MAP573 REPORT

Causal inference to assess the effect of a treatment on survival - HTE

Askinazi Eric, Feitz Jérémie, Lescure Charles, Resplandy Michaël, Zylberajch Hugo

ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

traumabase.eu

INTRODUCTION

Traumatic brain injuries affect about 120.000 people per year in France alone [1], and can often lead to death and life-long consequences [2]. Tranexamic acid is used in surgeries on traumatic brain injury patients in order to reduce surgical bleeding [3]. Our project consists in determining as precisely as possible the effect of tranexamic acid on patients suffering from traumatic brain injuries, using the TraumaBase database of patients [4].

Let us briefly formalize our problem. We observe $(W_i, Y_i, X_i)_i^N$, where W_i represents whether subject i was given Tranexamic Acid, Y_i is a binary variable representing whether he died or not, and X_i is a vector of covariates. We will denote by $Y_i(1)$ the random variable that represents the potential outcome of subject i had they received the treatment, and $Y_i(0)$ will represent the same potential outcome had they not received anything. The individual treatment effect for subject i can then be written as $Y_i(1) - Y_i(0)$.

In this project, we'll be using two concepts for calculating the effect of tranexamic acid on traumatic brain injury patients, which are Average Treatment Effect (ATE) and Heterogeneous Treatment Effect (HTE).

— ATE : $\tau = \mathbf{E}[Y_i(1) - Y_i(0)]$
— HTE : $\tau(x) = \mathbf{E}[Y_i(1) - Y_i(0)|X_i = x]$

A positive ATE means that our treatment decreases the survival rate on average while a negative ATE means that our treatment helps the victims surviving. Of course, since a patient is either treated or not treated, we can't measure its individual treatment effect $Y_i(1) - Y_i(0)$.
However, with some assumptions that we'll see, we'll be able to estimate it.

TABLE DES MATIÈRES

I
TRAUMABASE AND PREPROCESSING

I.1 The TraumaBase

The TraumaBase [4] is a private database composed of patients that have suffered major trauma. Its analysis is key for the doctors as it gives insights in how to predict trauma specific outcomes and decisions as well as improve patient care and survival. This is because, although every case should be adapted to a certain extent to the patient and the trauma system, procedures can be standardized and reproducible.

In its raw form, the database is composed of 20000 patients and 250 variables, both continuous and categorical, gathered from 16 hospitals and has a lot of missing values.

However, not all data will be of interest as we will be focusing exclusively on the effect of tranexamic acid on patients who have severe head trauma.

I.2 Preprocessing

I.2.1 • Covariate Selection

As explained previously, we will first filter the initial dataset by type of patients and by covariates.

Our causal inference analysis will only focus on the patients that have had a lesion visible on the CT scan (translated into `Trauma.cranien == 1`) and/or an `AIS.tete` score equal or higher than 2, given that the aim of our study is to analyse the (causal) effect of the administration of tranexamic acid in patients that have a brain injury.

Furthermore, we will also filter on the `ISS/Selection` variable, which stores the Abbreviated Injury Scale (AIS) code of a patient's injuries, by filtering out the patients that do not have one of the AIS codes that represent a traumatic brain injury.

Then, we will only keep the covariates that are related to the administration of Tranexamic acid and the death of the patient, which are shown in figure 1.

At this point, we have a dataset composed of 5337 patients and 39 covariates.

