

# Joint modeling compliance and outcome for causal analysis in longitudinal studies

Xin Gao,<sup>a</sup> Gregory K. Brown<sup>b</sup> and Michael R. Elliott<sup>a,c,\*†</sup>

This article discusses joint modeling of compliance and outcome for longitudinal studies when noncompliance is present. We focus on two-arm randomized longitudinal studies in which subjects are randomized at baseline, treatment is applied repeatedly over time, and compliance behaviors and clinical outcomes are measured and recorded repeatedly over time. In the proposed Markov compliance and outcome model, we use the potential outcome framework to define pre-randomization principal strata from the joint distribution of compliance under treatment and control arms, and estimate the effect of treatment within each principal strata. Besides the causal effect of the treatment, our proposed model can estimate the impact of the causal effect of the treatment at a given time on future compliance. Bayesian methods are used to estimate the parameters. The results are illustrated using a study assessing the effect of cognitive behavior therapy on depression. A simulation study is used to assess the repeated sampling properties of the proposed model. Copyright © 2013 John Wiley & Sons, Ltd.

**Keywords:** causal inference; potential outcome; principal stratification; noncompliance

## 1. Introduction

Because randomized treatment assignment removes both observed and unobserved confounding, randomized studies provide a means to estimate the causal effect of a treatment. However, because patients can choose whether or not to comply with their assigned treatment in many circumstances, noncompliance or partial compliance is very common in the randomized studies. To estimate causal treatment effects, traditional analysis methods include intent-to-treat (ITT) analysis, as-treated (AT) analysis, and per-protocol (PP) analysis. ITT analyses provide a causal estimate of the effect of randomization, which can differ from the causal effect of the treatment in the presence of noncompliance. AT analyses ignore the randomization assignment, and compare the outcomes by using the actual treatment received. PP analyses compare the outcomes for subjects who comply with the assigned treatment. However, because the latter two analysis methods condition on the treatment taken, which is a post-randomization variable, selection bias can affect the AT and PP estimates of the causal effect of treatment [1].

A large literature has developed in recent years to estimate causal effects via potential outcomes. Under the potential outcome framework, the causal estimands are comparisons of the potential outcomes that would have been observed under all possible assignments of treatments. The idea of describing causal effects in terms of potential outcomes dates back to Neyman in 1923 [2], and is now becoming widely used in the fields of economics, social and behavioral sciences, epidemiology, and statistics.

In this manuscript, we use the principal stratum approach formulated in [3]. For noncompliance problems, the principal strata are determined by the joint distributions of the compliance behaviors under each treatment arm. In the context of a two-arm randomized trial, they consist of compliers (subjects who take

<sup>a</sup>Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, U.S.A.

<sup>b</sup>Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, U.S.A.

<sup>c</sup>Survey Methodology Program, Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48106, U.S.A.

\*Correspondence to: Michael R. Elliott, Survey Methodology Program, Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48106, U.S.A.

†E-mail: mreliott@umich.edu

treatment if and only if assigned to it), always-takers (subjects who take treatment regardless of assignment), not-takers (subjects who do not take treatment regardless of assignment), and defiers (subjects who take the treatment if and only if assigned to control). Thus these principal strata are not affected by the randomization assignment, and stratum membership can be regarded as a pre-randomization variable. The treatment effect within each principal stratum thus has a causal interpretation. In particular, the treatment effect within the stratum of compliers, called the complier average causal effect (CACE) [4], is often of interest to the investigators, and can be interpreted as the causal effect of treatment, because it is the treatment assignment effect among the subpopulation who comply to the randomization assignment no matter to which treatment group they are assigned.

Recent research on the noncompliance problem in the randomized clinical studies by using principal stratification has focused on obtaining a valid estimate of the effect of treatment within principal strata [3–5]. Estimation of the effect of treatment in the presence of noncompliance has been extended to longitudinal studies in recent years. These longitudinal studies vary in several features. For example, subjects can be randomized once at baseline [6–8] or multiple times over time [9]. Treatment can be applied once [6] or repeatedly over time [7–9]. In this article, we focus on the situation in which subjects are randomized once at baseline, treatments are applied repeatedly over time, and subjects' compliance behavior may change over time.

An important feature of a longitudinal study of this design is that the outcomes and the compliance behaviors are measured and recorded repeatedly over time; and thus, the longitudinal data could reveal more information about the reason and mechanism of noncompliance. Possible reasons for noncompliance include forgetting to take the treatment, side effects, or small treatment effects. With an effective treatment, one might want to improve the compliance in future studies by means such as education about treatment or reduction in side effects. Motivated by this, we propose a Markov compliance and outcome model by using the principal stratification framework of [3]. This work extends that of Lin *et al.* [7], who used the concept of a 'superclass' in a similar design setting to define principal strata on the basis of trajectories of the (partially latent) compliance classes. Lin *et al.* provides a simple and efficient method to assess efficacy when compliance varies over time. Here, our goal is more ambitious: to use the longitudinal study to assess the impact of the causal effect of the treatment on the future compliance, as well as the causal effect of treatment within each principal stratum at each follow-up point in time.

We illustrate the model by using a study designed to assess the effect of cognitive behavioral therapy (CBT) on depression among a sample of suicide attempters.

The remainder of this manuscript is organized as follows. Section 2 describes our proposed Markov compliance and outcomes model. Section 3 applies our proposed model to the CBT study. Section 4 shows the results of a simulation study to the repeated sampling properties. Section 5 discusses the implications of our findings and future application and extensions.

## 2. A Markov compliance and outcome model for longitudinal compliance designs

We focus on the two-arm randomized control trial setting where subjects are randomized at baseline to either treatment or control but then are followed for  $t = 1, \dots, T$  time points after baseline, with the outcome of interest and compliance behavior assessed at each time point  $t$ . Our model is designed to focus on two goals, both of which are of clinical interest. Our first goal is to assess the treatment assignment (ITT) effects within principal strata at each time point – in particular the complier average causal effect, corresponding to the ITT effect with the principal stratum in which the treatment taken corresponds with the treatment assigned under both treatment arms. Our second goal, unique in the literature to our knowledge, is to determine how previous compliance behavior and causal effects may impact current compliance behavior. These goals can only be accomplished by considering the joint distribution of the potential outcomes and potential treatment received over all time points.

