

Appendix E. PCATS Online Application User Guide

PCATS: An Online App Implementing Bayesian Nonparametric Causal Inference

**Chen Chen and Bin Huang
Division of Biostatistics and Epidemiology
Cincinnati Children's Hospital Medical Center
Cincinnati, OH 45229**

Bin.huang@cchmc.org

**Version 1.1.
November 2, 2018**



Table of Contents

1. Introduction.....	3
2. Methods	4
2.1. Overview	4
2.2. GPMatch.....	5
2.3. BART.....	7
3. Before Running PCATS.....	8
4. User's Guide	9
4.1. Getting Started.....	9
Example 1.1: A Simple Binary Treatment Setting.....	9
4.2. Non-Adaptive Types of Treatments.....	15
Example 1.2: Binary Treatment with a Bounded Continuous Outcome	15
Example 1.3: Multi-level Treatment with a Post-treatment Confounder	19
Example 1.4: Binary Treatment with a Counting Outcome.....	25
Example 1.5: Subgroup Treatment Effect for a Binary Treatment.....	27
Example 1.6: Continuous Treatment in Dose-Response Setting.....	30
Example 1.7: Composite Binary and Continuous Treatments.....	34
4.3. Adaptive Types of Treatments	38
Example 2.1: Binary Treatment at Both Decision Point	38
Example 2.2: Subgroup HTE for a Binary Treatment at Both Decision Point	44
Example 2.3: Binary outcome at stage 1 and continuous outcome at stage 2	53

1. Introduction

The PCATS online App is designed to offer an online graphic user interface (GUI) that implements Bayesian causal inference methods for various types of treatment encountered in comparative effectiveness research. Here, the term treatment is used in a broader sense, and refers to actions or exposures that are expected to generate different outcomes. Some examples of a treatment include medication therapy to treat an illness, a diet plan to manage weight gain, or an educational program to improve certain knowledge/skills. The simplest type of treatment is a binary dichotomous treatment, where only two choices of the treatment are considered, such as a new intervention vs. a conventional control. However, most of the time, the treatment cannot be characterized in this simple way, rather, treatment is much more intricate. A treatment could have more than one category. For example, a diet intervention could be a vegetarian, Paleo, low carb or Atkins diet. It could be a combination of multiple treatments, such as a weight management program that includes both diet and exercise. Additionally, treatment could be continuous, such as the number of fruit servings or count of steps taken each day. Furthermore, a treatment could change over time. For example, a cancer patient may undergo a surgery first before going through chemotherapy. The time varying treatment can be **adaptive or non-adaptive**. The adaptive treatment is not determined ahead of time, but rather adapts to what happens as treatment progresses, thus treatment is constantly changing over time. The non-adaptive treatment is determined ahead of time, and adjustments over time are pre-planned regardless of what happens after the treatment. Most of the commonly used causal inference methods consider the simple non-adaptive binary type of treatment. The PCATS App implements Bayesian's nonparametric causal inference approaches which consider both adaptive and non-adaptive types of treatment. This can be applied to both experimental and nonexperimental observational data.

PCATS stands for patient centered adaptive treatment strategy. The PCATS project was motivated by comparative effectiveness researchers aiming to find better treatment strategies for patients with chronic illnesses. Routine clinical approaches seeking to treat patients with chronic or prolonged disease conditions are adaptive, where the patients often go through many courses or stages of treatment. Often, it is unclear as to what treatment plan works the best and how the treatment should be adjusted over time, yet, such knowledge could make an enormous difference in helping patients. Motivated by the patient centered comparative effectiveness questions, specifically "What is the best 1st line treatment option for a patient like me?", and "What is a better next treatment option given my past treatment history and disease progression?", the PCATS study is aimed to understand effectiveness of complex treatment types utilizing data collected from real clinical encounters by: 1) refining and improving Bayesian causal inference methods; and 2) evaluating effectiveness of different time-varying adaptive treatment strategies for treating children with newly diagnosed juvenile idiopathic arthritis (JIA).

Usually patient reported outcomes are obtained via questionnaires and are bound by the minimum and maximum possible values. Many clinical outcomes measured are also subject to ceiling or floor constraints, such as laboratory measures being limited by the lower and/or upper detection limits. Physician global evaluation of the patient disease activity is bounded between 0 and 1. The bounded feature of an outcome measure requires appropriate consideration in the CER. Failure to do so could bias the study results. Therefore, the PCATS App offers the option for users to specify outcome bounds and implement the methods properly in order to evaluate bounded outcomes.

This online application implements two Bayesian's nonparametric causal inference methods: GPMatch for continuous outcome, and BART for non-continuous outcomes, including binary and counting types of outcomes.

The PCATS App offers many advanced features for implementing robust causal inference with:

- 1) complex treatment, including binary, multilevel, continuous and composite, and time-varying adaptive treatment, as well non-adaptive treatment strategies
- 2) heterogeneous causal treatment effect
- 3) outcome measure bounded by ceiling and/or floor values
- 4) post-treatment confounding factors, such as adherence (under construction)

2. Methods

2.1. Overview

Statistical causal inference is specifically designed to answer causal questions. Unlike the prediction model, its goal is to understand how outcomes may differ by different decision. For this reason, causal inference is the standard method used for comparative effectiveness research. To better understand the difference between the causal inference and the predictive analyses, let's consider an example. If we know a patient's past premature birth history, family history, weight, age, smoking behavior, drinking behavior, general health and laboratory tests of a pregnant mother, we can predict very well her risk of a premature birth outcome. However, the prediction model does not inform if a life style intervention program may help lower her risk. To answer this causal question, we wish to know two potential outcomes, one is the birth outcome had the mother gone through the intervention (denoted as $Y(1)$), and the other is the birth outcome without intervention (denoted as $Y(0)$). If we have the knowledge of both potential outcomes, then the causal effect is simply the contrast between the two potential outcomes, i.e. $Y(1)-Y(0)$. However, we could never observe both potential outcomes, as there can be only one birth outcome for a baby, i.e. we could only observe $Y(1)$ if the mother took the intervention, observe $Y(0)$ if she did not .

A randomized controlled trial is one approach used to answer a causal question. With randomized and well controlled experiments, the results we observe in those randomized to the intervention group can be used to estimate the unobserved potential outcome for those randomized into the control group. In other words, when the two groups are exchangeable their potential outcomes can be considered exchangeable, and this is the fundamental idea behind the statistical causal inference. Therefore the key question, in causal inference, is to find the right exchangeable groups.

There are many statistical causal inference methods available which can broadly be classified into three categories: the propensity score method, outcome modeling and the combination methods. The propensity score (PS) method assumes we have the accurate knowledge of the propensity of treatment assignment for each individual then consider the individuals with the similar propensity score being exchangeable (Rosenbaum and Rubin, 1983). The outcome modeling approach supposes we can correctly model the scientific mechanisms underlying the generation of the potential outcomes. Individuals with the same risk of developing certain outcomes, or prognostic scores (PgS), are exchanges. Given the same PgS, the difference in

their outcomes between the treated and untreated groups are due to the treatment (Hansen, 2007). The combination methods use both PS and outcome modeling when estimating the treatment effect in order to provide additional protection against model miss-specification. Propensity score regression adjustment and augmented inverse probability of treatment weight (AIPTW, Bang & Robins, 2005) are two most commonly used combination methods. More recently, it was suggested that the two-staged approach by matching individuals first, followed by fitting an outcome model on the matched sample, performs better than many other methods, particularly when the model may not be correctly specified (Kang & Schafer, 2007).

Bayesian approach holds great promise in comparative effectiveness research, because it offers a coherent framework that allows synthesizing of evidences from different sources, it is highly flexible, and it's able to deal with many complex settings. Bayesian's nonparametric additive regression tree (BART, Hillman 2011) has proven to outperform many existing methods, particularly if the true model is a nonlinear function. Other studies have used Bayesian's nonparametric technique in order to ease the impact of potential model misspecification (Daniel et al, 2012; Roy et al 2017; Hanh & Murray, 2018). Up to date, the most of Bayesian' approach to the casual inference fall within the outcome modeling category. Much efforts are ongoing to incorporate knowledge of the treatment selection into Bayesian's causal inference. First, including PS as a covariate in outcome modeling may seem to be the most straightforward, but it does not address uncertainty in the estimates of PS. Joint modeling of treatment selection and outcome on the other hand introducing the model "feedback" issue (Zigler & Dominici, 2014). A proximate Bayesian approach was proposed as a Bayesian's version of AIPTW (Saarela & Liu, 2016). However, it suffers from the same limitations of AIPTW. Namely, the estimate treatment effect can be unstable as it subjects to undue influences from a few data points with PS close to 0 or 1. In an effort of searching for better Bayesian's causal inference methods, we proposed GPMatch, a full Bayesian's approach that allows for outcome modeling with accounted for confounding in treatment selection. We also extended BART for time-varying treatment. Extensive simulation studies examine the performances of GPMatch against many existing methods. Under the most realistic setting, when both outcome and propensity score modeling are miss-specified, these simulation studies have shown that our methods clearly outperform many existing causal inference methods.

At last, most of the patient and care provider reported outcomes are bounded measures where values are only between a minimum and a maximum possible value. In addition, many clinical outcome measures such as laboratory measures are also bounded. When the bounded nature of such outcome measures is not properly considering, the effectiveness estimation is biased. Different methods have been proposed to address the bounded outcome measures (Arostegui, Núñez-Antón, & Quintana, 2012). In this App, we implemented Tobit regression (Chib, 1992) to address the issue.

2.2. GPMatch

Gaussian process (GP) prior has been widely used to describe biological, social financial, and physical phenomena, due to its ability to model highly complex dynamic systems and its many desirable mathematical properties. The posterior consistency has been established in recent literatures for the Bayesian partially linear GP regression model (Choi & Woo, 2013).

The GPMatch method was proposed to offer a full Bayesian's approach, where matching is performed for each individual using the rest of data points in the study sample. Only those

individuals who are deemed similar to the given individual will contribute to the estimate of the potential outcome for a given individual. The degree of matching ranges from 100% matching to 0% matching and depends upon individual similarities or dissimilarities. The similarity is determined by the set of confounders expected to determine the treatment selection mechanism and the science mechanism underlying the outcome.

GPMatch utilizes GP prior as a matching tool. How the GP prior accomplishes the matching purpose can be understood from the weight-space point of view for the regression modeling. All regression modeling can be viewed as a model expressing outcomes by a weighted sum of the observed data over the given covariate space. In the case of the simple linear regression model, the weight is determined by the sample correlations between the covariates (X variable) and the observed responses (Y variable) for all individual data points. Unlike the linear regression, the GP regression is a nonparametric regression model, where the weight is determined for each individual data point by its relationship with the rest of the data points in the sample. The relationship is used to describe the similarity or the distance between these and other data points. In GPMatch, the distance can be viewed as a general version of the Mahalanobis distance within the p -dimensional covariate spaces. When the two individuals are identical in all values of the covariates, then their distance is 0 and the weight is assigned a value of 1, indicating that the two are completely matched. The weight decreases exponentially as the distance increases, which quickly declines to 0. The rate of decline is governed by the length scale parameters of the GP prior which is part of the model parameters to be determined by fitting the data. Therefore, the GPMatch imputes the unobserved outcome for any given individual, by a weighted sum of the observed outcomes from those matched (either completely or partially) individuals where the matching is determined by the observed covariates.

Since the matching and the outcome is done in a single step, GPMatch is more efficient than the two-staged approach suggested by Gutman and Rubin, (2017). With nonparametric flexible modeling, it mitigates the concerns over model misspecification. Extensive empirical studies have demonstrated that the method enjoys well calibrated frequentist properties. That is, it offers an unbiased point estimate and a nominal interval estimate. Further, the method outperforms many widely used causal inference methods, including variations of propensity score methods and matching methods, in terms of accuracy and robustness to potential model misspecification.

The GPMatch method requires a less stringent causal assumption than those widely adopted in many conventional causal inference methods. First, it assumes the observed outcome is a noisy version of the true potential outcome, i.e. stochastic consistency causal assumption (VanderWeele & Hernan, 2013). Second, it assumes that the true functional form of the scientific mechanisms that generate potential outcomes fall within the space of functions representable by a Gaussian process. And lastly, it requires the expected conditional mean of the marginal distribution of potential outcomes to satisfy ignorable assumptions. In other words, conditional on the observed covariates, the expected marginal mean of the potential outcomes stays the same despite the treatment actually assigned. Detailed discussions about how these assumptions compare with the widely adopted three causal inference assumptions: stable unit treatment value assumption (SUTVA), strong ignorable treatment assignment assumption and the positivity assumptions, are discussed in detail elsewhere (Huang et al, 2018).

2.3. BART

Bayesian's adaptive regression tree (BART) is a Bayesian's nonparametric modeling technique. It has been shown to offer more accurate estimates of averaged causal treatment effect when the true relationship between the outcome and the covariates are nonlinear, compared to many commonly used causal inference methods such as propensity score matching, propensity weighted estimates, and propensity regression adjustment (Hills, 2011). In the PCATS App, we extend the BART for time-varying adaptive treatment assignment. BART requires assumptions of SUTVA, strong ignorable treatment assignment assumption and the positivity assumptions.

References:

- Arostegui, I., Núñez-Antón, V., & Quintana, J. M. (2012). Statistical approaches to analyze patient-reported outcomes as response variables: An application to health-related quality of life. *Statistical Methods in Medical Research*, 21(2), 189–214. <http://doi.org/10.1177/0962280210379079>
- Bang, H., & Robins, J. M. (2005). Doubly Robust Estimation in Missing Data and Causal Inference Models. <http://doi.org/10.1111/j.1541-0420.2005.00377.x>
- Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, 51(1–2), 79–99. [http://doi.org/10.1016/0304-4076\(92\)90030-U](http://doi.org/10.1016/0304-4076(92)90030-U)
- Choi, T., & Woo, Y. (2013). On asymptotic properties of Bayesian partially linear models. *Journal of the Korean Statistical Society*, 42(4), 529–541.
- Daniels, M. J., Roy, J. A., Kim, C., Hogan, J. W., & Perri, M. G. (2012). Bayesian Inference for the Causal Effect of Mediation. *Biometrics*, 68(4), 1028–1036. <http://doi.org/10.1111/j.1541-0420.2012.01781.x>
- Gutman, R., & Rubin, D. B. (2017). Estimation of causal effects of binary treatments in unconfounded studies. *Statistical Methods in Medical Research*, 26(3), 1199–1215. <http://doi.org/10.1002/sim.6532>
- Hill, J. L. (2011). Bayesian Nonparametric Modeling for Causal Inference. *Journal of Computational and Graphical Statistics*, 20(1), 217–240. <http://doi.org/10.1198/jcgs.2010.08162>
- Hahn, P. R., Murray, J., & Carvalho, C. M. (n.d.). *BAYESIAN REGRESSION TREE MODELS FOR CAUSAL INFERENCE: REGULARIZATION, CONFOUNDING, AND HETEROGENEOUS EFFECTS*. Retrieved from <https://arxiv.org/pdf/1706.09523.pdf>
- Hansen, B. B. (2008). The prognostic analogue of the propensity score. *Biometrika*. <http://doi.org/10.1093/biomet/asn004>

- Kang, J. D. Y., & Schafer, J. L. (2007). Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statistical Science*, 22(4), 523–539. <http://doi.org/10.1214/07-STS227>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55.
- Roy, J., Lum, K. J., & Daniels, M. J. (2017). A Bayesian nonparametric approach to marginal structural models for point treatments and a continuous or survival outcome Bayesian nonparametric MSM. *Biostatistics*, 18(1), 32–47. <http://doi.org/10.1093/biostatistics/kxw029>
- Saarela, O., & Liu, Z. A. (2016). A flexible parametric approach for estimating continuous-time inverse probability of treatment and censoring weights. *Statistics in Medicine*, 35(23), 4238–4251. <http://doi.org/10.1002/sim.6979>
- Sivaganesan, S., Müller, P., & Huang, B. (2017). Subgroup finding via Bayesian additive regression trees. *Statistics in Medicine*, 36(15), 2391–2403. <http://doi.org/10.1002/sim.7276>
- VanderWeele, T. J., & Hernán, M. A. (2013). Causal Inference Under Multiple Versions of Treatment. *Journal of Causal Inference*, 1(1), 1–20. <http://doi.org/10.1515/jci-2012-0002>
- Xu, Y., Müller, P., Wahed, A. S., & Thall, P. F. (2016). Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times. *Journal of the American Statistical Association*, 115(October), 1–46. <http://doi.org/10.1080/01621459.2015.1086353>
- Zigler, C. M., & Dominici, F. (2014). Uncertainty in Propensity Score Estimation: Bayesian Methods for Variable Selection and Model Averaged Causal Effects. *Journal of the American Statistical Association*, 109(505), 95–107. <http://doi.org/10.1080/01621459.2013.869498>

3. Before Running PCATS

The PCATS app is an R Shiny application. It is hosted on the server at Cincinnati Children's Hospital Medical Center. The only requirement for users to use the app is to have access to an internet connection and an online browser, such as Firefox and Chrome. Users may choose to either use the app as an anonymous user, or as a registered user. The registration is free, and it

offers a way to directly communicate with the app developers as questions arise. If choosing to upload user specified data, please make sure the data is de-identified, does not contain patient or study subject identifiable information, such as social security number, names, phone number, address, zip code, etc., in order to protect your data and the personal information of your study participants. Users are encouraged to please go through some of the examples provided below as a tutorial before running your own data.

