Artificial Data Analysis

Florian Stijven

2023-08-24

Table of Contents

In this document, we examine the influence of two different data analytic choices on the estimation of optimal treatment regimes in the *modified Browne data*:

* The influence of the imputation model, i.e., global imputation versus imputation per arm.
* The influence of a misspecified outcome regression model on Q-learning versus value search estimation.

The *modified Browne data* are the Browne data where we artificially induce relationships between baseline covariates and the outcome variable. In what follows, we will only consider the CESD outcome measured at 6 months after randomization (denoted by ). For estimating the optimal regimes, we still consider the change scores, however. We further consider three modified versions of the Browne data, further referred to as the updated data 1, 2, and 3. The modifications are briefly explained next, but a more detailed description and justification is given in a separate document. In the following bullet points, a tilde indicates the updated value for CESD at 6 months. Treatment is “Sertraline alone” and treatment is “Sertraline and IPT”.

* **Update 1.** The outcome value is modified as follows,
* This update changes the interaction effects. Patients will thus benefit more from the addition of IPT to Sertraline if they have a more severe depression at baseline (larger baseline CESD and past MDD).
* **Update 2.** The outcome value is modified as follows where is now the outcome value after update 1,
* This update thus changes the main effects only. After this update, a linear regression model without quadratic age effect will thus be misspecified.

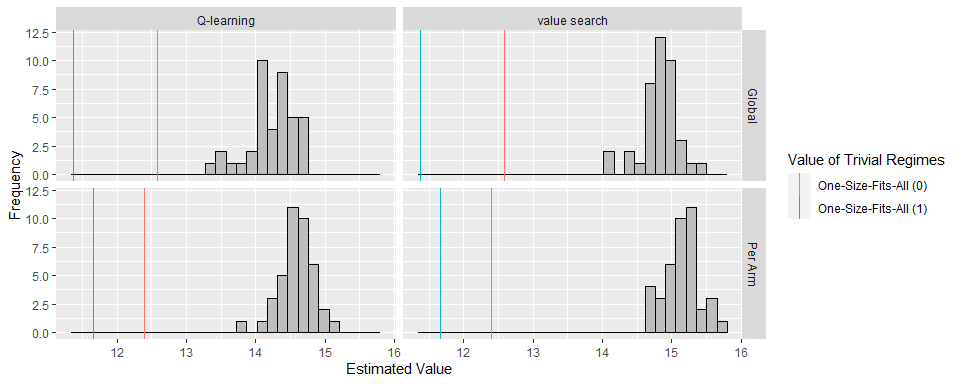
In the following analyses, we will only consider 40 imputations instead of the 200 imputations used for analyzing the original Browne data. This limits the computational burden. The *per arm* and *global* imputation models are the same as for the analyses of the original Browne data.

# 1 Choice of Imputation Model

In this section, we look at the influence of the imputation model on the estimation of regimes by Q-learning and value search estimation. We only look at the update 1 data. So, we do not consider the artificial settings where the outcome regression models are misspecified (update 2 and 3).

## 1.1 Estimated Values of Estimated Regimes

In the following histograms, the estimated values of the estimated regimes are summarized across the 40 imputations. These histograms show that imputation per arm leads to larger estimated values.



Frequency distribution of the estimated values of the estimated regimes across the imputations for the update 1 data. Note that each histogram represents 40 estimated values of 40, possibly different, estimated regimes. The value is estimated by the AIPW estimator which is explained at the end of the document for the original Browne data.

The following table summarizes the above results. This confirms our previous conclusions.

Average estimated value of the estimated regimes across the imputed data sets and the corresponding standard deviation. Note that each value is the average of 40 estimated values of 40, possibly different, estimated regimes. The value is estimated by the AIPW estimator which is explained at the end of the document for the original Browne data.

| imputation | OTR\_method | Mean Estimated Value | Between Imputation SD |
| --- | --- | --- | --- |
| Global | Q-learning | 14.230 | 0.347 |
| Per Arm | Q-learning | 14.578 | 0.246 |
| Global | value search | 14.831 | 0.265 |
| Per Arm | value search | 15.145 | 0.252 |

## 1.2 Distance from One-Size-Fits-All of Estimated Regimes

In the following figure and table, the same estimated regimes as above are summarized by the corresponding . This value measures the distance of the estimated regimes from a one-size-fits-all regime. The corresponding average distances are summarized in the following table.