### 2.1. Notation

For subject  $i = 1, \dots, n$ , we use  $Z_i$  to denote the randomization assignment (1 for treatment and 0 for control) and  $x_i$  to denote the baseline covariates. For subject  $i$  in the follow-up period  $t$ , we use  $Y_{i,t}(z_i)$  to denote the potential clinical outcomes under treatment assignment  $Z_i = z_i$ , and  $D_{i,t}(z_i)$  to denote the potential treatment received under treatment assignment  $Z_i = z_i$ . In the context of a two-arm

randomized trial without specific restriction, the principal strata consists of compliers, always-takers, not-takers, and defiers:

$$S_{i,t} = \begin{cases} c \text{ (complier)}, & D_{i,t}(z_i) = z_i \\ n \text{ (not-taker)}, & D_{i,t}(z_i) = 0 \\ a \text{ (always-taker)}, & D_{i,t}(z_i) = 1 \\ d \text{ (defier)}, & D_{i,t}(z_i) = 1 - z_i. \end{cases}$$

Thus, the joint distribution of  $D_{i,t}(1)$  and  $D_{i,t}(0)$  fully determines the principal strata  $S_{i,t}$  at time  $t$ . Because only one of  $D_{i,t}(1)$  or  $D_{i,t}(0)$  is observed,  $S_{i,t}$  is latent for all subjects.

## 2.2. Complete data distribution

Under the potential outcome framework, the complete data include all subjects' randomization statuses, principal strata memberships, and potential clinical outcomes under all possible treatment arms in all follow-up periods. The joint distribution of the complete data is given by

$$\begin{aligned} & f(Z_1, \dots, Z_n, Y_{1,1}(1), Y_{1,1}(0), \dots, Y_{n,T}(1), Y_{n,T}(0), S_{1,1}, \dots, S_{n,T}) \\ &= f(Z_1, \dots, Z_n) f(Y_{1,1}(1), Y_{1,1}(0), \dots, Y_{n,T}(1), Y_{n,T}(0), S_{1,1}, \dots, S_{n,T}) \\ &= \prod_{i=1}^n f(Z_i) f(Y_{i,1}(1), Y_{i,1}(0), \dots, Y_{i,T}(1), Y_{i,T}(0), S_{i,1}, \dots, S_{i,T}) \\ &= \prod_{i=1}^n \left\{ f(Z_i) f(S_{i,1}) f(Y_{i,1}(1), Y_{i,1}(0) | S_{i,1}) \right. \\ &\quad \times \prod_{t=2}^T \left[ f \left( Y_{i,t}(1), Y_{i,t}(0) | S_{i,t}, \dots, S_{i,t-K_c^y}, Y_{i,t-1}(1), Y_{i,t-1}(0), \dots, Y_{i,t-K_y^y}(1), Y_{i,t-K_y^y}(0) \right) \right. \\ &\quad \left. \left. \times f \left( S_{i,t} | S_{i,t-1}, \dots, S_{i,t-K_c^c}, Y_{i,t-1}(1), Y_{i,t-1}(0), \dots, Y_{i,t-K_c^c}(1), Y_{i,t-K_c^c}(0) \right) \right] \right\}. \end{aligned}$$

The three equalities follow respectively from these three assumptions:

- (1) Ignorable treatment assignment assumption [10].  
This assumption is that assignment is independent of all baseline variables (observed and unobserved) and potential outcomes. Under ignorability, we do not need to model the assignment mechanism. This assumption is reasonable in most clinical trial settings, where treatments are usually assigned at random.
- (2) Stable unit treatment value assumption [11].  
Stable unit treatment value assumption comprises two subassumptions. The first subassumption is that the potential compliance behavior of one individual is not affected by the randomization status of other individuals, and the potential outcomes of one individual are not affected by the randomization status and the potential compliance behaviors of other individuals. This assumption is reasonable in trials like the CBT study, where the disease (depression) is not infectious, and patients visit their doctors as individuals, instead of in groups. The second subassumption is that there is no 'hidden' version of the treatment, that is, there are no systematic differences in the treatments assigned within the treatment categories. This assumption is reasonable in the CBT study setting, given that care was taken to standardize treatment provided.
- (3) The longitudinal potential outcomes and principal compliance statuses follow a Markov process with possibly differing orders.  
In particular, we assume the potential outcomes at the end of the follow-up period  $t$  depend not only on the principal stratification membership at the current time but on the principal strata in previous  $t-1, \dots, t-K_c^y$  follow-up periods as well as on the potential outcomes at the end of previous  $t-1, \dots, t-K_y^y$  follow up periods. Similarly, we assume the principal strata in the follow-up period  $t$  depend on the principal strata in previous  $t-1, \dots, t-K_c^c$  follow-up periods and the potential outcomes at the end of previous  $t-1, \dots, t-K_y^c$  follow up periods. (We let  $K_c^y = 0$  imply independence between principal strata across time, and similarly for the other pairings.) Details of the Markov process assumed in the application are given in Section 2.3.

2.3. Parametric submodels

For each of the components in the complete data likelihood, we make the following additional modeling assumptions.

*Modeling principal stratum membership at  $t = 1$ .* We focus on the two-arm randomized trial in which subjects assigned to the control group could not access the treatment; therefore the principal strata only consists of compliers and not-takers. We use probit regression to model the (binary) baseline compliance strata, conditional on subject level baseline covariates  $\mathbf{x}_i$ :  $P(S_{i,1} = c | \mathbf{x}_i, \boldsymbol{\alpha}) = \Phi(\boldsymbol{\alpha}_0 + \mathbf{x}_i' \boldsymbol{\alpha}_1)$ .

*Modeling potential outcomes at  $t = 1$ .* Conditional on the principal strata in the first follow-up period, we assume the potential outcomes at the end of first follow-up period follow a bivariate normal distribution with correlation  $\rho$ . For individual  $i$ ,

$$(Y_{i,1}(1), Y_{i,1}(0))' | S_{i,1}, \boldsymbol{\beta}, \Sigma \sim MVN(\boldsymbol{\mu}_{i,1}, \Sigma),$$

$$\boldsymbol{\mu}_{i,1} = (\beta_1 + \beta_{c1}I(S_{i,1} = c), \beta_0 + \beta_{c0}I(S_{i,1} = c))', \Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}.$$

In cross-sectional settings where the causal parameters of interest are non-parametrically identified [12], estimation of the correlation between the potential outcomes of a given subject is not required and thus is usually ignored; when considered, a location shift between  $Y_{i,t}(1)$  and  $Y_{i,t}(0)$  ( $\rho = 1$ ) has sometimes been assumed [13, 14]. Even though only potential outcomes under the arm the subjects actually take are observable, in noncompliance settings where the outcome is binary, the within-correlation is partially identified [15]; however, in this continuous setting, there will be no data available to estimate this within-subject correlation  $\rho$ . Viewing  $\rho$  as a sensitivity parameter, we assume a variety of the correlations between 0 and 1. This correlation has little impact when making inferences about ITT effects in large finite populations and no impact on such inference in superpopulations [2]. But in this setting, this correlation may have considerable impact on the estimates of the parameters providing information about the impact of outcomes on compliance.