4. User's Guide

The rest of this document provides a step-by-step tutorial on how to use the app with examples. New users are recommended to start with example 1.1 first to get a general idea of the function and interface of the app. After example 1.1, one could skip to the specific example that is most similar to the features of his or her own research. These examples are designed to address some specific issues frequently encountered in comparative effectiveness research. In particular, example 1.2, 1.3 and 2.2 present our comparative effectiveness study with data obtained from routine clinical encounters. These results are part of our second aim from the PCATS project.

Each individual research project has its own unique features which require special attention to ensure unbiased conclusions. For example, in our JIA CER study, the outcome measure is a bounded summary score that requires special estimation methods for evaluating a treatment effect. This feature is offered in the app. However, there are many other special features one must carefully consider both in the design and analyses of the study. It is not possible for us to include all possible features in our app, and we strongly encourage users to consult with us while using the app. The PCATS project is funded by the Patient Centered Outcome Research Institute (PCORI ME-1408-19894). All publications, presentations, and any scholarship resulting from the use of the app must acknowledge and give appropriate credit to the funding support from PCORI and the scientific contributions made by authors.

4.1. Getting Started

Example 1.1: A Simple Binary Treatment Setting

This example considers a relatively simple setting using simulated data with a sample size of 400. The data are simulated by considering a single covariate $x \sim N(0,1)$. The potential outcome and the observed outcomes are generated by $y^{(0)} = e^x$, $y^{(a)} = y^{(0)} + a(1 + u)$ for binary treatment assignment, i.e. $a = 0, 1$. Thus, in this example, on the average, the treatment effect is 1, but the true treatment effect for each individual could be different, where $y_i^{(0)} - y_i^{(1)} = 1 + u$, with u standing for the unaccounted individual differences in their treatment effect. The treatment is assigned to a given unit following $A \sim \text{Bernoulli}(p(x))$, where $\text{logit } p(x) = -0.2 + (1.8x)^{\frac{1}{3}}$. At last, the observed outcome is generated by $y|x, a \sim y^{(a)} + \varepsilon$, $\varepsilon \sim N(0, 0.75)$. In other words, the observed outcome is a noisy version of the corresponding potential outcome. The dataset used for generating the following output can be downloaded by clicking [here](#).

Step 0. Import the Data

Once the data is downloaded, users should save the data in a CSV or Excel (both .xls and .xlsx) file and upload the data by clicking on the "Import File" button on the left panel. A "File Import" window pops up, and then users can choose a data file to be imported. After importing the data, a "Treatment select" window pops up ([Figure 1.1.1](#)). Users can choose the "Non-adaptive" option for this example ([Figure 1.1.2](#)), and proceed to the next page under the "Model" Tab. Once the

data finishes uploading, users may review the data by clicking the “Data” tab. [Figure 1.1.3](#) shows the list of variables in the data and the type of these variables which are categorical or numerical. Users can click on one of the variables here to show the statistical summary table in [Figure 1.1.4](#), and change the type of the variable by clicking the “Switching to Numeric” or “Switching to Categorical” button.

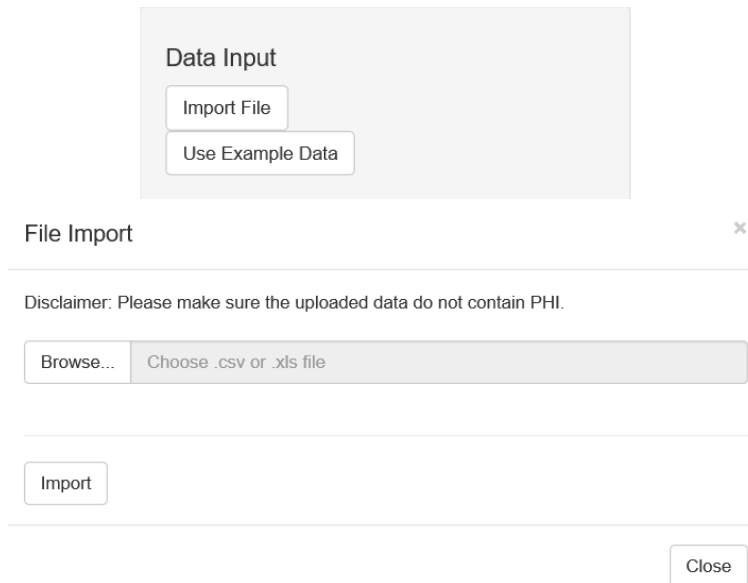


Figure 1.1.1 Import a data

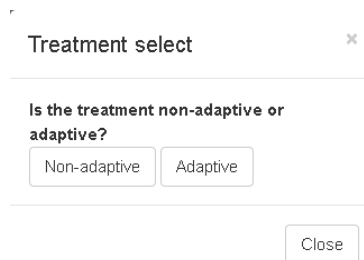


Figure 1.1.2 Treatment select

Variable list

subject categorical

C.A categorical

X numeric

Y numeric

Figure 1.1.3 Variable list

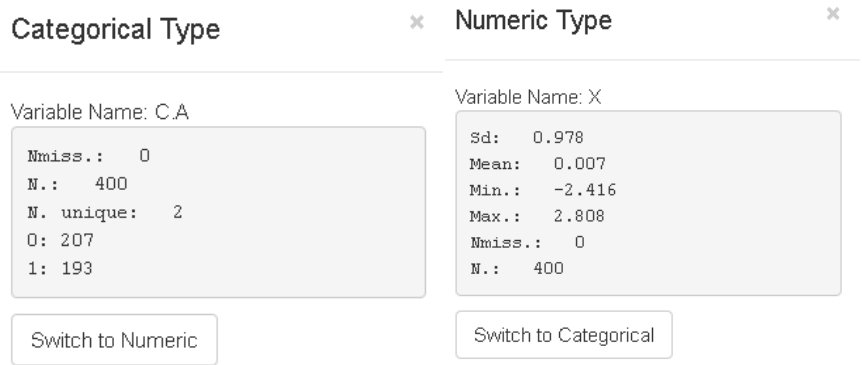


Figure 1.1.4 Statistical summary table of variables in the data

Step 1. Choose the outcome and treatment variables (Figure 1.1.5)

After choosing the type of treatment users are brought to the “Model” page. On this page they are prompted to go through steps for model building. First, users are asked to choose the response and treatment variables. In this example, “Y” is the outcome variable and “C.A” is the binary treatment.

Figure 1.1.5. Selected variables

Step 2. Select covariates in the model (Figure 1.1.6)

Two types of covariates are being specified. The first is the confounder variable. The confounders are those variables that relate to both treatment selection and the outcome model. Unless the data is obtained from a randomized experiment, treatment are often being selected for deliberate reasons, which are also related to the expected outcomes. Therefore, it is important to control for the confounders. Without adequately controlling for confounders, the difference between the treated and untreated groups on their outcomes is not an unbiased estimate of the treatment. In the clinical comparative study, one may wrongly conclude that there is no treatment effect when in fact there is, or wrongly conclude that is treatment effect that when the true effect is null. In GPMatch, the confounders are used for GP prior to determine the degree of match so that it can line up treated and un-treated individuals on their values of all confounders, subsequently removing the confounding effect.

The second type of covariate is the prediction variable. The prediction variables are those underlying the scientific mechanisms that lead to the potential outcomes. These variables are used to model the mean function of the outcome, much like linear regression models do.

In this simulated example, since “X” was the variable used for determining both potential outcome and treatment assignment “X” is selected for both types.

Step 2: Covariates

Confounding Variable(s): ?

X

Explanatory Variable(s): ?

X

Heterogeneous Treatment Effect ?

Figure 1.1.6. Selected covariates in the model

Step 3. Run and show results

The Markov chain Monte Carlo (MCMC) is a computing engine behind Bayesian statistical computing. Users should input the number of MCMC ‘burn-in’ samples, i.e. number of MCMC to be discarded and the number of MCMC samples after ‘burn-in’ (Figure 1.1.7). The larger number the better MCMC convergence, but this comes with the cost of a longer running time. The default values are set at 1000, but users should choose the number carefully for their given data. At last, users may hit the “Go” button to run the model. Once finished, the “Result” page will appear. The study results are shown on the page.

Step 3: Run

Number of MCMC samples to be discarded as ‘burn-in’ ?

5000

Number of MCMC samples after ‘burn-in’ ?

5000

Go!

Figure 1.1.7. Input the number of ‘burn-in’ and MCMC samples

Users can click the “MCMC trace plot” button to show the MCMC trace plot of the potential outcomes (Figure 1.1.8). This can be used to check the convergence of the MCMC. The plot can be downloaded and saved as a PNG file.

First, Table 1.1.1 shows the summary descriptive analysis by treatment groups and reports the number of observations per treatment group, count and percentage for categorical variables, and mean and standard deviation for continuous variables. Column “SMD” is the standardized mean difference. The definitions of SMD are available in Flury *et al.* (1986). This table allows users to inspect the treatment assignment mechanism and compare how covariates may differ between treatment groups.

The next tables (Table 1.1.2) shows the estimated ATE with SD and the 95% confidence interval. The estimate of ATE is -1.02 with the corresponding 95% equal tail credible interval (CI) of (-1.23, -0.81). Recalling that the data was simulated by setting the averaged treatment effect to -1 for the sample, it correctly estimated the ATE.

Further down, Table 1.1.3 presents estimated potential outcomes by treatment groups. It reports that the expected mean and its standard error for potential outcomes are $Y(0) = 1.64$ and $Y(1) =$

2.66. More detailed results are presented: MCMC posterior estimate of ATE (Figure 1.1.9) and of the estimates of potential outcome (Figure 1.1.10). The figures are interactive. By hovering the pointers, users are provided with the corresponding posterior estimates of the treatment effectiveness and potential outcomes. At last, showing in the side-by-side bar chart, are the estimated posterior probability of the potential outcomes being less than or equal to the chosen value under each of the treatment assignments (Figure 1.1.11). Users may set different values using the sliding bar.

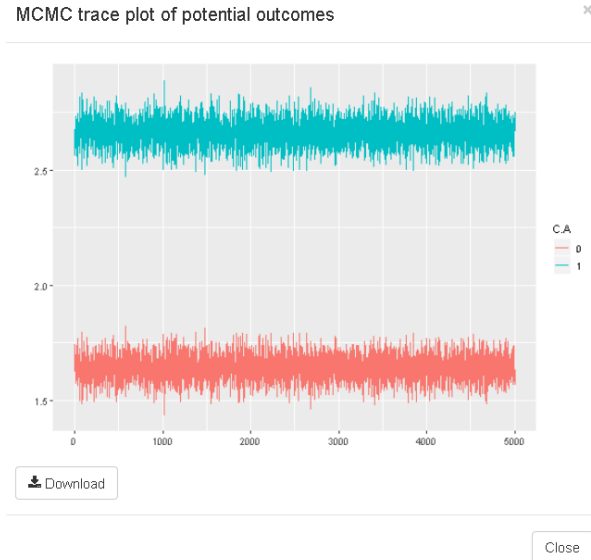


Figure 1.1.8. MCMC trace plot of potential outcomes

Descriptive statistics by treatment groups

	0	1	SMD
n	207.00	193.00	
X (mean (sd))	-0.38 (0.89)	0.42 (0.90)	0.885
Y (mean (sd))	1.09 (1.42)	3.26 (2.50)	1.067

Table 1.1.1. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	-1.0223	0.1077	-1.2334	-0.8111

Table 1.1.2. Estimated averaged causal treatment effect

Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	1.6403	0.0520	1.5384	1.7422
1	2.6626	0.0558	2.5532	2.7719

Table 1.1.3. Estimated potential outcomes

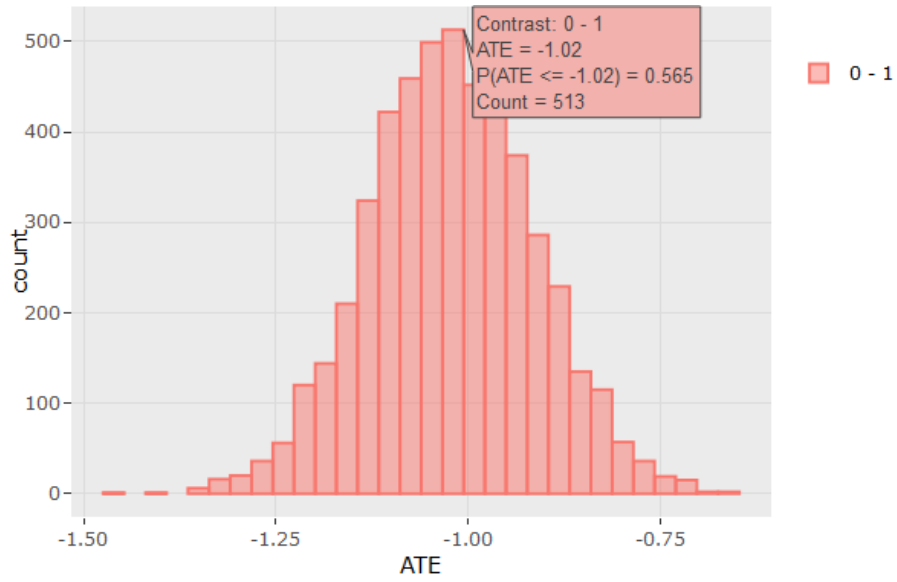


Figure 1.1.9. Histogram of averaged treatment effect estimates

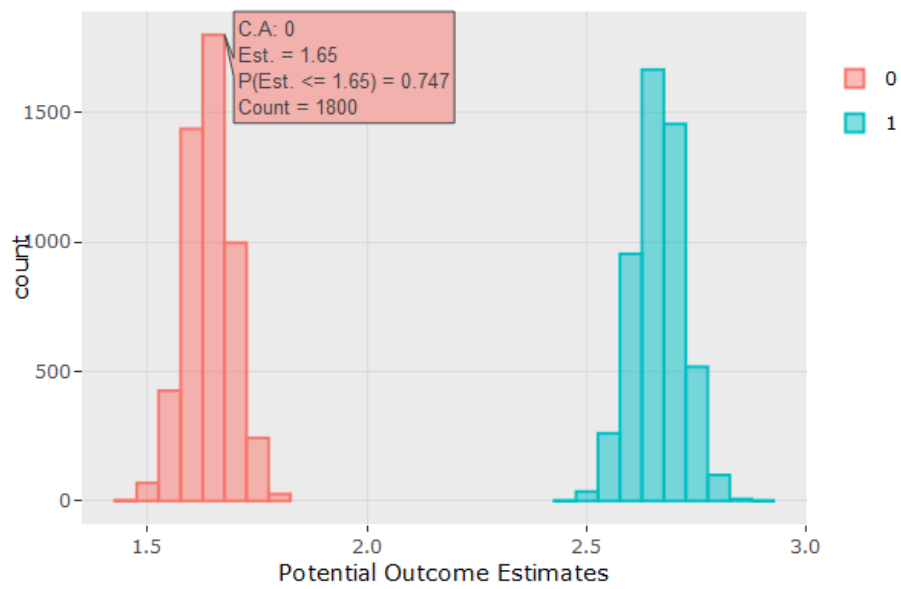


Figure 1.1.10. Histogram of potential outcome estimates

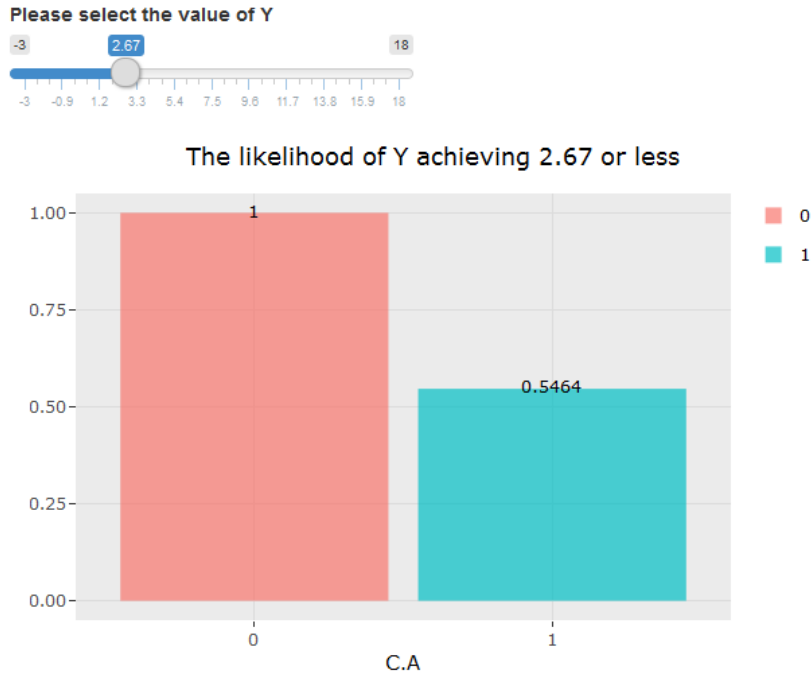


Figure 1.1.11. Bar chart of likelihood of $Y \leq 2.67$ by C.A

References

Flury, BK. and Riedwyl, H. (1986). Standard distance in univariate and multivariate analysis. *The American Statistician*, 40: 249-251