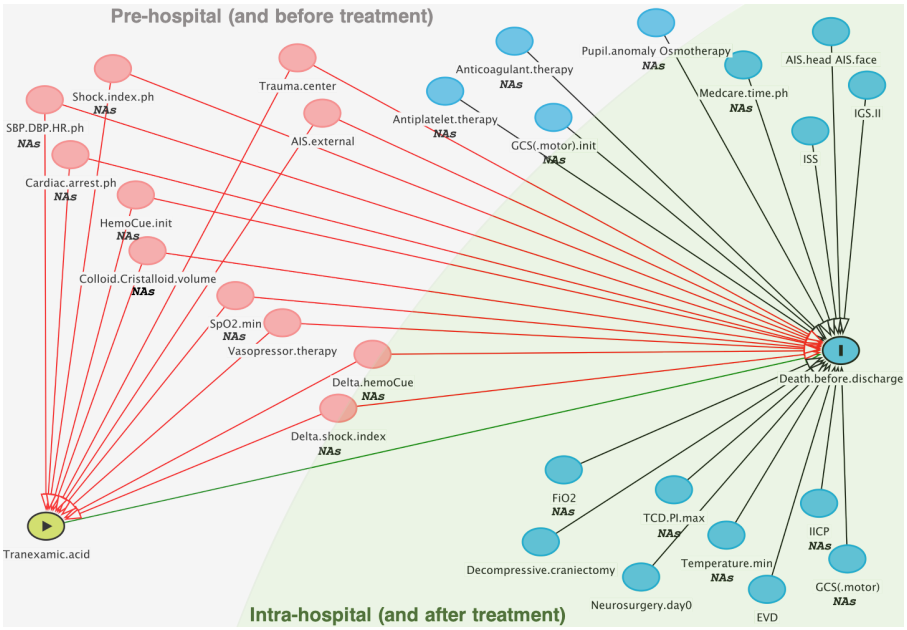


FIGURE 1 – Covariates useful for causal inference

I.2.2 • DATA CLEANING

A quick analysis of the database suggests that we should merge the two variables : `Osmotherapy.ph` and `Improv.anomaly.osmo` as the latter is a direct consequence of the former.

Moreover, we decided to remove the covariate `Temperature.min` altogether as it had more than 90 % of missing values.

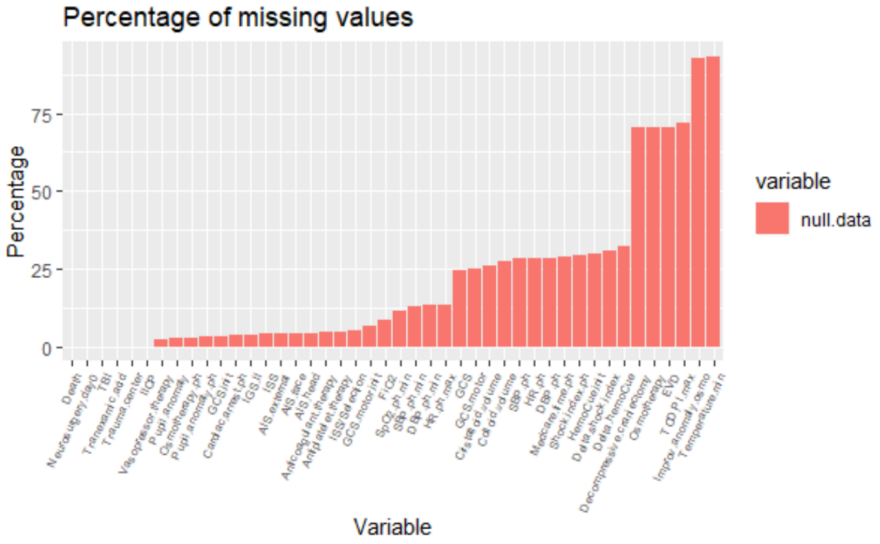


FIGURE 2 – Percentage of missing values for each covariate

I.2.3 • Missing Data Imputation

Our dataset contains a lot of missing data and imputing those efficiently is an important part of our project.

We decided to use 4 imputing methods in order to obtain 8 datasets to work on. We implemented :

— Simple Imputer : We simply impute missing numerical values by the mean of the distribution. For categorical variables, we create a new category called "missing".

— MissForest Imputation [5] : Data imputation using random forests.

— Mice [6] : Multiple imputation for multivariate missing data. It uses different methos such as Predictive Mean Matching, Logreg or Polyreg. The function returns 5 datasets.