*Modeling principal stratum membership at  $t > 1$ .* For the principal strata in follow-up period  $t$  ( $t > 1$ ), we assume a single-order Markov dependence ( $K_y^c = K_c^c = 1$ ). This Markov dependence allows for interactions between the effects of the previous potential outcomes and the compliance class status, and is assumed to be constant over time. Thus, for patient  $i$  in the follow-up period  $t$ ,

$$P(S_{i,t} = c | Y_{i,t-1}(1), Y_{i,t-1}(0), S_{i,t-1}, \boldsymbol{\theta})$$

$$= \Phi(\boldsymbol{\theta}_t + \boldsymbol{\theta}_0 Y_{i,t-1}(1) + \boldsymbol{\theta}_y (Y_{i,t-1}(0) - Y_{i,t-1}(1)) + \boldsymbol{\theta}_c I(S_{i,t-1} = c) + \boldsymbol{\theta}_{yc} (Y_{i,t-1}(0) - Y_{i,t-1}(1)) I(S_{i,t-1} = c)).$$

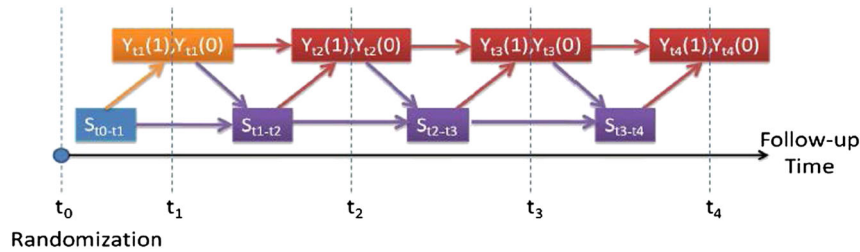
*Modeling potential outcomes at  $t > 1$ .* For the potential outcomes at the end of follow up periods  $t > 1$ , we assume a bivariate normal distribution conditional on principal strata and previous potential outcomes with a single-order Markov dependence for the potential outcomes and a zero-order Markov dependence on the principal strata, that is, depending only on the principal strata membership at time  $t$  ( $K_y^y = 1, K_y^c = 0$ ). This Markov relationship is assumed to be constant over time. Thus, for subject  $i$  and time  $t$ ,

$$(Y_{i,t}(1), Y_{i,t}(0) | Y_{i,t-1}(1), Y_{i,t-1}(0), S_{i,t}, \boldsymbol{\gamma}, \Sigma \sim MVN(\boldsymbol{\mu}_{i,t}, \Sigma),$$

$$\boldsymbol{\mu}_{i,t} = \begin{pmatrix} \gamma_{1t} + \gamma_0 Y_{i,t-1}(1) + \gamma_y (Y_{i,t-1}(0) - Y_{i,t-1}(1)) + \gamma_{c1t} I(S_{i,t} = c) \\ \gamma_{0t} + \gamma_0 Y_{i,t-1}(1) + \gamma_y (Y_{i,t-1}(0) - Y_{i,t-1}(1)) + \gamma_{c0t} I(S_{i,t} = c) \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}.$$

Figure 1 provides a graphical summary of the assumed Markovian relationships between the principal strata and the potential outcomes.



**Figure 1.** Assumed Markovian relationships between principal strata  $S$  and potential outcomes  $Y$ .

#### 2.4. Principal effects

Imbens and Rubin [4] defined the principal effect as the ITT effect of the treatment within each principal compliance stratum. Because of the conditioning structure of our model, we define the principal effects conditional on both the current principal stratum membership and the previous potential outcomes:

$$ITT_{c,t} = E(Y_{it}(1) - Y_{it}(0) | S_{it} = c, Y_{it-1}(1), Y_{it-1}(0)) = (\gamma_{1t} - \gamma_{0t}) + (\gamma_{c1t} - \gamma_{c0t})$$

$$ITT_{n,t} = E(Y_{it}(1) - Y_{it}(0) | S_{it} = n, Y_{it-1}(1), Y_{it-1}(0)) = (\gamma_{1t} - \gamma_{0t}).$$

The ITT effect within the complier stratum,  $ITT_c$ , is often termed the complier average causal effect (CACE), because it is the one principal stratum in which the treatment taken corresponds with the treatment assigned under both treatment arms.

The ITT effect of the treatment within the not-taker stratum  $ITT_n$  is often assumed zero in causal models for clinical studies. This is called the exclusion restriction (ER) assumption [12]. This assumption is plausible in many studies, but not always. For example, in the CBT study, the clinical outcome is a subject's depression severity. The not-takers randomized to the cognitive therapy may experience stress as a result of failing to participate in the therapy, and thus become more depressed, but may not be stressed about not participating in the therapy if they were assigned to usual care. Therefore not-takers may develop different levels of depression under different assignments; and thus, the ER assumption may not be met. Our proposed model allows us to estimate the ITT effect within the not-taker stratum,  $ITT_n$ .

Note that our assumptions about the Markov structure of the model allow us to obtain estimates of within-principal-stratum ITT effects that have a direct causal interpretation. Presence of substantial higher-order Markov relationships suggest that alternative approaches may be required; we provide more detail in Section 5.

#### 2.5. Estimation

A Bayesian approach with data augmentation [16] is a natural method to deal with the missing data structure. The posterior distributions of the parameters and full conditional distributions of the missing potential outcomes do not have simple closed forms, so we use a MCMC algorithm. To obtain draws from the posterior distributions of the parameters, a Gibbs sampling algorithm is used for  $\beta$ ,  $\gamma$ , and  $\sigma^2$ , and Metropolis algorithm is used for  $\alpha$  and  $\theta$ . A Metropolis algorithm is also used to obtain draws from the full conditional distributions of the missing potential outcomes. For the Metropolis algorithm, we propose random draws from a  $t$  distribution with three degrees of freedom. The variance is adjusted to give an acceptance rate of approximately 30% [17]. The detailed descriptions of the posterior distribution of the parameters and the full conditional distributions of missing potential outcomes and missing principal compliance are given in the Appendix. To check convergence, we calculate a measure of between and within-chain variance  $\hat{R}$  [17].  $\hat{R} < 1.1$  is considered acceptable.

In addition to the missing data in the unobserved treatment arm, 19% of subjects have one or more of their BDI outcomes not recorded after time  $t = 1$ . We assume a missing at random mechanism for these missing data, so that conditional on the observed outcomes and compliance behavior, missingness is assumed to be random. Thus, for missing outcome at time  $t > 1$ , draws from the joint distribution of  $(Y_{i,t}(0), Y_{i,t}(1) | Y_{i,1}(0), Y_{i,1}(1), S_{i,1}, \dots, Y_{i,t-1}(0), Y_{i,t-1}(1), S_{i,t-1}, S_{i,t}x_i, \alpha, \beta, \gamma, \theta, \Sigma)$  are imputed at each step of the Gibbs sampler to accommodate missingness under the missing at random assumption.



