Published in final edited form as:

Clin Trials. 2014 August ; 11(4): 494–502. doi:10.1177/1740774514531096.

A novel application of the *Intent to Attend* assessment to reduce bias due to missing data in a randomized controlled clinical trial

Dustin J Rabideau, M.S.,

Massachusetts General Hospital

Andrew A Nierenberg, M.D.,

Massachusetts General Hospital

Louisa G Sylvia, Ph.D.,

Massachusetts General Hospital

Edward S. Friedman, M.D.,

University of Pittsburgh School of Medicine

Charles L. Bowden, M.D.,

University of Texas Health Science Center, San Antonio

Michael E. Thase, M.D.,

University of Pennsylvania School of Medicine

Terence Ketter, M.D.,

Stanford University

Michael J. Ostacher, M.D., M.P.H., M.S.C.,

Stanford University

Noreen Reilly-Harrington, Ph.D.,

Massachusetts General Hospital

Dan V. Iosifescu, M.D.,

Mount Sinai School of Medicine

Joseph R. Calabrese, M.D.,

Case Western Reserve Univ. School of Medicine

Andrew C. Leon, Ph.D., and

Weill Cornell Medical College

David A Schoenfeld, Ph.D.

Corresponding author: Dustin J Rabideau, Biostatistics Center, Massachusetts General Hospital, 50 Staniford St, Suite 560, Boston, MA, USA, drabideau@partners.edu.

Clinical trial registration # and trial register: [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00667745) record number NCT00667745.

Dedication

This work is dedicated to the memory of Andrew C Leon, Ph.D. (died February 2012).

Conflict of Interest Statement

The authors receive consultancy fees, research funding, and other supporting income from various private and public sources. For a full declaration, please see Appendix 2 in the supplementary data.

Massachusetts General Hospital

Abstract

Background—Missing data are unavoidable in most randomized controlled clinical trials, especially when measurements are taken repeatedly. If strong assumptions about the missing data are not accurate, crude statistical analyses are biased and can lead to false inferences. Furthermore, if we fail to measure all predictors of missing data, we may not be able to model the missing data process sufficiently. In longitudinal randomized trials, measuring a patient's intent to attend future study visits may help to address both of these problems. Leon et al. developed and included the *Intent to Attend* assessment in the Lithium Treatment—Moderate dose Use Study (LiTMUS), aiming to remove bias due to missing data from the primary study hypothesis [1].

Purpose—The purpose of this study is to assess the performance of the *Intent to Attend* assessment with regard to its use in a sensitivity analysis of missing data.

Methods—We fit marginal models to assess whether a patient's self-rated intent predicted actual study adherence. We applied inverse probability of attrition weighting (IPAW) coupled with patient intent to assess whether there existed treatment group differences in response over time. We compared the IPAW results to those obtained using other methods.

Results—Patient-rated intent predicted missed study visits, even when adjusting for other predictors of missing data. On average, the hazard of retention increased by 19% for every one-point increase in intent. We also found that more severe mania, male gender, and a previously missed visit predicted subsequent absence. Although we found no difference in response between the randomized treatment groups, IPAW increased the estimated group difference over time.

Limitations—LiTMUS was designed to limit missed study visits, which may have attenuated the effects of adjusting for missing data. Additionally, IPAW can be less efficient and less powerful than maximum likelihood or Bayesian estimators, given that the parametric model is well-specified.

Conclusions—In LiTMUS, the *Intent to Attend* assessment predicted missed study visits. This item was incorporated into our IPAW models and helped reduce bias due to informative missing data. This analysis should both encourage and facilitate future use of the *Intent to Attend* assessment along with IPAW to address missing data in a randomized trial.

Keywords

intent to attend; inverse probability weighting; attrition; intermittent missing data; bipolar disorder; LiTMUS

Introduction

In an ideal randomized controlled clinical trial, a crude statistical analysis will accurately measure the causal effect of interest. Investigators seek effective randomization, full compliance with the assigned intervention, and no missing data. Any deviation from these conditions results in statistical complications. Among these conditions, missing data can bias results the most. Missing data will not only reduce statistical power, but also could lead to

false inferences regarding the study hypothesis if the data are missing not at random; thus, it is crucial to consider some form of adjustment in the analysis.

Andrew C. Leon et al. proposed a simple assessment, which measures a patient's intent to attend upcoming study visits [1]. The *Intent to Attend* item was first implemented in the Lithium Treatment—Moderate dose Use Study (LiTMUS), a prospective randomized comparative effectiveness study. We evaluate its use in LiTMUS as well as suggest a simple method of adjustment using the *Intent to Attend* measure along with inverse probability of attrition weighting (IPAW). Both the assessment and the adjustment method are broadly applicable and add little burden to a randomized trial.

In this paper, we first briefly discuss the complexities of missing data, the method of IPAW, and the rationale for measuring a patient's intent to attend future study visits. Second, we present the methods and results of implementing the *Intent to Attend* assessment in LiTMUS. We evaluate whether self-rated intentions predicted missed visits and detail a method to use intent in the context of IPAW. Finally, we discuss the limitations, strengths, and implications of this analysis.

Background

Missing data are inevitable in almost all randomized trials across medical research. If the data are missing *completely* at random, we do not have to adjust for the missing data mechanism. If the data are missing at random, it is common for investigators to use likelihood-based methods, such as mixed-effects regression, since the specification of the missing data mechanism can be ignored [2]. Unfortunately, the assumption of ignorable missing data cannot be empirically proven, and thus, there is no way to ascertain its validity. Therefore, it is important to carry out a sensitivity analysis to determine possible causes of missing data and the degree to which it may impact the results of a randomized trial.

Robins, Rotnitzky, and Zhao [3] as well as Hernán, Robins, and Brumback [4-5] have proposed inverse probability of censoring weights to adjust for informative censoring in longitudinal observational studies. These weights have also been referred to as inverse probability of attrition weights (IPAW) [6]. An advantage of IPAW is its conceptual simplicity. Specifically, each observation is weighted by its cumulative inverse probability of remaining in the study. For example, if a majority of patients from a subgroup (based on some common characteristics) dropped out before the second study visit, we would increase the weight of future observations of those remaining from this subgroup to account for those who were missing. These weights can be considered the number of copies of each observation that are necessary to create a pseudo-population in which all outcome measurements are observed [4].