4.2. Non-Adaptive Types of Treatments

Example 1.2: Binary Treatment with a Bounded Continuous Outcome

Juvenile idiopathic arthritis (JIA), a chronic inflammatory disease, is one of the most common childhood autoimmune diseases and a major cause of childhood disability. If managed well, children with JIA can be symptom free. However, if not managed well, patients may suffer from pain, deformation of the joints, and the inability to perform daily tasks as simple as walking upstairs or lifting things. The advent of disease modifying anti rheumatoid drugs (DMARDs), particularly the biologic DMARDs (bDMARDs), has made it possible for inducing an inactive disease status in children with JIA. However, it is not clear whether the early introduction of treatment using biologic DMARDs is more effective than the more conventional nonbiological DMARDs (nbDMARDs) therapy. There was limited evidence from randomized clinical trials suggesting that the early combination of biologic and nonbiologic DMARDs (n+nbDMARDs) could lead to better clinical outcomes than the nbDMARDs alone. The real world evidence of effectiveness however is lacking.

For the purpose of evaluating the clinical effectiveness of early aggressive b+nbDMARDs vs. the more conventional nbDMARDs approach in treating children with newly diagnosed poly articular course of JIA (pcJIA), this study utilized data collected from a completed prospectively followed up inception cohort research study (Seid et al 2014). The patients were cared for at a large pediatric rheumatology clinic center and were eligible to be enrolled into the study soon after being diagnosed with JIA. The patients were followed up at 6 and 12 months after baseline. All

study visits coincided with their regular clinical visits. Demographic and health related quality of life were collected via personal interview, and treatment and clinical variables were collected from reviewing patients' health records. Only a subset of pcJIA patients receiving a DMARDs medication prescription were included in this CER study.

A snap shot of the data structure is shown in the [Figure 1.2.1](#). [Table 1.2.1](#) provides the description of the data fields.

jadas6	C.treatment_group	jadas0	chaq_score	timediag	C.RF_pos	active_joints	active_joints_6
0	0	2	0	104	0	0	0
2	1	2	0.125	141	1	0	0
10.1	0	16.1	0	20	0	12	6
13.4	1	6	0.375	187	0	0	0
9.3	0	4	0.625	49	0	2	1
1	0	8	2.375	258	0	2	1

Figure 1.2.1. Data structure

<i>Variable</i>	<i>Name</i>	<i>Description</i>
<i>Outcome</i>	Jadas6	Juvenile Arthritis Disease Activity Score (cJADAS) at 6 months of follow up. Value ranges from 0-40, with higher score indicating more disease activities.
<i>Treatment Baseline Covariate</i>	C.treatment_group	Treatment group: 0 = nbDMARDs; 1 = b+nbDMARDs.
	jadas0	Juvenile Arthritis Disease Activity score at baseline. Value ranges from 0-40, with higher score indicating more disease activities.
	chaq_score	Functional ability measured by the Childhood Health Assessment Questionnaire, range from 0 to 3, with higher value indicting higher functional disability.
	timediag	Time since date of diagnosis of JIA (days) at the baseline. Baseline is the date when patient enrolled into the study.
	C.RF_pos	Positive Rheumatoid Factor (RF): 0 = no Vs. 1 = yes.
	C.Female	Gender: 0 = Male; 1 = Female
	C.private	Private Insurance: 0 = No; 1 = Yes
	age	Age at the baseline

Table 1.2.1. Data description

The baseline measures include demographic characteristics and disease specific clinical characteristics. The primary outcome is the Juvenile Arthritis Disease Activity Score (JADAS10) at the 6 months follow up visit. JADAS10 is a widely used outcome measure for children with JIA (Consolaro et al, 2009). This summary disease activity score is derived from four clinical outcomes: physician's global evaluation of patient's disease activity (0-10), patient's rating of his or her wellbeing (0-10), the number of active joint counts (truncated at 10) and the erythrocyte sedimentation rate or SED rate which is a marker for inflammation activity. As the result, JADAS is a bounded summary score that ranges between 0 and 40 with lower values indicating less disease activity.

Step 1. Choose outcome and treatment variables (Figure 1.2.2)

The outcome variable and its type should be chosen. Since JADAS is bounded between 0 and 40, the lower and upper thresholds are specified here. C.treatment_group is the treatment variable.

Step 1: Outcome and treatment variables

Outcome variable: [?](#)

jadas6

Left- or right- bounded [?](#)

0 : 40

Treatment variable(s): [?](#)

C.treatment_group

Figure 1.2.2. Selected Outcome and Treatment Variables

Step 2. Building the Model (Figure 1.2.3)

The patients' age, time since diagnosis, baseline JADAS score, and CHAQ score are identified confounders. These variables adding a positive Rheumatoid Factor are expected to explain the variation in JADAS outcome at 6 months thus are included into the explanatory variables.

Step 2: Covariates

Confounding Variable(s): [?](#)

age time diag jadas0 chaq_score

Explanatory Variable(s): [?](#)

age time diag jadas0 chaq_score C.RF_pos

Heterogeneous Treatment Effect [?](#)

Figure 1.2.3. Selected covariates in the model

Step 3. Run and show results

5000 'burn-in' and 5000 MCMC samples after 'burn-in' are considered here. The results of descriptive analysis by treatment groups are shown in Table 1.2.2. Table 1.2.3 shows the summary table of the estimated ATE evaluated for the expected JADAS score reduction for the early combination treatment vs. the nbDMARDs. The estimated ATE is 3.17, with the 95% interval estimate of (-0.15, 6.48) suggesting that, compared to nbDMARDs, b+nbDMARDs are expected to improve disease activities by a 3-point reduction in JADAS. The summary table of the estimated potential outcomes by treatment group is shown in Table 1.2.4. Figures 1.2.4 and 1.2.5 show the histograms of the estimated ATE and of the potential outcome estimates. Figure 1.2.6 shows that if treated on nbDMARDs, the patient has a 37% likelihood of achieving JADAS \leq 8.12; while if treated on an early combination of b+nbDMARDs, the patient has a 99% likelihood of achieving JADAS \leq 8.12. The cut-point of 8.12 is chosen as it corresponds to a twofold reduction of the sample mean JADAS from the baseline.

Descriptive statistics by treatment groups

	0	1	SMD
n	52.00	23.00	
age (mean (sd))	9.63 (4.32)	11.24 (3.70)	0.399
C.RF_pos = 1 (%)	3.0 (5.8)	7.0 (30.4)	0.676
chaq_score (mean (sd))	0.54 (0.76)	0.84 (0.79)	0.387
jadas0 (mean (sd))	11.00 (8.80)	17.54 (10.64)	0.670
timediag (mean (sd))	77.17 (55.05)	81.57 (53.95)	0.081
jadas6 (mean (sd))	7.38 (7.15)	7.12 (6.87)	0.038

Table 1.2.2. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	3.1681	1.6907	-0.1456	6.4818

Table 1.2.3. Estimated averaged causal treatment effect

Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	8.2725	0.5485	7.1974	9.3476
1	5.1044	1.1852	2.7813	7.4275

Table 1.2.4. Estimated potential outcomes

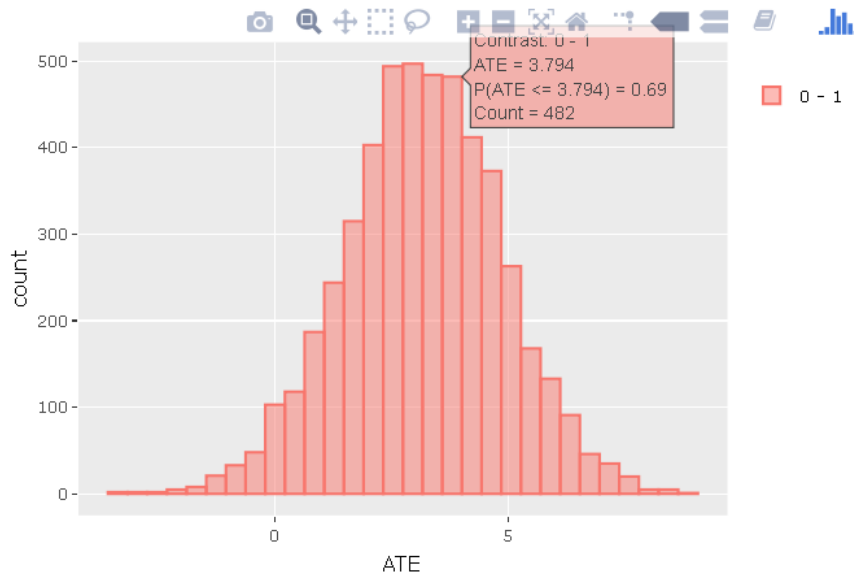


Figure 1.2.4. Histogram of averaged treatment effect estimates

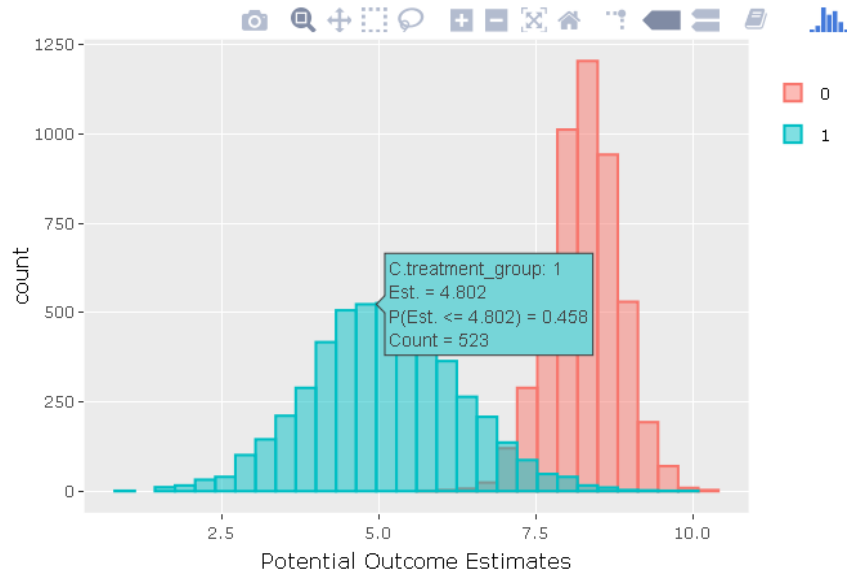


Figure 1.2.5. Histogram of potential outcome estimates



The likelihood of jadas6 achieving 8.12 or less

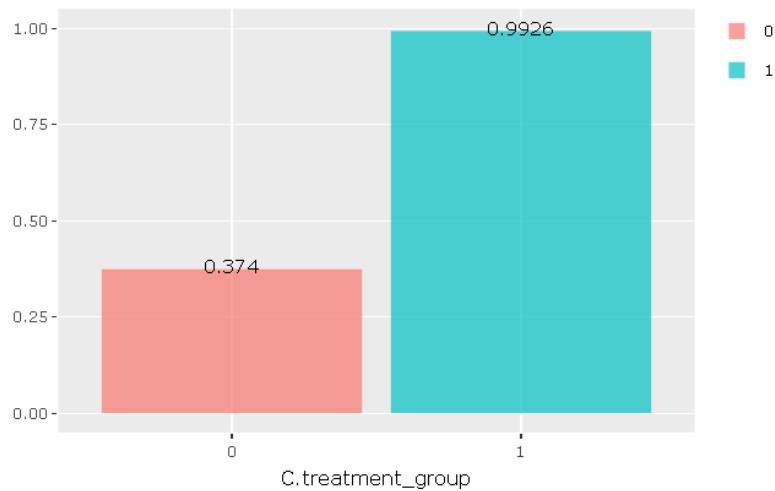


Figure 1.2.6. Bar chart of likelihood of jadas6 <= 8.12 by treatment group

Example 1.3: Multi-level Treatment with a Post-treatment Confounder

Here, we compare the effectiveness of three DMARD approaches by investigating whether an early combination of b+nbDMARD is more effective in reducing disease activity 6 months later compared to mono therapy with nbDMARDs or bDMARDs in treating children with newly diagnosed polyarticular course of Juvenile Idiopathic Arthritis (pcJIA). The dataset was derived from a single institute electronic health records Epic database dating from January 2009 to December 2017. Children 1-19 years of age, newly diagnosed with pcJIA, and receiving their 1st prescription of DMARDs soon after diagnosis were eligible. [Table 1.3.1](#) provides the description of the variables in the data.

Variable	Name	Description
<i>Outcome</i>	cJadas6	Clinical Juvenile Arthritis Disease Activity Score (cJADAS) after 6 months of treatment. Value ranges from 0-30, with higher scores indicating more disease activities.
<i>Treatment</i>	C.treatment_group	Treatment group: 0 = nbDMARDs; 1 = b+nbDMARDs; 2 = bDMARDs
<i>Baseline Covariate</i>	cjadas0	Clinical Juvenile Arthritis Disease Activity score at baseline, i.e. the initiation of the first DMARD. Value ranges from 0-30, with higher scores indicating more disease activities.
	timediag	Time since date of diagnosis of JIA (days) at the baseline. Baseline is the date that DMARDs were first prescribed for each patient.
	C.JIA_subtype	JIA Subtype: 1 = Poly RF-; 2 = Poly RF+; 3 = Oligo; 4 = Other
	C.Female	Gender: 0 = Male; 1 = Female
	C.private	Private Insurance: 0 = No; 1 = Yes
<i>Post treatment Covariate</i>	age	Age at the baseline, i.e. the 1 st DMARDs prescription.
	diffvisits	Days between baseline (i.e. the 1 st DMARDs) and the six month follow-up visit.

