HHS Public Access

Author manuscript

Biostat Epidemiol. Author manuscript; available in PMC 2018 September 17.

Published in final edited form as:

Biostat Epidemiol. 2018 ; 2(1): 23–33. doi:10.1080/24709360.2017.1407866.

The Kaplan-Meier Method for Estimating and Comparing Proportions in a Randomized Controlled Trial with Dropouts

Jarcy Zee[a] and **Sharon X. Xie**[b,*]

[a]Arbor Research Collaborative for Health, 340 E. Huron Street Suite 300, Ann Arbor, MI, USA, 734-369-9853

[b]Biostatistics, Epidemiology, & Informatics, University of Pennsylvania, 607 Blockley Hall, 423 Guardian Drive, Philadelphia, PA, USA, 215-573-3867

Abstract

We propose a method for estimating and comparing proportions of study participants who reached an event of interest during a randomized controlled trial. Standard methods for estimating this proportion include the intent-to-treat method, which counts the number who reached the event of interest divided by the total number of participants, and the completers-only method, which counts the number who reached the event only among those who completed the entire study. When participants drop out of the study early, however, these methods will either be biased or inefficient. We propose to use the Kaplan-Meier method from survival analysis to estimate the proportion of interest in this non-survival setting. We show through extensive simulation studies that the Kaplan-Meier method has less bias and is more efficient than the standard methods. We demonstrate the performance of all methods for estimating proportions in one sample and for comparing proportions across two samples. Finally, we apply the proposed method to a data set for estimating and comparing proportions of patients who achieved treatment response during a Parkinson's disease trial for the treatment of impulse control disorders.

Keywords

Kaplan-Meier; proportion; dropout; clinical trial; survival

[*]Contact Sharon X. Xie. sxie@mail.med.upenn.edu.

Notes on contributors

Dr. Jarcy Zee is a research scientist in biostatistics at Arbor Research Collaborative for Health, a non-profit health research organization. Her methodological research interests are in survival analysis and measurement error, and she conducts applied clinical research in kidney disease, organ transplant, mental health, and women's health. Dr. Zee is an awardee of the Nephrotic Syndrome Study Network (NEP-TUNE) Career Development Fellowship to research statistical methods for analyzing glomerular disease morphology data. She is also an awardee of a Chronic Kidney Disease Biomarkers Consortium Pilot and Feasibility program study to research statistical methods for predicting time-to-event outcomes with time-varying and high-dimensional biomarkers. Dr. Zee received her PhD degree in biostatistics from the University of Pennsylvania.

Dr. Sharon X. Xie is a professor of biostatistics at the University of Pennsylvania. She is currently the principal investigator of a National Institute of Health (NIH) funded R01 grant, which aims to develop novel statistical methods for assessing dementia risk in Parkinson's disease and other neurodegenerative diseases. She is also the principal investigator of the biostatistics and data management core for two NIH-funded neurodegenerative disease research centers: the Alzheimer's Disease Core Center and Morris K. Udall Parkinson's Disease Research Center of Excellence. Dr. Xie develops new statistical methods for survival analysis, measurement error problems, missing data, longitudinal analysis, and receiver operating characteristic (ROC) analysis. She also investigates various scientific problems in neurodegenerative diseases and other biomedical areas. Dr. Xie received her PhD degree in Biostatistics from the University of Washington.

1. Introduction

In randomized controlled trials, the outcome of interest is often the proportion of study participants who reached an event during the study period. For example, a trial testing the efficacy of naltrexone for treating impulse control disorders (ICDs) in Parkinson's disease (PD) may seek the occurrence of treatment response during the 8-week study period, e.g., a Clinical Global Impression-Change (CGI-C) score of 1 or 2 at any time during follow-up. Estimation of the proportion of patients with treatment response can be accomplished by counting the number of participants with clinically relevant CGI-C during the study period and dividing by the total number of participants. However, when participants drop out of (withdraw from) the study prior to eight weeks without achieving the event of interest, they cannot contribute to the numerator of the proportion while still contributing to the denominator. This intent-to-treat method thus results in a biased underestimation of the proportion of events during the study period.

