Censoring by Non-Response: Infant Single Ventricle (Enalapril vs Placebo Trial)

Malvika Rajeev

12/7/2019

"The NIH/NHLBI Pediatric Heart Network Infant Single Ventricle trial dataset was used in preparation of this work. Data were downloaded from https://www.pediatricheartnetwork.org//pud_login.asp?study_id=ISV on 12/03/2019."

Introduction

My objective was to download a real-world dataset and perform relevant analyses. On the Pedriatic Heart Network website, I came across the NHLBI Infant Single Ventricle (ISV) randomized double-blind placebo-controlled trial. It was conducted by the Pediatric Heart Network (PHN) at 10 centers in the United States and Canada, enrolling patients in 2003-2007. Patients were randomly assigned in a 1:1 ratio to the enalapril and placebo treatment groups using randomly permuted blocks within strata defined by the presence or absence of hypoplastic left heart syndrome, with dynamic balancing within center.

This trial was designed to determine whether ACE inhibition improves ventricular function and hemodynamic status in infants with single ventricle physiology, thereby improving a clinically important outcome, somatic growth. Improving survival and decreasing morbidity in children with single ventricle physiology is one of the most challenging problems in pediatric cardiology.

The PHN screened a total $1,245$ infants who were less than 45 days old with **single-ventricle anatomy** who had undergone surgical palliation, and enrolled 230 infants (115 in each group, 43% consent rate) who were randomized to receive enalapril (treatment level 1) or placebo (treatment level 2). Patients were followed until 14 months of age to allow assessment of the effects of ACE inhibitor therapy for at least 6 months after the volume-unloading that occurs after the Stage 2 superior cavopulmonary connection (SCPC) surgery, which occurs by age 8 months in most infants. A total of 185 infants completed the trial (45 withdrew, 17 of which were deaths or cardiac transplants). A 29% of patients (50/185) discontinued the study drug during the course of the study).

We have discussed the idea of censorship by death in observational studies, and I was interested in working with a dataset that included censorship by death' as well as non response. *Fully conditional models* stratify the longitudinal response (outcomes measured over time) trajectory by time of death. Fully conditional models are effective for describing individual responses, in terms of either time period between start of period to death or from death to end of program.

Causal models (principal stratification) as currently applied are fully conditional models, since group differences at one timepoint are described for a cohort that will survive past a later timepoint.

The primary outcome was weight-for-age z-score at 14 months of age. Secondary endpoints included other measures of somatic growth, ross heart failure class, brain natriuretic peptide (BNP) concentration, ventricular geometry and function obtained by two-dimensional echocardiography, and neurodevelopmental and functional status. The following paper by Hsu et Al[^1][^1]: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3692364/pdf/nihms-480850.pdf covers the impact, in detail, on the primary outcome of body measurement. They focused on the intention-to-treat principle, which makes sense since treatment assigned was treatment received: there were no discrepancies. I am also interested in the CAGE or complier average causal effect because of the high proportion (50/185, 29%) of patients who discontinued study drug during the course of the study - because of non-survival and other logistical reasons.

I concentrate on the somatic growth outcomes through body weight measurements which suffers one of the biggest impacts of single ventricle in infants. I create the variable BMI which I consider the of final outcome, along with a binary outcome for continued existence of risk of further health deterioration.

Missing Data vs Death By Censorship

Missing data can be very problematic for interpreting longitudinal data. Some solutions have been introduced, (Little and Rubin, 1987). But it is seen as a purely omitted variable: its absence itself is never assumed to mean much. This would make sense if during data collection, there was some human error and some values were simply not recorded. This can be viewed as a 'missing at random' assumption. For instance, in this study, there are 6 follow ups. If patients have only a few visits missing, then methods like Monte Carlo Markov Chains can be used to impute/estimate the missing follow up values.

In the case of death specially in medical studies, the missing values viz a viz absence of potential outcome can be viewed as a outcome of treatment itself - so it is not just a matter of omission. When we simply condition on survival status (like a covariate), we are treating the model as unconditional, that is, the outcome is not dependent on survival time. In this case death does not imply missing data - which is like saying that we compare the survivors in treatment and control. But as we have seen, this is a fundamentally flawed approach, because we are no longer comparing the 'same' set of units. In this case, we are 'estimating' data beyond death.

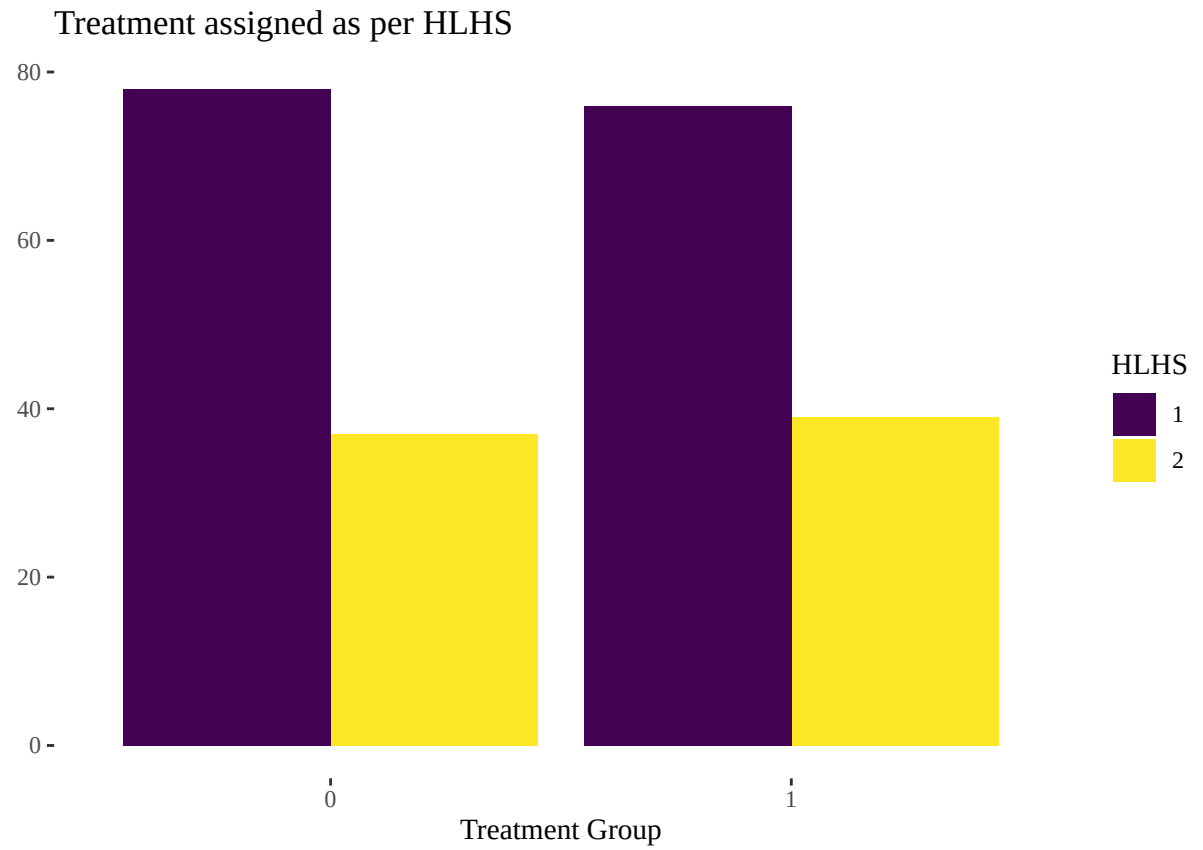
Principal Stratification, on the other hand, is a fully conditional model. We describe potential outcomes for units regardless of treatment effect.