To estimate the weights, Hernán et al. suggest using all relevant baseline and time-varying covariates that will help to predict missed visits [4]. The *Intent to Attend* item, or one's intention to remain in the study, seems an ideal assessment to predict one's probability of retention. Moreover, some randomized trials do not (or cannot) measure all predictors of missed study visits. In these cases, the *Intent to Attend* item could act as a surrogate for any

unmeasured covariates. It is this unique feature that sets the *Intent to Attend* item apart from other predictors and warrants its inclusion in a sensitivity analysis of missing data.

Several investigators have proposed using a participant's self-rated intent as a way to adjust for informative missing data [1, 7-8]. Leon et al. used a simulation study to evaluate the performance of a linear mixed model while accounting for a time-varying intent measure [1]. They found the group effect and group-by-time interaction to be more accurately estimated when previous intent was included as a covariate. Although their results have positive implications, covariate adjustment has been shown to lead to biased inferences in certain instances [6, 9]. Moreover, in contrast to IPAW, parametric likelihood-based methods, such as mixed-effects regression, are much more sensitive to model misspecification [3]. Therefore, using a patient's intent in the context of IPAW maintains the primary aim of measuring this variable, yet avoids certain shortcomings of covariate adjustment.

Methods

Study population and clinical assessments

Bipolar disorder is a psychiatric illness in which patients experience severe mood and behavioral swings, usually categorized as manic, hypomanic, or depressive episodes. Lithium has been a standard first-line medication option for bipolar disorder; however, its efficacy in combination with other medications remained uncertain.

LiTMUS was a six-month multi-site randomized trial designed to test the effectiveness of adding moderate doses of lithium to optimized treatment of patients with bipolar disorder. A total of 283 study participants were randomized to receive either lithium plus optimized treatment or optimized treatment only. Optimized treatment consisted of guideline-informed personalized treatments to address individuals' symptoms and side effects [10]. One of the co-primary hypotheses was that patients randomized to adjunct lithium would experience greater improvement as assessed with the Clinical Global Impressions scale for use in Bipolar Disorder Severity Scale (bipolar severity). At each study visit (9 visits including baseline), patients were assessed by trained raters, their clinicians, and self-rated symptoms, side effects, and functioning. In LiTMUS, patients were allowed to remain in the study even if visits were missed; thus, the frequency of missing data was expected to remain low compared to other clinical trials among bipolar patients [11]. The rationale and design of LiTMUS are discussed elsewhere [12]. In the primary analysis, there was no statistically significant advantage of adding moderate-dose lithium to optimized treatment regarding bipolar severity [13].

We briefly describe the clinical assessments to which we refer in our analyses. The Clinical Global Impressions scale for use in Bipolar Disorder [14] includes a severity rating for overall bipolar illness. The severity score ranges from 1 to 7, where 1 is "Normal, not ill" and 7 is "Very severely ill". Bipolar mood symptoms were measured using the Montgomery Asberg Depression Rating Scale [15] and the Young Mania Rating Scale [16]. These two scales range from 0 to 60, where a higher score represents more severe symptoms. All measures of symptom severity were administered by trained raters.

Other patient characteristics that were measured include quality of life (Quality of Life Enjoyment and Satisfaction Questionnaire) [17], psychosocial functioning (Range of Impaired Functioning Tool) [18], substance abuse comorbidity (Substance Use Disorder Module of the Structured Clinical Interview for DSM-IV) [19], and impulsivity (Impulsivity Rating Scale) [20]. The Quality of Life Enjoyment and Satisfaction Questionnaire is a self-reported measure that ranges from 0 to 100; higher scores reflect higher quality of life. The Range of Impaired Functioning Tool is a clinician-administered measure that ranges from 4 to 20, with higher scores indicating more impairment. The substance use interview assessed whether patients met certain threshold diagnosis criteria for different types of substance abuse (we dichotomized as any substance abuse versus none). The Impulsivity Rating Scale is a 7-item clinician-rated assessment that ranges from 0 to 21, wherein higher scores indicate higher impulsivity. We also, of course, measured a patient's intent to attend subsequent study visits.

Intent to Attend and missing data

In LiTMUS, the *Intent to Attend* item was implemented as designed by Leon et al. (see Figure 1) [1]. At each study visit, patients rated the likelihood that they would attend the next study visit on an ordinal scale from zero to nine: zero being a low likelihood, or intent, to return and nine corresponding to high intent.

We generated summary statistics and histograms to determine the distribution of intent among study participants. We also determined the frequency of missing outcome data in LiTMUS and whether it was monotonic (e.g. once missing, patient never returned) or intermittent (e.g. returned after missing a study visit). To test our hypothesis that the *Intent to Attend* would predict actual study attendance, we used Generalized Estimating Equations (GEEs) to fit the following marginal model:

$$\text{Logit} [E(R_{ij})] = \beta_0 + \beta_1 X_{i(j-1)} + \beta_2 T_j + \beta_3 X_{i(j-1)} T_j \quad (1)$$

The dependent variable was a binary indicator of whether a patient's outcome was observed ($R = 1$) or missing ($R = 0$). We included time, T_j , defined as the scheduled visit week, and a patient's intent at the previous study visit ($X_{i(j-1)}$). We assumed a logit link function, a Bernoulli variance, and an independent working correlation structure. Using the empirical standard error estimates, we tested whether the intent-by-time interaction was significant. If the interaction was not significant, it was removed, and the marginal effect of intent on missing the subsequent visit was tested.

Inverse probability of attrition weighting

We used IPAW to account for potential informative missing data in LiTMUS. This method involved fitting two separate models: the missing data model and the outcome model.

IPAW missing data model—The missing data model was used to calculate the observation-specific stabilized weights [4]. The general specification of the stabilized weights is the following:

$$sw_{ij} = \prod_{k=0}^j \frac{Pr [R_{ij}=1 | \bar{R}_{i(j-1)}, V_i, T_j]}{Pr [R_{ij}=1 | \bar{R}_{i(j-1)}, V_i, \bar{T}_{i(j-1)}, \bar{L}_{i(j-1)}, T_j]} \quad (2)$$

For each scheduled visit, R_{ij} indicates whether patient i 's outcome was observed at time j . The numerator of these weights is conditional on a patient's entire history of missed visits, up to and including the most recent scheduled visit ($\bar{R}_{i(j-1)}$), any baseline covariates thought to predict missed visits (V_i), and time (T_j). In addition to these three variables, the denominator is also conditional on a patient's entire intent history, up to the most recent visit ($i(j-1)$), and any other time-varying covariate histories thought to predict missed visits ($\bar{L}_{i(j-1)}$).