2.6. Model fit assessment

To assess the fit of the data, we consider the posterior predictive distribution (PPD)  $p$ -values [17]. Because the principal compliance of subjects randomized to the treatment group are equal to the observed compliance, we compare the PPD of the principal compliances and assess their fit to the observed principal compliances. We make posterior predictive checks of the fitted models by using the percentage of compliers in the treatment as the test statistics, and compare the observed percentage of compliers with their posterior predictive distribution to obtain PPD  $p$  values. A PPD  $p$  value close to 0.50 indicates good fit of the model to the data.

We use a chi-square type test of the form  $T = \sum_{i=1}^n \sum_{j=1}^{n_i} ((y_{i,t} - \mu_{i,t})^2 / \sigma^2)$ , where  $y_{i,t} = z_i * Y(1)_{i,t} + (1 - z_i)Y(0)_{i,t}$  and  $\mu_{i,t} = z_i * \mu(1)_{i,t} + (1 - z_i)\mu(0)_{i,t}$ . The distribution of this test statistics should be close to the chi-square distribution with  $\sum_i n_i$  degree of freedom. A Q-Q plot of the test statistics and the chi-square statistics around the 45 degree line indicates good fit of the model to the data.

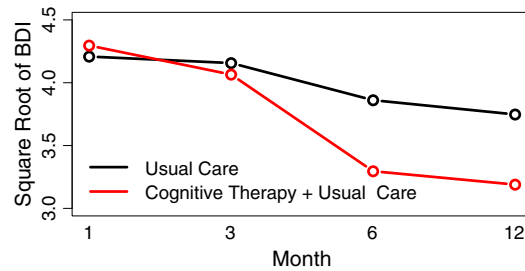
3. Application to the cognitive behavioral therapy trial

The CBT study was a randomized longitudinal study designed to test the effectiveness of a suicide prevention treatment [18]. This study was designed to determine whether a brief psychosocial intervention could reduce depression severity over an 18-month follow up interval in a sample of subjects who had attempted suicide. Cognitive therapy was the psychosocial intervention for this study. It is built on clinical investigations regarding the psychopathological characteristics of suicide behaviors. The study sample consisted of 120 individuals who attempted suicide. Individuals were initially identified in the emergency department following a suicide attempt at the Hospital of the University of Pennsylvania. After the subjects were medically cleared or stabilized in the emergency department, they were transferred to the psychiatric emergency department. Eligible individuals were randomized into two groups: the control group, where patients received usual care, and treatment group, where patients were offered outpatient cognitive therapy sessions specifically developed for preventing suicide attempts on a weekly or biweekly basis. The central feature of this psychotherapy was the identification of proximal thoughts, images, and core beliefs that were activated prior to the suicide attempt.

The clinical outcome measure was the beck depression inventory (BDI), which indicated the severity of depression. Subsequent in-person assessments of BDI were conducted at 1, 3, 6, 12, and 18 months following the baseline interview. Among the 120 subjects, 60 subjects were randomly assigned to the cognitive therapy group, and 60 subjects to the usual care group. We restricted our analysis to the 58 subjects in the treatment arm and the 56 subjects in the control arm with at least one BDI follow-up measurement. Subjects' age ranged from 18 to 66 years, and 61% of them are female. Subjects assigned to the CBT arm who met with the doctor to receive CBT at least once during a given follow up period are defined as having successfully received treatment. Table I shows that randomization achieved balance across the observed baseline covariates between two treatment arms. We apply a square root transformation on BDI to improve the normality approximation for the clinical outcomes, and summarize the observed  $\sqrt{BDI}$  of the study sample at each follow-up time in Figure 2. It shows that the subjects

**Table I.** Mean and standard deviation (in parenthesis) of baseline covariates of the cognitive behavior therapy study sample by treatment group.

Baseline covariates	Cognitive therapy	Usual care	$p$ value
Age	35 (10.0)	35 (10.3)	0.80
Gender(1) (%)	35 (60.3)	34 (60.7)	0.97
Beck depression index	33 (12.1)	31 (15.9)	0.53
Beck hopeless scale	11 (5.5)	12 (6.3)	0.79
Suicide ideation	28 (5.7)	29 (4.4)	0.48
Number of previous suicide attempts	4 (5.4)	6 (13.6)	0.47
Self-reported health status	2 (1.1)	2 (1.1)	0.60
Positive problem orientation	9 (4.3)	9 (5.0)	0.91
Rational problem style	9 (5.1)	8 (5.3)	0.73
Impulsive-careless style	10 (4.8)	9 (5.5)	0.22



**Figure 2.** Mean of the square root of the beck depression inventory (BDI) of the cognitive behavioral therapy study sample by treatment assignment.

randomized to the cognitive therapy group developed less severe depression on average than the subjects randomized to the usual care group from month 3 onward.

We define compliance as attending one or more cognitive therapy sessions during the follow-up period of interest. The noncompliance rate of subjects assigned to the cognitive therapy group are 7%, 9%, 36%, and 64% in follow-up periods of one month, three months, six months, and twelve months respectively. Data for the 18 month follow-up are not included in the analysis, because noncompliance rate is 96%.

We apply the proposed Markov compliance and outcome model to the CBT study. We assume relatively flat priors because we do not have strong prior knowledge for the CBT study. Specifically, we let  $\beta \sim MVN(\mathbf{0}, 10000I)$ ,  $\gamma \sim MVN(\mathbf{0}, 10000I)$ ,  $\theta \sim MVN(\mathbf{0}, 10I)$ ,  $\sigma^2 \sim \text{Inv-}\chi^2(1, 1)$ . Ten is chosen as the variance of prior of  $\theta$  because it reflects our belief that baseline covariates will not have extreme impacts on the probability of principal stratum membership, without being unduly informative on the probability scale away from 0 and 1. We ran two chains from different starting values, each with 100,000 iterations after a burn-in of 50,000. Our maximum value of  $\hat{R}$  is  $1.02 < 1.1$ , indicating the Markov chains have indeed converged.

We focus first on the situation where the correlation between the two potential outcomes under treatment and control is high ( $\rho = 0.9$ ). Table II shows the estimated CACE and  $ITT_n$  for cognitive therapy treatment. The results show that for compliers, the cognitive therapy consistently lowers depression severity more effectively than usual care. On average among compliers, the cognitive therapy lowers  $\sqrt{BDI}$  0.63 (95% CI .05, 1.27) more than that under usual care at 1 month, and this effect increases to 2.22 (95% CI 1.02, 3.35) at 12 months. For not-takers, because they would not participate in the cognitive therapy regardless the group they are assigned, we can view  $ITT_n$  as the effect of randomization to cognitive therapy without actually taking the therapy. The results show that subjects randomized to cognitive therapy who are unwilling or unable to complete treatment during the first month have average  $\sqrt{BDI}$  scores 1.58 (95% CI 0.38, 2.90) higher than if they had not been randomized to cognitive therapy. The results imply that being assigned to cognitive therapy – though not cognitive therapy itself – is harmful to not-takers at the beginning of the study. This effect decreases over time, with the 95% credible interval including 0 from 6 months on.

