Published in final edited form as:

Stat Med. 2014 September 10; 33(20): 3547–3555. doi:10.1002/sim.5702.

Estimating the Efficacy of an Interstitial Cystitis/Painful Bladder Syndrome Medication in a Randomized Trial with Both Non-adherence and Loss to Follow-up

Wei Yang, PhD, Kathleen J. Propert, ScD, and J. Richard Landis, PhD

Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA

Abstract

We are motivated by a randomized clinical trial evaluating the efficacy of amitriptyline for the treatment of interstitial cystitis and painful bladder syndrome in treatment-naïve patients. In the trial, both the non-adherence rate and the rate of loss to follow-up are fairly high. To estimate the effect of the treatment received on the outcome, we use the generalized structural mean model (GSMM), originally proposed to deal with non-adherence, to adjust for both non-adherence and loss to follow-up. In the model, loss to follow-up is handled by weighting the estimation equations for GSMM with one over the probability of not being lost to follow-up, estimated using a logistic regression model. We re-analyzed the data from the trial and found a possible benefit of amitriptyline when administered at a high-dose level.

Keywords

causal inference; non-adherence; loss to follow-up; inverse probability weighting; structural mean model

1. Introduction

Non-adherence and loss to follow-up are the two most common issues that complicate the application of causal inference models in randomized clinical trials. Various approaches have been proposed in the statistical literature to deal with them [1-7]. Principal stratification [8, 9] can be used in both cases, but most implementations require a categorical adherence measure. The structural mean model (SMM) with g-estimation [10, 11] is another method proposed for continuous outcomes that can handle non-adherence measured as a continuous or categorical variable.

The generalized structural mean model (GSMM), proposed by Vansteelandt and Goetghebeur [12], is an extension of the SMM and can be used to model continuous, binary and categorical outcomes. When modeling binary outcomes, GSMM with a logit link function guarantees that the fitted probability of the outcome is bounded between zero and

one. In addition to the structural model for modeling the effect of treatment on the outcome, GSMM requires a second model for the association of the treatment received and the outcome for the causal parameter identification.

In this paper, we use the GSMM with an inverse probability weighting (IPW) technique [13] to adjust for loss to follow-up. Specifically, we weight the estimating equations in the GSMM with the inverse of the probability of not being lost to follow-up. The probability of loss to follow-up can be estimated through a logistic regression model including both baseline and post-randomization predictors. However, unbiased estimation of the causal parameters has to rely on certain assumptions regarding the missing data mechanism.

We show through simulation the unbiased estimate of the causal parameter under the missing at random assumption. We also re-analyze the data from a randomized placebo-controlled trial, conducted by the Interstitial Cystitis Clinical Research Network (ICCRN). The outline of this paper is structured as follows. We first briefly describe the trial in section two. Section three describes the causal model followed by a simulation study in section four. The analytical results from the ICCRN trial are presented in section five. Finally, a discussion is provided in section six.

2. The ICCRN Trial

Our example utilizes data obtained from the first randomized clinical trial of the ICCRN conducted at 10 clinical centers in North America in treatment-naïve participants with interstitial cystitis/painful bladder syndrome (IC/PBS) [14]. Each participant was randomized to either oral amitriptyline or a matching placebo using a 1:1 randomized block design within each clinical center. The standardized protocol included a weekly dose-escalation strategy, beginning at 10 mg, and increasing weekly through 25 mg and 50 mg, reaching the target final dose of 75 mg to begin at the end of three weeks, subject to participant's tolerance.

A total of 271 participants met the eligibility criteria, and were randomized to either amitriptyline ($n = 135$) or matching placebo ($n = 136$) and followed for a total of 12 weeks. In both arms, participants were asked to follow a delineated dose-escalation strategy to maximum tolerance in the first six weeks, and stay on the dose achieved at week six through the end of the follow-up period. In the primary analysis of the trial, a dose of 50 mg or 75 mg was considered a high dose [14]. The primary endpoint was a seven-point self-reported global response assessment, ranging from "markedly worse" through a midpoint of "the same" to "markedly improved." Participants who indicated that they were "markedly" or "moderately" improved were considered responders. For further details of the trial, please see [14].