Our first missing data model was the simplest, and assumed that a patient's randomized treatment group, intent history, and history of bipolar severity predicted subsequent absence. Additionally, we assumed that a patient's presence or absence at the previous visit predicted whether the current visit was missed. This approach is similar to that of Preisser et al. [21], allowing us to address the intermittent missing data, though resulting in stronger assumptions about the missing outcomes (see Discussion). In the second model, we added the following baseline and time-varying covariates: gender, age, education status, quality of life, psychosocial functioning, substance abuse comorbidity, impulsivity, depressive severity, and manic severity. Specifically, we hypothesized that study participants with the following characteristics would be at a higher risk of missing visits: male gender, younger, no college education, low quality of life, high psychosocial functioning impairment, substance abuse comorbidity, high impulsivity, and more severe manic and depressive symptoms [22-24]. For each missing data model, we included a log transformation of time.

To estimate the weights, we modeled the numerator and denominator probabilities separately. We fit separate pooled logistic regression models to estimate the numerator and denominator probabilities, conditional on only the most recent covariate measurements. After fitting these models and determining the probability estimates, we calculated each participant's *cumulative* probability at any given study visit by multiplying the current probability by the probability at the previous study visit. Finally we divided the numerator probability by the denominator probability to obtain the stabilized weight estimate. This procedure is similar to that of Hernán et al. [4] and is further outlined in Appendix 1 (see supplementary data). The baseline weights were defined as one, since by design, every patient analyzed was observed at baseline. We assessed model fit using a Hosmer-Lemeshow Goodness of Fit Test.

IPAW outcome model—The primary hypothesis of LiTMUS was that patients randomized to the adjunct lithium group would experience greater improvement in mean bipolar severity over time compared to those not receiving lithium. Although the study protocol prespecified a linear change in mean bipolar severity over time, we determined that a log transformation of time fit the data much better. We used weighted GEEs to fit marginal models, which assumed an identity link function, a normal variance, and an independent

working correlation structure (see Appendix 1 in supplementary data). As opposed to unweighted GEEs, which are robust to covariance model misspecification, IPAW has been shown to produce biased estimates of longitudinal treatment effects unless one uses the independent working correlation structure [25]. Each observation was weighted using the estimates from the missing data models. The dependent variable was bipolar severity, while the independent variables were log-time and the randomized treatment group-by-log-time interaction. This model was used to estimate the difference in bipolar severity between treatment groups at 6 months as well as the associated 95% confidence interval (based on empirical standard error estimates).

A few other models were fit for comparison. First, we removed the weights from the analysis described above and fit the unweighted marginal model. We also fit two mixed-effects linear regression models with random intercepts and slopes: the first included time and the group-by-time interaction, while the second also included a patient's previous intent as a covariate (as suggested by Leon et al. [1]).

Assumptions—Some important assumptions underlie our methods. First, in our missing data models, we assume that every subject has a positive probability of being observed at each visit (i.e. no individual is considered a dropout). By design, LiTMUS patients were allowed to return to the study regardless of how many study visits they missed, so this modeling strategy is reasonable. The results of our weighted analysis are also based on the assumption that the probability of missing is independent of current bipolar severity as well as any previously unobserved variables, given the patient's entire observed history of outcomes and covariates. Although this assumption is not empirically testable, the *Intent to Attend* item acts as a surrogate for unmeasured covariates in the missing data models; thus, its inclusion strengthens the validity of this assumption.

Software

The data analysis was done using SAS 9.2 (SAS Institute, Cary, NC). The figures were created using R 2.14.2 (<http://www.R-project.org/>). See Appendix 1 in the supplementary data for select SAS code of this analysis.

Results

Intent to Attend and missing data in LiTMUS

Overall, a majority of patients reported high intentions of attending their next study visit. Over 90% of all responses were an 8 or above, suggesting a high likelihood that most patients would attend their next assessment session. Over the entire study, the mean intent was 8.3 with a standard deviation of 1.1. Despite this overall consistency, responses did span the entire range of the *Intent to Attend* scale, with 29 cases of intent lower than 5.

Table 1 shows the frequency of missing data (intermittent and monotonic) at each study visit. Out of a total of 2,547 possible outcome measurements (283 patients, 9 visits per patient), only 284 were missed (11%). A total of 91 patients (32%) contributed to these missed visits; thus, almost 70% of patients attended every study visit. Of those 91 patients who missed, 45 experienced only intermittent missing data, 38 experienced only monotonic

missing data, and 8 experienced both (i.e. returned after missing at least once, but were later censored). The high number of patients who returned after a missed study visit is not surprising, as this study allowed participants to miss a visit and still remain in the study.

Fitting the marginal model from Equation 1, we found that a patient's intent at the previous study visit was a significant predictor of missed study visits (p -value = 0.002). Specifically, every one-unit increase in a patient's intent was associated with a 25% increase in the odds of attending the subsequent visit. In other words, as a patient's intentions decreased, the probability of missing the next visit increased. The intent-by-time interaction was non-significant, indicating that the effect of intent did not change over time.

Probabilities of missed visits and stabilized weights

The parameter estimates from our missing data models (pooled logistic regressions) can be interpreted as adjusted hazard ratios [3, 6], such that a ratio greater than one reflects a higher probability of attending a study visit and a ratio less than one reflects a higher probability of not attending. The variables included in our missing data models and their corresponding hazard ratios are listed in Table 2. In both models, we confirmed that higher previous intent scores significantly increased a patient's hazard of being observed, adjusting for all other covariates. In addition, we found that a previously missed visit, male gender, and higher manic severity were significant predictors of missing outcome measurements. Both missing data models seemed to fit the data well, with non-significant goodness-of-fit statistics (see Table 2).

Figure 2 shows the distributions of the estimated probabilities for our second missing data model. The probabilities correspond to the denominator of the stabilized weights (see Equation 2). It is clear from Figure 2 that, in general, the probability of being observed decreases over time. The probabilities ranged from 0.02 to 0.98 (excluding the baseline probabilities, which are all defined as 1). Results were similar when considering the distributions of the stabilized weights. The bottom of Table 2 shows that the weights had a large range and their distributions were positively skewed. Similar to what was expected, the weights were primarily centered around 1 and increased in variability over time [4] (not shown).