**Table II.** Cognitive behavior therapy study: intent-to-treat effects in the complier stratum ( $ITT_c$  or CACE), intent-to-treat effects in the never-taker stratum ( $ITT_n$ ), overall ITT effects, and as-treated (AT) effects: posterior medians and 95% credible intervals (in parenthesis) at each follow-up time.  $\rho$ =assumed correlation among counterfactual potential outcomes.

	$\rho$	Month			
		1	3	6	12
CACE	0.9	-0.63(-1.27, -0.05)	-0.68(-1.48, 0)	-1.09(-1.87, -0.34)	-2.22(-3.35, -1.02)
	0.5	-0.66(-1.28, -0.03)	-0.74(-1.52, 0.02)	-1.42(-2.36, -0.50)	-2.40(-3.67, -0.90)
	0.1	-0.72(-1.32, -0.11)	-0.80(-1.54, -0.07)	-1.49(-2.43, -0.59)	-2.29(-3.72, -0.70)
$ITT_n$	0.9	1.58(0.38, 2.90)	1.87(0.44, 3.29)	0.45(-0.62, 1.56)	0.16(-0.82, 1.20)
	0.5	1.87(0.16, 3.62)	1.63(-0.77, 3.63)	0.56(-0.85, 1.89)	0.06(-1.10, 1.25)
	0.1	1.85(0.17, 3.59)	0.77(-1.44, 2.97)	0.12(-1.30, 1.58)	-0.26(-1.44, 0.93)
ITT		-0.01(-0.63, 0.60)	-0.22(-0.85, 0.40)	-0.66(-1.29, 0.03)	-0.81(-1.45, -0.17)
AT		0.12(-0.43, 0.66)	0.09(-0.64, 0.47)	-0.75(-1.33, -0.17)	-1.01(-1.70, -0.31)

To compare our proposed model with the traditional analysis methods, we show the results of ITT and AT analyses in Table II. For the ITT analyses, we fit a linear mixed model with a random intercept for each subject, with fixed-effect indicators for follow-up periods, treatment assignment, and their interaction. The analysis shows that assignment to cognitive therapy lowers depression severity on average in all follow up periods, although these effects are smaller than the CACE of cognitive therapy, and the CI does not exclude 0 until 6 months. Note that the effect of randomization is different from the effect of cognitive therapy in the CBT study because of the large proportion of noncompliance. For the AT analysis, we use a model similar to that for the ITT analysis but replace treatment assigned with treatment taken. The analysis shows that cognitive therapy lowers depression severity from 6 months after randomization. However, because of self-selection, subjects participating in the cognitive therapy and subjects participating in the usual care might no longer be comparable, and estimates of the effect of cognitive therapy from AT analysis appear to underestimate the CACE.

In Table III, we show the estimated probit regression parameters (elasticities) governing principal compliance in the follow-up period  $t$  ( $t > 1$ ) when within-subject correlation is assumed to be high ( $\rho = .9$ ). The coefficients  $\theta_y$  and  $\theta_y + \theta_{yc}$  estimate the impact of the effect of randomization to cognitive therapy among not-takers and compliers respectively at follow-up period  $t - 1$  on the probability of being a complier versus a never-taker at follow-up period  $t$ . The positive 95% credible intervals indicates that the probability of being a complier at follow up period  $t$  increases as the effect of randomization to cognitive therapy at the end of follow-up period  $t - 1$  increases. The quantity  $\theta_y + \theta_{yc}$  is larger, indicating this effect is stronger for compliers than not-takers at follow-up period  $t - 1$ . The parameter  $\theta_c$  shows the effect of being a complier at  $t - 1$  on that of being a complier at  $t$  among those for whom the treatment assignment was neither harmful nor beneficial. The results show that principal compliance status at time  $t - 1$  is not predictive of principal compliance status at time  $t$  when the treatment had no effect at time  $t - 1$ .

The parameters  $\alpha_{\text{gender}}$  and  $\alpha_{\text{BDI}}$  measure the impact of baseline covariates on the principal compliance in the first follow-up period. The 95% credible interval of  $\alpha_{\text{BDI}}$  is positive ( $\alpha_{\text{BDI}} = 0.79 (0.47, 0.87)$ ), indicating that the subjects have higher probability of being compliers in the first follow-up period if the subjects have more severe depression at the time of randomization.  $\alpha_{\text{gender}}$  is estimated to be positive, but the 95% credible interval covers zero ( $\alpha_{\text{gender}} = 0.62(-0.30, 0.84)$ ), indicating that females tend to be marginally more likely to comply in the first follow-up period. (Results not shown in Table III.)

To assess the fit of data, we check the PPD of the principal compliances and potential clinical outcomes. The  $p$  values of the PPD of principal compliances are 0.83, 0.69, 0.53, and 0.47 for follow-up periods 1, 3, 6, and 12 months, respectively, indicating a good fit for principal compliance. The Q-Q plot in Figure 3 shows the distribution of the potential clinical outcome test statistics is very close to the model-predicted chi-square distribution with 416 degrees of freedom, indicating that the normality assumption for the transformed BDI measure is reasonable.

### 3.1. Sensitivity to potential outcomes correlations

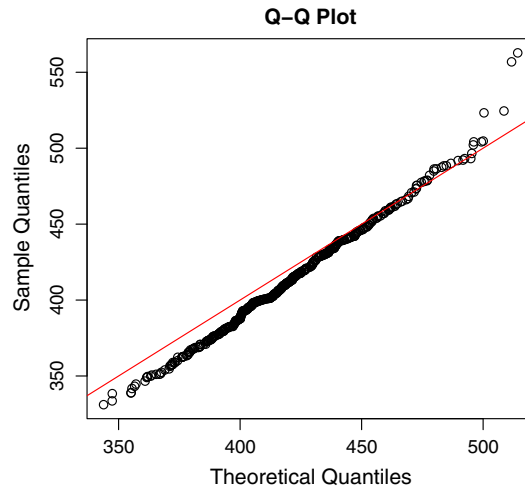
Because there is no data available to assess the correlation  $\rho$  of the potential outcomes, we fix this value at a constant and consider the impact of different values on our inferences of interest. The previous section assumed high correlation ( $\rho = 0.9$ ); here we consider a moderate correlation ( $\rho = 0.5$ ) or a small correlation ( $\rho = 0.1$ ).