Several statistical methods have been proposed for analyzing clinical trials data in the presence of dropouts. These include nonparametric methods, including the intent-to-treat method and completers-only analysis [1,2]; imputation methods, including last-observation-carried-forward and multiple imputation [1–4]; and semiparametric or parametric methods, including generalized estimating equation (GEE) models and mixed effects models [1–4]. Imputation methods, semiparametric methods, and parametric methods are sometimes advantageous, e.g., for dealing with missingness at random or non-monotonic missingness, but make parametric assumptions compared to nonparametric methods. Here, we focus on nonparametric methods to minimize assumptions and because our outcome of interest is the simple proportion of participants who reached an event during the trial. We also assume the case in which missingness is monotonic (i.e., once a participant drops out, they do not return) and missingness completely at random, which is required for Kaplan-Meier estimation.

Consider Table 1, which shows the number of study participants n_{ce} with completion status $c \in \{0, 1\}$ and event status $e \in \{0, 1\}$ in each cell. Values in the column margins represent the total number of participants who did not or did experience the event during the study, $n_{\cdot 0}$ and $n_{\cdot 1}$, respectively. Values in the row margins represent the total number of participants who did not or did complete the study, $n_{0\cdot}$ and $n_{1\cdot}$, respectively.

The intent-to-treat (ITT) method would estimate the proportion of events during the study period as

$$\hat{p}_{\text{ITT}} = \frac{n_{\cdot 1}}{N}$$

with estimated variance $\hat{V}(\hat{p}_{\text{ITT}}) = \hat{p}_{\text{ITT}}(1 - \hat{p}_{\text{ITT}})/N$. The reason why this proportion is biased is that some participants in the n_{00} cell could have experienced the event if they had completed the study. Although they are not counted as part of the numerator $n_{\cdot 1}$, they are always counted as part of the denominator, N.

To overcome this bias, an alternative method is to calculate the proportion only among participants who reached the end of the study,

$$\hat{p}_{\text{comp}} = \frac{n_{11}}{n_{1.}}$$

with estimated variance $\hat{V}(\hat{p}_{\text{comp}}) = \hat{p}_{\text{comp}}(1 - \hat{p}_{\text{comp}})/n_{1.}$. Those who drop out of the study prior to eight weeks in the PD example cannot contribute to the numerator nor to the denominator of the proportion. Since we assume missingness completely at random, the trimmed sample used in this completers-only method is a random sample of the full study sample. Therefore, the estimated proportion $\hat{p}_{\text{comp}}$ is a consistent and optimally efficient estimator of the true proportion asymptotically. In finite samples, the completers-only method still results in little bias, but the use of a trimmed sample makes it less efficient compared to using the full study sample [5, page 42].

Rather than discarding those who drop out of the study, we propose to include them as censored observations and borrow tools from survival analysis to estimate the proportion of interest in this non-survival setting. Specifically, we propose to use the Kaplan-Meier method to nonparametrically estimate a survival function that includes all participants, whether they drop out or not. The proportion of participants who reached the event during the study period can then be calculated from the estimated survival function and also used for comparison across groups.

The rest of this article is organized as follows. We first introduce the Kaplan-Meier method and explain how the method can be used to estimate and compare proportions for randomized controlled trials in Section 2. In Section 3, we show the results of simulation studies that demonstrate the performance of the Kaplan-Meier method as compared to the intent-to-treat and completers-only methods for estimating and comparing proportions in the presence of dropout. We then demonstrate the proposed method using a data example from a trial testing the efficacy of naltrexone for treating impulse control disorders in PD in Section 4. Finally, we summarize our findings and discuss applications for which the proposed method is particularly useful (Section 5).