— FAMD [7] : Imputation using the FAMD function from the FactoMineR [8] library (imputation method that support categorical variables).

II

AVERAGE TREATMENT EFFECT AND ITS STABILITY

The first value that we would like to estimate is the *Average Treatment Effect* (ATE), defined by :

$$\tau = \mathbf{E}[Y_i(1) - Y_i(0)]$$

II.1 Stability of the ATE in relation to the imputation method

As previously said, we used three imputation methods :

- Simple imputation : mean imputation
- MissForest : Imputation with Random Forests [5]
- Mice : Multiple imputation with several methods (*pmm*, *logreg*, *polyreg*) [6]

We would like to know if the ATE changes between the different imputed datasets. To compute it, we use the **R** package grf [9] for random forests developed by Stanford researchers. In table 1 are the results of the calculation on the 6 different datasets.

Base	ATE	std err.
MissForest	0.06201875	0.03421128
Mice 1	0.06599013	0.03428097
Mice 2	0.06591918	0.03269830
Mice 3	0.06515854	0.03049809
Mice 4	0.04747659	0.02389296
Mice 5	0.06897057	0.03042385

Table 1 – ATE of each imputed Database

This table seems to show that the ATE is relatively stable (around 0.06). Furthermore, since this result is positive, it shows that tranexamic acid has a negative average effect on patients.

In all further calculations, we focus on the database completed using MissForest.

II.2 Stability of the ATE in relation to the calculation method

To estimate the ATE, we first need these two assumptions :

— The **unconfoundness** assumption :

$$Y_i(1), Y_i(0) \perp W_i | X_i$$

This assumption simply says that the treatment was given considering the state of the patient only, given by the covariates X. In other words, knowing whether someone was treated doesn't give us more information than knowing the covariates, to estimate the potential outcomes.

— The **overlap** assumption which is needed for the two Inverse-Propensity-Weighting methods :

$$\forall x \in Supp(X), 0 < \mathbb{P}(W = 1 | X = x) < 1$$

Then, as we previously said, there are mainly three methods that allow to estimate our ATE.

II.2.1 • Ordinary Least Squares Regression Adjustment

The first method is the **Ordinary Least Squares Regression Adjustment** method.

$$\begin{aligned}
\tau(x) &= \mathbf{E}[Y_i(1) - Y_i(0) | X_i = x] \\
&= \mathbf{E}[Y_i(1) | X_i = x] - \mathbf{E}[Y_i(0) | X_i = x] \\
&= \mathbf{E}[Y_i(1) | W_i = 1, X_i = x] - \mathbf{E}[Y_i(0) | W_i = 0, X_i = x] \text{ } \mathbf{unconfoundness\ assumption} \\
&= \mu(1, x) - \mu(0, x) \\
\tau &= \mathbf{E}[\tau(x)] = \mathbf{E}\left[\mu\left(1, x\right) - \mu\left(0, x\right)\right]
\end{aligned}$$

$\mu(1, x)$ and $\mu(0, x)$ can be estimated by training two models respectively on the treated individuals and on the untreated individuals, to predict the survival outcome given the covariates X.

II.2.2 • Inverse-Propensity-Weighting

The second method is the **Inverse-Propensity-Weighting** Method. It uses $e(X_i)$, the propensity score which is the probability of getting treated given the covariate X. It can be estimated by training a model on the whole dataset.

$$\tau = \mathbf{E}\left[\frac{YiWi}{e(X_i)} - \frac{Yi(1 - Wi)}{(1 - e(X_i))}\right]$$

II.2.3 • Augmented-Inverse-Propensity-Weighting

The third method is the **Augmented-Inverse-Propensity-Weighting** Method, and is a combination of the first two methods.

$$\tau = \mathbf{E}\left[Wi\frac{Yi - \mu(1, X_i)}{e(X_i)} - (1 - Wi)\frac{Yi - \mu(0, X_i)}{(1 - e(X_i))} + \mu(1, X_i) - \mu(0, X_i)\right]$$

II.2.4 • Computing the ATE

In table 2, we show our results when calculating the ATE using for each of the 3 methods method 3 different types of predictive model : logistic regression, LGBoost model and XGBoost classification model.

