The Effect of Predicted Hemoglobin Trajectories on Anemia Outcomes Analysis

Shulei Yang
Statistical Consulting 7995 ,Dr. Karen Kafadar
University of Virginia
Mar 2, 2020

1 Executive Summary

This report consists of five parts and ultimately tries to evaluate the performance of the intervention. In the Introduction, I will briefly discuss why this question is important and specifically what we should do to answer this question. Then in Method, I will show all the technics that I used in my analysis. After that in the Results, I will present the results from my analysis and carefully show how I get to my conclusions. Also, there are some limitations in my analysis, and I will discuss them in the following Limitation section. Finally, I will conclude my whole analysis in the Conclusion section.

2 Introduction

End-stage renal failure is the final, permanent stage of chronic kidney disease. Despite the improvements in dialysis care, the mortality of patients with ESRD remains high. Patients with ESRDs typically suffer from ESRD-induced anemia, and the current paper protocol method has a lot of problems for making dosing decisions. Therefore, this project aims to build a dosing recommendation algorithm that can be used to provide the recommended dosing decision for anemia nurse manager. After the execution of this intervention for a while, we want to evaluate the performance of this intervention. There are two main aspects we want to evaluate, one is that whether this intervention improved patient outcomes and the second is whether this intervention decreased ESA usage. This is report ultimately tries to answer two questions: firstly, did the intervention increase patients' 'average percentage of time in range'. And secondly, did the intervention decrease the average weekly ESA dose patients received.

3 Method Overview

1. Paired t-test
The paired sample t-test is a statistical procedure used to determine whether the mean difference between two sets of observations is zero. For example, in our project, we are interested in evaluating the effectiveness of the intervention on decreasing ESA usage. Then we can consider to measure the average weekly ESA dose of our patients before and after the intervention, and analyze the differences using a paired sample t-test. In the two-tail paired-sample t-test, the null hypothesis will assume that the true mean difference is equal to zero. And the alternative hypothesis will assume that the true mean does not equal to zero. After performing the two-tailed paired sample t-test we can check the corresponding p-value. Because we are only interested in one direction (did the intervention decrease ESA usage), not both directions (did the intervention increase or decrease ESA usage). We need to divide the p-value we got from the two-tailed paired sample t-test to find the corresponding 'one-tail' p-value. If the 'one-tailed' p-value is significant (smaller than 0.05), then we can reject the null hypothesis and conclude that the intervention does decrease the ESA usage.
However, before performing the paired sample t-test, there are some assumptions we need to check.

(a) Normality assumption
Because our sample size is only 25, which is very small, we need to ensure that our data is approximately normally distributed. To test whether the differences are normally distributed, we can perform a Shapiro-Wilk's method. If the resulted p-value is larger than 0.05, then the distribution of the differences is not significantly different from a normal distribution. In other words, we can assume that the normality assumption holds. However, if the differences are not normally distributed, we cannot perform the paired sample t-test. Instead, we can consider a paired samples Wilcoxon test.

(b) No outliers assumption.
Outliers can bias the results and potentially lead to incorrect conclusions if not handled properly. Because we only have 25 patients, we can easily identify any potential outliers using a boxplot. If we do have outliers, there are two possible methods to deal with them. One method is to simply remove the outliers. However, some outliers are also influential points, it means that remove the outliers will have a lot of influence on the result. If this is the case, then we can consider using a nonparametric test such as the Wilcoxon Test instead of the paired sample t-test.

4 Results

1. Did the intervention improved patient outcomes?
In order to see whether the intervention improved patient outcomes, we can check if there exist a significant change in patients' 'percentage time in range' before and after the intervention. The following plot (Figure 1) shows the prior and post average 'percentage of time in range' for each patient. From this plot we can see that there isn't a great change in average percentage of time in range before and after the intervention.