Table 1.3.1. Data description

Step 1. Choose variables (Figure 1.3.1)

Clinical Juvenile Arthritis Disease Activity Score (cJADAS) is a brief version JADAS, which is a more suitable outcome measure that is directly available in clinic for children with JIA. This summary disease activity score is derived from physicians' global evaluation of each patient's disease activity (0-10), rating of his or her wellbeing (0-10), and the number of active joint counts (truncated at 10). Thus, this bounded summary score ranges between 0 and 30, with lower values indicating less disease activity.

For the purpose of evaluating our research question, C.treatment_group is coded into three categories: 0 indicating the patient received a prescription of nbDMARDs, 1 indicating prescription of b+nbDMARDs, and 2 indicating prescription of bDMARDs.

Step 2. Select covariates in the model (Figure 1.3.2)

The model includes the baseline variables of age, gender, disease subtype, insurance status, cJADAS0, and the timediag variables as confounder variables. These variables along with the treatment duration were entered as explanatory variables.

Step 3. Run and show results

After 2000 'burn-in' which allows the MCMC to achieve convergence the study results are derived from the additional 2000 MCMC.

Table 1.3.2 presents the descriptive statistics of the baseline patient characteristics by their treatment assignment. As expected, patients assigned to b+nbDMARDs have a more severe presentation of disease activity at baseline than the other two groups.

Step 1: Outcome and treatment variables

Outcome variable: [i](#)

cjadas6

Left- or right- bounded [i](#)

0 : 30

Treatment variable(s): [i](#)

C.treatment_group

Figure 1.3.1. Selected variables

Step 2: Covariates

Confounding Variable(s): [i](#)

age C.Female C.JIA_subtype C.private cjadas0 timeddiag

Explanatory Variable(s): [i](#)

age C.Female C.JIA_subtype C.private cjadas0 timeddiag diffvisits

Heterogeneous Treatment Effect [i](#)

Figure 1.3.2. Selected covariates in the model

Descriptive statistics by treatment groups

	0	1	2	SMD
n	147.00	68.00	31.00	
age (mean (sd))	9.85 (5.09)	9.97 (4.72)	13.20 (4.50)	0.474
C.Female = 1 (%)	101.0 (68.7)	48.0 (70.6)	14.0 (45.2)	0.354
C.JIA_subtype (%)				0.409
1	48.0 (32.7)	26.0 (38.2)	8.0 (25.8)	
2	11.0 (7.5)	8.0 (11.8)	2.0 (6.5)	
3	46.0 (31.3)	19.0 (27.9)	6.0 (19.4)	
4	42.0 (28.6)	15.0 (22.1)	15.0 (48.4)	
C.private = 1 (%)	103.0 (70.1)	46.0 (67.6)	25.0 (80.6)	0.200
cjadas0 (mean (sd))	12.68 (5.81)	15.69 (7.09)	12.45 (5.76)	0.335
diffvisits (mean (sd))	0.41 (0.14)	0.43 (0.12)	0.40 (0.13)	0.124
timeddiag (mean (sd))	0.10 (0.16)	0.10 (0.17)	0.11 (0.17)	0.051
cjadas6 (mean (sd))	7.46 (6.02)	6.65 (5.81)	6.28 (5.16)	0.138

Table 1.3.2. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	1.7154	0.7744	0.1975	3.2333
0 - 2	0.8629	1.0534	-1.2016	2.9275
1 - 2	-0.8524	1.1638	-3.1335	1.4287

Table 1.3.3. Estimated averaged causal treatment effect

Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	7.6584	0.3466	6.9791	8.3376
1	5.9430	0.5900	4.7866	7.0994
2	6.7954	0.9286	4.9753	8.6156

Table 1.3.4. Estimated potential outcomes

Table 1.3.4 provides the estimated expected potential outcomes. Table 1.3.3 provides the estimated average treatment effects and their 95% confidence intervals comparing among the three treatment groups. The results suggest that if treated on b+nbDMARDs, patients are expected to achieve lower disease activity at six months on average than the other two treatment groups, with estimated cJADAS score of 5.9 and 95% CI of (4.8, 7.1), compared to 7.7, 95% CI of (6.8, 8.3) on nbDMARDs and 6.8, 95% CI of (5.0, 8.6) on the bDMARDs. The results suggest that, on average, the nb+bDMARD treatment is expected to lead to a 1.7-point (95%CI of 0.20, 3.23) reduction in cJADAS after six months compared to the nbDMARD treatment.

Users can choose to show the histogram of the estimated ATEs (Figure 1.3.3) and the estimated potential outcomes (Figure 1.3.5). Users can hide any traces in the plot by clicking on the legend on the right side (Figure 1.3.4). In each of the histograms, the user could obtain a more detailed report of the results by hovering the mouse over the specific histogram. For example, if a user is interested in finding if the probability of 0-1 is greater than 0, then (s)he could click on the pink histogram in Figure 1.2.3, then move the mouse to the bar at ATE=0. The user could also save the entire figure into a PNG file, to be used for dissemination of the results.

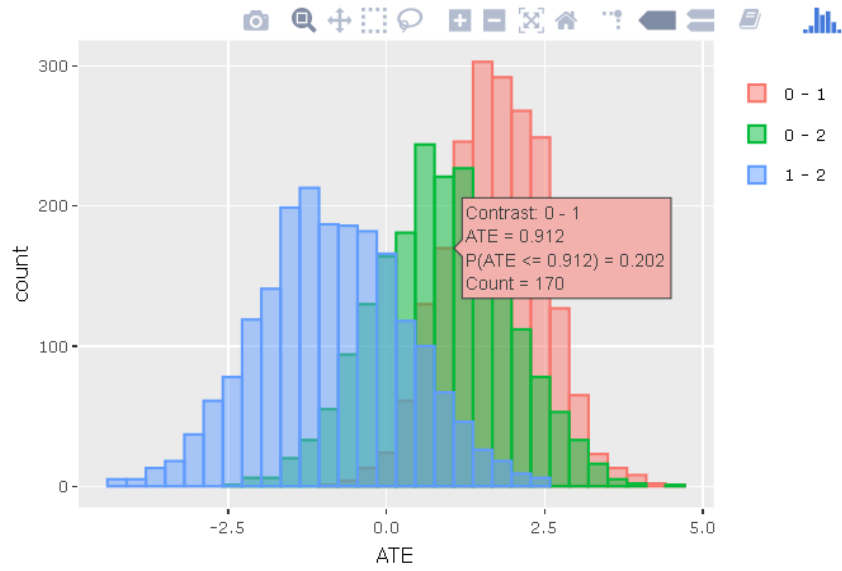


Figure 1.3.3. Histogram of averaged treatment effect estimates

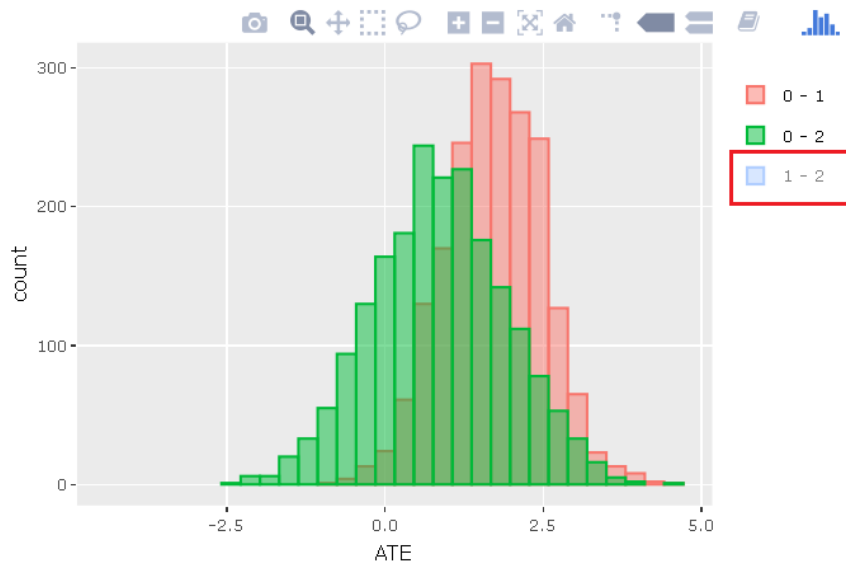


Figure 1.3.4. Histogram of averaged treatment effect estimates without 1-2 group

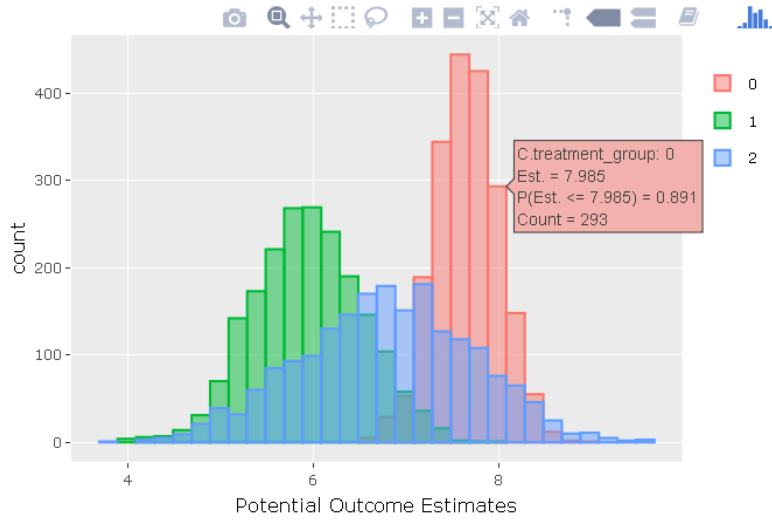


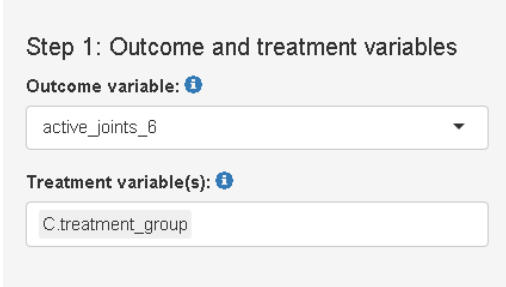
Figure 1.3.5. Histogram of potential outcome estimates

Example 1.4: Binary Treatment with a Counting Outcome

This example uses the same data from [Example 1.2](#) and chooses the “Active Joints Count” as an example with the counting outcome. To treat it as the counting outcome, please set its type as categorical in the variable list panel.

Step 1. Choose variables ([Figure 1.4.1](#))

Since the outcome is not continuous Bayesian additive regression tree (BART; Hill 2011) instead of GPMatch is used for the counting outcome. In the example, the heterogeneous treatment effect is not considered.



Step 1: Outcome and treatment variables

Outcome variable: ⓘ

active_joints_6

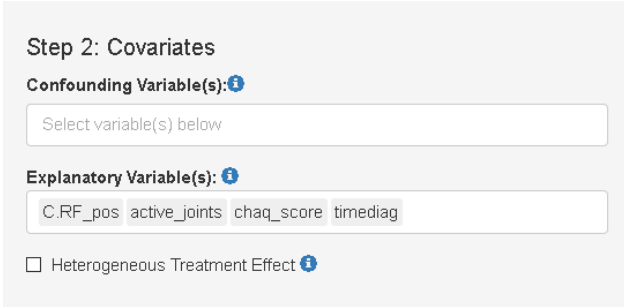
Treatment variable(s): ⓘ

C.treatment_group

Figure 1.4.1. Selected variables

Step 2. Select covariates in the model ([Figure 1.4.2](#))

Since BART does not distinguish the confounder from the explanatory variable, users were asked to select the covariates in the outcome model. Similar covariates are chosen here as [Example 1.2](#). Since the BART approach does not distinguish the confounding and explanatory variables in modeling the outcome you may leave the confounding variable selection empty.



Step 2: Covariates

Confounding Variable(s): ⓘ

Select variable(s) below

Explanatory Variable(s): ⓘ

C.RF_pos active_joints chaq_score timediag

Heterogeneous Treatment Effect ⓘ

Figure 1.4.2. Selected covariates in the model

Step 3. Run and show results

5000 ‘burn-in’ and 5000 MCMC samples after ‘burn-in’ are considered here. [Table 1.4.1](#) shows the summary table of descriptive analysis by treatment group. The estimated ATE is 1.14 ([Table 1.4.2](#)) and its 95% confidence interval is (-0.96, 3.25). The results suggest that b+nbDMARDs is more effective than nbDMARDs. If treated on the early combination, patients are expected to reduce one more active joint by six months, as opposed to being treated on the nbDMARDs. . [Table 1.4.3](#) shows the estimated potential outcomes by treatment group. Users can also choose to show the histogram of ATE and the potential outcomes.

Descriptive statistics by treatment groups

	0	1	SMD
n	52.00	23.00	
active_joints (mean (sd))	5.19 (6.07)	9.13 (8.57)	0.530
C.RF_pos = 1 (%)	3.0 (5.8)	7.0 (30.4)	0.676
chaq_score (mean (sd))	0.54 (0.76)	0.84 (0.79)	0.387
timediag (mean (sd))	77.17 (55.05)	81.57 (53.95)	0.081
active_joints_6 (mean (sd))	2.62 (4.80)	2.52 (4.73)	0.020

Table 1.4.1. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	1.1435	1.0744	-0.9622	3.2492

Table 1.4.2. Estimated averaged causal treatment effect

Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	3.0310	0.5493	1.9544	4.1077
1	1.8875	0.8636	0.1949	3.5802

Table 1.4.3. Estimated potential outcomes

References

- Consolaro, A., Ruperto, N., Bazso, A., Pistorio, A., Magni-Manzoni, S., Filocamo, G., ... Ravelli, A. (2009). Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis & Rheumatism*, *61*(5), 658–666. <http://doi.org/10.1002/art.24516>
- Hill JL (2011). Bayesian Nonparametric Modeling for Causal Inference. *J Comput Graph Stat*, *20*: 217–240.
- Seid, M., Huang, B., Niehaus, S., Brunner, H. I., & Lovell, D. J. (2014). Determinants of Health-Related Quality of Life in Children Newly Diagnosed With Juvenile Idiopathic Arthritis. *Arthritis Care & Research*, *66*(2), 263–269. <http://doi.org/10.1002/acr.22117>

Example 1.5: Subgroup Treatment Effect for a Binary Treatment

Heterogeneous treatment effects (HTE) refer to the fact that treatment effects may vary across sub-populations or change as a function of a continuous factor. This example shows how to use the app to estimate the heterogeneous effects. The example data in [Example 1.2](#) is considered here. Potential heterogeneous treatment effects by Rheumatoid Factor (RF) and Juvenile Arthritis Disease Activity scores (JADAS) at baseline are evaluated.

Step 1. Choose variables

The similar selections are chosen as [Example 1.2](#).

Step 2. Select covariates in the model ([Figure 1.5.1](#))

The covariates are selected the same way as in the [Example 1.2](#). By checking on the “Heterogeneous Treatment Effect” box, an additional variable selection panel for HTE is shown. Users can select the variables which may have heterogeneous effects. In this example, Rheumatoid Factor (RF) and JADAS at baseline are considered.

Step 2: Covariates

Confounding Variable(s):

age timediag jadas0 chaq_score

Explanatory Variable(s):

age timediag jadas0 chaq_score C.RF_pos

Heterogeneous Treatment Effect

Variable(s) with HTE:

C.RF_pos jadas0

Figure 1.5.1. Selected covariates

Step 3. Run and show results

5000 ‘burn-in’ and 5000 MCMC samples are input here. After clicking on the “Go” button, the results are shown in the “Result” tab.