The analysis of the primary endpoint from this trial was complicated primarily by two issues: adherence to protocol (dose escalation) and loss to follow-up. The trial was designed so that the participant could self-select the study drug dose for the duration of the follow-up period, based on symptom tolerability during the dose-escalation phase. As can be noted in Table 1, 66% of the participants in the active drug arm attained the high dose, whereas 87% in the placebo arm attained the high dose. In this paper, we do the primary analysis

assuming that amitriptyline has an effect only when a high dose is administered and a sensitivity analysis assuming a linear dose-response relationship. The loss to follow-up rate was 18% in the treatment arm, compared to 13% in the placebo arm (Table 1). Median loss to follow-up time was 5.5 (interquartile range: 3-6) and 5 (interquartile range: 2.5-6) weeks in the treatment and placebo arms, respectively. In the original paper in which an intention-to-treat (ITT) analysis was done, participants who were lost to follow-up were considered to be non-responders. In this paper, we use an alternative approach, i.e., inverse probability weighting (IPW), to handle loss to follow-up.

3. The Causal Model

3.1 Notation and Assumptions

Let R denote a randomized treatment assignment. $R = 1$ indicates the treatment arm and $R = 0$ indicates the placebo arm. Z denotes the actual treatment received. $Z = 1$ indicates active treatment is received, i.e., a high dose of amitriptyline is administered in the trial. Given that participants in the placebo arm have no access to the treatment, $Z = 0$ for all participants in the placebo arm. Let Y denote the binary outcome variable. $Y = 1$ if the participant is classified as a responder and $Y = 0$ otherwise. V denotes a vector of both baseline variables and variables that are measured after randomization. X denotes the baseline covariates only, which is a subset of V . Let C denote the indicator of loss to follow-up where $C = 1$ if a participant is lost to follow-up and $C = 0$ otherwise. We use Z^r to denote the potential treatment received when $R = r$, i.e., the treatment received for a participant randomized to the $R = r$ arm. We also use double index notation $Y^{r,z}$ to denote the potential outcome [15, 16] when $R = r$ and $Z = z$, i.e., the outcome for a participant randomized to the $R = r$ arm who receives a treatment of $Z = z$.

We make the following assumptions that are typically assumed in causal inference literature.

1. Under the Stable Unit Treatment Value Assumption (SUTVA) [17], the potential treatment-free outcome for one participant does not depend on the treatment received by other participants in the study population. SUTVA also requires that a participant's potential outcome under her observed treatment level is the same as her observed outcome, i.e., $Y^{r,z} = Y^{obs}$ if $R = r$ and $Z = z$, which is also called the consistency assumption. This assumption is crucial in linking the potential outcomes to the observed outcome so that parameters in the model for potential outcomes can be identified from the observed data.
2. Under randomization, the distributions of the potential outcomes do not differ between the treatment and placebo arms, i.e., $P(Y^{r,z} | R) = P(Y^{r,z})$. A less restrictive assumption is that conditional on the baseline covariates X , the distributions of the potential outcomes do not differ between the treatment and placebo arms, i.e., $P(Y^{r,z} | R, X) = P(Y^{r,z} | X)$. An even more relaxed assumption is the mean independence assumption, which says the means of the potential outcomes do not differ between the treatment and placebo arms, i.e., $E(Y^{r,z} | R) = E(Y^{r,z})$ or $E(Y^{r,z} | R, X) = E(Y^{r,z} | X)$. The estimating equation approach used in the GSMM only requires the mean independence assumption. However, in randomized clinical trials, all these assumptions are usually satisfied.