2. The Kaplan-Meier Method for Estimating and Comparing Proportions in the Presence of Dropouts

The Kaplan-Meier method is a nonparametric estimator of the survival function, the probability of being event-free at any time point t in the study [6],

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{t_i \le t} \left[1 - \frac{d_i}{Y_i}\right] & \text{if } t_1 \le t \end{cases}$$

where $t_1, t_2, \cdots t_D$ are ordered times at which events occur, d_i is the number of individuals who experience the event at time t_i, and Y_i is the number of individuals who are 'at-risk' of experiencing the event at time t_i. When participants drop out of the study before experiencing the event or when participants are still event-free at the end of the study, they are considered right-censored in the survival analysis framework. The Kaplan-Meier estimator takes into account right-censored observations by counting them in the at risk set only until they drop out of the study. To apply this survival method to a randomized controlled trial with dropouts, we first calculate each participant's follow-up time in the study. For those who experience the event by the time point of interest, their follow-up time is the time from study entry to the event. For those who do not experience the event by the time point of interest, their follow-up time is the time from study entry until dropout or the time point of interest, whichever comes first. We assume that dropouts occur completely at random. We can then calculate the Kaplan-Meier survival estimate using the event times for those who experience the event and treating the dropouts as right-censored observations. To estimate the proportion of study participants who experienced the event by a specific time point τ, we calculate one minus the Kaplan-Meier survival estimate at that time point,

$$\hat{p}_{\mathrm{KM}} = 1 - \hat{S}(\tau).$$

Using the Kaplan-Meier method to estimate a proportion is similar to Maller and Zhou's method for estimating the proportion of immunes in a sample [7]. However, the proportion of immunes is estimated only at the end of the study period and estimates the proportion who did *not* reach the event. Here, the question of interest is the proportion who did reach the event and the time point of interest may not be at the end of the study period. Since the Kaplan-Meier method estimates a survival function over time, we can easily change the time point of interest. This is another advantage of the Kaplan-Meier method compared to the completers-only method, since the completers-only method can result in a different study sample at each time point.

Several methods have been proposed to estimate variances for the Kaplan-Meier survival estimate at time t, one of the most common of which uses Greenwood's formula [8],

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \leq t} \frac{d_i}{Y_i(Y_i - d_i)}.$$

We can similarly use Greenwood's formula to estimate the variance of the proportion of interest, since $\hat{V}(\hat{p}_{\mathrm{KM}}) = \hat{V}[\hat{S}(\tau)]$. To compare proportions across groups at a particular time of interest τ, we can calculate a Wald test statistic,

$$Z = \frac{\hat{p}_{1\mathrm{KM}} - \hat{p}_{2\mathrm{KM}}}{\sqrt{\hat{V}(\hat{p}_{1\mathrm{KM}}) + \hat{V}(\hat{p}_{2\mathrm{KM}})}} = \frac{\hat{S}_1(\tau) - \hat{S}_2(\tau)}{\sqrt{\hat{V}[\hat{S}_1(\tau)] + \hat{V}[\hat{S}_2(\tau)]}},$$

and conduct a Wald-type test [9], where $\hat{S}_j(\tau)$ is the Kaplan-Meier survival estimate for group j at time τ and $\hat{p}_{j\text{KM}}$ is the estimate of the proportion who experienced the event in group j by time τ.

3. Simulation Studies

3.1. Estimating Proportions in One Sample

We conducted a series of simulation studies to evaluate the performance of the Kaplan-Meier method compared to the intent-to-treat and completers-only methods for estimating and comparing proportions in the presence of dropout. For estimating a proportion in a single group, we set the trial sample size to $N \in \{30, 60, 120, 200\}$. We first sampled an event time and dropout time for each participant from exponential distributions, $T \sim \text{Exp}(\lambda_T)$ and $U \sim \text{Exp}(\lambda_U)$, respectively. The true proportion of events was determined by $p = 1 - \exp\{\lambda_T \cdot t\}$ at time t. Exponential rate parameters were varied to simulate low, moderate, and high amounts of dropout $(\lambda_T, \lambda_U) \in \{(4, 12), (8, 8), (12, 4)\}$, respectively. We then sampled event times and dropout times from uniform distributions, $T \sim \text{Unif}[0, \theta_T]$ and $U \sim \text{Unif}[0, \theta_U]$. For the uniform distribution, the true proportion of events was determined by $p = t/\theta_T$ at time t. To simulate low, moderate, and high amounts of dropout, we set the maximum of the uniform distributions to $(\theta_T, \theta_U) \in \{(16, 20), (16, 16), (20, 16)\}$, respectively. Observed event times were calculated by $X = \min(T, U)$ and event indicators by $\delta = I(T \leq U)$. Using the observed data, (X, δ), we estimated the proportion of events using the intent-to-treat method, the completers-only method, and the Kaplan-Meier method at each of three time points of interest, $t \in \{4, 8, 12\}$. We also estimated the proportion of events using the simulated event times, T, which yields the true proportion estimate if there was no dropout. This true proportion estimate would not be available in real data.

