.

To calculate standard errors for SACE, we use the delta method. See the Appendix for further details on the entire estimation algorithm.

4. Data analysis

We use the approach proposed in Section 3 to analyze the QoL data described in Section 2. The data consists of the 466 subjects randomized to either IFL or FOLFOX. The observation time of the response was defined as the (12-week) time window during which the survey was filled out (0=baseline, 1 = 1–84 days after going on study, 2 = 85–168 days after going on study,..., 5 = 337–420 days after going

Table I. Break down of time of progression/death (in 12-week windows) by treatment groups for QoL data. Proportions in parentheses. The last 12-week window corresponds to 'after study.'

	12-week window of progression/death						
Treatment	1	2	3	4	5	6	Total
IFL	15 (0.064)	18 (0.077)	25 (0.107)	21 (0.090)	29 (0.124)	126 (0.539)	234
FOLFOX	7 (0.030)	8 (0.034)	21 (0.091)	17 (0.073)	22 (0.095)	157 (0.677)	232
Total	22 (0.047)	26 (0.056)	46 (0.099)	38 (0.082)	51 (0.109)	283 (0.607)	466

Table II. Maximum likelihood estimates of the regression coefficients, ψ_{tx} for the model in (4). Standard errors are in parentheses. $AGE = I\{\text{age} < 70\}$; $PS = I\{\text{performance status} \in (2,3,4)\}$; $NOSITE1 = I\{\#\text{ of sites} = 2\}$; $NOSITE2 = I\{\#\text{ of sites} > 2\}$. * indicates significance with 95% confidence level.

	IFL	FOLFOX
Intercept	0.51 (0.23)	1.32 (0.27)
AGE	−0.39 (0.33)	−0.29 (0.35)
PS	−2.41* (1.06)	−1.53* (0.66)
NOSITE1	−0.04 (0.32)	−0.55 (0.35)
NOSITE2	−0.67 (0.34)	−0.91* (0.36)

Table III. Maximum likelihood estimates of $P(Y_{T1} \leq k_1, Y_{T2} \leq k_2 | D_T = 0, \text{Tx})$. $P(Y_{T1} \leq k_1, Y_{T2} \leq k_2 | D_T = 0, \text{Tx})$ is the joint cumulative probability of insomnia and outlook for those who did not progress/die before the end of study.

		IFL k_2				FOLFOX k_2			
		1	2	3	4	1	2	3	4
k_1	1	0.023	0.031	0.055	0.061	0.028	0.039	0.070	0.077
	2	0.059	0.084	0.169	0.193	0.069	0.099	0.205	0.234
	3	0.125	0.188	0.460	0.572	0.137	0.206	0.509	0.632
	4	0.161	0.251	0.719	1.000	0.168	0.262	0.735	1.000