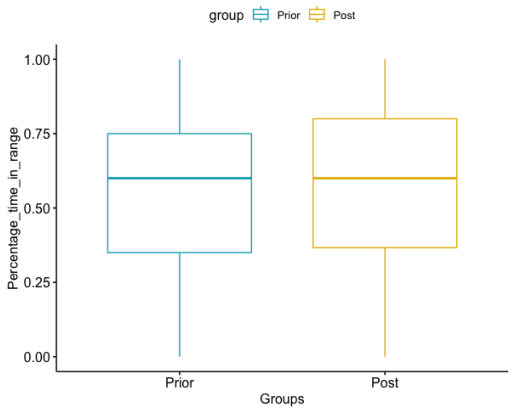

Figure 1

And from the summary statistics (Figure 2)below we find that also there is slightly decrease in the mean of 'average percentage time in range', the standard deviation becomes larger.

	group	count	mean	sd
	<fct>	<int>	<dbl>	<dbl>
1	Post	23	0.560	0.315
2	Prior	23	0.572	0.294

Figure 2

In order to further evaluate whether the intervention improve patients' outcomes, we cannot reach to any conclusion base on these week evidences. We need to perform a formal test on our hypothesis. We can first find out the 'average percentage of time in range' before and after the intervention for each patient and then perform a paired sample t-test on these data. But firstly, we need to check the assumptions of paired sample t-test.

(a) Check for outliers.
From the boxplot (Figure 3) we can see that there is no outlier in our data.

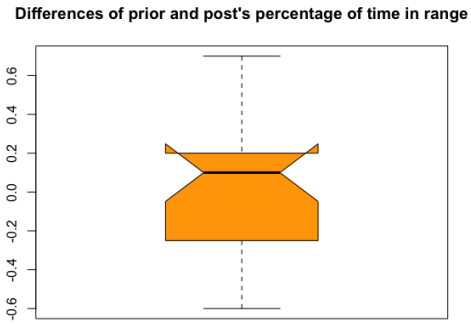

Figure 3

(b) Check for normality assumption.
After performing the Shapiro-Wilk normality test, we found that the p-value is 0.8604. This p-value is larger than 0.05, which indicates that our data is not significantly different from normal distribution. Therefore, the normal assumption also holds. (Result shown below in Figure 4)

```
        Shapiro-Wilk normality test

data:  range$diff
W = 0.97754, p-value = 0.8604
```

Figure 4

(c) Perform paired sample t-test.
After checking these assumptions, we can perform the paired sample t-test. The null hypothesis assumes that the true mean difference is equal to zero, and the alternative hypothesis assumes that the true mean difference does not equal to zero. The result is shown below in Figure 5.

```
        Paired t-test

data:  percentage_time_in_range by group
t = -0.16481, df = 22, p-value = 0.8706
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1510128  0.1287780
sample estimates:
mean of the differences
            -0.01111739
```

Figure 5

Therefore, the corresponding p-value for a one-tail test will be 0.8706/2=0.4353. 0.04353 is larger than 0.05. In conclusion, there is no evidence supports that the intervention increases patients' 'percentage of time in range'

2. Did the intervention decrease ESA usage?
In order to check whether the intervention decrease ESA usage, we can try to find whether the average weekly ESA dose decreased after the intervention. The following plot (Figure 6)show the average weekly ESA dose for each patient before and after the intervention. From the plot we can see that there is a decrease in weekly ESA dose. And the spread of weekly ESA dose became narrower.

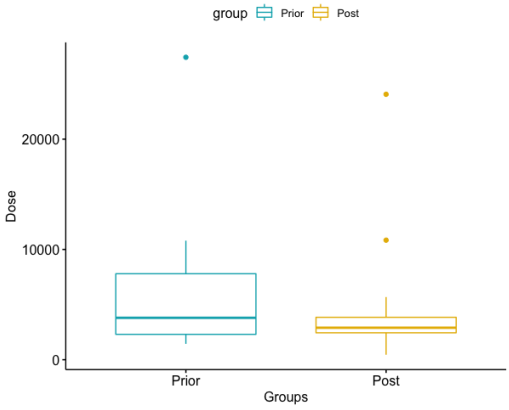

Figure 6

The following plot(Figure 7) lines up the two observations (weekly ESA dose before and after the intervention) for each patient. From these plot we can see for most of the patients, there seems to be a decrease in weekly ESA dose after the intervention.