3. The positivity assumption states that the probability of assigning a participant to the active treatment arm should be greater than zero, i.e., $P(R = 1 | X) > 0$ and is satisfied in randomized trials. In addition, the probability of receiving the active treatment in the treatment arm should also be greater than zero, i.e., $P(Z = 1 | A = 1) > 0$.
4. The exclusion restriction assumption states that the effect of the treatment assignment on the outcome is completely mediated by the actual treatment a participant receives. In other words, the treatment assignment has no direct effect on the outcome, i.e., $Y^{r=1,z=1} = Y^{r=0,z=1}$ and $Y^{r=1,z=0} = Y^{r=0,z=0}$. In a double-blinded randomized trial, given that both participants and physicians are unaware of the treatment assignment, the exclusion restriction assumption is typically satisfied. Further, in a trial when the participant in the placebo arm has no access to the active treatment, only the second equation is relevant, which can be rewritten as $Y^{r=1,z=0} = Y^{r=0,z=0} = Y^0$. Consequently, in this paper we use single notation Y^z to denote the potential outcomes. For example, Y^0 denotes the treatment-free potential outcome, i.e., the outcome when a participant does not receive any active treatment.
5. The Missing at Random (MAR) [18] or random loss to follow-up assumption states that the probability that a participant is lost to follow-up depends only on the observed variables. This assumption is crucial to justify the use of the IPW approach to adjust for loss to follow-up. However, it is an un-testable assumption. Compared to the other assumptions that are likely true in a double-blinded randomized trial, the MAR assumption is often not true. However, careful selection of the predictors of loss to follow-up by clinical knowledge may yield an approximation to the true dropout mechanism and minimize the bias due to the informative loss to follow-up.

3.2 Review of the Logistic Structural Mean Models

Following Vansteelandt and Goetghebeur [12] and Tan [19], the GSMM for a binary outcome with a logit link function is

$$\text{logit} \{P(Y^z | Z^r = z, X = x)\} - \text{logit} \{P(Y^0 | Z^r = z, X = x)\} = g_s(z, x) \psi \quad (1)$$

in which the parameter ψ is the log causal odds ratio.

When $r = 1$, the model (1) is

$$\text{logit} \{P(Y^z | Z^1 = z, X = x)\} - \text{logit} \{P(Y^0 | Z^1 = z, X = x)\} = g_s(z, x) \psi. \quad (2)$$

Due to SUTVA, equation (2) is the same as

$$\text{logit} \{P(Y | Z, X, R=1)\} - \text{logit} \{P(Y^0 | Z, X, R=1)\} = g_s(Z, X) \psi. \quad (3)$$

Notice that the model cannot be fit for the $R = 0$ arm in a randomized trial when the participant in the placebo arm has no access to the active treatment.

Following Vansteelandt and Goetghebeur [12], a second model for $P(Y|Z, X, R = 1)$ needs to be specified for the identification of the causal parameter ψ . It is an association model and has the following form:

$$\text{logit} \{P(Y|Z, X, R=1)\} = g_a(Z, X) \beta. \quad (4)$$

The association parameter β in model (4) can be estimated in a standard way using the maximum likelihood estimation principle. The causal parameter in model (3) can be estimated by testing the mean independence between the potential treatment-free outcomes and the treatment assignment. In a randomized trial in which participants in the placebo arm have no access to the active treatment, one can define an expected putative potential treatment-free outcome based on model (3) and (4), $Y^0(\beta, \psi) = \text{expit}\{g_a(Z, X)\beta - g_s(Z, X)\psi\}R + Y(1 - R)$, in which $\text{expit}(t) = \exp(t) / \{1 + \exp(t)\}$. When β and ψ equal the true values of the parameters, $\beta = \beta_0$ and $\psi = \psi_0$, $Y^0(\beta, \psi)$ is mean independent of the treatment assignment due to randomization. One way of estimating the causal parameter ψ is to invert the test of mean independence between $Y^0(\beta, \psi)$ and R . Vansteelandt and Goetghebeur [12] provided the general forms of the estimating equations:

$$U = \begin{pmatrix} d_s(X, R) \{Y^0(\beta, \psi) - q(X)\} \\ d_a(Z, X) R [Y - \text{expit}\{g_a(Z, X)\beta\}] \end{pmatrix} = 0 \quad (5)$$

in which $d_s(X, R)$ is an arbitrary function of X and R with the restrictions that $E\{d_s(X, R)|X\} = 0$, $q(X)$ is an arbitrary function of X and $d_a(Z, X)$ is an arbitrary function of Z and X .