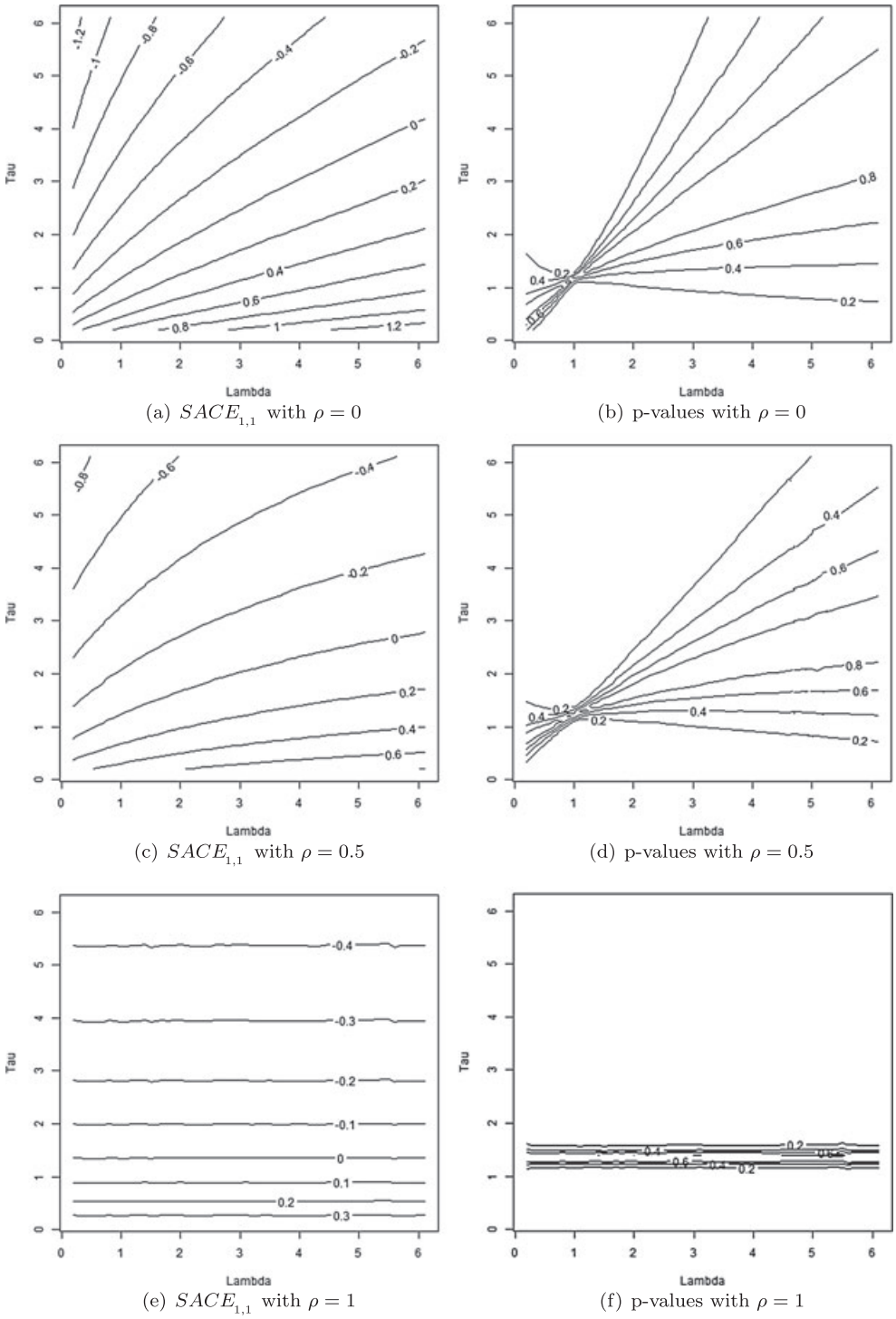
on study). We have the actual progression/death times for subjects who dropped out because of progression/death and also for those who dropped out for other reasons. The number of subjects in each 12-week progression/death window by treatment group is given in Table I. We fit our models to only those subjects who had not died or progressed by the end of the study (the sixth 12-week window in Table I).

Because very few patients reported category 1 for both QOL measures (insomnia and outlook), we collapsed category 1 and 2 into one category (category 1) and then categories 3–5 were relabeled as categories 2–4, respectively. As covariates for the models in (9) and (10), we included treatment (tx=1 for FOLFOX; tx=0 for IFL), (re-scaled) visit number ($TIME = 0.0, 0.1, \cdots, 0.5$), and baseline covariates that were age (*AGE*), patient performance status (*PS*), and number of sites of metastatic disease (NOSITE). These covariates were categorized as in (Lee *et al.* [11]) (age < 70 ($AGE = 0$) and ≥ 70 ($AGE = 1$); performance status is high $\{0, 1\}$ ($PS = 0$) and low $\{2, 3, 4\}$ ($PS = 1$); number of sites of metastatic disease was 1 $(NOSITE1, NOSITE2) = (0, 0)$, 2 $(NOSITE1, NOSITE2) = (1, 0)$, and > 2 $(NOSITE1, NOSITE2) = (0, 1)$). The correlation matrix R_i in (11) is assumed to have an AR(1) structure.

As indicated in Table I for each subject, we know when they died or progressed if it happened before the study ended. To estimate survival probabilities, $g_T(\text{tx}, w)$ in (4), we fitted a logistic regression for each treatment arm with baseline covariates as indicated above. The estimated regression coefficients, $\hat{\psi}$ in (4), are given in Table II. On IFL treatment, the coefficient of *PS* was significant. This indicates that the estimated survival probability on the IFL treatment was less for the low performance group than for

the high performance group. On FOLFOX treatment, the coefficients of *PS* and *NOSITE2* were significant. This indicates that the estimated survival probability on the FOLFOX treatment was less for the low performance group than for the high performance and was less for patients with more than 2 sites of metastatic disease than those with only one.