Kurland et al provide a simulated example of a partly conditional model, by estimating the marginal probability of survival for different time independently from data.

Why stratified sampling?

Broadly speaking, the presence or absence of left heart syndrome gives rise to two **distinct** symptoms of having a single ventricle. We can think of them as causing hypo and hyper blood pressure. They have different symptoms but their treatment is still the same.

Thus, we can consider '*HLHS*' as an important covariate upon which to stratify on for our analysis.

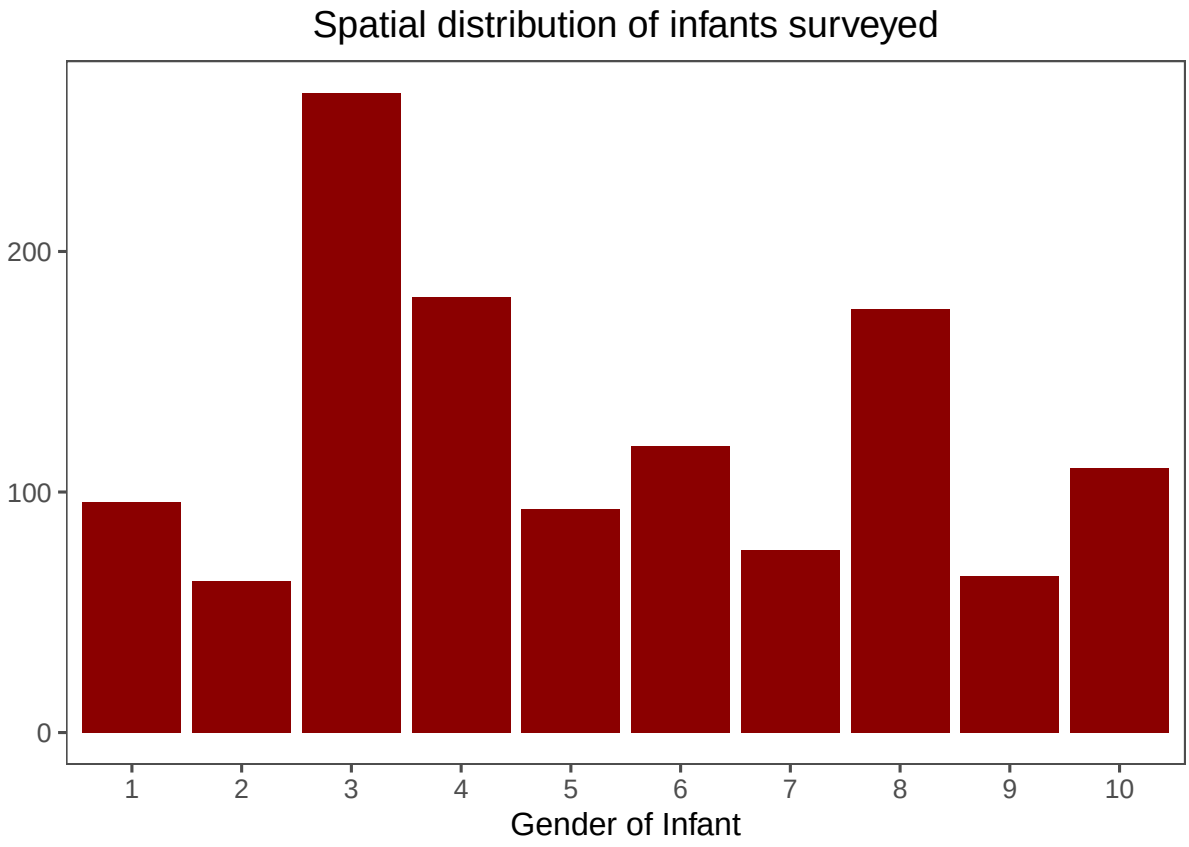


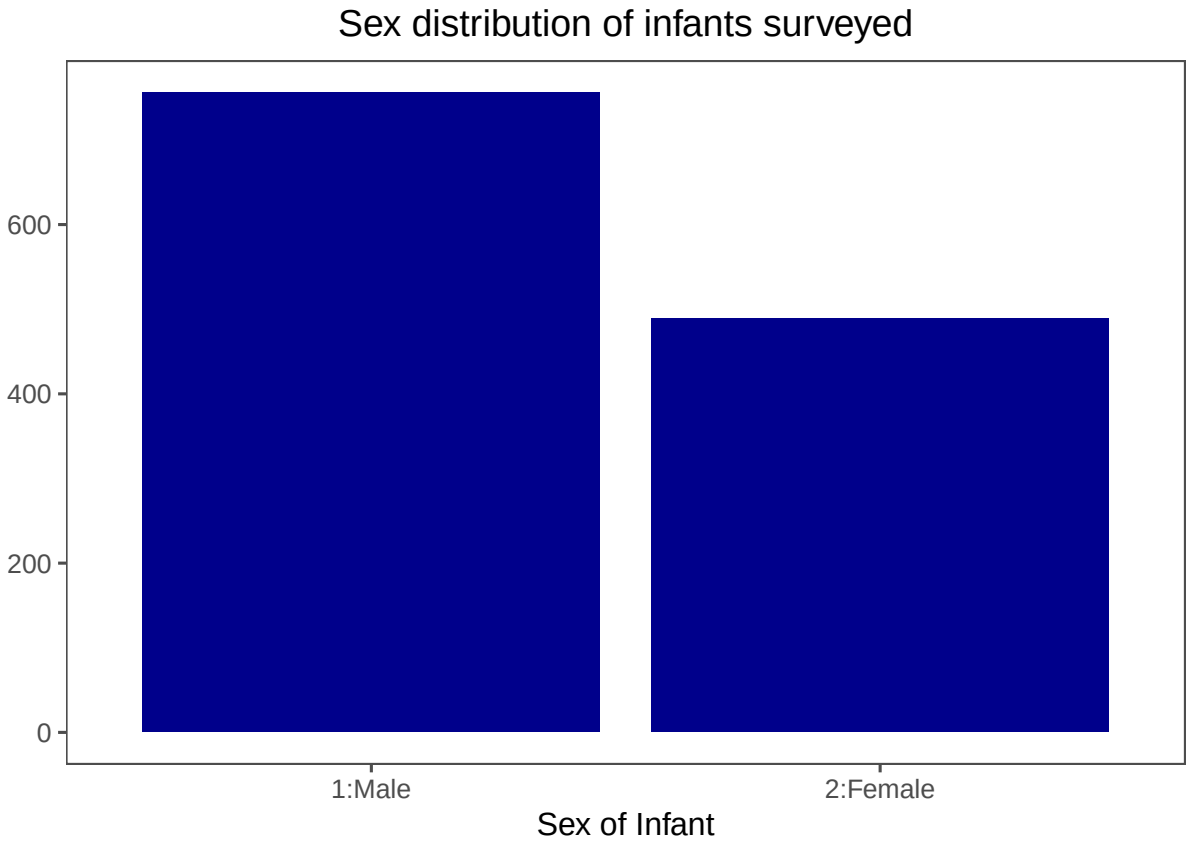
Stratified random sampling designs divide the population into homogeneous strata, and the relevant number of participants are chosen at random from each strata. .This method can be used to improve the sample's representativeness of the population, by ensuring that characteristics (and their proportions) of the study sample reflect the characteristics of the population.

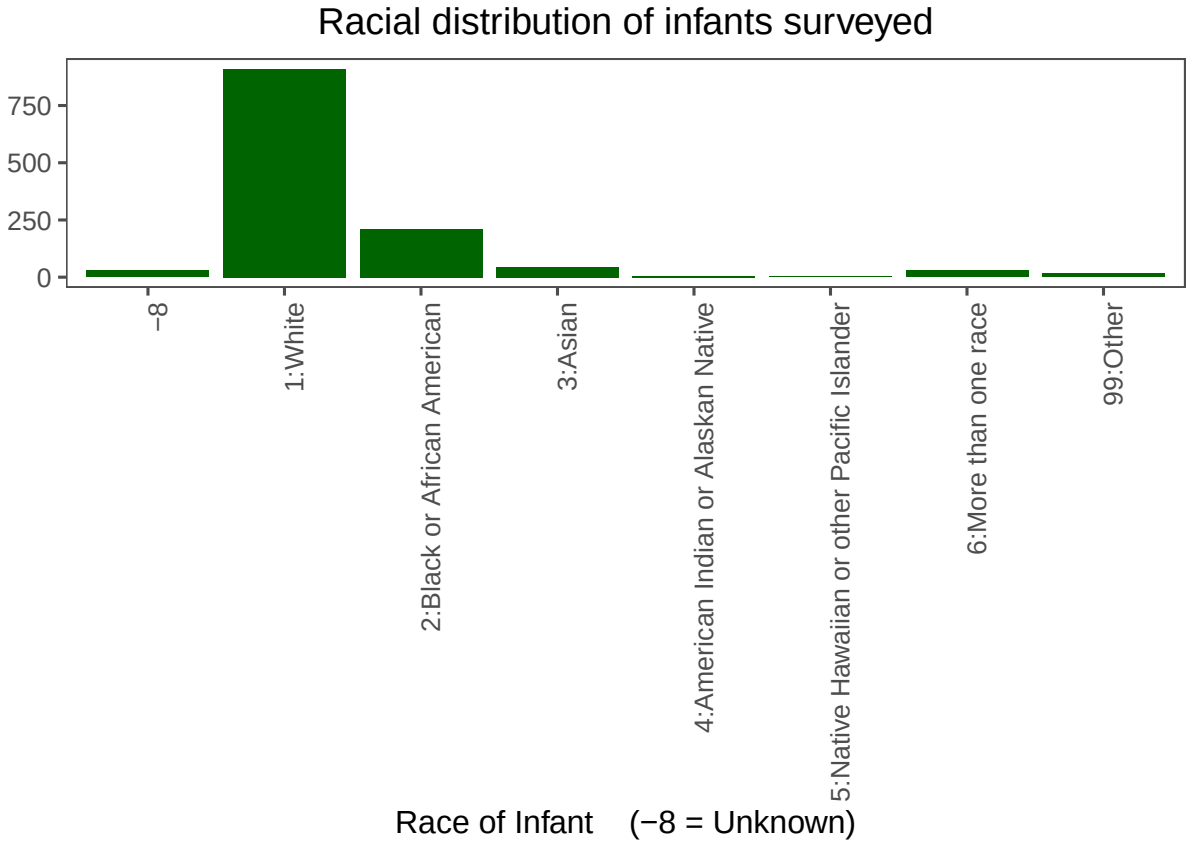
Clearly, this study assumes that the presence of HLHS is more prevalent in single ventricle infants.