Each combination of parameters constituted a simulation, which was repeated 1000 times. To compare across methods, we estimated bias (proportion estimate – true proportion), estimated standard errors, observed sample standard deviations, relative efficiency compared to the true proportion estimate (variance/variance of true proportion estimate), and 95% confidence interval coverage. The 95% confidence interval for the Kaplan-Meier method was calculated using log-log transformation, as is standard in practice [10]. Simulation results for estimating proportions using data simulated from the exponential distribution, $N = 60$, and $t = 8$ are shown in Table 2. We first note that all simulation results show that observed sample standard deviations are close to standard error estimates for each estimator.

Bias was lowest for the true proportion estimate, as expected. The intent-to-treat method resulted in large negative bias which increased as the amount of dropouts increased. Both the completers-only method and Kaplan-Meier method had little bias in estimating proportions. Compared to the completers-only method, the Kaplan-Meier method had lower standard errors, relative efficiency closer to 1, and coverage closer to 95%. As the amount of dropouts increased, the efficiency gains of the Kaplan-Meier method compared to the completers-only method were more pronounced. Similar results were obtained under all other simulation parameters, including with uniformly distributed event and dropout times, smaller and larger sample sizes, and other time points of interest (not shown; available upon request).

3.2. Comparing Proportions across Two Samples

To compare proportions across two groups, we assumed equal group sizes and $N_1 = N_2 \in \{20, 40, 60, 80\}$ in each group. Data for group 1 were sampled as described above for estimating proportions in one group. We then set the difference between proportions of events in group 2 and group 1 to be $\Delta_T \in \{0, 0.1, 0.2, \cdots, \min(1 - p_1, 0.5)\}$ at the time point of interest, where p_1 is the true proportion of events in group 1 and the maximum value of Δ_T is dependent on p_1 to avoid proportions of events that are outside of the range [0, 1] for group 2. When the group 1 event times were sampled from an exponential distribution with rate parameter λ_{1T}, group 2 event times were sampled from an exponential distribution with rate parameter $\lambda_{2T} = -1/t \cdot \ln[\exp\{-\lambda_{1T} \cdot t\} - \Delta_T]$. When the group 1 event times were sampled from a uniform distribution with maximum θ_{1T}, group 2 event times were sampled from a uniform distribution with maximum $\theta_{2T} = t/[\Delta_T + t/\theta_{1T}]$. We also allowed the dropout distributions across the two groups to differ by setting the difference between proportions of dropouts to be $\Delta_U \in \{0, 0.1, 0.2\}$. Dropout distribution parameters for group 2 were calculated similarly to group 2 event distribution parameters.

For each combination of parameters, we estimated proportions and their variances in each group using the true proportion estimator, intent-to-treat method, the completers-only method, and the Kaplan-Meier method. We then conducted a Wald-type test for differences in proportions across the two groups, H_0: $\Delta_T = 0$, using each method. To compare across methods, we calculated bias (estimated difference – true difference), standard error of the estimated difference, and 95% confidence interval coverage. Simulation results for comparing proportions using data simulated from the uniform distribution, $N_1 = N_2 = 40$, $t = 8$, and no differences in dropout rates across the two groups are shown in Table 3.

Bias in estimating the difference in proportions between two groups was lowest for the true proportion estimate and highest for the intent-to-treat method. As the difference in proportions increased, the bias from the intent-to-treat method was increased. Across all simulations, both the completers-only method and Kaplan-Meier method had little bias in estimating the difference in proportions. Compared to the completers-only method, the Kaplan-Meier method had lower standard errors, which resulted in coverage closer to 95%. Similar results were obtained under all other simulation parameters, including with exponentially distributed event and dropout times, lower and higher sample sizes, other time points of interest and effect sizes, and with a difference in dropout rates across the two groups (not shown; available upon request).