ATE calculation method	IPW	OLS	AIPW
Logitic regression	-1.827	-0.761	0.056
LGBoost	-0.209	-0.010	0.051
XGBoost	-0.054	0.094	0.060

Table 2 – ATE with each method

Clearly, there is a method that is more stable than the others : the AIPW. Furthermore, it is quite reassuring to find an ATE around 0.06 which is the result we had using grf.

However, all values are very sensitive to the models used, and even changing slightly a hyper-parameter has big influence on the ATE (even with the AIPW method).

II.3 Stability of the ATE in relation to the predictors

To test its stability even more, we tried to mix the models used for our predictors *ie* for the propensity score and for the $\mu(1, x)$ and $\mu(0, x)$ scores. For each combination of models, we then used bootstrapping to build confidence intervals.

$\mu(1/0, x)$ | propensity	Logitic regression	LGBoost	XGBoost
Logitic regression	0.066 +/- 0.011	0.047 +/- 0.002	0.012 +/- 0.004
LGBoost	0.091 +/- 0.006	0.052 +/- 0.001	0.037 +/- 0.001
XGBoost	-0.011 +/- 0.013	0.091 +/- 0.001	0.062 +/- 0.004

Table 3 – AIPW estimator with each predictors

This table shows that even the AIPW-estimator varies a lot depending on the method used to estimate our parameters. This also shows that we should keep a critical eye on the results we'll show in the next part.

III
ATE ON CLUSTERS

In the previous section, we noted that tranexamic acid has an overall neutral or slightly positive average treatment effect on patients suffering from traumatic brain injuries, depending on the method used to compute this effect.

Our main goal is to find if some patients with traumatic brain have a negative treatment effect to tranexamic acid, which would indicate that tranexamic acid is useful for these patients.

This is why we start by dividing the set into clusters of patients with similar features, and compute the ATE on these clusters, in order to evaluate the effect of tranexamic acid on different kinds of patients independently.

III.1 Clustering by type of injury

Doctors already differentiate 3 different types of traumatic brain injuries :

- "Lésion axonale diffuse"
- "Lésion extra-axiale"
- "Lésion intra-axiale"

Not all patients lie in one cluster or another, when asked about it doctors told us to filter out these patients, leaving us with a bit more than 5000 patients, as seen in I.2.1.

Our first approach was to use differentiate each one of these three types, and compute the ATE on each of these types as separate databases, to evaluate the effect of tranexamic acid for each one of these injury types. The only injury type that gave us interesting results was the cluster 2 with the grade 0 ("Lésion intra-axiale").

$\mu(1/0, x)$ | propensity	Logitic regression	LGBoost	XGBoost
Logitic regression	-0.457 +/- 0.078	-0.001 +/- 0.004	0.0016 +/- 0.0022
LGBoost	-0.173 +/- 0.004	0.011 +/- 0.002	0.011 +/- 0.002
XGBoost	-0.586 +/- 0.098	0.035 +/- 0.007	0.055 +/- 0.004

Table 4 – ATE for cluster 2 and grade 0, using different calculation methods

Even though there are still some positive values, most of them are negative (which wasn't the case with the other clusters/grades where there was at most one slightly negative value).

However, that still depends on the method used and we didn't manage to get the same results when using causal forests as we got a positive ATE for all clusters.

III.2 Hierarchical clustering

We tried to get interesting clusters using k-means but our efforts weren't successful. Then, we tried to use a hierarchical clustering and to compare the relevance of the results. The hierarchical clustering turned out to be more efficient since we managed to get clusters with negative ATE with only 15 clusters repeatedly.