For Q-learning, estimated regimes under imputation per arm tend to be further away from one-size-fits-all. However, there is no considerable difference for the regimes estimated by value search estimation.



Frequency distribution of the distance from one-size-fits-all for the estimated regimes across the imputations.

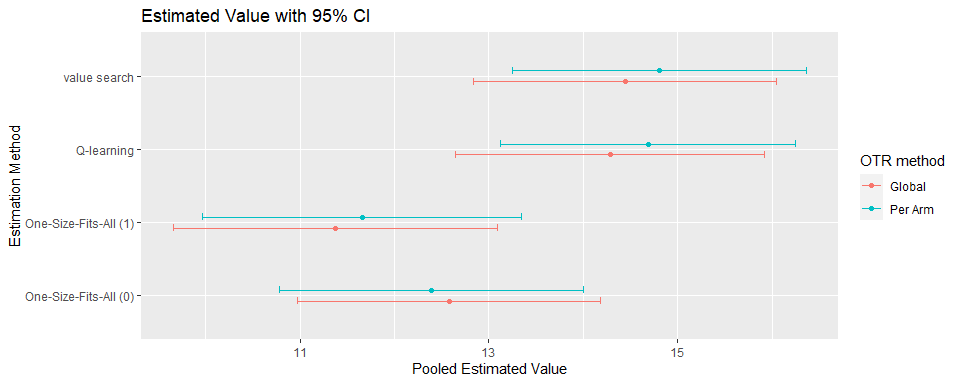
Average distance from one-size-fits-all of the estimated regimes across the imputed data sets. Note that each value is the average of 40 distances of 40, possibly different, estimated regimes.

| OTR\_method | Global | Per Arm |
| --- | --- | --- |
| Q-learning | 0.414 | 0.449 |
| value search | 0.409 | 0.398 |

## 1.3 Aggregated Regimes

Finally, we aggregate the 40 estimated regimes in each setting (imputation and regime estimation method) with the circular mean. We consider the pooled estimate of the value and the corresponding 95% CIs in the following figure (by applying Rubin’s rules). The results for the trivial regimes are added for reference.

The figure below confirms the previous results. Imputation per arm leads to a better aggregated regime in terms of the pooled estimate of its value. This holds for both Q-learning and value search estimation. Even though the difference for these data is likely of little clinical relevance, it is still a considerable difference when taking into account that only about 10% of the patients have missing values (at 6 months).



Pooled estimates for the value of the aggregated regimes together with 95% confidence intervals for the update 1 data. The pooled estimates and confidence interval are obtained by applying Rubin’s rules.

In the next table, the aggregated regime parameters are summarized. The estimated parameters for CESD and past MDD tend to be closer to zero under global imputation. Although the difference in estimated parameters is relatively small, there is a considerable difference in the corresponding pooled estimated value (as was shown above).

Linear regime parameter estimates for the aggregated regimes (without additional modificiations) for the update 1 data. The respective parameter vectors have unit norm.

| imputation | OTR\_method | outcome\_model | constant | sex | age | famfun | cesd |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Global | Q-learning | Correct | -0.954 | -0.031 | 0.004 | 0.076 | 0.014 |
| Per Arm | Q-learning | Correct | -0.946 | -0.022 | 0.004 | 0.059 | 0.016 |
| Global | value search | Correct | -0.964 | 0.051 | 0.003 | 0.056 | 0.017 |
| Per Arm | value search | Correct | -0.957 | 0.028 | 0.003 | 0.048 | 0.017 |