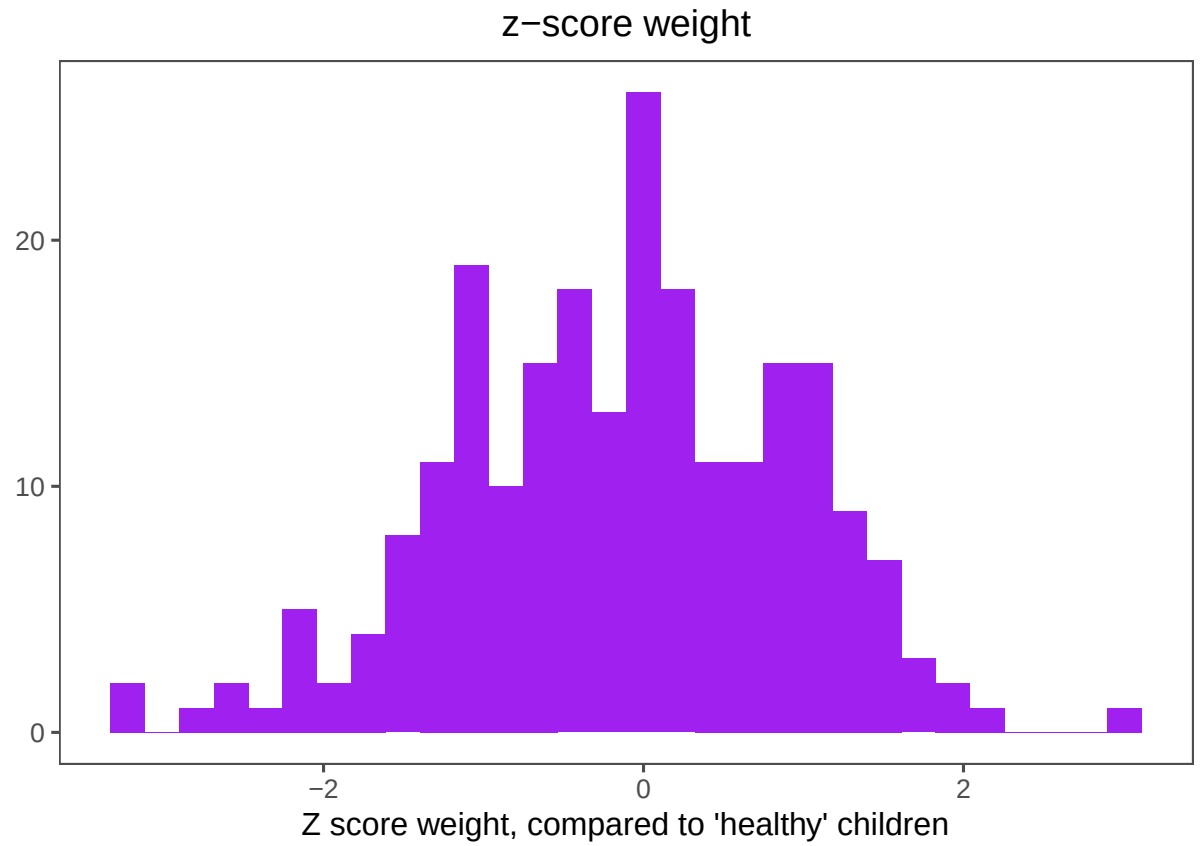
Initial EDA

I begin my report with first exploring the patients who were surveyed and then contrast them with the patients who were finally enrolled into the different treatments.


Spatial distribution of infants surveyed
200
100
0
1
2
3
4
5
6
7
8
9
10
Gender of Infant


Sex distribution of infants surveyed
600
400
200
0
1:Male
2:Female
Sex of Infant


Racial distribution of infants surveyed
750
500
250
0
−8
1:White
2:Black or African American
3:Asian
4:American Indian or Alaskan Native
5:Native Hawaiian or other Pacific Islander
6:More than one race
99:Other
Race of Infant (−8 = Unknown)



Conditional Density of Age of Eligible Infants given their acceptance into program