LiTMUS primary study hypothesis

In Figure 3, we plotted the crude sample bipolar severity means by treatment group and study visit over the entire range of the scale. On average, patients in both groups seem to improve (decrease) in bipolar severity over the study; however, it is not clear from the crude means whether the lithium group experienced greater improvement compared to those not receiving lithium.

In Table 3, we report the model-based change from baseline in each group as well as the 6 month differences in bipolar severity. In Figure 4, we compare the mean estimates of three of our models: the unweighted marginal model (GEE: unweighted), the mixed model with previous intent as a covariate (MIXED: with intent), and the weighted marginal model including all covariates (GEE: IPAW model 2). We focus in on a smaller range of the scale to show the model differences.

Regardless of the model, we confirm that both treatment groups improve significantly over the course of the study (see Table 3). Testing each of our models, we did not find statistically significant differences between the treatment groups at 6 months; however, the weighted models resulted in much larger differences than the unweighted model and both mixed models (see Table 3 and Figure 4). We found that, although non-significant, the adjunct lithium group seemed to experience less improvement over the 24 weeks.

Discussion

Missing data can be accounted for in many ways and the best strategy depends on the reason why patient outcomes are not being observed. The *Intent to Attend* assessment is a simple yet effective way to address missing data in most areas of clinical research. Likewise, IPAW is an intuitive and broadly applicable method that can be used to address many types of missing data. We illustrated how a patient's intent could be used in the context of IPAW to adjust for missed visits in a randomized trial. A patient's time-varying intent was found to be a necessary predictor of study attendance, even when adjusting for other predictors of missing data. Compared to an unweighted analysis, IPAW yielded larger differences in disease severity change over time between treatment groups.

IPAW is not the only way to use self-rated intent to address missing data. We also considered including intent as a covariate in this analysis. In certain cases, covariate adjustment may introduce bias into the effect of interest [6, 9]; however, it may be easier to implement into standard statistical software. Another approach could involve multiple imputation. In one simulation, Carpenter and Kenward found that although multiple imputation was more efficient than IPAW under a correctly specified missing data model, the estimates were extremely biased under model misspecification, whereas IPAW estimates remained virtually unaffected [26]. We believe that there are many benefits of using IPAW along with the *Intent to Attend* assessment, but there are other available methods. Thus, further research into these alternatives is warranted and encouraged.

We must consider the limitations of this study. Given that LiTMUS had a relatively small proportion of missing outcome data (11%), the effects of adjusting for it may have been attenuated. This method would have a larger effect in studies where missing data are more prevalent. In addition, semi-parametric methods, like IPAW, can be less efficient and less powerful than maximum likelihood or Bayesian estimators, given that the parametric model is well-specified [2, 7]. Our results are based on the assumption that the IPAW missing data models were correctly specified. Although correct model specification is a strong assumption, it is likely less restrictive than assuming the missing data are ignorable without any type of adjustment. Another concern is that there may be other predictors of study attendance not included in the analysis; however, our models included those supported by the literature and those that would be most relevant among this patient population. Also, we included the *Intent to Attend* item, which acts as a surrogate for unmeasured predictors of study visit attendance. Another possible drawback is that we assumed the probability of missing depended on the patient's presence (or absence) at the previous visit, allowing this method to incorporate intermittent missing data. This approach does not differentiate between dropouts and those who simply miss one visit; instead, it assumes that each subject

has a positive probability of being observed at each study visit. In LiTMUS, patients were allowed to return to the study at any time, regardless of how many previous visits they missed. Thus, in this case, we believe this approach was reasonable.

Despite its limitations, this analysis has a number of strengths. This is the first implementation and analysis of the *Intent to Attend* assessment in a randomized trial. We confirmed that this item predicted a patient's actual study attendance and could be used in a sensitivity analysis of missing data. This article is also one of the few descriptions and implementations of IPAW in the setting of a psychiatric randomized trial. Another strength of this study is that the LiTMUS database included nearly complete data on numerous variables, allowing us to incorporate various predictors into our missing data models.

Furthermore, psychiatric research is an area in which the *Intent to Attend* assessment may prove to be especially useful. Even among bipolar patients, who frequently experience severe mood swings, we found that their self-rated intentions helped to predict their actual study attendance. Compared with other questionnaires that are often employed in psychiatric studies, this item adds little burden both to investigators and study participants. In addition, the benefits of this assessment go beyond its use in a sensitivity analysis. In LiTMUS, for example, if a patient reported a low intent score, investigators inquired into the patient's motivation for early termination. This helped investigators understand the causal mechanisms behind missed visits, as well as allowed for adjustments to be made to prevent the patient from dropping out of the study.

In summary, we demonstrated how a simple self-report measure, the *Intent to Attend* assessment, coupled with IPAW could be used in a sensitivity analysis of missing data in LiTMUS. To the best of our knowledge, this was the first time that a participants' intent was used to adjust for missing data in a randomized trial. We hope this study encourages and facilitates future use of the *Intent to Attend* assessment along with IPAW in randomized trials.

Acknowledgments

Funding

LiTMUS was funded by National Institute of Mental Health (contract # NO1MH80001) and is registered with ClinicalTrials.gov (record number NCT00667745).

Grant Support: LiTMUS was funded by National Institute of Mental Health (contract # NO1MH80001)

Appendix 1

Select SAS code for IPAW analysis

Below, we briefly outline the SAS code to carry out IPAW missing data model 2 and the IPAW outcome model used in this paper. Similar code was used to fit the first missing data model. Our approach was based on the code outlined in Hernán et al. [4].

The rows of our data file (litmus) corresponded to every possible study visit of all patients. Baseline covariates were recorded at every time point for each patient. Time-varying

covariates were manipulated to align the most recent covariate measurement (prior to the current visit) with a dichotomous indicator (r) of whether the outcome was observed ($r = 1$) or missing ($r = 0$) at the current visit. The data was sorted by subject identification (ptnumber) and study visit (visitid).