Similar tables and figures such as [Example 1.2](#) are provided. Users can choose “Yes” for the question of “Do you want to show HTE?” to show the results of heterogeneous effects. Users can select the values of the treatment variable as the control or treatment group ([Figure 1.5.2](#)). Then users should choose one of the variables which may have the heterogeneous effect. The available variables of HTE here are from the input of [Figure 1.5.1](#). After selection, users can hit the “Go” button to calculate the heterogeneous effects. The corresponding results are shown in an interactive graph. Except for the selected variable with HTE, the values of other covariates in the model are the observed values for the calculation of HTE. Please note that the calculation may need some time. For a categorical variable, violin plots overlaid with box plots by the categorical variable are shown in [Figure 1.5.3](#). Violin plots show the kernel probability density of the estimated HTEs. Box plots show the median of the estimated HTEs and a box indicates the interquartile range. The corresponding values are shown in the hover text. The plot shows that the effectiveness of b+nbDMARDs compared to nbDMARDs is bigger on patients with positive Rheumatoid Factor (RF) than patients with non-positive RF after 6 months of treatment. For a continuous variable, the estimated HTE and its 95% confidence interval by the values of the continuous variable are shown in [Figure 1.5.4](#). The corresponding values are also shown in the hover text. The plot shows that the effectiveness of b+nbDMARDs compared to nbDMARDs is

bigger on the patients, but the effect is not related to the JADAS values at baseline. Users can also specify the values of variables which may have the heterogeneous effects (Figure 1.5.5). The histogram of estimated potential outcomes is shown in Figure 1.5.6. The results suggest RF+ patients are more likely to benefit more from the nb+bDMARD treatment approach than the nbDMARD only.

HTE on <C.treatment_group>:

Choose < C.treatment_group > value as control group to calculate HTE:

0
 1

Choose < C.treatment_group > value as treatment group to calculate HTE:

0
 1

Choose a variable with HTE

C.RF_pos
 jadas0

Go!

Figure 1.5.2. Selection in HTE panel

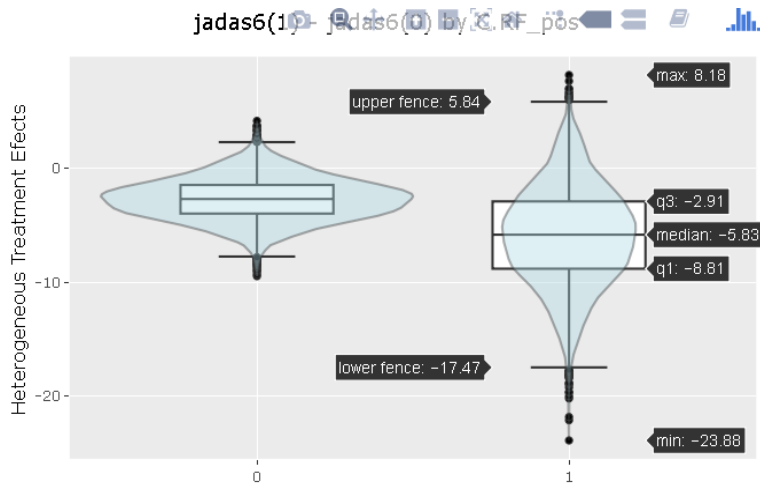


Figure 1.5.3. Estimated HTEs by Rheumatoid Factor

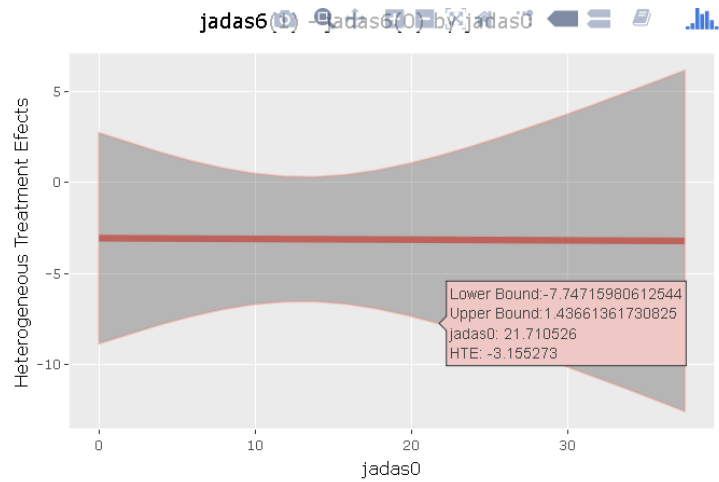


Figure 1.5.4. Estimated HTEs by JADAs at baseline

Do you want to show HTE on specified values of covariates?

Yes

No

C.RF_pos

jadas0

Figure 1.5.5. Specify the covariates

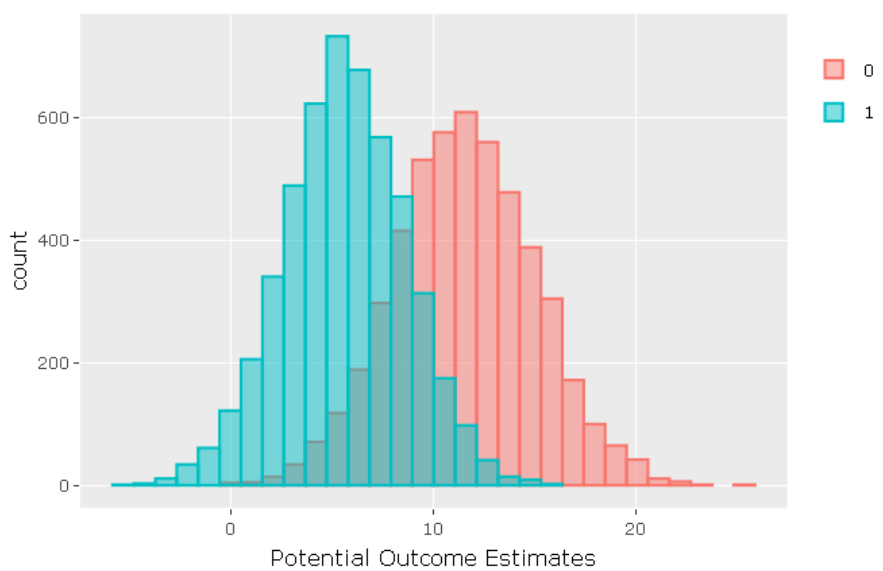


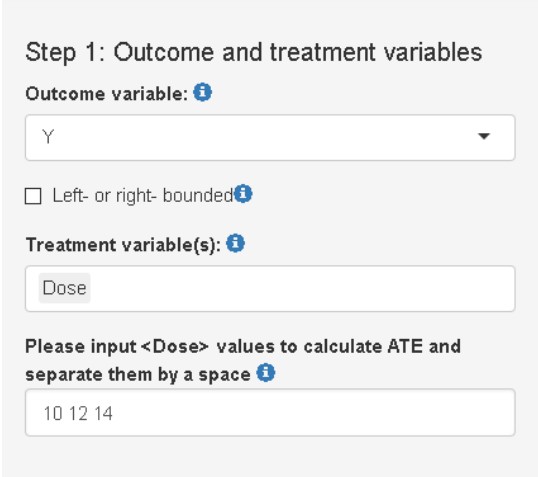
Figure 1.5.6. Potential outcome estimates on specified covariates

Example 1.6: Continuous Treatment in Dose-Response Setting

Simulated data are used here to show how to use the app to handle the data with a continuous treatment. The sample size is 200. Here, the observed data are $(X_i, Dose_i, Y_i)$ and the generative models are

$$\begin{aligned} X &\sim N(0,1), \\ Dose &\sim N(12,1), \\ Y &\sim N(2 + 2X + Dose, 1). \end{aligned}$$

Figures 1.6.1 and 1.6.2 show how to choose outcome and treatment and covariates in the model. 2000 ‘burn-in’ and 2000 MCMC samples are used to run the analyses. Since the treatment variable is continuous, users need to input the specified values of treatment variables (Figure 1.6.1) which are used to calculate the ATE. A space should be added to separate these values. In the example, 10, 12 and 14 for “Dose” are input here. Estimated average treatment effects (Table 1.6.2) and estimated potential outcomes (Table 1.6.3) are calculated based on these specified values of “Dose”. In the Table 1.6.2, the estimated ATE effects of “10-12”, “10-14” and “12-14” are close to the true values -2, -4 and -2. The 95% confidence intervals cover the true values. The estimated potential outcomes of “Dose” at “10”, “12” and “14” are also close to the true values 12, 14 and 16 in the Table 1.6.3. The corresponding histograms are shown in Figures 1.6.3 and 1.6.4. The histograms of ATE for “10-12” and “12-14” are overlapped in Figure 1.6.3. Users can click on the legend on the right side of the figure to show or hide histograms of ATE (Figure 1.6.5). Table 1.6.1 shows the results of descriptive analysis by treatment groups. Since “Dose” is a continuous variable, user-defined cut points can be used to categorize it. And the descriptive analysis is done based on the new generated treatment groups. The default cut point is the median value of the treatment variable. Users can choose “Yes” for the question of “Do you want to input cut points to categorize <Dose>?”. After “Yes” is chosen, users can input the specified values of “Dose” in Figure 1.6.6. After clicking on the “Show Table 1” button, the results of descriptive analysis by users-specified treatment groups are shown in Table 1.6.4.



Step 1: Outcome and treatment variables

Outcome variable: **i**

Y

Left- or right- bounded **i**

Treatment variable(s): **i**

Dose

Please input <Dose> values to calculate ATE and separate them by a space **i**

10 12 14

Figure 1.6.1. Selected variables

Step 2: Covariates

Confounding Variable(s):

Explanatory Variable(s):

Heterogeneous Treatment Effect

Figure 1.6.2. Selected covariates in the model

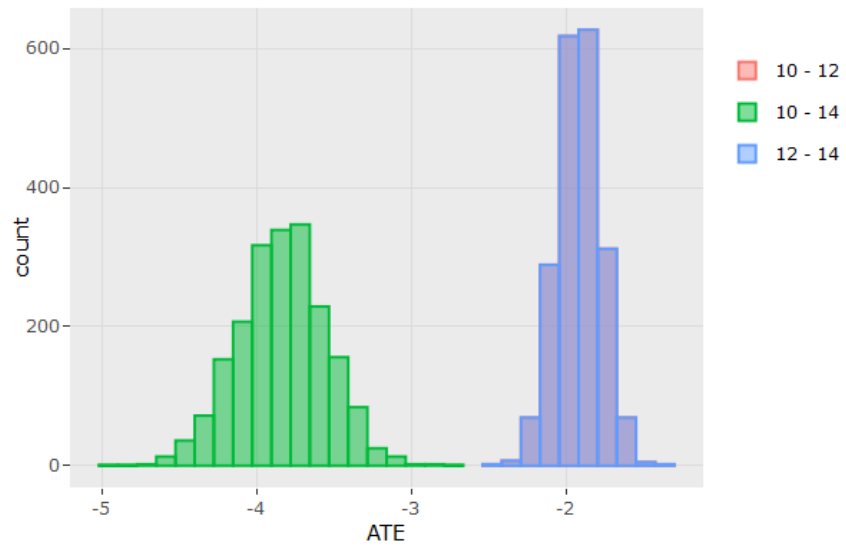


Figure 1.6.3. Histogram of averaged treatment effect estimates

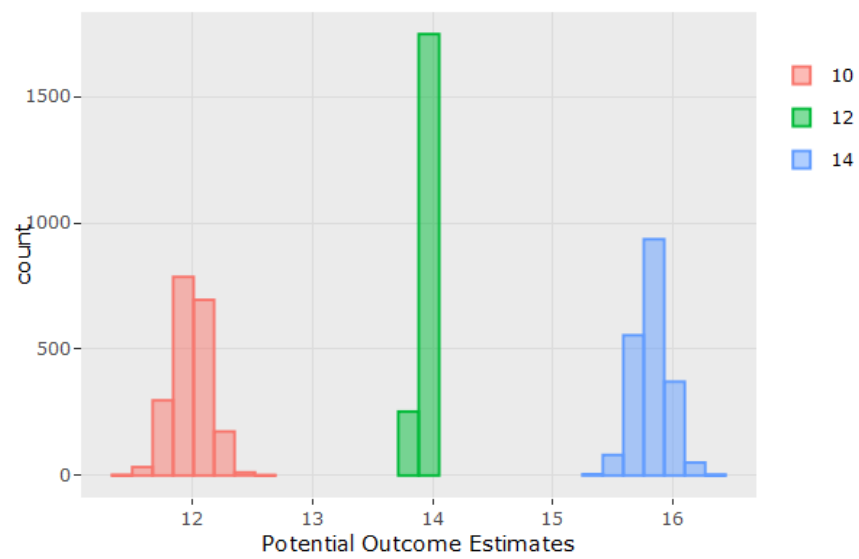


Figure 1.6.4. Histogram of potential outcome estimates

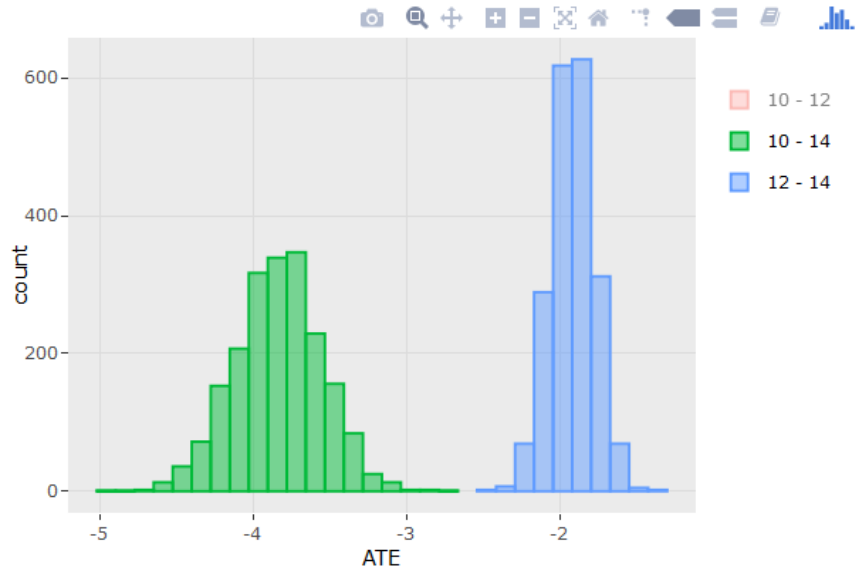


Figure 1.6.5. Histogram of ATE with hiding histogram of “10-12”

Descriptive statistics by treatment groups

	(-Inf,12.1]	(12.1, Inf]	SMD
n	100.00	100.00	
Dose (mean (sd))	11.31 (0.59)	12.81 (0.59)	2.535
X (mean (sd))	-0.02 (0.89)	-0.06 (0.95)	0.043
Y (mean (sd))	13.30 (2.09)	14.62 (2.05)	0.641

Note: Median value of <Dose> is used to categorize it.

Do you want to input cutpoints to categorize < Dose >?

Yes

No

Table 1.6.1. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
10 - 12	-1.9184	0.1422	-2.1971	-1.6398
10 - 14	-3.8368	0.2843	-4.3941	-3.2795
12 - 14	-1.9184	0.1422	-2.1971	-1.6398

Table 1.6.2. Estimated averaged causal treatment effect

Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
10.0000	11.9848	0.1470	11.6967	12.2730
12.0000	13.9033	0.0177	13.8686	13.9379
14.0000	15.8217	0.1394	15.5484	16.0949

Table 1.6.3. Estimated potential outcomes

Please separate the cutpoints by a space

11 13

Show Table 1

Figure 1.6.6. Input specified values of treatment to do a descriptive analysis

Descriptive statistics by treatment groups

	(-Inf,11]	(11,13]	(13, Inf]	SMD
n	27.00	138.00	35.00	
Dose (mean (sd))	10.51 (0.35)	12.00 (0.51)	13.49 (0.42)	4.755
X (mean (sd))	0.06 (0.99)	-0.05 (0.84)	-0.09 (1.14)	0.102
Y (mean (sd))	12.52 (2.31)	13.91 (1.88)	15.26 (2.42)	0.816

Table 1.6.4. Descriptive statistics by user specified treatment groups

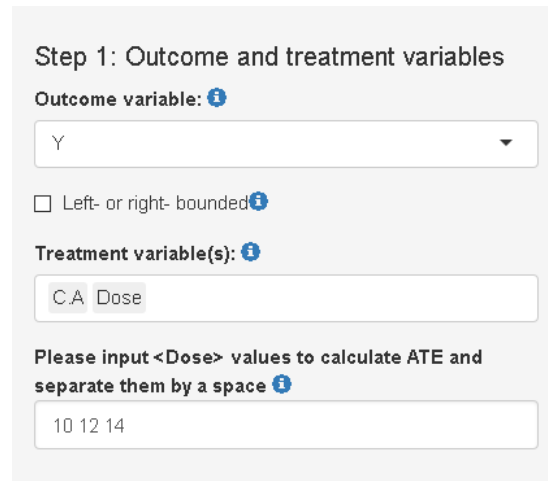
Example 1.7: Composite Binary and Continuous Treatments

Often the treatment consists of more than one intervention. For example, given the same medication treatment, different doses may be used. Thus, patients who are treated on the same medication in fact are receiving different treatments if the dose assignment is different. Here, we demonstrate the use of the PCATS App to handle such a composite treatment. Simulated data are used with a sample size of 200. Here, the observed data are $(X_i, A_i, Dose_i, Y_i)$ and the generative models are

$$\begin{aligned} X &\sim N(0,1), \\ A &\sim \text{Bernoulli}(0.5) \\ Dose &\sim N(12,1), \\ Y &\sim N(2 + 2X + 3A + Dose + A * Dose, 1). \end{aligned}$$

Step 1. Choose variables

As shown in [Figure 1.7.1](#), the outcome “Y” is continuous. Variables “C.A” and “Dose” are chosen as treatment variables. Please note that, at most, 2 treatment variables can be chosen in the app.