3.3 Weighting the Estimating Equations to Adjust for Loss to Follow-up

We adjust for loss to follow-up through the inverse probability weighting (IPW) approach, which is to weight the estimating equations for GSMM with the inverse of the probability of not being lost to follow-up. The loss to follow-up probability is estimated by specifying a logistic regression model for the indicator of loss to follow-up, which includes both baseline and post-randomization predictors:

$$\text{logit} \{P(C=1|V)\} = V\theta. \quad (6)$$

The estimating equations in (5) are then weighted so that those who are not lost to follow-up

have weight equal to $w = \frac{1}{\hat{P}(C=0|V)}$ and those who are lost to follow-up have weight equal to zero.

In the ICCRN trial, we are interested in estimating the causal effect of receiving a high dose of amitriptyline on the outcome. In the primary analysis, we assume that the causal effect is zero for those who receive a low dose of the active treatment. The models used in the analyses corresponding to model (3) and (4) above are

$$\begin{aligned} \text{logit} \{P(Y|Z, X, R=1)\} - \text{logit} \{P(Y^0|Z, X, R=1)\} &= Z\psi \text{ and} \\ \text{logit} \{P(Y|Z, X, R=1)\} &= \beta_0 + Z\beta_1. \end{aligned}$$

Given that no participant in the placebo arm received any active treatment, $Z = 1$ only if a participant receives a high dose in the treatment arm. Accordingly,

$$Y^0(\beta, \psi) = \text{expit} \{ \beta_0 + Z\beta_1 - Z\psi \} R + Y(1 - R).$$

Vansteelandt and Goetghebeur [12] provided the near optimal forms of $d_s(X, R)$, $q(X)$ and $d_a(Z, X)$ in (5) so that the efficiency of the causal parameter estimate is close to the semi-parametric efficiency bound [12]. A simple version of the estimating equations can be

achieved by setting $d_s(X, R) = \{R - E(R|X)\}$, $q(X) = 0$ and $d_a(Z, X) = \left(\frac{1}{Z}\right) \cdot E(R|X)$ is the propensity of assigning the active treatment. In 1:1 randomized trials, $P(R = 1 | X) = 0.5$ by the study design. However, it often improves the efficiency of the causal parameter estimates by specifying a model for $P(R = 1 | X; \alpha)$ [12]. To further improve efficiency, we

use $d_s(X, R) = \{R - E(R|X)\} E\left(\frac{\partial Y^0(\beta, \psi)}{\partial \psi} | X\right)$ and $q(X) = E\{Y^0(\beta, \psi) | X\}$ when analyzing the data from the trial. $E\{Y^0(\beta, \psi) | X\}$ is the expectation of treatment-free outcome and is estimated using the data from the placebo arm as

$\hat{E}\{Y^0(\beta, \psi) | X\} = \delta X \cdot E\left(\frac{\partial Y^0(\beta, \psi)}{\partial \psi} | X\right)$ is the expectation of the first-order derivative of the treatment-free outcome with respect to the causal parameter ψ and is estimated as

$\hat{E}\left(\frac{\partial Y^0(\beta, \psi)}{\partial \psi} | X\right) = \mu X$. Both $E\{Y^0(\beta, \psi) | X\}$ and $E\left(\frac{\partial Y^0(\beta, \psi)}{\partial \psi} | X\right)$ help to improve the efficiency of the causal parameter estimate. Misspecification of neither of them affects the consistency of the causal parameter estimate. The final estimating equations that were used are

$$U_w = wU = w \begin{pmatrix} \{R - E(R|X)\} E\left(\frac{\partial Y^0(\beta, \psi)}{\partial \psi} | X\right) [Y^0(\beta, \psi) - E\{Y^0(\beta, \psi) | X\}] \\ \left(\frac{1}{Z}\right) R \{Y - \text{expit}(\beta_0 + Z\beta_1)\} \end{pmatrix} = 0.$$