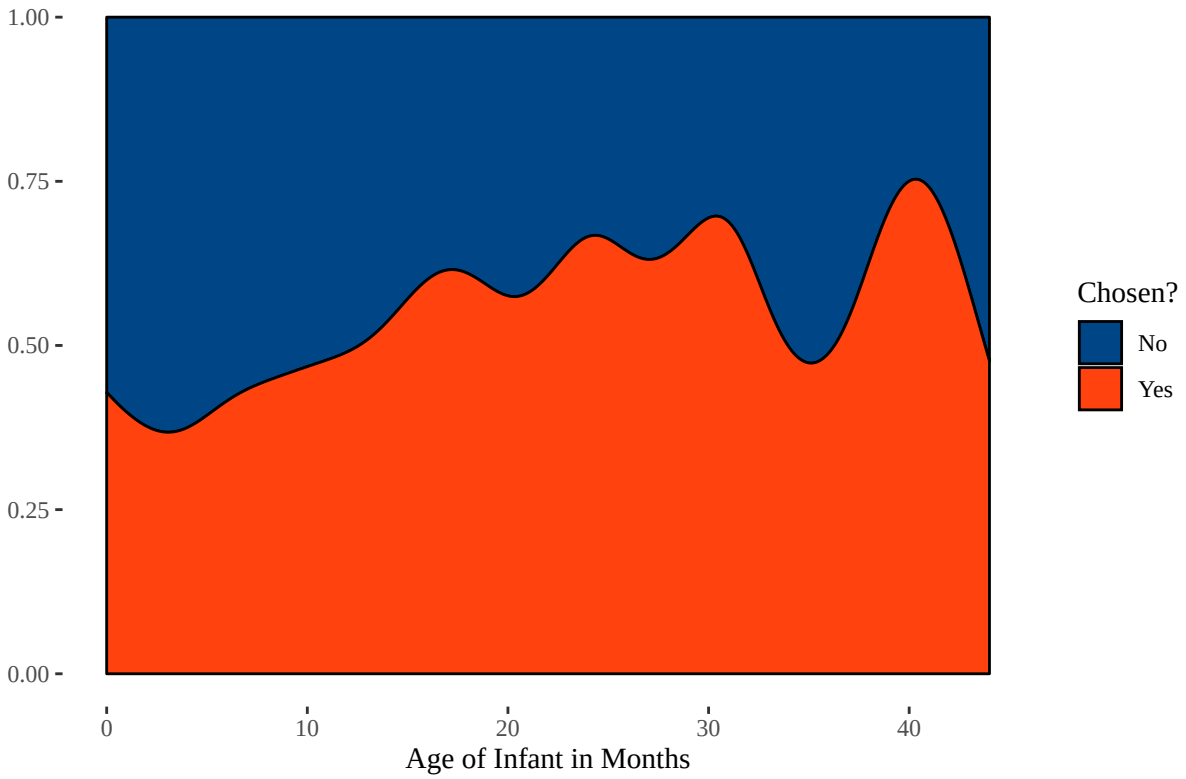
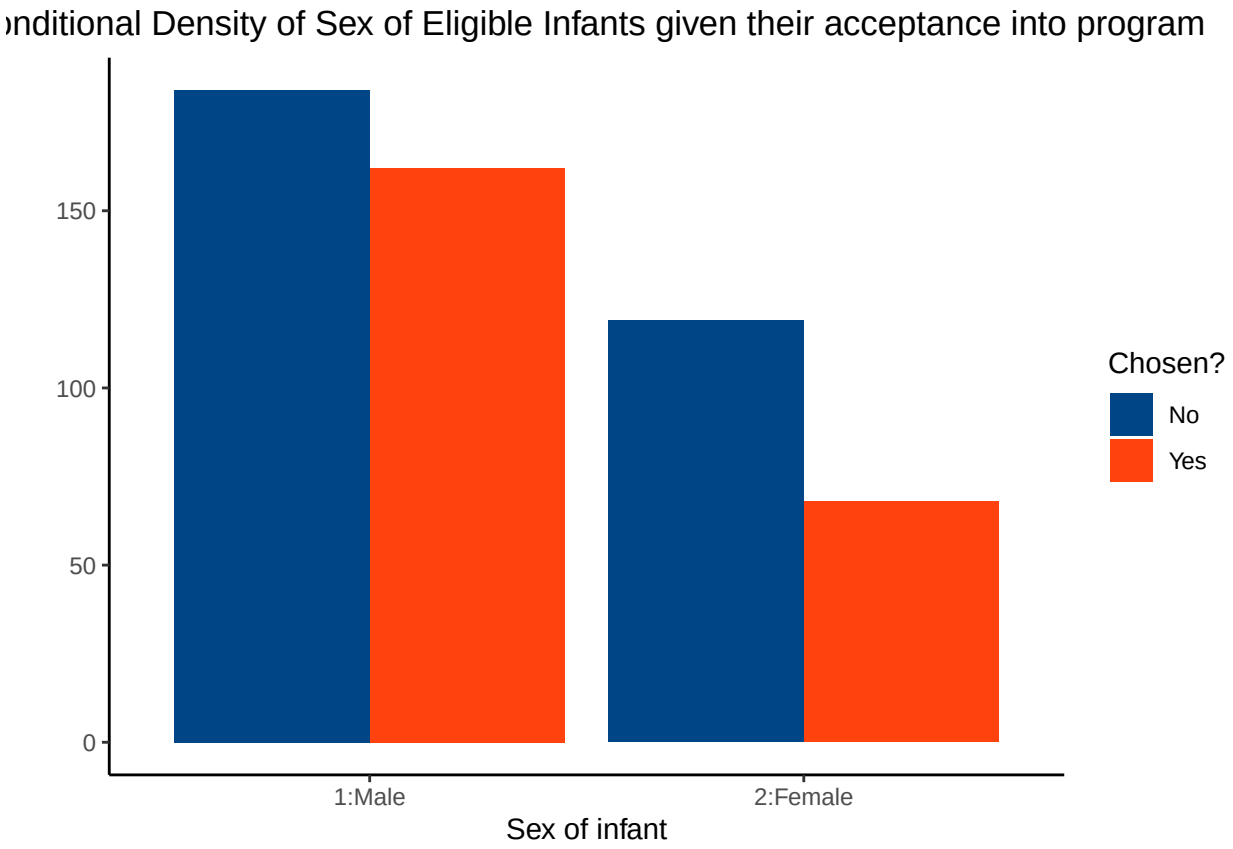


Table 1: Results of t-test of rate of enrollments across different sexes

estimate	estimate1	estimate2	statistic	p.value	parameter	conf.low	conf.high	method
-0.1045717	0.3636364	0.4682081	-2.358535	0.009418	393.0747	-Inf	-0.0314706	Welch Two Sample t-te

Table 2: Results of t-test of rate of enrollments across age of infants

estimate	estimate1	estimate2	statistic	p.value	parameter	conf.low	conf.high	method
2.556507	13.83043	11.27393	3.304889	0.0005114	469.9787	1.281613	Inf	Welch Two Sample t-test




```
## [1] 3.607898e-05
```