Once the data was in this format, we fit the following pooled logistic regression models:

```
/*IPAW missing data model 2, num prob*/
```

```
proc logistic data = litmus;
```

```
model r = log_week tx gender age educ subst;
```

```
output out=m2_num p=p_0;
```

```
run;
```

```
/*IPAW missing data model 2, denom prob*/
```

```
proc logistic data = litmus;
```

```
model r = log_week tx gender age educ subst
```

```
qlesq liferift irs madsr ymrs intent cgi_prev r_prev;
```

```
output out=m2_den p=p_w;
```

```
run;
```

Baseline covariates included randomized treatment group (tx), gender, age, education status (educ), and substance abuse comorbidity (subst). Time-varying covariates included the previous measurement of quality of life (qlesq), psychosocial functioning (liferift), impulsivity (irs), depressive severity (madsr), manic severity (ymrs), intent to attend (intent), bipolar severity (cgi_prev) and whether the previous visit was missed (r_prev). We also included a log transformation of the current visit week (log_week). We output a new data set for each model, which included the original data as well as the model-based probabilities of being observed (p_0 and p_w).

Next, we used the following data step to calculate the stabilized weights:

```
/* Create weights for IPAW missing data model 2*/
```

```
data weights2;
```

```
merge m2_num m2_den;
```

```
where not(missing(cgi));
```

```
by ptnumber visitid;
```

```
/*reset prob estimates for a new patient*/
```

```

if first.ptnumber then do;

k_0=1; k_w=1;

end;

retain k_0 k_w;

/*cumulative prob estimates*/

if not(first.ptnumber) then do; *keep as 1 for baseline;

k_0 = k_0*p_0;

k_w = k_w*p_w;

end;

/*stabilized weights*/

stabw = k_0/k_w;

run;

```

Starting with a baseline probability of one, we calculated the numerator (k_0) and denominator (k_w) *cumulative* probabilities of a patient being observed at the given study visit (see Equation 2). Then, we divided to obtain the observation-specific stabilized weight. Our new data set (weights2) included all visits at which bipolar severity (cgi) was observed as well as the stabilized weights (stabw).

Lastly, we fit our weighted GEE marginal models to assess whether there existed treatment group differences in response over time.

```

/*IPAW outcome model 2*/

proc genmod data=weights2;

class ptnumber visitid;

model cgi = log_week tx*log_week;

repeated subject=ptnumber / withinsubject=visitid type=ind;

weight stabw;

estimate "Li at 6mo" log_week 3.178 tx*log_week 3.178;

estimate "OPT at 6mo" log_week 3.178;

estimate "Li-OPT at 6mo" tx*log_week 3.178;

run;

```

The outcome variable was the current bipolar severity (cgi). The independent variables included log-time the interaction between treatment group and log-time. We used the “weight” option to fit the model using the stabilized weights. The “estimate” statements estimated the model-based 6-month change in bipolar severity for each treatment group and the difference in bipolar severity between treatment groups at 6 months (3.178 corresponds to the natural logarithm of 24 weeks).

Appendix 2

Detailed conflict of interest statement

Mr. Rabideau reports no competing interests. **Dr. Nierenberg** is a consultant for Abbott Laboratories, Astra Zeneca, Basilea, BrainCells Inc., Brandeis University, Bristol-Myers Squibb, Cephalon, Corcept, Eli Lilly & Co., Forest, Genaissance, GlaxoSmithKline, Innapharma, Janssen Pharmaceutica, Jazz Pharmaceuticals, Lundbeck, Merck, Novartis, PamLabs, PGx Health, Pfizer, Ridge Diagnostics, Roche, Sepracor, Schering-Plough, Shire, Somerset, Sunovion, Takeda, Targacept, and Teva. He is a stakeholder in Appliance Computing, Inc. (MindSite); Brain Cells, Inc., InfoMed (potential share of income). He receives research support from AHRQ, Bristol-Myers Squibb, Cederroth, Cyberonics, Elan, Forest Pharmaceuticals, GlaxoSmithKline, Janssen Pharmaceutica, Lichtwer Pharma, Eli Lilly, Mylin (formerly Dey Pharmaceuticals), NARSAD, NIMH, PamLabs, Pfizer, Shire, Stanley Foundation, and Wyeth-Ayerst. Honoraria include MGH Psychiatry Academy in the past 3 years (Prior to 3 years ago, honoraria from Bristol-Myers Squibb, Cyberonics, Forest Pharmaceuticals, GlaxoSmithKline, Eli Lilly, Shire, Wyeth-Ayerst). Dr. Nierenberg receives other income from legal case reviews for CRICO, MBL Publishing for past services as Editor-in-chief of CNS Spectrums, Slack Inc. for services as Associate Editor of Psychiatric Annals, and Editorial Board, Mind Mood Memory, Belvior Publications. He has copyright joint ownership with MGH for Structured Clinical Interview for MADRS and Clinical Positive Affect Scale and additional honoraria from ADURS, American Society for Clinical Psychopharmacology and Zucker Hillside Hospital and Forest and Janssen, Biomedical Development, Boston Center for the Arts, University of Pisa, University of Wisconsin at Madison, University Texas Southwest at Dallas, Health New England and Harold Grinspoon Charitable Foundation and Eli Lilly and AstraZeneca, Brandeis University, International Society for Bipolar Disorder, 2nd East Asian Bipolar Forum, Mid-Atlantic Permanente Research Institute. **Dr. Sylvia** was a shareholder in Concordant Rater Systems and serves as a consultant for United Biosource Corporation and Clintara. She receives royalties from New Harbinger. **Dr. Friedman** receives grant support from Novartis, St. Jude Medical, Medtronics, Repligen, Astra-Zeneca, Roche, and Takeda. He receives royalties from Springer. **Dr. Bowden** is a research collaborator with Elan and a consultant with Teva, He has no participation with speaker bureaus, nor does he or his wife hold any equity position in any biomedical or pharmaceutical corporation. **Dr. Thase** has, during the past three years, been an advisor/consultant: to Alkermes, Allergan, AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Forest Laboratories, GlaxoSmithKline, Janssen Pharmaceuticals, Lundbeck, MedAvante, Merck, Mylan, Neuronetics, Otsuka, PamLab, Pfizer, PharmaNeuroboost, Roche, Shire, Sunovion, Takeda, and Teva. During the same time frame, he received honoraria for continuing education talks that were supported by AstraZeneca, Bristol-Myers