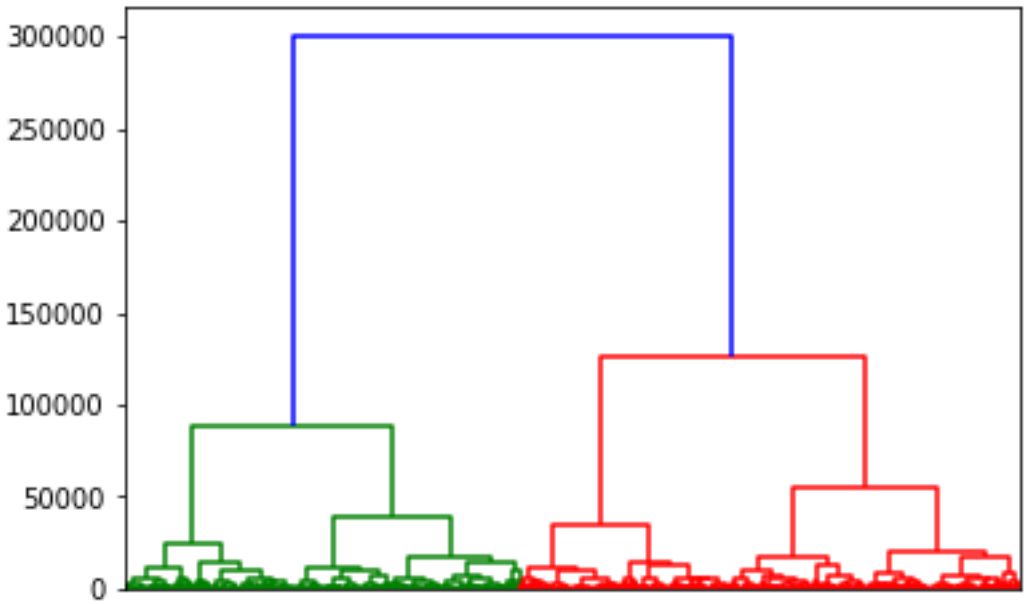


Figure 3 – Dendrogram of the hierarchical clustering

In table 5, we can see that the ATE on cluster 9 is negative, meaning that the patients in this cluster benefit from being given tranexamic acid. We decided to investigate how this cluster is different from the others.

It seems that the interesting covariates would be the ones that have a significant difference between the cluster and the global database. We decided to calculate the difference in means for each variable between the values in the cluster and in the overall database, which we would then divide by the extent of the variable, to get a "percentage of difference" between the cluster and the TraumaBase for each covariate.

Cluster	Number of individuals	ATE	Std. err
0	304	0.44	0.22
1	122	0.083	0.074
2	80	0.062	0.075
3	182	0.58	0.39
4	327	-0.20	0.20
...	...	...	...
9	92	-0.127	0.071
...	...	...	...

TABLE 5 – ATE on 15 clusters

$$importance_{var} = \frac{|\text{mean}\,(var|_{cluster}) - \text{mean}\,(var)|}{\max(var) - \min(var)}$$

We found out that in the example above, cluster 9 had 5 variables whose "percentage of difference" with the TraumaBase was at least 15%.

Covariate	GCS.motor .init	Cristalloid .volume	Vasopressor .therapy	GCS	GCS .motor	FiO2
Intra-cluster mean	3.478	1789	0.47	6.80	3.18	0.87
Global mean	4.243	793	0.19	9.11	4.12	0.63

TABLE 6 – Interesting features of cluster 9

Table 6 shows an overview of the difference between the average patient of cluster 9 and the average patient in the TraumaBase overall.

IV
HETEROGENEOUS TREATMENT EFFECT

If we want to have a more granular view on the efficiency of the drug, we cannot rely solely on clusters.

The *Heterogeneous treatment effect* (HTE) is defined as : $\tau(x) = \mathbf{E}(Y(1) - Y(0)|X = x)$ which means that we compute the expectancy of the efficacy of the drug on the subset of patients sharing the same exact features.

We will now introduce a new tool : causal forests [10], which can be used for computing the heterogeneous treatment effect.