4. Data Example

PD is an age-related neurodegenerative disease with both motor and non-motor symptoms. Treatment for non-motor symptoms such as ICD is complex and understudied. To test the efficacy of naltrexone, an opioid antagonist, for ICDs in PD, a randomized controlled trial was conducted [11]. A total of $N = 50$ participants were randomly assigned to a treatment (naltrexone) or placebo group and evaluated every two weeks for a total of 8 weeks. The CGI-C was used to measure the change in ICD symptoms, which ranges from 1–7 with 1 indicating very much improved and 7 indicating very much worse compared to baseline. For the current study, the outcome of interest is the occurrence of a treatment response at any

time during the 8-week study period, with treatment response defined as a score of 1 (very much improved) or 2 (much improved) on the CGI-C.

To estimate and compare the proportions of participants who achieved treatment response during the study period in each of the treatment and placebo groups, we used the intent-to-treat method, the completers-only method, and the Kaplan-Meier method. For the intent-to-treat method, we counted the number of participants who achieved treatment response at any time during the 8-week study period and divided by the total sample size, $N = 50$. For the completers-only method, we calculated the proportion who achieved treatment response at any time during the study only among the 45 participants who completed the entire study. For the Kaplan-Meier method, we calculated the time to first treatment response for each participant. Those who did not achieve a treatment response but completed the study were censored at eight weeks. The five participants who did not achieve a treatment response and did not complete the entire study were censored at their last follow-up visit for the Kaplan-Meier method. Two of these five participants dropped out of the study before week two, so they are censored at time 0 and therefore do not contribute to the Kaplan-Meier estimate.

Table 4 shows the estimated proportions using each method, the differences across the treatment and placebo groups, and the result of the hypothesis test for no difference across groups.

The ITT method resulted in smaller proportion estimates in each group and a smaller difference compared to the completers-only and Kaplan-Meier methods. The completers-only method and Kaplan-Meier method gave similar proportion estimates, standard errors, and differences. This is likely due to the fact that there was little dropout and two of the five who dropped out did so before week two. Therefore, there were only three additional participants included in the Kaplan-Meier calculation compared to the completers-only calculation. The potential efficiency gains of the Kaplan-Meier method are therefore not apparent.

To better illustrate the advantages of the Kaplan-Meier method, we randomly chose an additional 10 participants in the study to drop out between weeks two and six. Therefore, our modified data had a total of 35 participants who completed the study. Table 5 shows the estimated proportions using each method, the differences across treatment and placebo groups, and the result of the hypothesis test for no difference across groups.

Again, the ITT method likely underestimated each proportion and the difference in proportions. The completers-only and Kaplan-Meier methods estimated similar proportions. However, with more dropout, we can clearly see that the Kaplan-Meier method was more efficient, since standard errors of proportion estimates were lower than those from the completers-only method and the 95% confidence interval is narrower.

5. Discussion

Through extensive simulation studies and a data example, we showed that the Kaplan-Meier method can be used to accurately and efficiently estimate proportions of participants who experience an event during a randomized controlled trial with dropouts. By using each

participant's follow-up time and event status, we were able to apply a survival analysis method in a non-survival setting to estimate proportions. We demonstrated that the Kaplan-Meier method outperformed the standard intent-to-treat and completers-only methods. It does so by using information on those who do not experience the event of interest before dropping out of the study—namely, that these participants were event-free until their dropout time. In contrast, the intent-to-treat method assumes these participants will never experience the event during the study even if they had not dropped out. The completers-only method excludes all participants who drop out of the trial early, regardless of whether they experience the event or not.

The benefits of the Kaplan-Meier method are most apparent when there are many dropouts from the trial. However, even with few dropouts, the Kaplan-Meier method still produces proportion estimates with little bias and close to optimal efficiency. The Kaplan-Meier method would suffer from the same inefficiency as the completers-only method if dropouts occur before the first follow-up time, since they similarly cannot be included in analyses. This highlights the fact that the Kaplan-Meier method is most useful when there are multiple follow-up times throughout the trial to assess whether participants achieved the event of interest. For trials that are designed only to evaluate participants at the beginning and end of the trial, e.g., for trials only interested in the proportion of events at the end of study, the Kaplan-Meier method would give similar results as the completers-only method.