The results in Table II imply that both CACE and  $ITT_n$  are not sensitive to the correlation of the potential outcomes. Under both moderate and small correlations, the cognitive therapy consistently has better effects on lowering depression severity than the usual care for compliers. The magnitudes of the

**Table III.** Posterior medians and 95% credible intervals of probit regression parameters (elasticities) governing the cognitive behavior therapy compliance classes for follow-up period  $t > 1$ .  $\rho =$  assumed correlation among counterfactual potential outcomes.

$\rho$	Causal effect at $t - 1$ in not-takers ( $\theta_y$ )	Causal effect at $t - 1$ in compliers ( $\theta_y + \theta_{yc}$ )	Effect of compliance at $t - 1$ ( $\theta_c$ )
0.9	3.19(1.15, 6.26)	4.59(1.73, 9.05)	-2.03(-5.22, 0.53)
0.5	1.26(-0.52, 4.05)	1.43(-0.65, 3.60)	0.37(-2.68, 2.96)
0.1	0.41(-0.74, 2.01)	0.49(-0.31, 1.75)	1.29(-0.74, 2.87)





**Figure 3.** Q-Q plot of outcome test statistic with posterior predictive chi-square distribution.

difference are very similar to those under high correlation of potential outcomes. For not-takers, under both moderate and small correlations, assignment to cognitive therapy makes subjects' depression more severe than assignment to usual care at the beginning of the study. This effect decreases as time increases and disappears at 1 year after randomization.

The estimated parameters governing the principal compliance are shown in Table III. The results show that, as the correlation of potential outcomes becomes smaller, the impact of the causal effect of cognitive therapy among not-takers at the previous time period ( $\theta_y$ ) and similarly the impact of the causal effect of cognitive therapy among compliers at the previous time ( $\theta_y + \theta_{yc}$ ) shrink toward 0, whereas the impact of being a complier at the previous time ( $\theta_c$ ) expands simultaneously. The variance of the difference of bivariate normal random variables becomes larger as their correlation becomes smaller. Because the treatment effect is the difference of the bivariate random variables, as the correlation of within-subject potential outcomes becomes smaller, the variance of the treatment effect will become larger; thus, the association between the treatment effect and future compliance will become weaker, and other predictors will play a more important role in the model.

#### 4. Simulation results

Although we utilize a Bayesian framework for our analysis, we are still interested in the repeated sampling properties of our proposed model. To assess this, we simulate 50 datasets of 100 subjects each with four follow-up periods. To make subjects' compliance behaviors and clinical outcomes similar to the CBT study data, each dataset are simulated under  $\alpha = (\alpha_0, \alpha_{\text{gender}}, \alpha_{\text{BDI}}) = (-3.0, 0.6, 0.8)$ ,  $\beta = (\beta_1, \beta_0, \beta_{c1}, \beta_{c0}) = (3.1, 1.7, 1.2, 3.3)$ ,  $\gamma = (\gamma_{12}, \gamma_{13}, \gamma_{14}, \gamma_{02}, \gamma_{03}, \gamma_{04}, \gamma_0, \gamma_y, \gamma_{c12}, \gamma_{c13}, \gamma_{c14}, \gamma_{c02}, \gamma_{c03}, \gamma_{c04}) = (2.5, 1.4, 1.0, 1.0, 1.3, 1.1, 0.6, 1.0, -2.0, -1.9, -1.9, 0.4, -0.6, 0.2)$ ,  $\theta = (\theta_2, \theta_3, \theta_4, \theta_0, \theta_y, \theta_c, \theta_{yc}) = (0.7, -2.6, -4.7, 0.6, 2.1, -1.2, 1.1)$ ,  $\rho = .9$ , and  $\sigma^2 = 2.1$ . We analyze the simulated data by using the proposed Markov compliance and outcome model.

We summarize the simulation results of the estimated CACE and  $ITT_n$  in each follow up-period along with other parameters of interest in Table IV and Table V. Given the modest number of simulations due to the relatively long time necessary to fit the model to a single simulated dataset, very accurate assessments of the repeated sampling properties of our Bayesian model are not possible; however, all parameters appear to have relatively low amounts of bias and either approximately correct coverage or modest undercoverage.

#### 5. Discussion

In the setting of a two-arm randomized control trial setting, where subjects are randomized at baseline to either treatment or control, and both outcomes and compliance behavior are assessed at multiple time points after baseline, we develop a model by using a principal stratification approach that assesses a

**Table IV.** Summary of estimated complier average causal effect (CACE) and intent-to-treat effects in the never-taker stratum ( $ITT_n$ ) from 50 simulations.

Month	Coverage of 95% C.I.		Median (true value)	
	CACE	$ITT_n$	CACE	$ITT_n$
1	94%	92%	-0.6(-0.7)	1.3(1.4)
3	92%	90%	-0.9(-0.9)	1.6(1.5)
6	92%	94%	-1.2(-1.2)	0(0.1)
12	90%	90%	-1.7(-2.2)	-0.2(-0.1)

**Table V.** Summary of parameters governing principal compliance estimated from 50 simulations.

	$\alpha_{\text{gender}}$	$\alpha_{\text{BDI}}$	$\theta_y$	$\theta_c$	$\theta_{yc}$
Coverage of 95% C.I.	94%	92%	90%	88%	96%
Median (true value)	0.8 (0.6)	0.8 (0.8)	2.1 (2.1)	-1.1 (-1.2)	1.6 (1.1)

both a complier average causal effect for the outcome at each time point as well as an estimate of how previous compliance behavior and causal effects may impact current compliance behavior. This latter aspect of our model is unique in the literature to our knowledge. We apply this model to a CBT study, showing that the stronger the individual-level effect of randomization to cognitive therapy is at the end of follow-up period  $t$ , the greater the probability that subjects will be compliers in the follow up period  $t + 1$ . This association is stronger for compliers than not-takers at time  $t$ . Our findings imply that subjects are sensing whether the treatment is effective for them, and adapting their compliance behavior accordingly. It is important to note that this result refers to the unobservable potential effect of the treatment on a given subject.

Our proposed model accommodates time-varying latent compliance classes, similar to [7, 9], and [8]. Our approach differs from [7, 8], because our proposed model does not have a ‘super’ principal compliance class, which summarizes the longitudinal pattern of compliance behaviors. There is a trade-off between this superclass approach and the Markovian approach in this manuscript – the former allows one to consider the cumulative impact of compliance behavior in a simple and direct fashion without, however, providing a true complier average causal effect, whereas the latter allows one to both ‘unpack’ the effect of potential outcomes at time  $t - 1$  on compliance at time  $t$ , as well as provide a valid interpretation of the causal effect within the complier principal stratum at each follow-up time  $t$  among a possibly differing group of compliers at each time point.