Squibb, Merck, and Pfizer, and he has received research grants from Alkermes, AstraZeneca, Eli Lilly, Forest, GlaxoSmithKline, Otsuka, PharmaNeuroboost, and Roche. **Dr. Ketter** receives research support from the Agency for Healthcare Research and Quality, AstraZeneca Pharmaceuticals LP, Cephalon Inc., Eli Lilly and Company, National Institute of Mental Health, Pfizer Inc., and Sunovion Pharmaceuticals. He receives consultant fees from Allergan, Inc., Avanir Pharmaceuticals, Bristol-Myers Squibb Company, Cephalon Inc., Forest Pharmaceuticals, Janssen Pharmaceutica Products, LP, Merck & Co., Inc., Sunovion Pharmaceuticals, and Teva Pharmaceuticals. He has lecture honoraria from Abbott Laboratories, Inc., AstraZeneca Pharmaceuticals LP, GlaxoSmithKline, and Otsuka Pharmaceuticals and receives publication royalties from American Psychiatric Publishing, Inc. In addition, Dr. Ketter's spouse is an employee of and holds stock in Janssen Pharmaceuticals. **Dr. Ostacher** has served as a consultant for Alexza Pharmaceuticals, Janssen, Bristol Myers Squibb, Eli Lilly, Otsuka Pharmaceuticals, and Sunovion. **Dr. Reilly-Harrington** receives royalties from Oxford University Press and the American Psychological Association and serves as a consultant for United Biosource Corporation. **Dr. Iosifescu** is a consultant for CNS Response and Servier. **Dr. Calabrese** receives federal funding from the Department of Defense, Health Resources Services Administration, and NIMH; he receives research funding or grants from the following private industries or nonprofit funds: Cleveland Foundation, NARSAD, and Stanley Medical Research Institute; he receives research grants from Abbott, AstraZeneca, Cephalon, GlaxoSmithKline, Janssen, Eli Lilly, and Lundbeck; he serves on the advisory boards of Abbott, AstraZeneca, Bristol-Myers Squibb, Dainippon Sumitomo Pharma, Forest, France Foundation, GlaxoSmithKline, Janssen, NeuroSearch, OrthoMcNeil, Repligen, Schering-Plough, Servier, Solvay/Wyeth, Takeda, and Supernus Pharmaceuticals; and he reports CME activities with AstraZeneca, Bristol-Myers Squibb, France Foundation, GlaxoSmithKline, Janssen, Johnson & Johnson, Schering-Plough, and Solvay/Wyeth. **Dr. Leon** served on independent Data and Safety Monitoring Boards for AstraZeneca, Sunovion, and Pfizer. He served as a consultant and advisor to MedAvante and Roche and had equity in MedAvante. **Dr. Schoenfeld** has an ongoing consulting relationship with the following pharmaceutical companies, Biogen, Neuronova, Cytokinetics, Edison Pharmaceuticals, Galaxo Smith Kine, Merck, Agennix, Alexion and Pfizer. He also has a consulting relationship with Aptiv a contract research organization.

A novel application of the Intent to Attend assessment to reduce bias due to missing data in a randomized controlled clinical trial

Manuscript ID:	CT-13-0182.R2																					
Funding Institution:	National Institute of Mental Health, NO1MH80001																					
Submitting Author:	<div> <div>Rabideau, Dustin John</div> <div>Save</div> <div>Rabideau, Dustin John (proxy)</div> </div> <hr/> <p> <i>primary affiliation</i> Massachusetts General Hospital - Biostatistics 50 Staniford St Suite 560 Boston Massachusetts 02114 United States T: 6176436235 </p>																					
Authors & Institutions:	<table border="0"> <tr> <td> Rabideau, Dustin John proxy </td> <td> <ul style="list-style-type: none"> Massachusetts General Hospital - Biostatistics 50 Staniford St Suite 560 , Boston, Massachusetts 02114 United States </td> </tr> <tr> <td> Nierenberg, Andrew A. Alan proxy </td> <td> <ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States </td> </tr> <tr> <td> Sylvia, Louisa proxy </td> <td> <ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States </td> </tr> <tr> <td> Friedman, Edward S proxy </td> <td> <ul style="list-style-type: none"> University if Pittsburgh - Psychiatry 3811 O'Hara St Pittsburgh PA, Pittsburgh, Pennsylvania 15213 United States </td> </tr> <tr> <td> Bowden, Charles L proxy </td> <td> <ul style="list-style-type: none"> University of Texas - Psychiatry San Antonio, Texas United States </td> </tr> <tr> <td> Thase, Michael proxy </td> <td> <ul style="list-style-type: none"> University of Pennsylvania - Psychiatry Philadelphia, Pennsylvania United States </td> </tr> <tr> <td> Ketter, Terence A proxy </td> <td> <ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States </td> </tr> <tr> <td> Ostacher, Michael Joshua proxy </td> <td> <ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States </td> </tr> <tr> <td> Reilly-Harrington, Noreen proxy </td> <td> <ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States </td> </tr> <tr> <td> Iosifescu, Dan proxy </td> <td> <ul style="list-style-type: none"> Mount Sinai School of Medicine - Psychiatry New York, New York United States </td> </tr> </table>		Rabideau, Dustin John proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Biostatistics 50 Staniford St Suite 560 , Boston, Massachusetts 02114 United States 	Nierenberg, Andrew A. Alan proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 	Sylvia, Louisa proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 	Friedman, Edward S proxy	<ul style="list-style-type: none"> University if Pittsburgh - Psychiatry 3811 O'Hara St Pittsburgh PA, Pittsburgh, Pennsylvania 15213 United States 	Bowden, Charles L proxy	<ul style="list-style-type: none"> University of Texas - Psychiatry San Antonio, Texas United States 	Thase, Michael proxy	<ul style="list-style-type: none"> University of Pennsylvania - Psychiatry Philadelphia, Pennsylvania United States 	Ketter, Terence A proxy	<ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States 	Ostacher, Michael Joshua proxy	<ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States 	Reilly-Harrington, Noreen proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 	Iosifescu, Dan proxy	<ul style="list-style-type: none"> Mount Sinai School of Medicine - Psychiatry New York, New York United States
Rabideau, Dustin John proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Biostatistics 50 Staniford St Suite 560 , Boston, Massachusetts 02114 United States 																					
Nierenberg, Andrew A. Alan proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 																					
Sylvia, Louisa proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 																					
Friedman, Edward S proxy	<ul style="list-style-type: none"> University if Pittsburgh - Psychiatry 3811 O'Hara St Pittsburgh PA, Pittsburgh, Pennsylvania 15213 United States 																					
Bowden, Charles L proxy	<ul style="list-style-type: none"> University of Texas - Psychiatry San Antonio, Texas United States 																					
Thase, Michael proxy	<ul style="list-style-type: none"> University of Pennsylvania - Psychiatry Philadelphia, Pennsylvania United States 																					
Ketter, Terence A proxy	<ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States 																					
Ostacher, Michael Joshua proxy	<ul style="list-style-type: none"> Stanford University - Psychiatry and Behavioral Sciences Stanford, California United States 																					
Reilly-Harrington, Noreen proxy	<ul style="list-style-type: none"> Massachusetts General Hospital - Psychiatry 50 Staniford St Suite 580, Boston, Massachusetts 02114 United States 																					
Iosifescu, Dan proxy	<ul style="list-style-type: none"> Mount Sinai School of Medicine - Psychiatry New York, New York United States