IV.1 Causal Forests

Causal forests are a type of random forests used to make predictions not only on the outcome of an experiment but also on a covariate, in our case the treatment W.

The idea behind causal forests is that when training a tree the splits between nodes will be made so that the difference of the average treatment effect between each node will be maximal.

Causal forests are implemented in the **R** package grf (generalized random forests) [9]. These forests can be used to predict ATE and HTE.

We trained our causal forests by using $n = 2000$ trees, and in figure 4 we can see an extract of a tree.

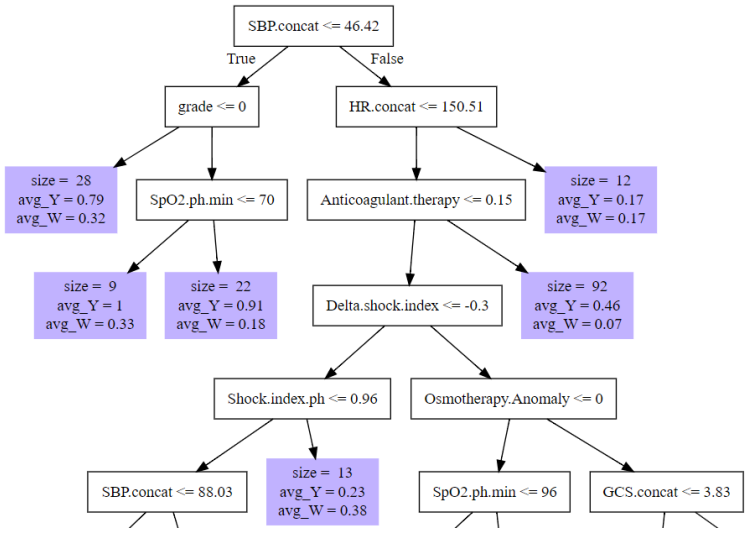


FIGURE 4 – Top nodes of a tree from the random forest

IV.2 COMPUTING THE HTE

To compute the treatment effect on a patient we look at the leaves in which it "falls down". Then we get the mean of predictions for the cases where $W = 1$ and subtract the mean of predictions for the cases where $W = 0$. The result of this difference is the estimated individual treatment effect for this patient.

We can then use our forest in order to plot the histogram of the HTE for each patient, as seen in figure 5.

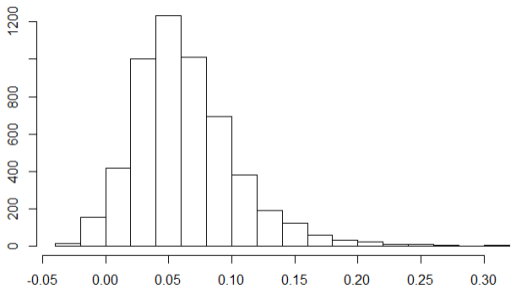


FIGURE 5 – Histogram of the treatment effect

Figure 5 is quite concerning for very few individuals appear to have a negative estimated treatment effet. This result can be explained in two ways :

— The drug is useful only for a particular subset of patients and is harmful for most

— The causal forests fail to capture the heterogeneity of patients and how to drug manages to save the life of patients, they might be "underpowered".

IV.3 Looking for relevant variables

We used two ways to find the most relevant variables :

— by looking at the variables each tree based its splits on

— by looking at clusters for which the covariate distribution differs greatly from the global covariate distribution.

The first method is motivated by the intuition that if many trees split their leaves based on the values of a variable, it is more likely to be important whereas if a variable is ignored by the trees, its effect on the HTE might be negligible.

Variable name	Percentage of splits
Shock.index.ph	10.7%
Cristalloid.volume	9.3%
Delta.shock.index	9.0%
HR.ph	8.2%
Delta.hemoCue	7.7%
ISS	6.6%
IGS.II	6.2%

Table 7 – A measure of the importance of some variables

We can see that 6 variables account for almost 50% of all the decisions made by the tree.