Let $\gamma = (\beta, \psi)$ and $\gamma_0 = (\beta_0, \psi_0)$ be the true parameter values; it can be shown that $n^{1/2}(\hat{\gamma} - \gamma)$ is asymptotically normally distributed with mean 0 and variance

$$E' \left\{ w \frac{\partial U}{\partial \gamma}(\gamma_0) \right\}^{-1} \text{var} \{ wU(\gamma_0) \} E \left\{ w \frac{\partial U}{\partial \gamma}(\gamma_0) \right\}^{-1}. \quad (7)$$

The variance for the parameter estimate can be estimated using the standard sandwich estimator in (7).

4. Simulation

We evaluate the performance of the weighted estimating equations through simulations. The full data (without missing data) are simulated as follows. We first simulate a binary covariate X with $P(X = 1) = 0.5$ and randomization indicator R with $P(R = 1) = 0.5$.

Adherence status Z follows the model $P(Z = 1 | X, R = 1) = 0.5 + 0.4X$ in the treatment arm and is zero for all participants in the placebo arm. The model for outcome Y is $P(Y = 1 | X, Z) = \text{expit}(2X + Z)$, where $\text{expit}(t) = \exp(t) / \{1 + \exp(t)\}$. Note the log causal odds ratio, i.e., the coefficient of Z in the model for Y , is 1 in the simulation. The missing indicator C for the outcome Y follows the model $P(C = 1) = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 Y$. The parameters $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ in this model vary in the simulation. The sample size for the simulation is 2,000 with 1,000 replications.

In the first simulation, the parameters in the model for missingness are $\beta_0 = 0.1, \beta_1 = 0.3, \beta_2 = 0.3, \beta_3 = 0$. The missing data mechanism is missing at random. We fit three models. Model 1 is the GSMM using the full data (without missing data). Model 2 is the GSMM using only those participants with complete data and ignores the missing data problem. Model 3 is the weighted GSMM assuming the outcome Y is missing at random. The results are shown in Table 2. Model 1 has essentially no bias (-0.004) with coverage probability of 96.1%. The bias for model 2 is -0.425 with the coverage probability of 77.8%. After accounting for missingness by IPW in model 3, the bias reduces to almost zero (0.002) and the coverage probability is 94.9%.

In the second simulation (Table 3), the parameters in the model for missingness are $\beta_0 = 0.1, \beta_1 = 0.2, \beta_2 = 0.2, \beta_3 = 0.2$. It is missing not at random (MNAR) because the missingness is associated with the outcome Y ($\beta_3 = 0.2$). We fit the same models as in the first simulation. Model 1, using the full data, has essentially no bias (-0.004) and good coverage probability (96.1%). By doing a complete case analysis, model 2 performs poorly with a bias of -0.399 and coverage probability of 75.8%. There is still bias in model 3 because IPW cannot adjust for the MNAR scenario. However, the bias (-0.086) is substantially smaller with a much better coverage probability (93.8%) compared to model 2.

5. Analysis of the ICCRN Trial

The potential impact of loss to follow-up on the primary outcome comparison is a major concern in this trial. In the active treatment arm, 24/135 (18%) are lost to follow-up, whereas the loss to follow-up is 17/136 (13%) in the placebo arm. In the original paper [14] in which an intention-to-treat analysis was done, participants who were lost to follow-up were considered as non-responders. With a higher proportion of missing data in the active treatment arm than in the placebo arm, such an adjustment for missing data resulted in an intent-to-treat assessment favoring the placebo arm. Instead, we use the IPW approach to adjust for loss to follow-up under the missing at random assumption. We fit a logistic regression model to estimate the probability of loss to follow-up for each participant. In the model, candidate predictors include demographic variables (age, sex and race), baseline covariates (time since onset of initial symptoms, pain score, urgency score, 24-hour voiding frequency, IC symptom index and IC problem index), the indicator of treatment assignment,

treatments that are actually received and the average grade of adverse events reported during the follow-up period. A reduced model after backward model selection to eliminate non-significant predictors retains only the baseline urgency score, treatment assignment and the treatment that is actually received. The C-statistics for the full and reduced model are 0.831 and 0.805, respectively. The predicted probabilities of being lost to follow-up are very similar between the full and reduced models. We choose the reduced model to estimate the probability of not being lost to follow-up and weight the participant based on the estimated probability. Participants who are not lost to follow-up are given a weight of one over the probability of not being lost to follow-up. Participants who are lost to follow-up are given a weight of zero, i.e., these participants are excluded from the analysis.