Step 1: Outcome and treatment variables

Outcome variable: [?](#)

Y

Left- or right- bounded [?](#)

Treatment variable(s): [?](#)

C.A Dose

Please input <Dose> values to calculate ATE and separate them by a space [?](#)

10 12 14

Figure 1.7.1. Selected variables

Step 2. Select covariates in the model

Variable “X” is chosen in the covariance function of GP term and in the mean function ([Figure 1.7.2](#)). The heterogeneous treatment effect box is checked because there is an interaction of the two treatment variables. For the HTE with “C.A”, “Dose” is chosen.

Step 3. Run and show results

2000 ‘burn-in’ and 2000 MCMC samples are considered here. Since one treatment variable, “Dose”, is continuous, users need to specify the values of “Dose” as [Example 1.6](#). 10, 12 and 14 for “Dose” are also considered here. The results of descriptive analysis are shown in [Table 1.7.1](#). [Table 1.7.2](#) shows the summary table of the estimated averaged treatment effect and [Table 1.7.3](#) shows the summary table of the estimated potential outcomes. In [Table 1.7.3](#), the estimated potential outcomes are very close to the true values 12, 14, 16, 25, 29 and 33. Users can choose to show the results of heterogeneous effects ([Figure 1.7.3](#)). After selection, the estimated HTE and its 95% confidence interval by the values of the continuous variable are shown in [Figure 1.7.4](#).

The results suggest that the treatment effect of A differs by the dosage with the higher dosage corresponding to the stronger treatment effect.

Step 2: Covariates

Confounding Variable(s):

Explanatory Variable(s):

Heterogeneous Treatment Effect

Variable(s) with HTE (C.A.):

Variable(s) with HTE (Dose):

Figure 1.7.2. Selected covariates in the model

HTE on <C.A>:

Choose < C.A > value as control group to calculate HTE:

0

1

Choose < C.A > value as treatment group to calculate HTE:

0

1

Choose a variable with HTE (C.A):

Dose

Figure 1.7.3. Selection in HTE panel

Descriptive statistics by treatment groups

	C.A = 0; Dose in (-Inf,12.1]	C.A = 0; Dose in (12.1, Inf]	C.A = 1; Dose in (-Inf,12.1]	C.A = 1; Dose in (12.1, Inf]	SMD
n	44.00	60.00	56.00	40.00	
Dose (mean (sd))	11.32 (0.52)	12.78 (0.59)	11.30 (0.64)	12.85 (0.61)	1.732
X (mean (sd))	-0.29 (0.89)	-0.24 (1.07)	0.19 (0.85)	0.20 (0.65)	0.364
Y (mean (sd))	12.99 (2.03)	14.25 (2.29)	27.84 (2.44)	31.04 (1.67)	5.424

Note: Median value of <Dose> is used to categorize it.

Table 1.7.1. Descriptive statistics by treatment groups

Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0, 10 - 0, 12	-1.5874	0.2108	-2.0005	-1.1743
0, 10 - 0, 14	-3.1748	0.4216	-4.0010	-2.3485
0, 10 - 1, 10	-12.2272	0.3305	-12.8751	-11.5794
0, 10 - 1, 12	-16.4068	0.2689	-16.9339	-15.8796
0, 10 - 1, 14	-20.5863	0.3433	-21.2592	-19.9134
0, 12 - 0, 14	-1.5874	0.2108	-2.0005	-1.1743
0, 12 - 1, 10	-10.6399	0.2430	-11.1160	-10.1637
0, 12 - 1, 12	-14.8194	0.1450	-15.1035	-14.5353
0, 12 - 1, 14	-18.9989	0.2559	-19.5004	-18.4974
0, 14 - 1, 10	-9.0525	0.3125	-9.6650	-8.4399
0, 14 - 1, 12	-13.2320	0.2420	-13.7063	-12.7577
0, 14 - 1, 14	-17.4115	0.3192	-18.0372	-16.7858
1, 10 - 1, 12	-4.1795	0.2031	-4.5775	-3.7815
1, 10 - 1, 14	-8.3591	0.4061	-9.1550	-7.5631
1, 12 - 1, 14	-4.1795	0.2031	-4.5775	-3.7815

Table 1.7.2. Estimated averaged causal treatment effect

Potential Outcome Estimates

C.A	Dose	Estimation	SD	95% Lower Bound	95% Upper Bound
0	10	12.4190	0.2387	11.9510	12.8869
0	12	14.0064	0.0731	13.8630	14.1497
0	14	15.5938	0.2063	15.1895	15.9981
1	10	24.6462	0.2111	24.2324	25.0600
1	12	28.8258	0.0771	28.6747	28.9768
1	14	33.0053	0.2231	32.5680	33.4425

Table 1.7.3. Estimated potential outcomes

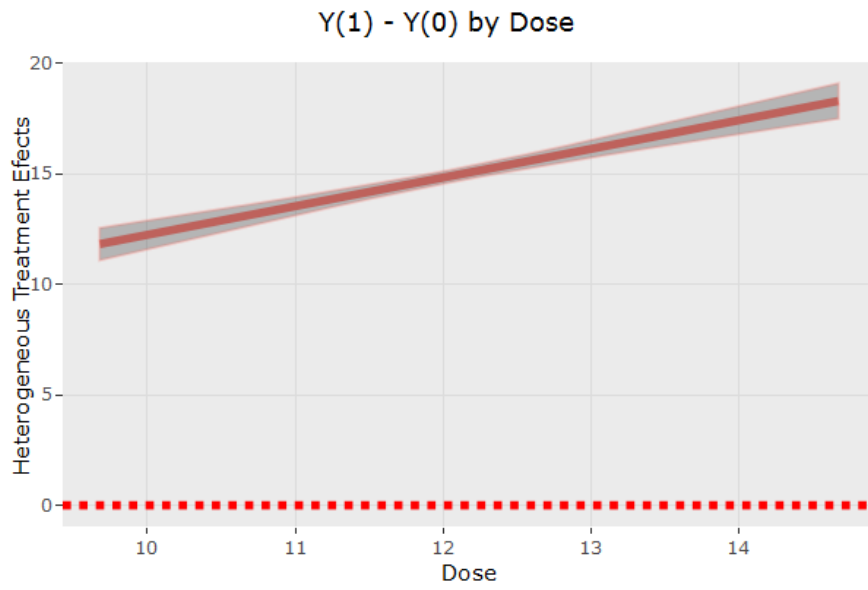


Figure 1.7.4. Estimated HTEs by Dose

4.3. Adaptive Types of Treatments

Example 2.1: Binary Treatment at Both Decision Point

In clinical practice, patients may be treated with a series of treatments for a long time. At some time points physicians may change the treatment according to the patients' disease characteristics. This example shows how to use the app to estimate the average treatment effect (ATE) for data with two time points. Simulated data are considered here. The simulation setting is listed as follows:

- $X \sim N(0,1)$
- $A_1 \sim \text{Bernoulli}(0.5)$
- $L_1 | A_1, X \sim N(0.25 + 0.3A_1 - 0.2X, 1)$
- $A_2 | L_1, A_1, X \sim \text{Bernoulli}(\text{expit}(-0.2 - 0.38A_1 + L_1))$
- $Y | A_1, X, L_1, A_2 \sim N(-2 + 2.5A_1 + 3.5A_2 + 0.5A_1A_2 - 0.6L_1, 2^2)$

The expectation of the potential outcomes $Y_i(a_1, a_2)$ can be derived as

$$E(Y(a_1, a_2)) = -2.15 + 2.32a_1 + 3.5a_2 + 0.5a_1a_2.$$

A step-by-step instruction is shown below:

Step 1. Choose variables and covariates at stage 1 (Figure 2.1.1)

"L1" is the continuous outcome and "C.A1" is the treatment at the first stage. "X" is chosen as confounding and explanatory variables. There is no heterogeneous treatment effect at any stage.

Step 2. Choose variables and covariates at stage 2 (Figure 2.1.2)

Y is the continuous outcome and "C.A2" is the treatment at the second stage. "L1" and "X" are chosen as confounding and explanatory variables.

Step 3. Run and show results

2000 'burn-in' and 2000 MCMC samples after 'burn-in' are considered here. The results are shown stage by stage.

For stage 1, [Figure 2.1.3](#) lists three tables for the descriptive analysis by the treatment, the estimated ATE, and the estimated potential outcomes. Table 2 in [Figure 2.1.3](#), shows that the estimated ATE is -0.34 which is very close to the true value 0.28. The estimated potential outcomes are also very good. [Figure 2.1.4](#) shows the histogram of the estimated ATE, and [Figure 2.1.5](#) shows the histogram of the estimated potential outcomes and the bar chart of the likelihood of L1 by the treatment group "C.A1" less than or equal to a value that can be chosen by users.

[Figure 2.1.6](#) shows the summary table of the results for stage 2. All 95% confidence intervals of the estimated potential outcomes in Table 6 cover the true values. [Figure 2.1.7](#) shows the histogram of the estimated ATE, and [Figure 2.1.8](#) shows the histogram of the estimated potential outcomes and the bar chart of the likelihood of $Y \leq 1.5$ by the treatment group "C.A1" and "C.A2".

Stage 1

Step 1: Outcome and treatment variables

Outcome variable: ⓘ

L1 ▼

Left- or right- bounded ⓘ

Treatment variable: ⓘ

C.A1 ▼

Step 2: Covariates

Confounding Variable(s): ⓘ

X

Explanatory Variable(s): ⓘ

X

Heterogeneous Treatment Effect ⓘ

Figure 2.1.1. Selected variables and covariates at stage 1

Stage 2

Step 1: Outcome and treatment variables

Outcome variable: ⓘ

Y ▼

Left- or right- bounded ⓘ

Treatment variable: ⓘ

C.A2 ▼

Step 2: Covariates

Confounding Variable(s): ⓘ

X L1

Explanatory Variable(s): ⓘ

X L1

Heterogeneous Treatment Effect ⓘ

Figure 2.1.2. Selected variables and covariates at stage 1

Stage 1:

Table 1: Descriptive statistics by treatment groups

	0	1	SMD
n	173.00	227.00	
X (mean (sd))	-0.04 (0.98)	-0.02 (0.92)	0.023
L1 (mean (sd))	0.20 (1.11)	0.54 (1.00)	0.323

Table 2: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	-0.3444	0.1028	-0.5459	-0.1430

Table 3: Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0.2000	0.0590	0.0845	0.3156
1	0.5445	0.0454	0.4555	0.6335

Figure 2.1.3. Results at stage 1

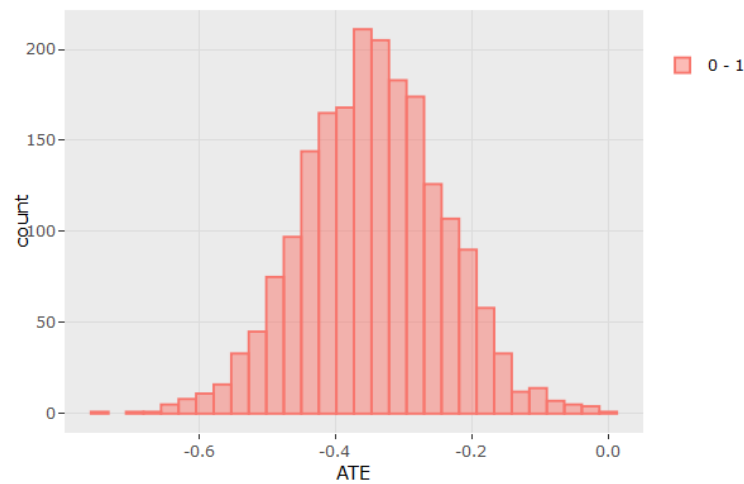


Figure 2.1.4. Histogram of averaged treatment effect estimates at stage 1

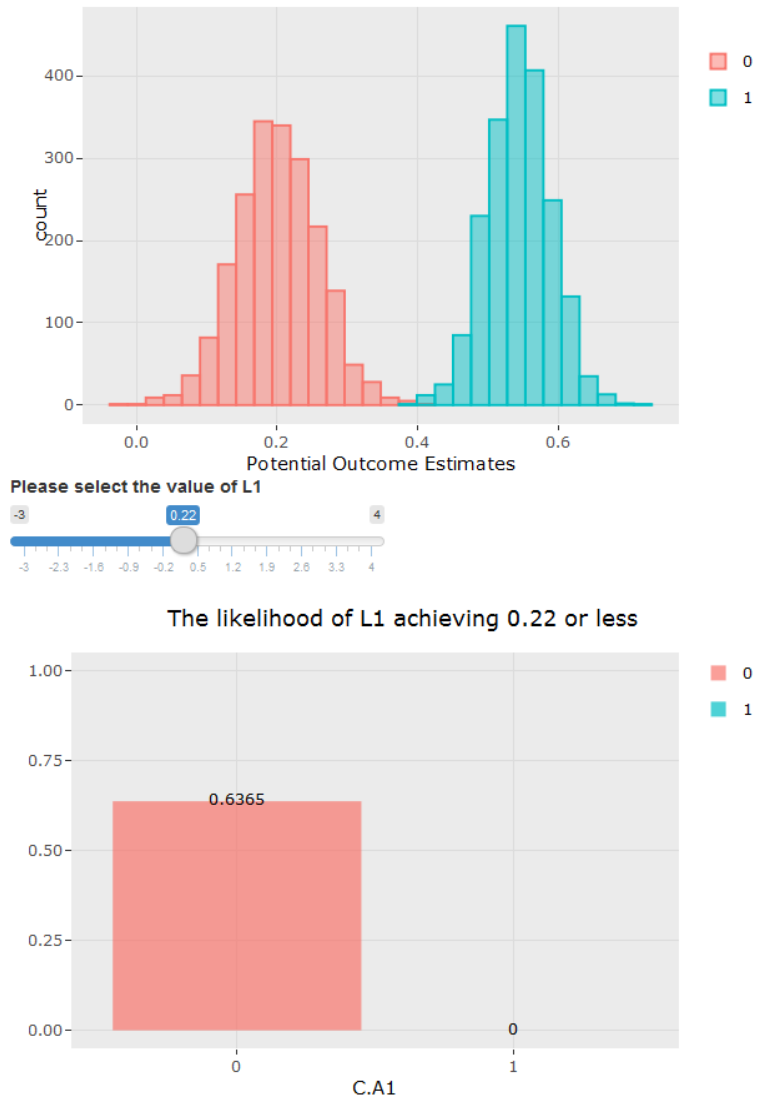


Figure 2.1.5. Histogram of potential outcome estimates and Bar chart of likelihood of $L1 \leq 0.22$ by C.A1 at stage 1

Stage 2:

Table 4: Descriptive statistics by treatment groups

	C.A1 = 0; C.A2 = 0	C.A1 = 0; C.A2 = 1	C.A1 = 1; C.A2 = 0	C.A1 = 1; C.A2 = 1	SMD
n	88.00	85.00	112.00	115.00	
L1 (mean (sd))	-0.21 (1.03)	0.62 (1.03)	0.07 (0.87)	1.01 (0.91)	0.729
X (mean (sd))	0.04 (1.05)	-0.13 (0.91)	0.18 (0.92)	-0.22 (0.89)	0.243
Y (mean (sd))	-1.54 (2.00)	0.86 (2.16)	0.28 (1.97)	3.98 (2.06)	1.399

Table 5: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0, 0 - 0, 1	-2.9144	0.3117	-3.5254	-2.3035
0, 0 - 1, 0	-1.8240	0.2869	-2.3862	-1.2618
0, 0 - 1, 1	-6.0829	0.3019	-6.6747	-5.4912
0, 1 - 1, 0	1.0904	0.3071	0.4885	1.6924
0, 1 - 1, 1	-3.1685	0.2894	-3.7358	-2.6012
1, 0 - 1, 1	-4.2589	0.2861	-4.8196	-3.6983

Table 6: Potential Outcome Estimates

C.A1	C.A2	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0	-1.7597	0.2014	-2.1544	-1.3650
0	1	1.1547	0.1971	0.7684	1.5410
1	0	0.0643	0.1641	-0.2573	0.3859
1	1	4.3232	0.1770	3.9764	4.6700

Figure 2.1.6. Results at stage 2

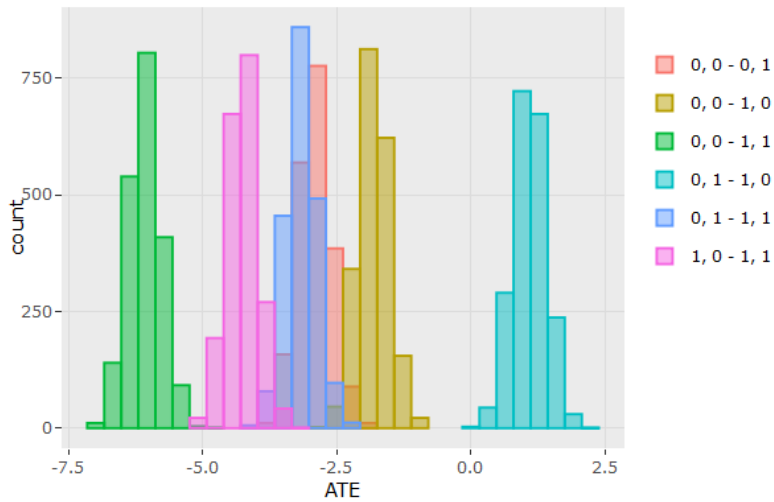


Figure 2.1.7. Histogram of averaged treatment effect estimates at stage 2

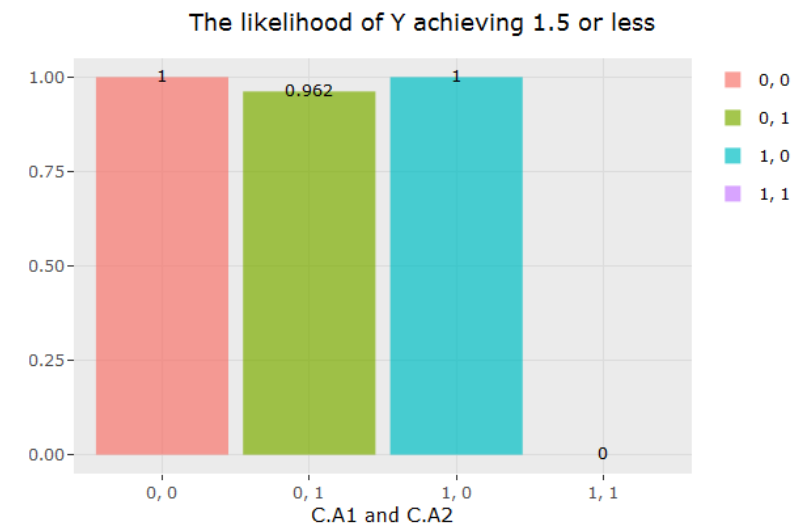
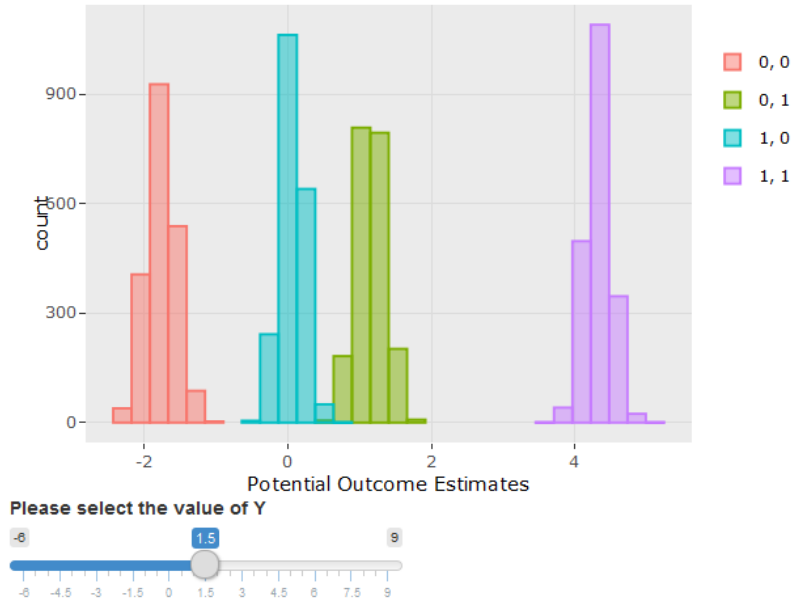


Figure 2.1.8. Histogram of potential outcome estimates and Bar chart of likelihood of $Y \leq 1.5$ by C.A1 and C.A2 at stage 2

Example 2.2: Subgroup HTE for a Binary Treatment at Both Decision Point

This example shows how to use the app to estimate the subgroup heterogeneous treatment effects (HTE). The data is simulated from the following generative models:

$$\begin{aligned} X &= -5,0,5 \text{ with Probability} = \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \\ A_0 &\sim \text{Bernoulli}(\text{expit}(0.3 - 0.05X)) \\ L_1 &\sim N(0.75X - 0.75A_0 - 0.25A_0X, 1) \\ A_1 &\sim \text{Bernoulli}(\text{expit}(0.05X + 0.2A_0 - 0.05L_1 - 0.1L_1 * A_0 - 0.01L_1^2)) \\ Y &\sim N(3 + 0.5A_0 + 0.4A_0X - L_1 - L_1^2 + 2A_1 - A_0A_1 + A_1L_1, 1) \end{aligned}$$

Figure 2.2.1 and Figure 2.2.2 show how to choose the outcomes and treatments in both stages, and the covariates in the model. 2000 ‘burn-in’ and 2000 MCMC samples after ‘burn-in’ are considered here. The results are shown stage by stage.

Figure 2.2.3 lists three tables for the descriptive analysis by the treatment at stage 1, the estimated ATE and the estimated potential outcomes at stage 1. The estimated ATE is 0.78 and its 95% confidence interval covers the true value “0.75”. Figure 2.2.4 shows the histogram of the estimated ATE and Figure 2.2.5 shows the histogram of the estimated potential outcomes at stage 1 and the bar chart of the likelihood of the outcomes “L1” <= -0.8 over the treatment group “C.A0”. The likelihood is 0 for the group C.A0=0 and 0.11 for the group C.A0=1. As shown in Figure 2.2.6, users can choose to show the estimated HTE and then control and treatment groups. In the example, C.A0=0 is chosen as the control group and C.A0=1 is the treatment group. After clicking on the “Go” button, the estimated HTE with its 95% confidence interval over X will be shown in Figure 2.2.7. The plot shows that C.A0=1 has a positive treatment effect and the effect decreases with the increase in X when X<-2.4, while C.A0=1 has a negative treatment effect and the effect increases with the increase in X when >2.4. Users can choose to show HTEs which are calculated on the user-specified value of HTE variables (Figure 2.2.8). Figure 2.2.9 plots the histogram of the estimated potential outcomes on these specified values. For other covariates included in the models, the observed values are used.

For the second stage, Figure 2.2.10 shows the summary tables of descriptive analysis by the treatments, the estimated ATE and the estimated potential outcomes at stage 2. For table 6 in Figure 2.2.10, the 95% confidence intervals of the estimated potential outcomes cover the true values -6.375, -4.375, -0.479 and -0.229. Figure 2.2.11 shows the histogram of the estimated ATEs, and Figure 2.2.12 shows the histogram of the estimated potential outcomes. The HTE panel at stage 2 is shown in Figure 2.2.13. Figure 2.2.14 shows the estimated HTE with its 95% confidence interval over L1.

Stage 1

Step 1: Outcome and treatment variables

Outcome variable: [i](#)

L1 ▼

Left- or right- bounded [i](#)

Treatment variable: [i](#)

C.A1 ▼

Step 2: Covariates

Confounding Variable(s): [i](#)

X

Explanatory Variable(s): [i](#)

X

Heterogeneous Treatment Effect [i](#)

Variable(s) with HTE: [i](#)

X

Figure 2.2.1. Selected variables and covariates at stage 1

Stage 2

Step 1: Outcome and treatment variables

Outcome variable: [?](#)

Y

Left- or right- bounded [?](#)

Treatment variable: [?](#)

C.A2

Step 2: Covariates

Confounding Variable(s): [?](#)

X L1

Explanatory Variable(s): [?](#)

X L1

Heterogeneous Treatment Effect [?](#)

Baseline variable(s) with HTE (C.A1): [?](#)

X

Variable(s) with HTE (C.A2): [?](#)

L1

Figure 2.2.2. Selected variables and covariates at stage 2

Stage 1:

Table 1: Descriptive statistics by treatment groups

	0	1	SMD
n	81.00	119.00	
X (mean (sd))	0.00 (4.40)	0.13 (4.09)	0.030
L1 (mean (sd))	-0.01 (3.53)	-0.71 (2.18)	0.239

Table 2: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	0.7777	0.1344	0.5143	1.0411

Table 3: Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0.0461	0.0819	-0.1144	0.2066
1	-0.7316	0.0556	-0.8407	-0.6226

Figure 2.2.3. Results at stage 1

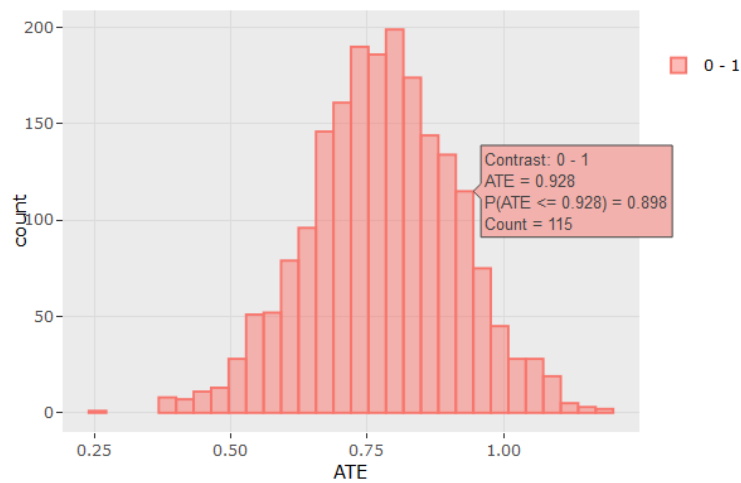


Figure 2.2.4. Histogram of averaged treatment effect estimates at stage 1

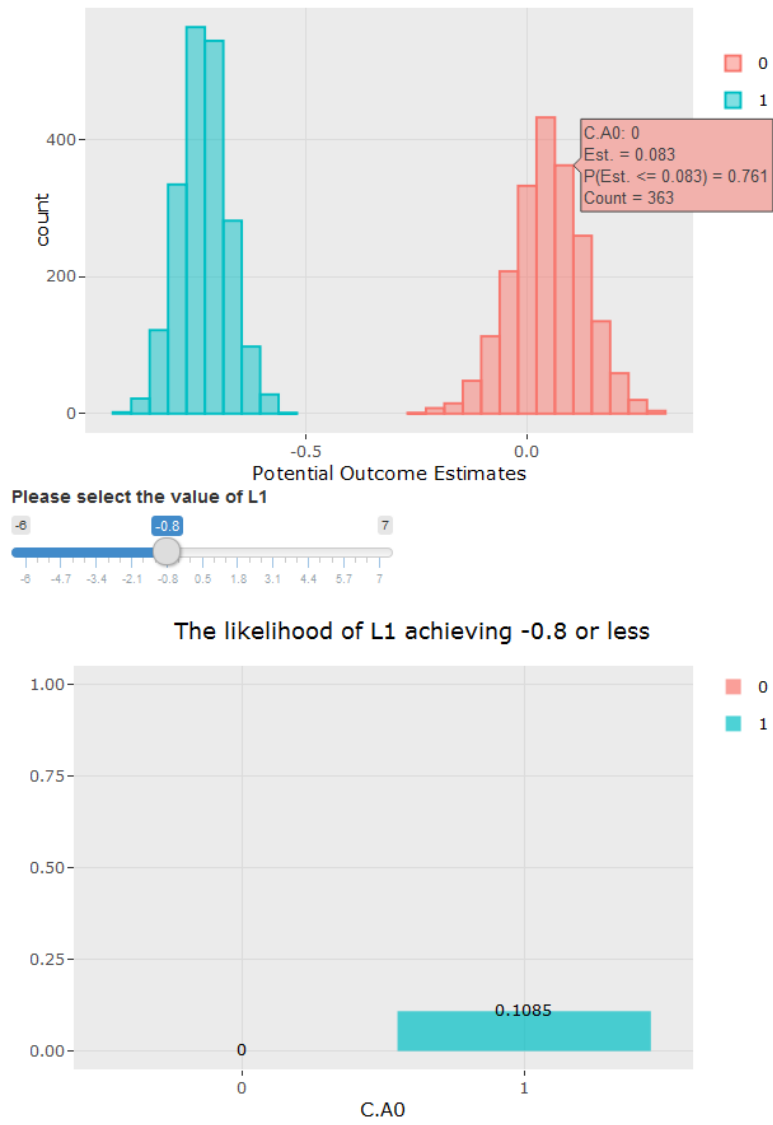


Figure 2.2.5. Histogram of potential outcome estimates and Bar chart of likelihood of L1 ≤ -0.8 by C.A0

Do you want to show HTE?

- Yes
- No

Choose < C.A0 > value as control group to calculate HTE:

- 0
- 1

Choose < C.A0 > value as treatment group to calculate HTE:

- 0
- 1

Choose a variable with HTE

- X

Go!

Figure 2.2.6. HTE panel at stage 1

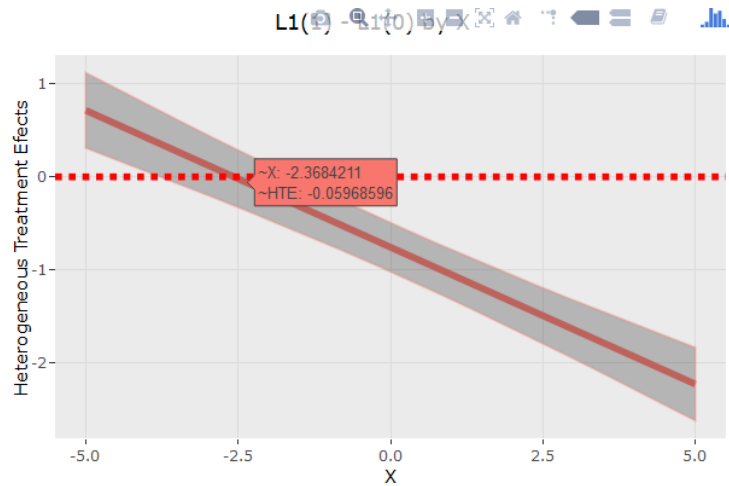


Figure 2.2.7. HTE over X at stage 1

Do you want to show HTE on specified values of covariates?

- Yes
- No

X

0

Go!