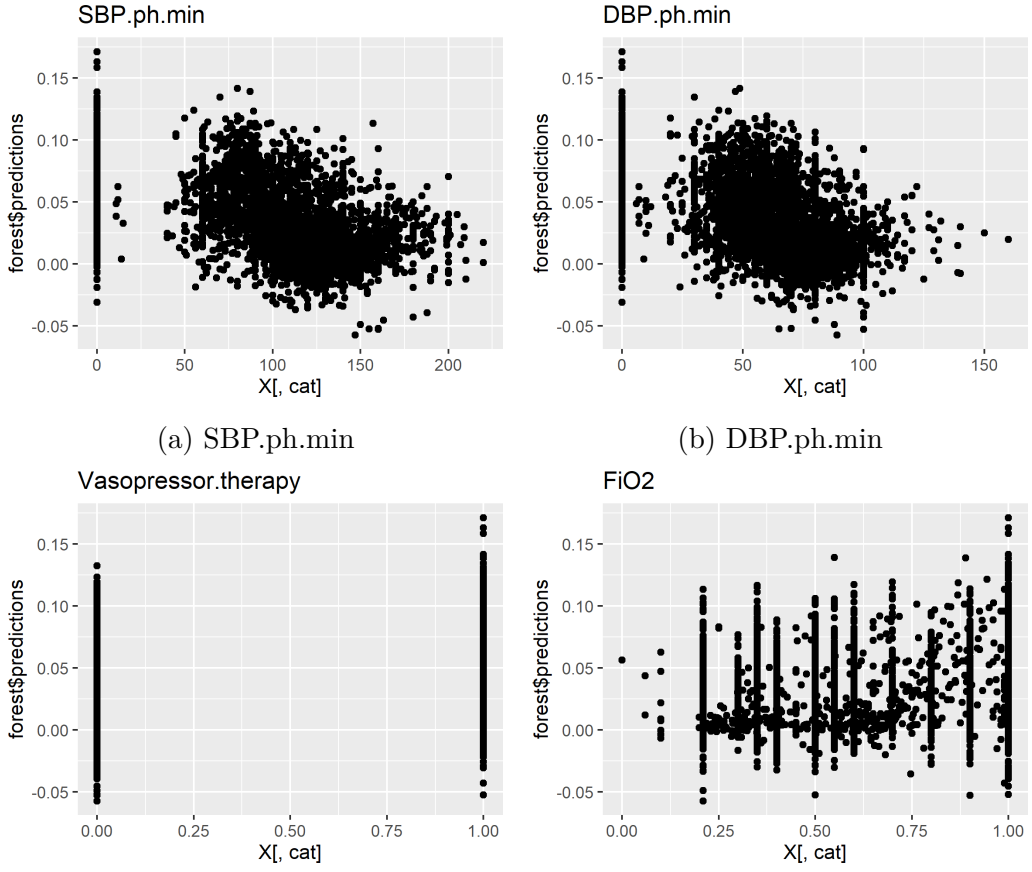
The other method for determining which variables are relevant for the HTE is to see which variables are very different from the average patient, using a similar approach as in III.2 : we compute the difference in means for each variable between the values for individuals with negative HTE and in the overall database, which we would then divide by the extent of the variable, to get a "percentage of difference" between the patients with a negative HTE and the average TraumaBase patient for each covariate.

$$importance_{var} = \frac{|\text{mean}\,(var|_{\text{HTE}<0}) - \text{mean}\,(var)|}{\max(var) - \min(var)}$$

Covariate	SBP.ph.min	DBP.ph.min	Vasopressor.therapy	FiO2
Mean for negative HTE patients	124.0	73.8	0.085	0.572
Global mean	108.4	63.6	0.19	0.631

TABLE 8 – Key features of individuals with a negative HTE

In table 8, we can see the 4 covariates that vary more than 5% from the average patient with a negative HTE to the average patient in the TraumaBase.