Table III presents the estimated target probabilities (12) on the two treatment arms computed using the longitudinal model for the observed data specified in Section 3.3. Overall, the estimated cumulative joint



(a) $SACE_{1,1}$ with $\rho = 0$ (b) p-values with $\rho = 0$

(c) $SACE_{1,1}$ with $\rho = 0.5$ (d) p-values with $\rho = 0.5$

(e) $SACE_{1,1}$ with $\rho = 1$ (f) p-values with $\rho = 1$

Figure 1. Maximum likelihood estimates of $\log SACE_{1,1}$ with $\rho = 0, 0.5, 1.0$, and corresponding p-values for $H_0 : \log SACE_{1,1} = 0$.

probabilities on FOLFOX were higher than those on the IFL. These quantities were used to estimate the SACE via the algorithm in Subsection 3.3.

The estimated correlation between insomnia and outlook from Σ_2 in (11) was 0.352, which indicates a positive relation (i.e., high insomnia is positively correlated with negative outlook). Also, the estimated correlation (ρ) for serial dependence in (11) was 0.794 that indicates a strong serial correlation.

We can estimate the $SACE_{k_1,k_1}$ for all $k_1, k_2 \in \{1, \ldots, 4\} \bigotimes \{1, \ldots, 4\}$. However, because the treatment effect on high insomnia and negative outlook ($k_1 = 1$, $k_2 = 1$) was of interest, we estimated $SACE_{1,1}$ that corresponds to high insomnia and negative outlook ($k_1 = 1, k_2 = 1$). Because the identification of *SACE* relies on untestable assumptions, we implemented a sensitivity analysis to draw inference about *SACE* by varying the sensitivity parameters τ and λ in Assumption 3. We chose a very wide range for the sensitivity parameters, $[0.1, 6.0]$. The estimated values of $SACE_{1,1}$ and associated p-values as a function of λ and τ under $H_0 : \log SACE = 0$ are given in Figure 1 for various values of ρ.

When $\rho = 0$ (cf: assumption 2), which corresponds to dying on FOLFOX being independent of dying on IFL, the estimated values of $SACE_{1,1}$ increased as λ and τ respectively increased and decreased (Figure 1(a)). The same pattern happened for $\rho = 0.5$ that corresponds to a moderate correlation between dying on the two treatments (Figure 1(c)). In contrast, when $\rho = 1$ (high correlation), the estimated values of $SACE_{1,1}(1, 0)$ decreased slowly as τ increased (Figure 1(e)). So when τ is less than 1, the probability of high insomnia and negative outlook on the FOLFOX treatment was greater than the corresponding probability on the IFL treatment. When τ was greater than 1, the probability of high insomnia and negative outlook on the FOLFOX treatment was lower. However, for all values of ρ there were *no significant difference* between these probabilities (Figure 1(b), 1(d), 1(f)).

5. Conclusion

In this paper, we have proposed a causal inference approach for bivariate ordinal outcomes in the presence of dropout because of progression or death. To do this we proposed a novel causal estimand and an appropriate set of assumptions to identify it. We used all available longitudinal data (not just the data at the time point of interest) via a longitudinal model and a partial ignorability assumption. We implemented a sensitivity analysis for nonidentifiable assumptions that were parameterized parsimoniously.

On the basis of this approach, we concluded that patients' insomnia and outlook levels were not influenced by treatment. However, we note that the power was low for this comparison with few patients still on study at one year. In our approach, we somewhat offset the small sample size at year one via modeling assumptions that allowed us to use all available longitudinal data.

Different choices of the SACE and/or examining several SACE at once would require different and/or more complex alternatives to Assumption 3.

Acknowledgements

We would like to thank Dr. Daniel Sargent and Ms. Erin Green of the Mayo Clinic for providing the data and for their help in data collection and clarifying some issues with data. This project was supported by NIH grant CA85295 and Basic Science Research Program through the National Research Foundation of Korea (KRF) funded by the Ministry of Education, Science and Technology (2012-004002).