Figure 2.2.8. Show HTE on specified values at stage 1

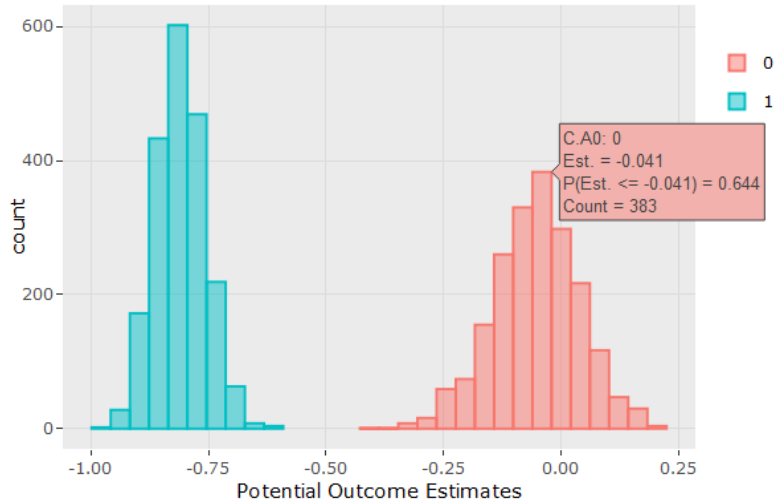


Figure 2.2.9. Show HTE on specified values at stage 1

Stage 2:

Table 4: Descriptive statistics by treatment groups

	C.A0 = 0; C.A1 = 0	C.A0 = 0; C.A1 = 1	C.A0 = 1; C.A1 = 0	C.A0 = 1; C.A1 = 1	SMD
n	48.00	33.00	50.00	69.00	
L1 (mean (sd))	-0.40 (3.75)	0.56 (3.15)	-0.37 (2.09)	-0.95 (2.23)	0.274
X (mean (sd))	-0.62 (4.45)	0.91 (4.23)	0.60 (4.12)	-0.22 (4.06)	0.213
Y (mean (sd))	-10.44 (10.50)	-4.81 (8.00)	-0.21 (4.48)	-1.32 (7.05)	0.708

Table 5: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0, 0 - 0, 1	-2.0725	0.2658	-2.5934	-1.5515
0, 0 - 1, 0	-7.3575	0.7346	-8.7973	-5.9177
0, 0 - 1, 1	-7.5972	0.7388	-9.0452	-6.1491
0, 1 - 1, 0	-5.2850	0.7373	-6.7301	-3.8400
0, 1 - 1, 1	-5.5247	0.7465	-6.9878	-4.0617
1, 0 - 1, 1	-0.2397	0.2083	-0.6480	0.1687

Table 6: Potential Outcome Estimates

C.A0	C.A1	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0	-7.6429	0.5255	-8.6729	-6.6129
0	1	-5.5705	0.5248	-6.5992	-4.5418
1	0	-0.2854	0.2736	-0.8218	0.2509
1	1	-0.0457	0.2846	-0.6035	0.5120

Figure 2.2.10. Results at stage 2

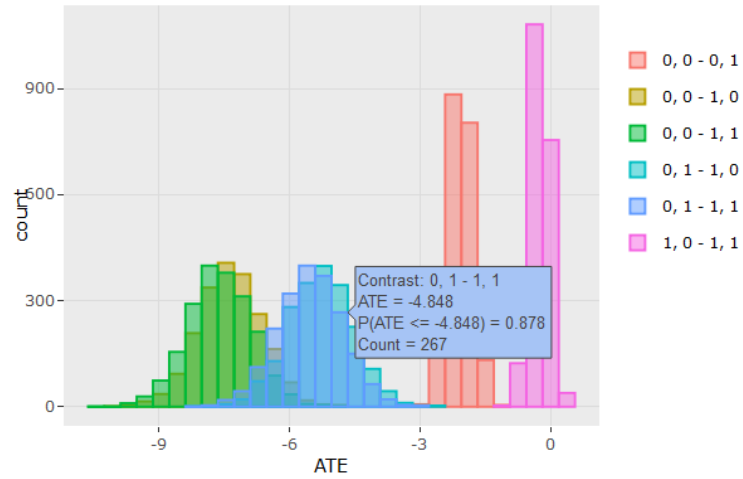


Figure 2.2.11. Histogram of averaged treatment effect estimates at stage 2

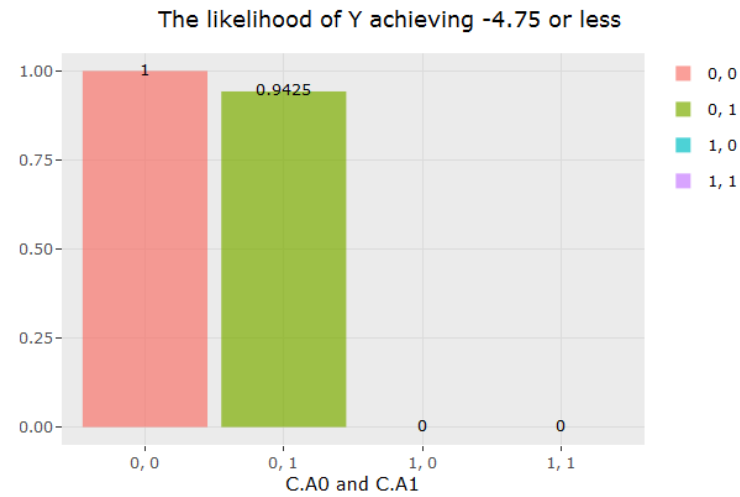
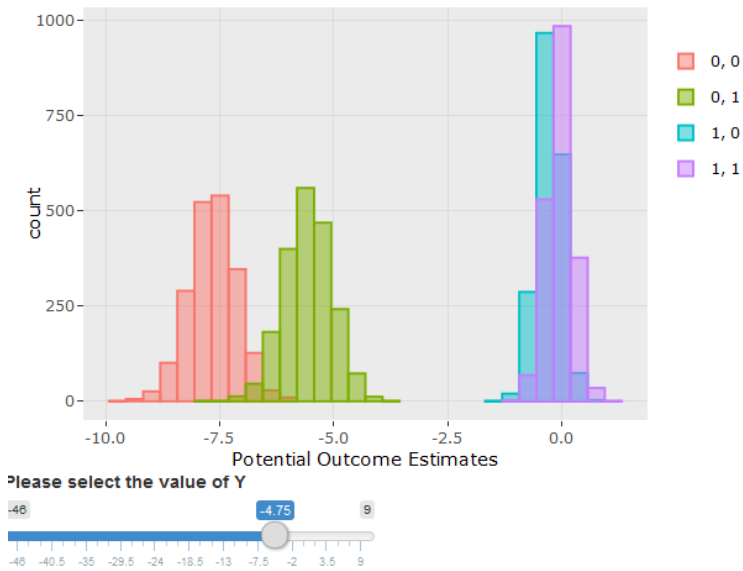


Figure 2.2.12. Histogram of potential outcome estimates and Bar chart of likelihood of $Y \leq -4.75$ by C.A0 and C.A1

Do you want to show HTE?

Yes

No

Choose < C.A0 > value as control group to calculate HTE:

0

1

Choose < C.A1 > value as control group to calculate HTE:

0

1

Choose < C.A0 > value as treatment group to calculate HTE:

0

1

Choose < C.A1 > value as treatment group to calculate HTE:

0

1

Choose a variable with HTE

L1

X

Go!

Figure 2.2.13. HTE panel at stage 2

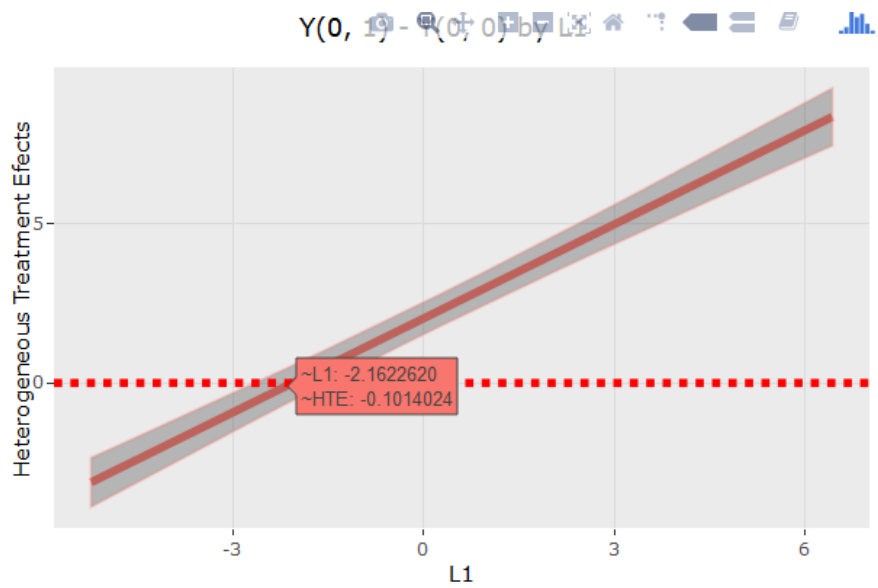


Figure 2.2.14. HTE over L1 at stage 2

Example 2.3: Binary outcome at stage 1 and continuous outcome at stage 2

This example shows how to use the app to estimate the average treatment effect (ATE) for data with two different types of outcome at time points. Simulated data are considered here based on the following simulation setting (Daniel *et al.*, 2013):

- $U_0 \sim \text{Bernoulli}(0.4)$
- $A_0 \sim \text{Bernoulli}(0.5)$
- $L_1 | A_0, U_0 \sim \text{Bernoulli}(\text{expit}(0.25 + 0.3A_0 - 0.2U_0 - 0.05A_0U_0))$
- $A_1 | L_1, A_0, U_0 \sim \text{Bernoulli}(\text{expit}(0.4 + 0.5A_0 - 0.3L_1 - 0.4A_0L_1))$
- $Y | A_0, U_0, L_1, A_1 \sim N(0.25 - 0.5A_0 - 0.75A_1 + 0.2A_0A_1 - U_0, 0.2)$

The expectation of the potential outcomes $Y_i(a_1, a_2)$ can be derived as

$$E(Y(a_1, a_2)) = 2.1 - 0.5a_0 - 0.75a_1 + 0.2a_0a_1.$$

Step 1. Choose variables and covariates at stage 1 (Figure 2.3.1)

“L1” is the binary outcome and “A0” is the treatment at the first stage. “U0” is chosen as confounding and explanatory variables. Since there is an interaction of A_0 and U_0 , U_0 is selected as the variable which has the heterogeneous treatment effect.

Stage 1

The screenshot displays the 'Stage 1' configuration interface, divided into two main sections: 'Step 1: Outcome and treatment variables' and 'Step 2: Covariates'.

Step 1: Outcome and treatment variables

- Outcome variable:** A dropdown menu with 'L1' selected.
- Treatment variable:** A dropdown menu with 'A0' selected.

Step 2: Covariates

- Confounding Variable(s):** A text input field containing 'U0'.
- Explanatory Variable(s):** A text input field containing 'U0'.
- Heterogeneous Treatment Effect:** A checkbox that is checked.
- Variable(s) with HTE:** A text input field containing 'U0'.

Figure 2.3.1. Selected variables and covariates at stage 1

Step 2. Choose variables and covariates at stage 2 (Figure 2.3.2)

Y is the continuous outcome and “A1” is the treatment at the second stage. “L1” and “U0” are chosen as confounding and explanatory variables.

Stage 2

The figure shows a two-step configuration interface for Stage 2. Step 1, titled "Step 1: Outcome and treatment variables", contains a dropdown menu for the "Outcome variable" set to "Y", an unchecked checkbox for "Left- or right- bounded", and a dropdown menu for the "Treatment variable" set to "A1". Step 2, titled "Step 2: Covariates", contains a text input for "Confounding Variable(s)" with "U0 L1" entered, another text input for "Explanatory Variable(s)" also with "U0 L1" entered, and an unchecked checkbox for "Heterogeneous Treatment Effect".

Figure 2.3.2. Selected variables and covariates at stage 2

Step 3. Run and show results

2000 ‘burn-in’ and 2000 MCMC samples after ‘burn-in’ are considered here. The results are shown stage by stage.

Figure 2.3.3 shows the tables for the descriptive analysis by the treatment at stage 1, the estimated ATE, and the estimated potential outcomes at stage 1. Table 2 in Figure 2.3.3, shows that the estimated ATE is 0.06 and its 95% confidence interval is (-0.10,0.17). Figure 2.3.4 is the violin plots overlaid with box plots by U0.

Figure 2.3.5 shows the summary table of the results for stage 2. The estimated potential outcomes in Table 6 are close to the true values.

Stage 1:

Table 1: Descriptive statistics by treatment groups

	0	1	SMD
n	94.0	106.0	
U0 = 1 (%)	38.0 (40.4)	49.0 (46.2)	0.117
L1 = 1 (%)	54.0 (57.4)	64.0 (60.4)	0.060

Table 2: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0 - 1	0.0352	0.0690	-0.1000	0.1704

Table 3: Potential Outcome Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0.4326	0.0498	0.3350	0.5301
1	0.3974	0.0472	0.3049	0.4898

Figure 2.3.3. Results at stage 1

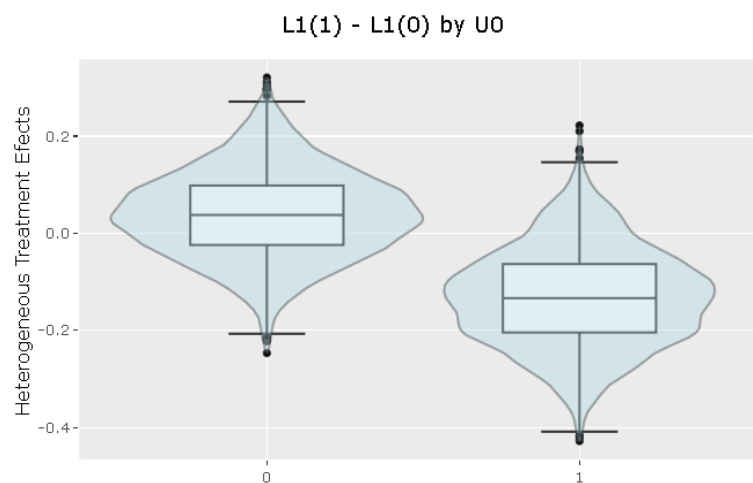


Figure 2.3.4. Violin plot of L1 by U0 at stage 1

Stage 2:

Table 4: Descriptive statistics by treatment groups

	A0 = 0; A1 = 0	A0 = 0; A1 = 1	A0 = 1; A1 = 0	A0 = 1; A1 = 1	SMD
n	44.0	50.0	40.0	66.0	
L1 = 1 (%)	27.0 (61.4)	27.0 (54.0)	28.0 (70.0)	36.0 (54.5)	0.190
U0 = 1 (%)	18.0 (40.9)	20.0 (40.0)	19.0 (47.5)	30.0 (45.5)	0.091
Y (mean (sd))	2.11 (0.53)	1.35 (0.54)	1.53 (0.57)	0.98 (0.54)	1.103

Table 5: Averaged Causal Treatment Effect Estimates

	Estimation	SD	95% Lower Bound	95% Upper Bound
0, 0 - 0, 1	0.7736	0.0420	0.6914	0.8559
0, 0 - 1, 0	0.5135	0.0447	0.4260	0.6011
0, 0 - 1, 1	1.0930	0.0405	1.0136	1.1724
0, 1 - 1, 0	-0.2601	0.0452	-0.3487	-0.1714
0, 1 - 1, 1	0.3194	0.0399	0.2413	0.3975
1, 0 - 1, 1	0.5794	0.0413	0.4985	0.6604

Table 6: Potential Outcome Estimates

A0	A1	Estimation	SD	95% Lower Bound	95% Upper Bound
0	0	2.0852	0.0281	2.0300	2.1403
0	1	1.3116	0.0260	1.2607	1.3624
1	0	1.5716	0.0321	1.5087	1.6345
1	1	0.9922	0.0214	0.9502	1.0342

Figure 2.3.5. Results at stage 2

References

Daniel, R., Cousens, S., De Stavola, B., Kenward, M. G. and Sterne, J. A. (2013), Methods for dealing with time - dependent confounding. *Statist. Med.*, 32: 1584-1618.