	Calabrese, Joseph R proxy	• Case Western Reserve University Cleveland, Ohio United States
	Leon, Andrew C proxy	• Weill Medical College of Cornell University - Psychiatry and Public Health PWC Box 140 525 East 68th Street, New York, New York 10021 United States
	Schoenfeld, David Alan proxy	• Massachusetts General Hospital - Biostatistics Center 50 Staniford Street , bpston, Massachusetts 02114 United States
Contact Author (populates the ##PROLE_AUTHOR_..# # e-mail tags):	<input type="text" value="Rabideau, Dustin John"/>	☒ Save
Running Head:	Intent to Attend in a randomized trial	
Keywords:	intent to attend ✱, inverse probability weighting ✱, attrition ✱, intermittent missing data ✱, bipolar disorder ✱, LiTMUS ✱	

Current Contact Author:
Rabideau, Dustin John
([proxy](#))

References

1. Leon AC, Demirtas H, Hedeker D. Bias reduction with an adjustment for participants' intent to dropout of a randomized controlled clinical trial. *Clinical Trials*. 2007; 4:540–47. [PubMed: 17942469]
2. Laird NM. Missing Data in Longitudinal Studies. *Statistics in Medicine*. 1988; 7:305–15. [PubMed: 3353609]
3. Robins JM, Rotnitzky A, Zhao LP. Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association*. 1995; 90(429):106–21.
4. Hernán MA, Brumback B, Robins JM. Marginal Structural Models to Estimate the Casual Effect of Zidovudine on the Survival of HIV-Positive Men. *Epidemiology*. 2000; 11(5):561–70. [PubMed: 10955409]
5. Robins JM, Hernán MA, Brumback B. Marginal Structural Models and Casual Inference in Epidemiology. *Epidemiology*. 2000; 11(5):550–60. [PubMed: 10955408]
6. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, Evans DA, Mendes de Leon CF. Accounting for Bias Due to Selective Attrition: The Example of Smoking and Cognitive Decline. *Epidemiology*. 2012; 23(1):119–27. [PubMed: 21989136]
7. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychological Methods*. 2002; 7(2):147–77. [PubMed: 12090408]
8. Demirtas H, Schafer JL. On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine*. 2003; 22:2553–75. [PubMed: 12898544]
9. Hernán MA. Hernandez-Diaz S and Robins JM. A Structural Approach to Selection Bias. *Epidemiology*. 2004; 15(5):615–25. [PubMed: 15308962]
10. Suppes T, Dennehy EB, Hirschfeld RM, Altshuler LL, Bowden CL, Calabrese JR, Crismon ML, Ketter TA, Sachs GS, Swann AC, Texas Consensus Conference Panel on Medication Treatment of Bipolar Disorder. The Texas Implementation of Medication Algorithms: Update to the Algorithms for Treatment of Bipolar I Disorder. *Journal of Clinical Psychiatry*. 2005; 66(7):870–886. [PubMed: 16013903]
11. Sylvia LG, Reilly-Harrington NA, Leon AC, Kansky CI, Ketter TA, Calabrese JR, Thase ME, Bowden CL, Friedman ES, Ostacher MJ, Iosifescu DV, Severe J, Keyes M, Nierenberg AA.

Methods to limit attrition in longitudinal comparative effectiveness trials: lessons from the Lithium Treatment - Moderate dose Use Study (LiTMUS) for bipolar disorder. *Clinical Trials*. 2012; 9:94–101. [PubMed: 22076437]

12. Nierenberg AN, Sylvia LG, Leon AC, Reilly-Harrington NA, Ketter TA, Calabrese JR, Thase ME, Bowden CL, Friedman ES, Ostacher MJ, Novak L, Iosifescu DV, LiTMUS Study Group. Lithium Treatment—Moderate dose Use Study (LiTMUS): Rationale and design. *Clinical Trials*. 2009; 6:637–48. [PubMed: 19933719]
13. Nierenberg AN, Friedman ES, Bowden CL, Sylvia LG, Thase ME, Ketter T, Ostacher MJ, Leon AC, Reilly-Harrington NA, Iosifescu DV, Pencina M, Severe JB, Calabrese JR. Lithium Treatment Moderate-Dose Use Study (LiTMUS) for Bipolar Disorder: A Randomized Comparative Effectiveness Trial of Optimized Personalized Treatment With and Without Lithium. *The American Journal of Psychiatry*. 2013; 170(1):102–110. [PubMed: 23288387]
14. Spearing MK, Post RM, Leverich GS, Brandt D, Nolen W. Modification of the Clinical Global Impressions (CGI) scale for use in bipolar illness (BP): the CGI-BP. *Psychiatry Research*. 1997; 73(3):159–171. [PubMed: 9481807]
15. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*. 1979; 134:382–89. [PubMed: 444788]
16. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry*. 1978; 133:429–35. [PubMed: 728692]
17. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacology Bulletin*. 1993; 29(2):321–26. [PubMed: 8290681]
18. Leon AC, Solomon DA, Mueller TI, Turvey CL, Endicott J, Keller MB. The Range of Impaired Functioning Tool (LIFE–RIFT): a brief measure of functional impairment. *Psychological Medicine*. 1999; 29(04):869–78. [PubMed: 10473314]
19. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV). American Psychiatric Press, Inc.; Washington, D.C.: 1996.
20. Lecrubier Y, Braconnier A, Said S, Payan C. The impulsivity rating scale (IRS): preliminary results. *European Psychiatry*. 1995; 10(7):331–38. [PubMed: 19698364]
21. Preisser JS, Galecki AT, Lohman KK, Wagenknecht LE. Analysis of smoking trends with incomplete longitudinal binary response. *J Am Statist Assoc*. 2000; 95:1021–31.
22. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, Smith M. The ins and outs of an online bipolar education program: a study of program attrition. *J Med Internet Res*. 2010; 12(5):e57. [PubMed: 21169169]
23. Graff FS, Griffin ML, Weiss RD. Predictors of dropout from group therapy among patients with bipolar and substance use disorders. *Drug Alcohol Depend*. 2008; 94(1-3):272–75. [PubMed: 18162331]
24. Prisciandaro JJ, Rembold J, Brown DG, Brady KT, Tolliver BK. Predictors of clinical trial dropout in individuals with co-occurring bipolar disorder and alcohol dependence. *Drug Alcohol Depend*. 2011; 118(2-3):493–96. [PubMed: 21549529]
25. Tchetgen Tchetgen EJ, Glymour MM, Weuve J, Robins JM. Specifying the Correlation Structure in Inverse-Probability-Weighting Estimation for Repeated Measures. *Epidemiology*. 2012; 23(4): 644–46. [PubMed: 22659551]
26. Carpenter JR, Kenward MG, Vansteelandt S. A comparison of multiple imputation and doubly robust estimation for analyses with missing data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2006; 169:571–584.