For trials in which the outcome of interest is the proportion of events at—rather than by—a specific time point, other methods should be considered. For example, the original analyses of the PD trial data focused on the proportion of participants who achieved a treatment response at eight weeks [11]. To answer this question, a GEE model was used, particularly since it allows and can account for participants achieving the event multiple times during the trial. However, the GEE requires a good model fit and cannot estimate the proportion who had an event during the entire study period. The Kaplan-Meier method proposed in this study is nonparametric, thereby making few assumptions, and is able to estimate the proportion who experienced an event during the entire trial.

One assumption we do make is that dropouts occur completely at random, or equivalently that the missing data mechanism due to dropout is completely at random. This is an assumption required by the Kaplan-Meier method as well as the completers-only methods, since they are both nonparametric. For situations in which data are missing at random, alternative methods that make additional assumptions may be needed, like mixed effects models.

Given the costs of conducting a randomized controlled trial, it is important to maximize efficiency. Although most trials are designed to prevent as much dropout as possible, it is inevitable that some patients will drop out of trials early due to uncontrollable circumstances. To mitigate the consequences of these dropouts, we recommend the use of the Kaplan-Meier method over the intent-to-treat or completers-only methods to estimate the proportion of participants who experience an event during a randomized controlled trial.

Acknowledgments

We thank Daniel Weintraub, MD, and the naltrexone trial study team for providing the data used for illustration in this study.

Funding

Dr. Xie received support from NIH grants R01-NS102324, AG10124, AG17586, and NS053488 to conduct this research.

References

1. Myers WR. Handling missing data in clinical trials: An overview. Drug Information Journal. 2000; 34(2):525–533.
2. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. New England Journal of Medicine. 2012; 367(14):1355–1360. [PubMed: 23034025]
3. Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the locf and mmrm approaches. Pharmaceutical Statistics. 2008; 7(2):93–106. [PubMed: 17351897]
4. DeSouza CM, Legedza ATR, Sankoh AJ. An overview of practical approaches for handling missing data in clinical trials. Journal of Biopharmaceutical Statistics. 2009; 19(6):1055– 1073. [PubMed: 20183464]
5. Little RJ, Rubin DB. Statistical analysis with missing data. 2. Hoboken: John Wiley & Sons, Inc; 2002. 42
6. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association. 1958; 53(282):457–481.
7. Maller RA, Zhou S. Estimating the proportion of immunes in a censored sample. Biometrika. 1992; 79(4):731–739.
8. Greenwood M. The natural duration of cancer. London: His Majesty's Stationery Office; 1926. 1–26.
9. Klein JP, Moeschberger ML. Survival analysis: Techniques for censored and truncated data. 2. New York: Springer-Verlag, Inc; 2003. Statistics for Biology and Health
10. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2. Hoboken: John Wiley & Sons, Inc; 2002.
11. Papay K, Xie SX, Stern M, et al. Naltrexone for impulse control disorders in parkinson disease: A placebo-controlled study. Neurology. 2014; 83(9):826–833. [PubMed: 25037206]

Table 1

2 × 2 table of study completion status and event status

		Event		
		0	**1**	**Total**
Completed Study	0	n_{00}	n_{01}	$n_{0.}$
	1	n_{10}	n_{11}	$n_{1.}$
	Total	$n_{.0}$	$n_{.1}$	N

Table 2

Simulation Results for Estimating Proportions, $N = 60$, $t = 8$

Dropout Amount	*p*	Method	Bias	Standard Error	Standard Deviation	Relative Efficiency	95% CI Coverage
low	0.865	True	0.000	0.043	0.044	1.000	0.910
		ITT	−0.166	0.059	0.060	1.961	0.160
		Comp	0.002	0.059	0.063	1.930	0.891
		KM	0.002	0.054	0.059	1.578	0.931
moderate	0.632	True	0.003	0.062	0.063	1.000	0.934
		ITT	−0.197	0.063	0.066	1.068	0.136
		Comp	0.002	0.101	0.103	2.711	0.936
		KM	0.000	0.083	0.084	1.851	0.952
high	0.487	True	−0.001	0.064	0.063	1.000	0.955
		ITT	−0.256	0.054	0.054	0.709	0.008
		Comp	0.002	0.165	0.191	7.123	0.879
		KM	−0.006	0.115	0.131	3.352	0.907

p: true proportion of events; CI: confidence interval; True: true proportion estimate; ITT: Intent-to-Treat; Comp: Completers-Only; KM: Kaplan-Meier