References

1. Hogan JW, Laird NM. Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine* 1997; **16**:239–257.
2. Pauler DK, McCoy S, Moinpour C. Pattern mixture models for longitudinal quality of life studies in advanced stage disease. *Statistics in Medicine* 2003; **22**:795–809.
3. Kurland BF, Heagerty PJ. Directly parameterized regression conditioning on being alive: analysis of longitudinal data truncated by deaths. *Biostatistics* 2005; **6**:241–258.
4. Lee K, Daniels M. A class of Markov models for longitudinal ordinal data. *Biometrics* 2007; **63**:1060–1067.
5. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974; **66**:688–701.
6. Rubin DB. Bayesian inference for causal effects. *Annals of Statistics* 1978; **6**:34–58.
7. Holland P. Statistics and causal inference. *Journal of the American Statistical Association* 1986; **81**:945–961.
8. Frangakis CE, Rubin D. Principal stratification in causal inference. *Biometrics* 2002; **58**:21–29.

9. Egleston BL, Scharfstein DO, Freeman ER, West SK. Causal inference for non-mortality outcomes in the presence of death. *Biostatistics* 2007; **8**:526–545.
10. Lee K, Daniels M. Marginalized models for longitudinal ordinal data with application to quality of life studies. *Statistics in Medicine* 2008; **27**:4359–4380.
11. Lee K, Daniels MJ, Sargent DJ. Causal effects of treatments for informative missing data due to progression/death. *Journal of the American Statistical Association* 2010; **105**:912–929.
12. Cheng J. Estimation and inference for the causal effect of receiving treatment on a multinomial outcome. *Biometrics* 2009; **65**:96–103.
13. Williamson J, Kim K. A global odds ratio regression model for bivariate ordered categorical data from ophthalmologic studies. *Statistics in Medicine* 1996; **15**:1507–1518.
14. Harel O, Schafer JL. Partial and latent ignorability in missing-data problems. *Biometrika* 2009; **96**:37–50.
15. Goldberg RM, Sargent DJ, Morton RF, Fuchs CS, Ramanathan RK, Williamson SK, Findlay BP, Pitot HC, Albets SR. A Randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *Journal of Clinical Oncology* 2004; **22**:23–30.
16. Redeker NS, Lev EL, Ruggiero J. Insomnia, fatigue, anxiety, depression, and quality of life of cancer patients undergoing chemotherapy. *Research and Theory for Nursing Practice* 2000; **14**:275–290.
17. Williamson J, Kim K, Lipsitz SR. Analyzing bivariate ordinal data using a global odds ratio. *Journal of the American Statistical Association* 1995; **90**:1432–1437.
18. Liu I. Describing ordinal odds ratios for stratified $t \times c$ tables. *Biometrical Journal* 2003; **6**:730–750.
19. Heibron D. The analysis of ratios of odds ratios in stratified contingency tables. *Biometrics* 1981; **37**:55–66.
20. Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees*. Wadsworth: Belmont, CA, 1984.
21. Roy J, Hogan JW, Marcus BH. Principal stratification with predictors of compliance for randomized trials with 2 active treatments. *Biostatistics* 2008; **9**:277–289.
22. Marshall AW, Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* 1997; **84**:641–652.
23. Sankaran PG, Jayakumar K. On proportional odds models. *Statistical Papers* 2008; **49**:779–789.
24. Egleston BL, Scharfstein DO, MacKenzie E. On estimation of the survivor average causal effect in observational studies when important confounders are missing due to death. *Biometrics* 2009; **65**:497–504.
25. Lee K, Daniels MJ, Joo Y. Flexible marginalized models for bivariate longitudinal ordinal data. *Biostatistics* 2013. DOI: 10.1093/biostatistics/kxs058.
26. Sjolander A, Humphreys K, Vansteelandt S, Bellocco R, Palmgren J. Sensitivity analysis for principal stratum direct effects, with an application to a study of physical activity and coronary heart disease. *Biometrics* 2009; **65**:514–520.
27. Heagerty PJ. Marginally specified logistic-normal models for longitudinal binary data. *Biometrics* 1999; **55**:688–698.
28. Joe H. Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis* 2006; **97**:2177–2189.
29. Daniels MJ, Pourahmadi M. Modeling covariance matrices via partial autocorrelations. *Journal of Multivariate Analysis* 2009; **100**:2352–2363.