SUBJECT ID	DATE OF VISIT	VISIT NUMBER
_____	____/____/____ Month Date Year	Week: _____

INTENT TO ATTEND

Please answer the questions by circling the correct answer or the answer which seems the most appropriate.

How likely is it that you will attend the next assessment session?

0	1	2	3	4	5	6	7	8	9
Not at all		Slightly		Moderately		Markedly		Extremely	

Figure 1.
Intent to Attend assessment.

Figure 2.

Estimated probabilities of retention by study week (GEE: IPAW model 2).

GEE: generalized estimating equations; IPAW: inverse probability of attrition weighting.

Figure 3.
Crude mean Clinical Global Impressions Bipolar Disorder Severity by treatment group.

Figure 4.

Model-based mean Clinical Global Impressions Bipolar Disorder Severity by treatment group.

MIXED: mixed effects regression; GEE: generalized estimating equations; IPAW: inverse probability of attrition weighting

Table 1

Missing outcome data in LiTMUS by study week.

	<u>Scheduled study week</u>								
	0	2	4	6	8	12	16	20	24
Frequency of intermittent missing	0	8	9	17	12	11	10	11	0
Frequency of monotonic missing	1	8	14	19	23	25	32	38	46

Intermittent missing: a patient who missed a visit and returned later. Monotonic missing: a patient who missed a visit and all subsequent visits.

Table 2

Adjusted hazard ratios and 95% confidence intervals of retention.

Variable	Mean $\pm$ SD or % at baseline	Adjusted hazard ratio [95% CI]	
		Model 1	Model 2
Log-time	.	0.91 [0.70 to 1.18]	0.88 [0.67 to 1.15]
Intent to Attend	.	1.16 [1.03 to 1.31]	1.19 [1.05 to 1.35]
Missed previous visit	.	0.02 [0.02 to 0.03]	0.03 [0.02 to 0.04]
Lithium group (vs. no lithium)	49.8%	0.89 [0.64 to 1.24]	0.95 [0.68 to 1.33]
CGI-BP Severity	4.3 $\pm$ 0.9	0.95 [0.83 to 1.08]	0.95 [0.76 to 1.18]
Age	39.1 $\pm$ 12.3	.	1.01 [0.99 to 1.02]
Female (vs. male)	56.5%	.	1.42 [1.01 to 2.00]
Some college (vs. none)	72.8%	.	1.11 [0.77 to 1.60]
Substance abuse (vs. none)	38.5%	.	0.85 [0.60 to 1.21]
IRS	6.0 $\pm$ 4.6	.	0.97 [0.94 to 1.01]
MADRS	22.5 $\pm$ 10.1	.	1.02 [0.99 to 1.04]
YMRS	12.5 $\pm$ 8.2	.	0.97 [0.95 to 0.99]
Q-LES-Q	45.8 $\pm$ 18.7	.	1.00 [0.99 to 1.02]
LIFE-RIFT	13.4 $\pm$ 3.6	.	1.01 [0.95 to 1.08]
Hosmer-Lemeshow GOF Test: p-value		0.65	0.47
Stabilized weights: mean $\pm$ SD [min, max]		1.05 $\pm$ 1.08 [0.43, 22.09]	1.07 $\pm$ 1.19 [0.23, 27.97]

CGI-BP: Clinical Global Impressions scale for use in Bipolar Disorder; IRS: Impulsivity Rating Scale; MADRS: Montgomery-Asberg Depression Rating Scale; YMRS: Young Mania Rating Scale; Q-LES-Q: Quality of Life Enjoyment and Satisfaction Questionnaire; LIFE-RIFT: Range of Impaired Functioning Tool; GOF: goodness of fit. Hazard ratio greater than 1 reflects a higher probability of retention; hazard ratio less than 1 reflects a higher probability of attrition.

Table 3

Model-based treatment effect estimates: Clinical Global Impressions Bipolar Disorder Severity.

Outcome model	Estimated change from baseline		Estimated 6-month difference	
	Lithium plus optimized treatment mean [95% CI]	Optimized treatment only mean [95% CI]	mean [95% CI]	P-value
GEE: unweighted	−1.12 [−1.35 to −0.88]	−1.20 [−1.42 to −0.97]	0.08 [−0.22 to 0.38]	0.60
GEE: IPAW model 1	−0.87 [−1.29 to −0.45]	−1.19 [−1.48 to −0.90]	0.32 [−0.10 to 0.73]	0.13
GEE: IPAW model 2	−0.77 [−1.20 to −0.35]	−1.16 [−1.45 to −0.86]	0.38 [−0.03 to 0.79]	0.07
MIXED: without intent	−1.19 [−1.41 to −0.97]	−1.25 [−1.47 to −1.04]	0.06 [−0.20 to 0.33]	0.64
MIXED: with intent	−1.22 [−1.44 to −0.99]	−1.30 [−1.52 to −1.08]	0.08 [−0.18 to 0.35]	0.55

MIXED: mixed effects regression; GEE: generalized estimating equations; IPAW: inverse probability of attrition weighting.