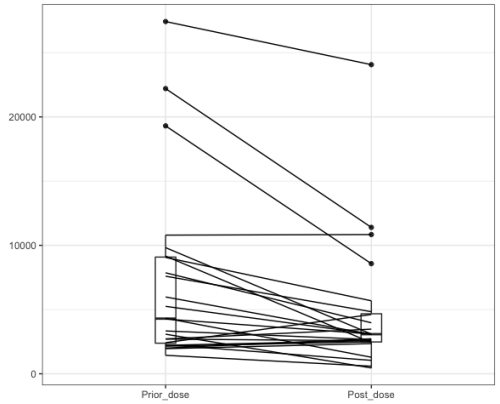

Figure 7

The following statistical summaries (Figure 8 and Figure 9) also demonstrate this phenomenon. Weekly dose's mean, median, standard deviation and interquartile range (IQR) all decrease.

	group	count	mean	sd
	<fct>	<int>	<dbl>	<dbl>
1	Post	24	4626.	5015.
2	Prior	24	7056.	6876.

Figure 8

	group	count	median	IQR
	<fct>	<int>	<dbl>	<dbl>
1	Post	24	3083.	2193.
2	Prior	24	4295.	6705.

Figure 9

Despite these evidences, to fully support our hypothesis, we still need to perform a formal test. To achieve this, we can first find the average weekly ESA dose before and after the intervention for each patient and then perform a paired t-test. Still, we need to check the assumptions first.

(a) Check for outliers.
From the boxplot (Figure 10), we can clearly see two outliers. They are patient 7 and patient 13. Patient 7's weekly ESA dose decreased from 22,204 to 11,397. Patient 13's weekly ESA dose decreased from 19,303 to 8,573. Further investigation needs to make to find out the reason behind these dramatic changes. In my analysis, I deleted the data of these two patients.

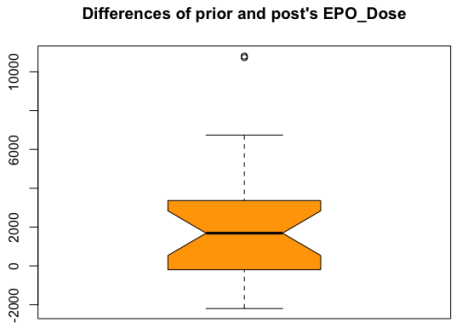

Figure 10

(b) Check for normality assumption.
After performing the Shapiro-Wilk normality test, we found that the p-value is 0.1673. This p-value is larger than 0.05, which indicates that our data is not significantly different from normal distribution. Therefore, the normal assumption also holds. (Result shown in Figure 11)

```
        Shapiro-Wilk normality test

data:  range$diff
W = 0.97754, p-value = 0.8604
```

Figure 11

(c) Perform paired sample t-test.
After checking these two assumptions, we ca perform the paired sample t-test. The null hypothesis assumes that the true mean difference is equal to zero, and the alternative hypothesis assumes that the true mean difference does not equal to zero. The result is shown below in Figure 12.

```
        Paired t-test

data:  dose by group
t = -3.3606, df = 21, p-value = 0.002959
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2706.4004  -637.2635
sample estimates:
mean of the differences
              -1671.832
```

Figure 12

Therefore, the corresponding p-value for a one-tail test will be 0.002959/2=0.0014795. 0.0014795 is smaller than 0.05. Then, we can conclude that the intervention did decrease the ESA usage.

5 Limitations

We should notice that there are some limitations about my analysis. Firstly, we do not know whether the anemia nurse manager made

the dosing decisions based on our recommendation or not. The anemia nurse manager could still make dosing decisions very different from our recommendations. In the future, we can record the recommended dosing usage along with the actual dosing usage the patients got. And performing analysis only on those patients who are receiving recommended dosing usage.
Secondly, there are many potential confounding variables embedded in our analysis. For example, we do not know whether the patients are taking their medicine appropriately. Or is it possible that some of the patients are receiving treatments like dialysis for their kidney disease during this time period? We need to collect more data on these potential problems and carefully control their influences.

6 Conclusion

In conclusion, we do have enough evidence to show there is a decrease in average weekly ESA dose after the intervention. However, because of all the limitations I mentioned above, we cannot fully prove that this decrease is attributed to the intervention. We need more data on those potential confounding variables to perform some robustness tests. Also, based on current data, there is no evidence suggests that the intervention increase the 'percentage of time in range' for patients.