We show the results from five different models in Table 4: 1) the ITT analysis using logistic regression on completers; 2) the ITT analysis using logistic regression assuming those who are lost to follow-up are all non-responders, as was done in the original paper [14]; 3) the analysis using the GSMM approach on completers; 4) the analysis using the GSMM approach assuming those who are lost to follow-up are all non-responders; and 5) the analysis using the GSMM approach with IPW to adjust for loss to follow-up. We did not do the ITT analysis using IPW because the weights are estimated from a model including post-randomization predictors, which will in turn be adjusted in the weighted ITT analysis through the weighting process.

Both the ITT and GSMM on completers give statistically significant results. The odds ratios are 1.90 [1.12, 3.24] for the ITT analysis and 2.33 [1.16, 4.69] for the GSMM analysis. As a comparison, the results are not statistically significant using either the ITT (1.49 [0.92, 2.41]) or GSMM (1.88 [0.90, 3.90]) approaches, assuming those who are lost to follow-up are all non-responders. Using IPW to adjust for loss to follow-up, the causal odds ratio from the GSMM is 2.21 (95% CI: [0.98, 4.99]).

We did two sensitivity analyses for model five in Table 4. One is to estimate the probability of loss to follow-up taking into account the time of loss to follow-up. The other one is to assume a linear dose-response relationship without the assumption of any treatment effect in the low-dose group. The results are very similar (data not shown).

6. Discussion

We re-analyzed the data from the amitriptyline trial with two major differences compared to the approach used in the original paper [14]. First, the estimand in this paper is the effect of the actual treatment that a participant received (as-treated analysis), rather than the effect of being assigned to the active treatment. Second, we used the IPW approach to handle the issue of participants lost to follow-up. The use of IPW provided a larger estimate for the effect of amitriptyline; however, it did not reach statistical significance.

One key assumption of using IPW to adjust for loss to follow-up is missing at random, i.e., after controlling for measured baseline and post-randomization variables, participants are lost to follow-up at random. This is an un-testable assumption. However, a careful consideration of the variables that may predict loss to follow-up at the design phase may

help to make the assumption more plausible. Notice that post-randomization variables, e.g., actual treatment received, that are predictive of loss to follow-up should also be included in the model for loss to follow-up. In our data analysis, the dose level that was administered was the strongest predictor of loss to follow-up.

There is at least one other approach that can be used to handle loss to follow-up. Fischer et al. [7] used multiple imputations to fill in the missing outcome data before they fit the structural mean model. Multiple imputations require the same missing at random assumption as is assumed in the IPW approach. However, the implementations for the two approaches are quite different. IPW requires a model for the dropout process. After the probability of dropout is estimated from the model, all participants who are in the study throughout the follow-up period are weighted based on the estimated probability so that they are representative of the original study population. Multiple imputations require a specification for the outcome model and the missing outcome is then imputed directly from the model. The uncertainty of the missing outcome is taken into account by imputing the data multiple times. As a special case, assuming those who are lost to follow-up as non-responders can be considered as an ad hoc single imputation that may not be realistic in many scenarios. The idea of building a doubly robust estimator by specifying a model for the outcome in addition to the model for the loss to follow-up has been suggested [20]. One benefit of a doubly robust approach is that the inference is correct if either the model for loss to follow-up or the model for outcome is correctly specified. We will explore the possible forms of doubly robust estimators under the framework of the GSMM in follow-up research.