(a) SBP.ph.min

(b) DBP.ph.min

(c) Vasopressor.therapy

(d) FiO2

FIGURE 6 – HTE as a function of the important variables

We can then plot the HTE as a function of these variables, as seen in figure 6, in order to observe the effect of each of the important covariates on the HTE. For instance, in figure 6d, we can observe that individuals with smaller values of `FiO2` will be less likely to have a negative HTE, which is why the average `FiO2` is higher in individuals with a negative HTE.

CONCLUSION

Our work has been divided into several successive stages.
First, we selected the relevant patients and covariates of the TraumaBase and cleaned the corresponding data. We managed to reduce the number of covariates from 250 to 39 and patients from 20000 to 5337. We then had to impute a large part of the data, and tested several methods of doing it.

In a second step, we had to choose a way of computing the ATE on our dataset. To do so, we compared several methods of ATE-computing, and tested their stability. The most efficient method turned out to be the Augmented Inverse-Propensity Weighting (AIPW). However, we realized in this step that ATE is a difficult to estimate, and we didn't manage to compute a stable estimator of it. It helped us realizing that there is a gap between theory and practice and that we need to be cautious regarding the results we have.

Once we had chosen our method, we computed the ATE on several clusters from our database in order to find cluster inside of which the treatment had a positive effect. We then tried to characterize these clusters using various features.

Finally, we estimated the HTE on the patients of the TraumaBase and identified relevant variables to find patients with a negative HTE.

A
HOW TO USE OUR CODE

Our code relies on several files using both `R` and `Python`, here is a brief summary of them :

- `create_classes.ipynb` : This Python Notebook is used to add columns to the csv containing the grades and the type of lesions as classified by the doctors
- `hierarchical_clustering.py` : This code is used for hierarchical clustering, it adds columns containing the computed classes to a csv file
- `ATE-finale_version.ipynb` : This file contains our own implementation of the functions and estimators needed to compute the ATE.
- `analysis_causal.Rmd` : This file is used to do some analysis of the database, we also introduce causal forests to compute HTE and to gain more insights on the data
- `impute_data.Rmd` : This file is used to impute missing values using MissForest.
- `analyze_clusters.py` : This file is used to analyze the clusters computed in III.2.
- `causal_forest.Rmd` : This file is used to train a causal forest and plot the computed HTE.
- `Mice.Rmd` : This file is used to impute data using the mice library
- `Simple Imputer + Mice.ipynb` : This file is used to impute data using simple mean imputer. This file also formats mice datasets into a good shape.
- `analyze_causal_forest.py` : This file is used to analyze the profile of patients with a negative HTE in IV.3.

RÉFÉRENCES

[1] UNAFTC. Mieux comprendre le Traumatisme Crânien et les Lésions Cérébrales. *traumacranien.org*, 2019.

[2] CDC. TBI: Get the Facts. *cdc.gov*, 2019.

[3] Haleema Shakur-Still and Jorge Mejía-Mantilla. Effects of tranexamic acid on death, disability, vascular occlusive events and other morbidities in patients with acute traumatic brain injury (CRASH-3): a randomised, placebo-controlled trial. 10 2019.

[4] Groupe TraumaBase. TraumaBase. *traumabase.eu*, 2019.

[5] Daniel Stekhoven and Peter Bühlmann. MissForest - Non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England)*, 28 :112–8, 01 2012.

[6] Stef Buuren and Catharina Groothuis-Oudshoorn. MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45, 12 2011.

[7] Francois Husson, Jérôme Pagès, and Sébastien Lê. *Exploratory Multivariate Analysis by Example Using R*. CRC, 10 2014.

[8] François Husson, Julie Josse, and Sébastien Lê. FactoMineR. *factominer.free.fr*, 2019.

[9] grf labs. Generalized random forests. *GitLab*, 2019.

[10] Susan Athey and Stefan Wager. Estimating Treatment Effects with Causal Forests: An Application, 2019.