Table 3

Simulation Results for Comparing Proportions, $N_1 = N_2 = 40$, $t = 8$, $\Delta_U = 0$

Dropout Amount	Δ_T	Method	Bias	Standard Error	95% CI Coverage
low	0	True	−0.004	0.110	0.954
		ITT	−0.003	0.108	0.954
		Comp	−0.005	0.143	0.942
		KM	−0.004	0.128	0.939
	0.2	True	−0.002	0.106	0.940
		ITT	−0.043	0.109	0.924
		Comp	−0.002	0.136	0.947
		KM	−0.003	0.124	0.946
	0.4	True	0.001	0.091	0.936
		ITT	−0.083	0.104	0.896
		Comp	0.004	0.116	0.933
		KM	0.003	0.108	0.955
moderate	0	True	−0.003	0.110	0.945
		ITT	0.000	0.107	0.948
		Comp	−0.002	0.155	0.940
		KM	−0.002	0.135	0.940
	0.2	True	0.001	0.106	0.949
		ITT	−0.049	0.109	0.915
		Comp	−0.001	0.149	0.950
		KM	−0.001	0.131	0.949
	0.4	True	−0.001	0.091	0.936
		ITT	−0.102	0.105	0.838
		Comp	−0.003	0.128	0.926
		KM	−0.002	0.115	0.943
high	0	True	−0.002	0.108	0.943

Dropout Amount	Δ_T	Method	Bias	Standard Error	95% CI Coverage
0.2		ITT	−0.002	0.101	0.942
		Comp	0.000	0.152	0.925
		KM	−0.005	0.131	0.939
		True	0.000	0.108	0.934
		ITT	−0.049	0.105	0.922
		Comp	0.002	0.152	0.930
		KM	0.000	0.132	0.941
0.4		True	0.000	0.099	0.949
		ITT	−0.098	0.104	0.844
		Comp	0.001	0.139	0.918
		KM	0.002	0.122	0.932

Δ_T: true difference in proportions of events; CI: confidence interval; True: true proportion estimate; ITT: Intent-to-Treat; Comp: Completers-Only; KM: Kaplan-Meier

Table 4

Comparison of Proportions with Treatment Response (CGI-C of 1 or 2) across Naltrexone Treatment and Placebo Groups

	Treatment		Placebo		Difference				
Method	$\hat{p}$	**SE**	$\hat{p}$	**SE**	**Δ**	**SE**	***Z***	**p-value**	**95% CI**
ITT	0.500	0.098	0.417	0.101	0.083	0.141	0.59	0.553	(−0.19,0.36)
Comp	0.591	0.105	0.435	0.103	0.156	0.147	1.06	0.289	(−0.13,0.45)
KM	0.564	0.105	0.430	0.103	0.134	0.147	0.91	0.363	(−0.15,0.42)

$\hat{p}$: estimated proportion; SE: standard error; *Z*: Wald test statistic; CI: confidence interval; ITT: Intent-to-Treat; Comp: Completers-Only; KM: Kaplan-Meier

Table 5

Comparison of Proportions with Treatment Response (CGI-C of 1 or 2) across Naltrexone Treatment and Placebo Groups with Increased Dropout

	Treatment		Placebo		Difference				
Method	$\hat{p}$	**SE**	$\hat{p}$	**SE**	**Δ**	**SE**	***Z***	**p-value**	**95% CI**
ITT	0.500	0.098	0.417	0.101	0.083	0.141	0.59	0.553	(−0.19,0.36)
Comp	0.625	0.121	0.474	0.115	0.151	0.167	0.91	0.364	(−0.18,0.48)
KM	0.623	0.112	0.460	0.109	0.163	0.156	1.05	0.296	(−0.14,0.47)

$\hat{p}$: estimated proportion; SE: standard error; *Z*: Wald test statistic; CI: confidence interval; ITT: Intent-to-Treat; Comp: Completers-Only; KM: Kaplan-Meier