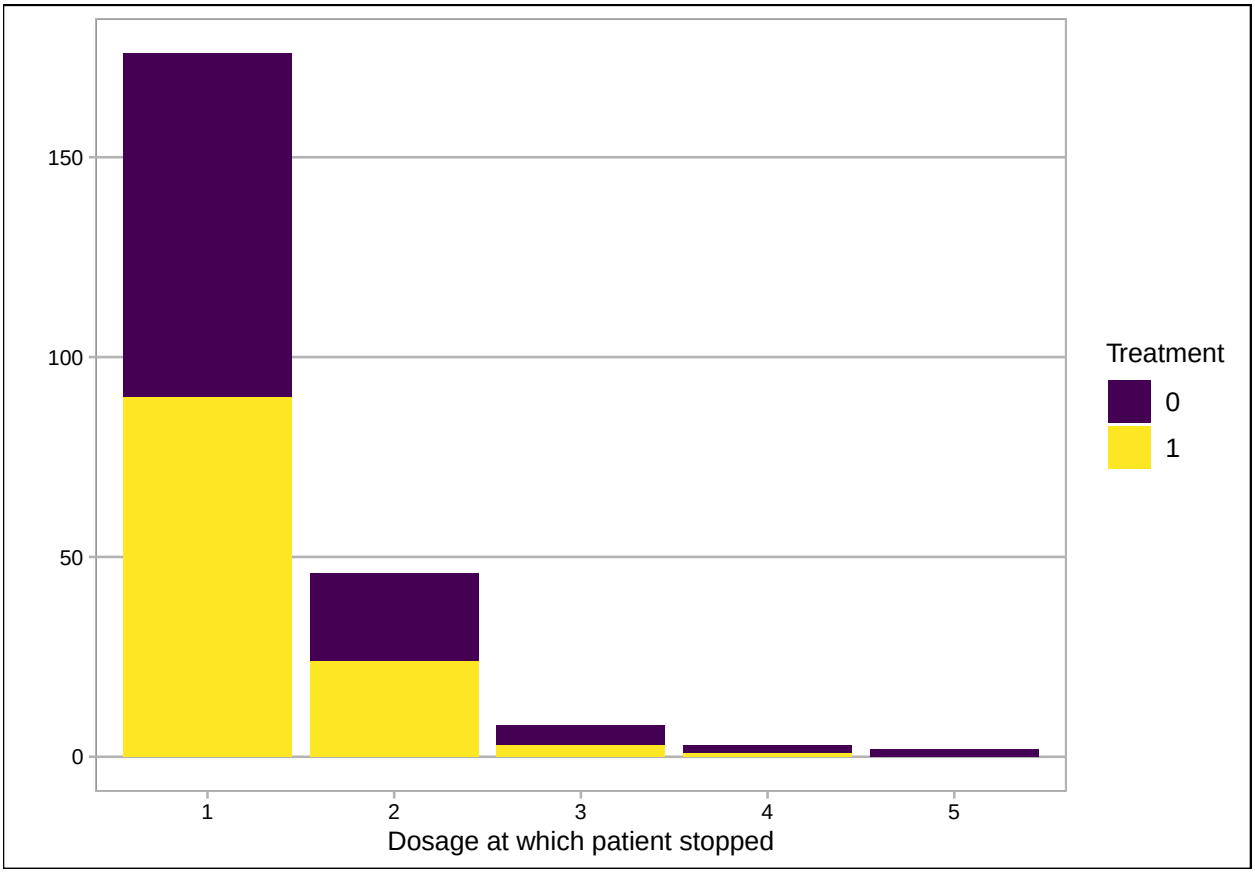
- Out of the infants surveyed and then enrolled, more female patients were *eligible but not enrolled* than male patients in proportion. A one sided difference-in-means test gave a statistical significant result. (36% for female vs 46% for males, p.val = 0.009)
- Out of the infants surveyed and then enrolled, older infants tended to be *eligible but not enrolled*. A one sided difference-in-means test gave a statistical significant result. (p.val = 0.005)

Covariates

Most of the covariates in this study: Gender, Race, Presence of HLHS, Hospital site number are categorical. To deal with this, we create dummy variables for each non-binary categorical variable. To make this more efficient in R, I use the package `fastDummies`.

Longitudinal Dosage

Now, a problem arises because even though all patients were required to have 6 dosages, few arbitrarily stopped, and some didn't take any at all. **(missing data)**



term	estimate	std.error	statistic	p.value
(Intercept)	1.3931624	0.0634314	21.963300	0.0000000
treatment1	-0.1135014	0.0895153	-1.267956	0.2060794

It is evident that the dosage at which they stop is uniform among treatment groups, i.e, treatment doesn't affect the dosage at which they stop, but it does effect the final outcome. So it cannot be considered a post-treatment variable.

Heart Transplants

Seven infants received heart transplants.
3 of them were under treatment and 4 of them were under placebos. In further analysis, I will discard them.

Death

46 people 'officially' withdrew from the study. I will only register the infants who didn't survive. The rest, for example, whose families withdrew because of social or logistic reasons, I eliminate from my analysis. It is essentially missing data. This is because there is a distinction between withdrawing from a study and death

during the study.

Dying during a medical study is more strongly related to the treatment level than 'truncation by non-response' as discussed earlier.

BNP at Visit 6

B-type natriuretic peptide (BNP) is a hormone produced by the heart. BNP is released in response to changes in pressure inside the heart. These changes can be related to heart failure and other cardiac problems. Levels goes up when heart failure develops or gets worse, and levels goes down when heart failure is stable. BNP level is higher in single ventricle infants than normal infants.

- We assume monotonicity. Treatment does not put infants at higher risk of death than placebo.

- We work with the concept of principal scores (Ding et al; 2016). Through our covariates, we fit a logistic regression model for treatment and covariates separately, lets say $p_1(X)$ and $p_0(X)$.

Under the assumption of Monotonicity, always survivors, healthy infants (survive with treatment and don't survive with control), and always martyrs (don't survive regardless) exist.