Comparing the results from the GSMM (model 4) to the results from the ITT analysis (model 2) by treating those who were lost to follow-up as non-responders, the as-treated effect estimated using the GSMM was larger than the ITT effect, consistent with the general rule that the ITT estimand is typically biased towards the null. However, the standard error from the GSMM approach was much larger than the one from the ITT analysis. One way to improve the efficiency is to use the optimal estimating equations in the GSMM. However, the estimator for the effect of treatment received is typically less efficient than the estimator for the ITT effect. And the efficiency is primarily driven by the correlation between the treatment assignment and the actual treatment received. Nonetheless, neither approach showed statistically significant results. When we used IPW to adjust for loss to follow-up in the GSMM analysis (model 5), the estimated treatment effect was even bigger than both the ITT (model 2) and GSMM (model 4) analyses treating those who were lost to follow-up all as non-responders. This is consistent with the fact that losses to follow-up were treated as non-responders and because loss to follow-up was more common in the treatment arm, which was more adversely affected.

It is a common practice to treat loss to follow-up participants conservatively as non-responders in randomized clinical trials. However, from the clinical point of view, participants who are lost to follow-up may have a variety of reasons. The assumption of treating loss to follow-up participants as non-responders is reasonable if they were lost to follow-up because they did not feel well and so were more likely to be non-responders if they decided to stay in the study. It is easy to see that this can happen in an opposite way, i.e., participants were lost to follow-up because they felt well and so were more likely to be

responders had they stayed in the study. The assumption of treating loss to follow-up participants as non-responders is clearly not correct if the latter is the case. A nice feature of using IPW is that it can handle both scenarios in the same model under the missing at random assumption, so long as variables sufficient to render loss to follow-up ignorable are measured (i.e., factors predicting both loss to follow-up and outcome) and included in the model for loss to follow-up. Unfortunately, neither assumption is testable so that we cannot tell in an absolute way which one is better. Clinical experience is more useful to determine the plausibility of various assumptions and helps in the selection of the best modeling strategy.

We re-analyzed the data from the amitriptyline trial using generalized structural mean models. The estimated causal effect was bigger compared to the ITT analysis and suggested a possibly beneficial effect of amitriptyline when administered at a high-dose level.

Acknowledgments

We are grateful to Marshall Joffe and the late Tom Ten Have for their helpful comments and discussions. We also thank Stijn Vansteelandt for providing the original R code for the GSMM model. The two reviewers provided valuable comments that helped to greatly improve the quality of the manuscript.

References

1. Frangakis C, Rubin D. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*. 1999; 86:365–379.
2. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*. 1996; 91:444–455.
3. Imbens GW, Rubin DB. Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. *The Annals of Statistics*. 1997; 25:305–327.
4. Chen H, Geng Z, Zhou X-H. Identifiability and Estimation of Causal Effects in Randomized Trials with Noncompliance and Completely Nonignorable Missing Data. *Biometrics*. 2009; 65:675–682. [PubMed: 18759847]
5. Imai K. Statistical analysis of randomized experiments with non-ignorable missing binary outcomes: an application to a voting experiment. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2009; 58:83–104.
6. Mealli F, Imbens GW, Ferro S, Biggeri A. Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Biostat*. 2004; 5:207–222.
7. Fischer K, Goetghebeur E, Vrijens B, White IR. A structural mean model to allow for noncompliance in a randomized trial comparing 2 active treatments. *Biostatistics*. 2011; 12:247–257. [PubMed: 20805286]
8. Frangakis CE, Rubin DB, An MW, MacKenzie E. Principal stratification designs to estimate input data missing due to death. *Biometrics*. 2007; 63:641–649. discussion 650–662. [PubMed: 17824995]
9. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002; 58:21–29. [PubMed: 11890317]
10. Robins, J. Correcting for non-compliance in randomized trials using structural nested mean models. Taylor & Francis; Philadelphia, PA, ETATS-UNIS: 1994.
11. Robins J, Rotnitzky A. Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models. *Biometrika*. 2004; 91:763–783.
12. Vansteelandt S, Goetghebeur E. Causal inference with generalized structural mean models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*. 2003; 65:817–835.
13. Joffe MM, Brensinger C. Weighting in instrumental variables and G-estimation. *Stat Med*. 2003; 22:1285–1303. [PubMed: 12687655]