In terms of the assumptions, our proposed causal model is different from previous work in a variety of ways. First, we relax the exclusion restriction, which is a common assumption [4–6, 12]. In the CBT study, the main clinical outcome is depression severity. It is special, because even the randomization to treatment may affect it. The analysis provides evidence that assignment to the cognitive therapy group is harmful for not-takers at the beginning of the study, though this effect eventually disappeared at the end of 1 year. Because avoiding the enrollment of non-compliant individuals can be very difficult, especially in a mental health study, these results emphasize the need to monitor carefully noncompliers for harmful effects of CBT assignment, especially those who are noncompliant with treatment during the early period of the study.

Our research also differs in the way we model the potential outcomes. Previous research typically assumed a location shift between the potential outcomes under treatment and control, as in [13]. Such models assume a perfect correlation between the potential outcomes in all cases. In reality, it is reasonable to consider a correlation between 0 and 1 [19]. As in [19], we consider the correlation as a sensitivity parameter, and assess the sensitivity of analysis results to this correlation. Our results show that the principal effects are relatively insensitive to the choice of this correlation. There is more sensitivity to the choice of correlation for the prediction of future compliance behavior. If we assume the within-subject correlation of potential outcomes to be high, the effectiveness of the treatment for a given

subject is a much more important predictor for future compliance than for previous compliance. If the within-subject correlation of potential outcomes is assumed to be low, then within-subject treatment effectiveness has little association with future compliance, and previous compliance better predicts future compliance. (There is little to guide us on the exact choice of this correlation; although we might assume it is non-negative but less than perfect; perhaps a moderately high assumption in the range of 0.3–0.7 is reasonable, depending on how one conceptualizes the sources of the residual ‘within-subject’ error unique to each counterfactual outcome.)

We build our Markov compliance class and outcome model on the basis of a single-order Markov relationship. This model has the advantage of clear interpretation of model parameters, and the posterior predictive checks for compliance classes and potential outcomes show that the single order Markov relationship model provides sufficient fit for the CBT study. If fitting a higher order Markov model suggests that this relationship does not hold, for example, if the dependence between the  $t$  and  $t-1$  potential outcomes were to differ by treatment assignment, then we need to replace  $\gamma_0$  and  $\gamma_1$  with  $\gamma_{0z}$  and  $\gamma_{1z}$ , yielding

$$\begin{aligned} ITT_{c,t} &= E(Y_{it}(1) - Y_{it}(0) | S_{it} = c, Y_{it-1}(1), Y_{it-1}(0)) \\ &= (\gamma_{1t} - \gamma_{0t}) + (\gamma_{c1t} - \gamma_{c0t}) + (\gamma_{01} - \gamma_{00})Y_i(0) + (\gamma_{11} - \gamma_{10})(Y_i(1) - Y_i(0)), \\ ITT_{n,t} &= E(Y_{it}(1) - Y_{it}(0) | S_{it} = n, Y_{it-1}(1), Y_{it-1}(0)) \\ &= (\gamma_{1t} - \gamma_{0t}) + (\gamma_{01} - \gamma_{00})Y_i(0) + (\gamma_{11} - \gamma_{10})(Y_i(1) - Y_i(0)). \end{aligned}$$

One could view this model as assuming a sort of second-order failure of the exclusion restriction, namely that treatment assignment by itself affects dependence between the  $t$  and  $t-1$  potential outcomes. This seems rather implausible, although even in this situation  $\gamma_{c1t} - \gamma_{c0t}$  still encapsulates the difference between the  $ITT_{c,t}$  effect and the  $ITT_{n,t}$  effect. A more serious situation would be that in which the previous principal stratum membership still carries information about the  $ITT$  effect even after conditioning on the current principal stratum membership and the previous potential outcomes. In this setting, the principal strata at time  $t$  would be defined in terms of vectors of principal strata, for example,  $\{(S_{i,t-1} = c, S_{it} = c), (S_{i,t-1} = c, S_{it} = n), (S_{i,t-1} = n, S_{it} = c), (S_{i,t-1} = n, S_{it} = n)\}$  if a one-degree Markov relationship with principal compliance held. Simple causal interpretation of these strata is no longer possible. If this is the case, alternative approaches such as those developed in [7, 8] that use a clustering algorithm to develop relatively interpretable strata such as high or low compliers may be used, at the cost of both clean causal interpretability of the principal stratum treatment effect, as well as the ability to estimate the link between the previous causal treatment effect and current compliance behavior. We also note that estimating such higher-order Markov models will likely require large sample sizes to obtain stable estimates of effects.

Other extensions of the approach developed here are possible. The mixture model for the potential outcomes is not nonparametrically identified in the absence of the exclusion restriction assumption [20]. Although allowing for the exclusion restriction would provide for nonparametric identification, such an assumption may not be warranted in this setting, where the act of noncompliance on the treatment arm may have an effect on mental states that would not occur if assigned to the control arm. Instead, we rely on the normality assumption to identify the mixture components associated with the complier and not-taker groups in the control arm. Although the marginal distributions of  $Y_i(1)$  and  $Y_i(0)$  appear approximately normally distributed, the assumption of joint normality conditional on compliance stratum cannot be assessed, given that complier status is partially latent and that only one of the potential outcomes is observed. An alternative to this approach would be to weaken or eliminate this parametric assumption and rely on either observed predictors of compliance [21] or prior distributional assumptions to induce posterior modes. Assessing the effect of departures from the normality assumption in these longitudinal potential outcome models remains an open research area.

## Appendix

Let  $Y_{i,t}$  denote  $(Y_{i,t}(Z = 1), Y_{i,t}(Z = 0))'$  of subject  $i$  at the end of follow-up period  $t$ . Let  $M_{i,t}$  denote the design matrix of subject  $i$  for the potential outcomes at the end of the follow-up period  $t$ . The posterior distributions of  $\beta$ ,  $\gamma$ , and  $\sigma^2$  are