# 2 Misspecification of Main Effects

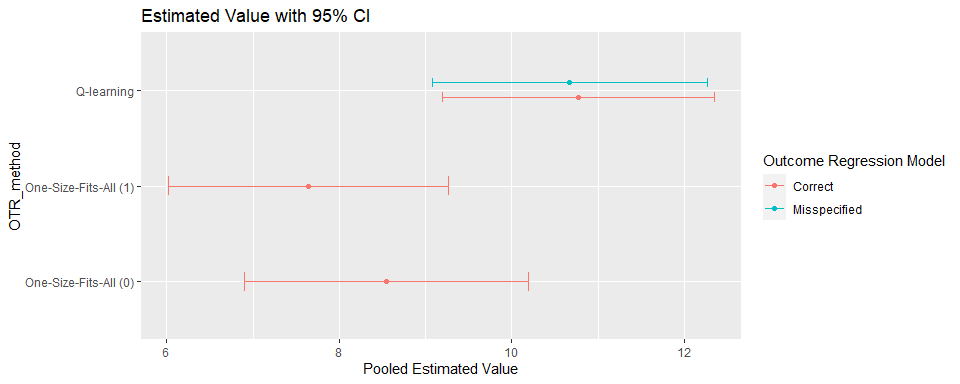
In this section, we will look at the influence of misspecification of the outcome regression model on Q-learning versus value search estimation. Q-learning relies on a correctly specified outcome regression model while the value search estimator does not. However, this does not mean that Q-learning will perform much worse under all types of misspecifications. We have artificially added a quadratic effect of age on CESD at 6 months to the data.

In this section, we only consider imputation per arm. Note that a quadratic age effect has also been added to the imputation model. This is consistent with proper imputation: the imputation model and analysis model should not contradict each other. Since we use a quadratic age effect in the analyses with correctly specified outcome regression model, we should also include it in the imputation model.

We conduct four analyses: Q-learning and value search estimation, each with and without a correctly specified outcome regression model.

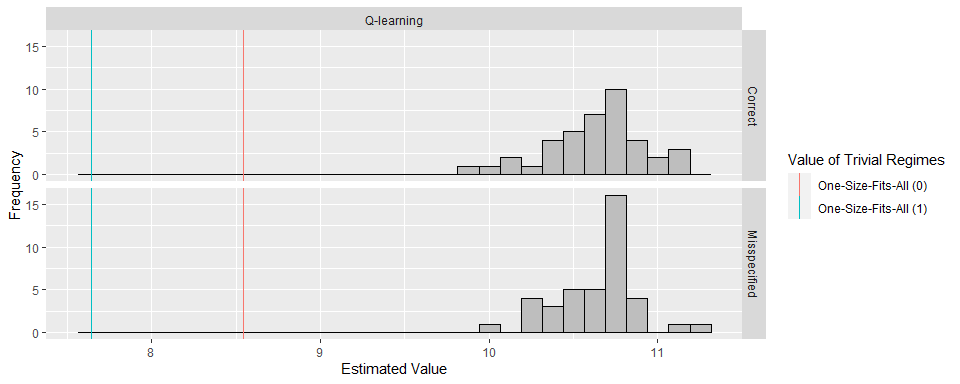
In the next figure, the pooled estimated value of the aggregated regimes are summarized together with 95% CI’s. The values of the aggregated regimes are estimated with the AIPW estimator that uses the same outcome regression model as the estimator for the optimal regimes.

It is clear that failing to include a quadratic effect of age in the outcome regression model for Q-learning does not have an important effect on the estimated regime’s value. The same is true for value search estimation.



Pooled estimates for the value of the aggregated regimes together with 95% confidence intervals for the update 2 data. The pooled estimates and confidence interval are obtained by applying Rubin’s rules.

We next look at the same comparisons, but *before* aggregation, i.e., the estimated values of the estimated regimes in each imputed data set. The advantage of comparing both methods *before aggregation* is that the aggregation method cannot influence the comparison.



Frequency distribution of the estimated values of the estimated regimes across the imputations for the update 2 data. Note that each histogram represents 40 estimated values of 40, possibly different, estimated regimes.