14. Foster HE Jr, Hanno PM, Nickel JC, Payne CK, Mayer RD, Burks DA, Yang CC, Chai TC, Kreder KJ, Peters KM, Lukacz ES, FitzGerald MP, Cen L, Landis JR, Probert KJ, Yang W, Kusek JW, Nyberg LM. Effect of amitriptyline on symptoms in treatment naive patients with interstitial cystitis/painful bladder syndrome. *J Urol*. 2010; 183:1853–1858. [PubMed: 20303115]
15. Rubin DB. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*. 1974; 66:688–701.
16. Rubin DB. Bayesian-Inference for Causal Effects - Role of Randomization. *Annals of Statistics*. 1978; 6:34–58.
17. Rubin DB. Formal mode of statistical inference for causal effects. *Journal of Statistical Planning and Inference*. 1990; 25:279–292.
18. Little, RJA.; Rubin, DB. *Statistical analysis with missing data*. 2nd edn. Wiley; Hoboken, N.J.: 2002.
19. Tan Z. Marginal and Nested Structural Models Using Instrumental Variables. *Journal of the American Statistical Association*. 2010; 105:157–169.
20. Tsiatis AA, Davidian M, Cao W. Improved doubly robust estimation when data are monotonely coarsened, with application to longitudinal studies with dropout. *Biometrics*. 2011; 67:536–545. [PubMed: 20731640]

Table 1

Dose level and loss to follow-up pattern in the two groups.

Dose Level	Loss to follow-up	Treatment arm (N=135)	Placebo arm (N=136)
Low	No	26	11
	Yes	20	7
High	No	85	108
	Yes	4	10

Table 2

Simulation I: $\psi=1$ and $\beta=c(0.1,0.3,0.3,0)$ in the model for lost to follow-up.

	Model 1	Model 2	Model 3
Bias	-0.004	-0.425	0.002
Empirical SE	0.202	0.332	0.281
Model-based SE	0.207	0.335	0.282
Coverage probability	0.961	0.778	0.949

Model 1: GSMM with full data; model 2: GSMM with complete data only, ignoring the missing data problem; model 3: GSMM with complete data, using IPW to adjust for lost to follow-up. SE: standard error. Sample size was 2000 with 1000 replications

Table 3

Simulation II: $\psi=1$ and $\beta=c(0.1,0.2,0.2,0.2)$ in the model for lost to follow-up.

	Model 1	Model 2	Model 3
Bias	−0.004	−0.399	−0.086
Empirical SE	0.202	0.304	0.267
Model-based SE	0.207	0.307	0.266
Coverage probability	0.961	0.758	0.938

Model 1: GSMM with full data; model 2: GSMM with complete data only, ignoring the missing data problem; model 3: GSMM with complete data, using IPW to adjust for lost to follow-up. SE: standard error. Sample size was 2000 with 1000 replications

Table 4

Odds Ratios and 95% confidence intervals for the effect of amitriptyline estimated from ITT analyses and GSMM under different strategies of handling loss to follow-up.

Models	Sample Size	Parameter Estimate Log(Odds Ratio)	Standard Error Log(Odds Ratio)	Odds Ratio [95% Confidence Interval]
Model 1: ITT, complete case analysis	230	0.64	0.27	1.90 [1.12, 3.24]
Model 2: ITT, loss to follow-up as non-responder	271	0.40	0.24	1.49 [0.92, 2.41]
Model 3: GSMM, complete case analysis	230	0.85	0.36	2.33 [1.16, 4.69]
Model 4: GSMM, Loss to follow-up as non-responder	271	0.63	0.37	1.88 [0.90, 3.90]
Model 5: GSMM, using IPW to adjust for loss to follow-up	230	0.79	0.42	2.21 [0.98, 4.99]

OR: odds ratio; CI: confidence interval.