$$\begin{aligned} \beta | \cdot &\sim MVN(\hat{\mu}_{n\beta}, \hat{\Sigma}_{n\beta}), \\ \hat{\mu}_{n\beta} &= \left( \Sigma_{\beta}^{-1} + \sum_{i=1}^n M_{i,1} \Sigma^{-1} M'_{i,1} \right)^{-1} \left( \Sigma_{\beta}^{-1} \mu_{\beta} + \sum_{i=1}^n M_{i,1} \Sigma^{-1} Y_{i,1} \right), \\ \hat{\Sigma}_{n\beta} &= \left( \Sigma_{\beta}^{-1} + \sum_{i=1}^n M_{i,1} \Sigma^{-1} M'_{i,1} \right)^{-1} \\ \hat{\gamma} | \cdot &\sim MVN(\hat{\mu}_{n\gamma}, \hat{\Sigma}_{n\gamma}), \\ \hat{\mu}_{n\gamma} &= \left( \Sigma_{\gamma}^{-1} + \sum_{i=1}^n \sum_{t=2}^T M_{i,t} \Sigma^{-1} M'_{i,t} \right)^{-1} \left( \Sigma_{\gamma}^{-1} \mu_{\gamma} + \sum_{i=1}^n \sum_{t=2}^T M_{i,t} \Sigma^{-1} Y_{i,t} \right), \\ \hat{\Sigma}_{n\gamma} &= \left( \Sigma_{\gamma}^{-1} + \sum_{i=1}^n \sum_{t=2}^T M_{i,t} \Sigma^{-1} M'_{i,t} \right)^{-1} \\ \sigma^2 | \cdot &\sim \text{Inv} - \chi^2(v_n, \psi_n), \\ v_n &= 2nT + v, \\ &\text{and} \\ \psi_n &= \left\{ v\psi + \frac{1}{1-\rho^2} \sum_{i=1}^n \sum_{t=1}^T [(Y_{i,t} - \mu_{i,t})'(Y_{i,t} - \mu_{i,t}) - 2\rho(Y_{i,t} - \mu_{i,t})'(Y_{i,t} - \mu_{i,t})] \right\} \frac{1}{2nT+v}. \end{aligned}$$

The full conditional distribution of principal compliance is

$$P(S_{i1} = c | \cdot) = \begin{cases} \frac{f(Y_{i1}(1), Y_{i1}(0), S_{i1}=c) \Phi(\alpha'_0 + \alpha_1 x_i)}{f(Y_{i1}(1), Y_{i1}(0) | S_{i1}=c, \beta) \Phi(\alpha'_0 + \alpha_1 x_i) + f(Y_{i1}(1), Y_{i1}(0) | S_{i1}=n, \beta) (1 - \Phi(\alpha_0 + \alpha'_1 x_i))}; & Z_i = 0, \\ 1; & Z_i = 1, D_{i1}(1) = 1 \\ 0; & Z_i = 1, D_{i1}(1) = 0 \end{cases}$$

$$P(S_{it} = c | \cdot) = \begin{cases} \frac{p_c}{p_c + p_n} & Z_i = 0 \\ 1, & Z_i = 1, D_{it} = 1 \\ 0, & Z_i = 1, D_{it} = 0, \end{cases}$$

$$p_c = (\Phi(\mu_{\theta}^{t+1}) I(S_{i,t+1} = c) + (1 - \Phi(\mu_{\theta}^{t+1})) I(S_{i,t+1} = n)) f(Y_{it}(1), Y_{it}(0) | S_{it} = c, \gamma) \Phi(\mu_{\theta}^t),$$

$$p_n = (\Phi(\mu_{\theta}^{t+1}) I(S_{i,t+1} = c) + (1 - \Phi(\mu_{\theta}^{t+1})) I(S_{i,t+1} = n)) f(Y_{it}(1), Y_{it}(0) | S_{it} = n, \gamma) (1 - \Phi(\mu_{\theta}^t)),$$

$$\mu_{\theta}^t = \theta_t + \theta_0 Y_{i,t-1}(1) + \theta_y (Y_{i,t-1}(0) - Y_{i,t-1}(1)) + \theta_c I(S_{i,t-1} = c) + \theta_{yc} (Y_{i,t-1}(0) - Y_{i,t-1}(1)) I(S_{i,t-1} = c).$$

## Acknowledgements

The authors thank Dr. Aaron T. Beck of the University of Pennsylvania School of Medicine, who was the primary investigator of the cognitive behavioral therapy trial. The authors also thank Thomas Ten Have, Peter Yang, Roderick Little, Jeremy Taylor, and Bendick Hansen for their helpful comments and review. The National Institutes of Health grant R01 MH078016, R01 MH60915, and P20 MH71905, and Centers for Disease Control and Prevention grant R37 CCR316866 funded this research.

## References

- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; **70**:41–55.
- Rubin DB. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* 1990; **5**(4):472–480.
- Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002; **58**(1):21–29.
- Imbens GW, Rubin DB. Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics* 1997; **25**(1):305–327.
- Cheng J, Small D. Bounds on causal effects in three-arm trials with non-compliance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2006; **68**(5):815–836.
- Yau LHY, Little RJ. Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for the unemployed. *Journal of the American Statistical Association* 2001; **96**(456):1232–1244.
- Lin JY, Ten Have TRT, Elliott MR. Longitudinal nested compliance class model in the presence of time-varying noncompliance. *Journal of the American Statistical Association* 2008; **103**(482):462–473.

8. Lin JY, Have TRT, Elliott MR. Nested Markov compliance class model in the presence of time-varying noncompliance. *Biometrics* 2009; **65**(2):505–513.
9. Frangakis CE, Brookmeyer RS, Varadhan R, Safaeian M, Vlahov D, Strathdee SA. Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association* 2004; **99**(465):239–249.
10. Rubin DB. Bayesian inference for causal effects: the role of randomization. *The Annals of Statistics* 1978; **6**(1):34–58.
11. Rubin DB. Randomization analysis of experimental data: the Fisher randomization test comment. *Journal of the American Statistical Association* 1980; **75**(371):591–593.
12. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996; **91**(434):444–455.
13. Efron B, Feldman D. Compliance as an explanatory variable in clinical trials. *Journal of the American Statistical Association* 1991; **86**(413):9–17.
14. Elliott MR, Joffe MM, Chen Z. A potential outcomes approach to developmental toxicity analyses. *Biometrics* 2006; **62**(2):352–360.
15. Elliott M, Raghunathan T, Li Y. Bayesian inference for causal mediation effects using principal stratification with dichotomous mediators and outcomes. *Biostatistics* 2010; **11**:353–372.
16. Tanner MA, Wong WH. An application of imputation to an estimation problem in grouped lifetime analysis. *Technometrics* 1987; **29**(1):23–32.
17. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. CRC press: Boca Raton, Florida, 2004.
18. Brown GK, Ten Have T, Henriques GR, Xie SX, Hollander JE, Beck AT. Cognitive therapy for the prevention of suicide attempts: a randomized controlled trial. *Journal of the American Medical Association* 2005; **294**(5):563–570.
19. Jin H, Rubin DB. Principal stratification for causal inference with extended partial compliance. *Journal of the American Statistical Association* 2008; **103**(481):101–111.
20. Long Q, Little RJA, Lin X. Estimating causal effects in trials involving multitreatment arms subject to non-compliance: a Bayesian framework. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2010; **59**(3):513–531.
21. Roy J, Hogan JW, Marcus BH. Principal stratification with predictors for compliance for randomized trials with 2 active treatments. *Biostatistics* 2008; **9**(2):277–289.