So,

$$p_0(X) = e_{ss}(X)$$

and

$$1 - p_1(X) = e_{\bar{s}\bar{s}}(X)$$

and

$$e_{s\bar{s}} = p_1(X) - p_0(X)$$

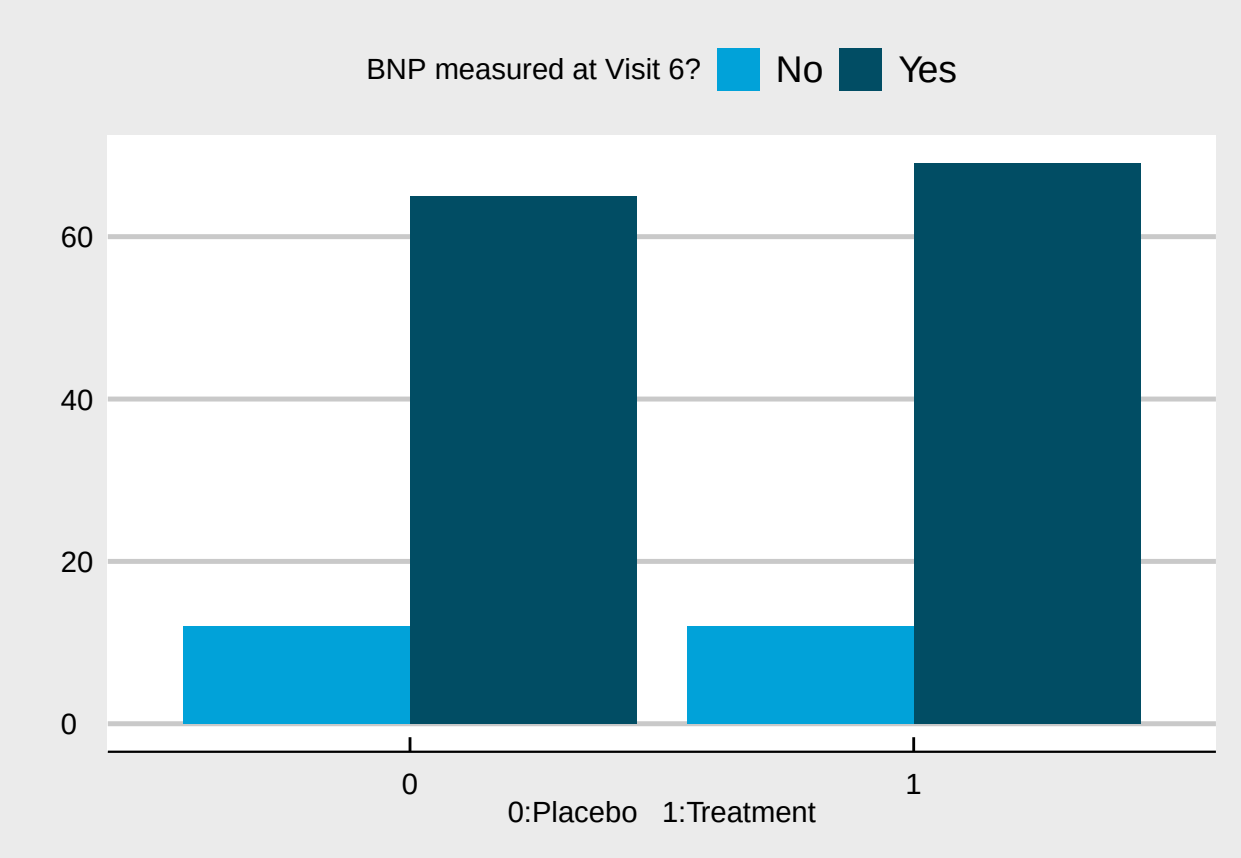
And we can empirically estimate p1 and p0, which are the probabilities of surviving under treatment and control respectively by the sample mean for each treatment group.

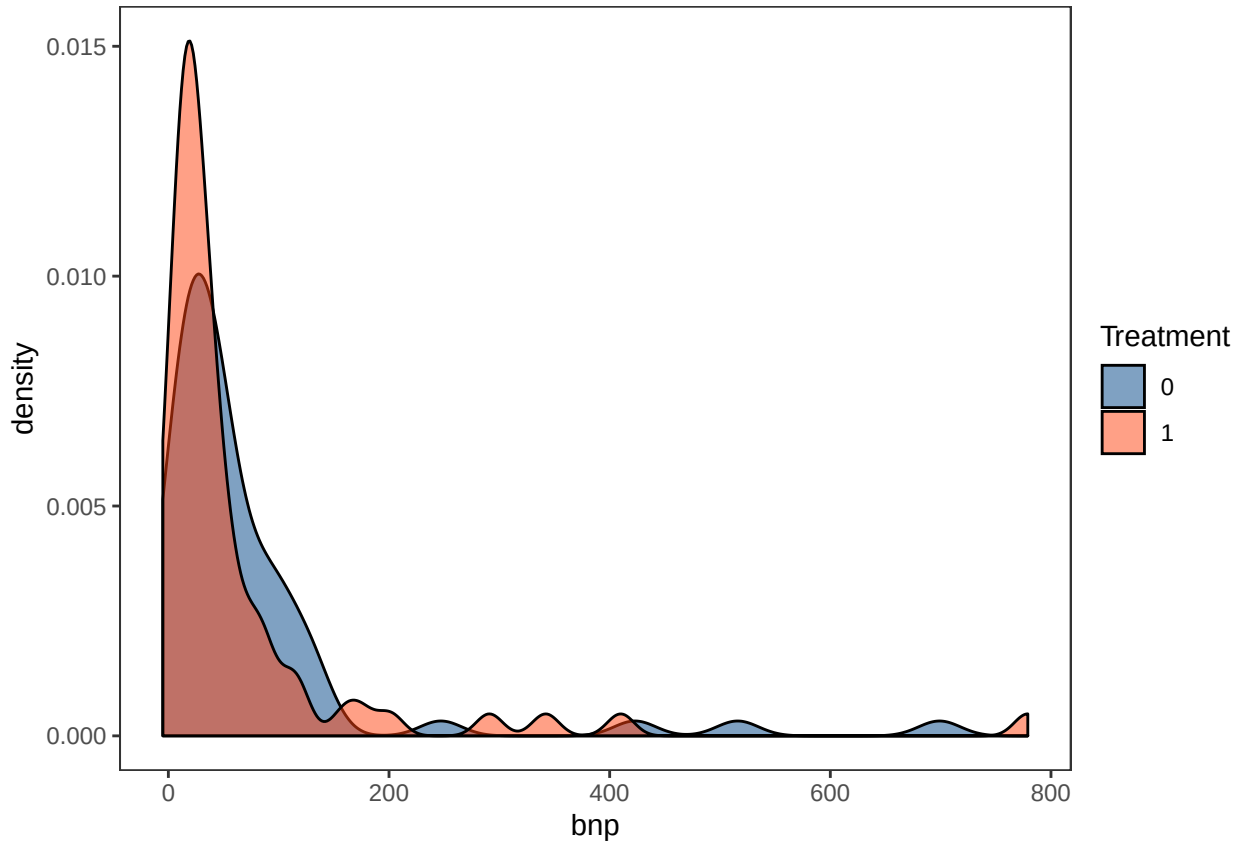
- We can then use the *General Ignorability Principle* which assumes independence between outcome and principal Strata given covariates.

-The intuition is as follows: To estimate the mean of BNP under assignment to enalaprin in those who would survive t under assignment to either enalaprin or placebo, we only want to include those subjects who we think are quite likely to have survived had they been randomized to placebo. Thus subjects with high values of fitted values under control are given a large weight, while those we predict to have a lower chance of surviving under randomization to placebo are given a lower weight.

-To estimate the mean of BNP under assignment to placebo, we do the same, but now weight using of probabilities computed under treatment group.

-I use bootstrapping to get the standard error.





According to bootstrapping, the SD is around 5.9111526. The estimate is around -13.75 which indicates that the treatment reduces BNP which is a positive trend and the results are significant.

When **I didn't account for death**, and simply dropped the deceased from the analysis hence considering survivors in each group, I got very contradictory results:

-According to the Neyman potential outcome framework, there is no statistically significant causal effect on BNP, 95% confidence interval includes zero.

-Additionally, the Fisher exact P.Value is also insignificant.

High Risk Binary Outcome

One of the binary outcomes I consider are the evidence of high risk in infants. But, there are conducted in the **stages after the surgery (mid-program)** and thus do not account for the infants who died, or whose parents withdrew them from the treatment, may be because of lack of improvement. Approximately 19 infants died in this stage. This is SACE for binary outcome.

Table 3: Results of Principal Stratification

Upper.Bound	Lower.Bound	Quantity
0.539	-0.673	Potential Outcomes
0.299	-0.912	SACE

To estimate the *Survival Average Causal Effect*, I use the tight bounds:

- **Principal Stratification**: Assume Monotonicity and exclusive restriction to get upper and lower bounds on SACE.

From table 3, we can say that infants under treatment and placebo are under the same high risk.

BODY MEASUREMENTS: Final Outcome: Body Mass Index

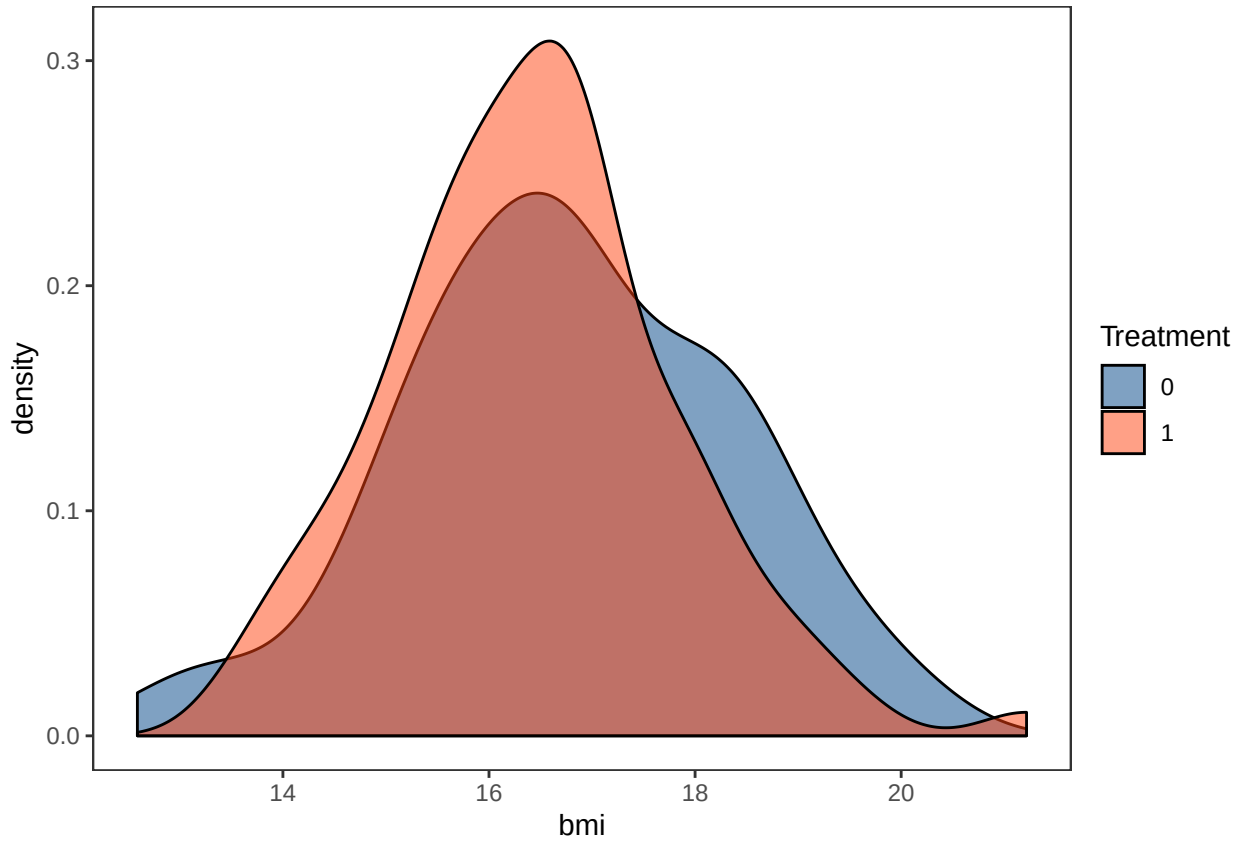
For the final outcome I assess the body mass index as on Visit 6, that is, 14 months.

The outcome is continuous so it will not be straightforward to get bounds on survival average causal effect like in the previous section - but using principal scores, we can get get an estimate and then bootstrap the process to get a standard error.

I make the following assumptions:

- Taking enalapril doesn't increase infant's health risk than placebo.

- Treatment does not effect the infant's health in any other way than the actual drug administered.



According to bootstrapping, the SD is around 0. The estimate is around 0 which indicates that the effect is not significant.

Conclusion

Using principal scores, I observed that SACE for BNP is favoring treatment and is significant. Same cannot be said for risk index or BMI.

References:

1. Dealing with the difference between truncation by non response and by death: https://projecteuclid.org/download/pdfview_1/euclid.ss/1263478382
2. A Simple Method for Principal Strata Effects When the Outcome Has Been Truncated Due to Death Yasutaka Chiba* and Tyler J. VanderWeele
3. An Estimator for Treatment Comparisons among Survivors in Randomized Trials - Douglas Hayden, Donna K. Pauler, and David Schoenfeld
4. Ding, Peng & Lu, Jiannan. (2017). Principal stratification analysis using principal scores. Journal of The Royal Statistical Society, Series B. 79. 757-777. 10.1111/rssb.12191.