Section I. Introduction

Study Description:

A study enrolling 24 patients diagnosed with keratistis sicca(KCS) was conducted to determine the effect of a new ocular wetting agent, Oker-Rinse, for improving ocular dryness based on Rose-Bengal staining scores. Patients were instructed to use Oker-Rinse in eye and Hypotears, as a control, in the other eye 4 times a day for 3 weeks. Rose-Bengal staining score were measured as 0,1,2,3 or 4 in each of four areas of the eye: cornea, limbus, lateral conjunctiva, and medial conjunctiva. Higher scores represent a great number of devitalized cells, a conditional associated with KCS. The overall scores for each eye (see table) were obtained by adding the scores over the four areas.

Our research question is :is there evidence of a difference between the two test preparations in the overall Rose-Bengal staining scores?In this study, our variable is the difference of overall Rose-Bengal staining scores between patients who were instructed to use Oker-Rinse in eye, and patients who were instructed to use Oker-Rinse in the other eye.

Our data for the study is following:

Patient Number	Hypotears	OkerRinse
1	15	8
2	10	3
3	6	7
4	5	13
5	10	2
6	15	12
7	7	14
8	5	8
9	8	13
10	12	3
11	4	9
12	13	3
13	8	10
14	10	2
15	11	4
16	13	7
17	6	1
18	6	11
19	9	3
20	5	5
21	10	2
22	9	8
23	11	5
24	8	8

We create the Hypothesis:

H0: $\mu=0$

Ha: $\mu\neq0$

where μ means the difference of overall Rose-Bengal staining score between patients who were instructed to use Oker-Rinse in eye, and patients who were instructed to use Oker-Rinse in the other eye.

Section II.

1. **Summary statistics**

   We use four method to test the null hypothesis is true or not, that are T-test, paried t-test, sign-test, and Wilcoxon signed Rank test.

2. **Frequency table**

From the following table, we know that count is a new variable that means the difference between the between patients scores who were instructed to use Oker-Rinse , and patients who were instructed to use Oker-Rinse. X is 'YES' if count greater than 0, is 'NO' if count is less than 0, is to be deleted from table if count is 0.

Obs	patientnum	hypotears	Okerrinse	count	X
1	1	15	8	7	Yes
2	2	10	3	7	Yes
3	3	6	7	-1	No
4	4	5	13	-8	No
5	5	10	2	8	Yes
6	6	15	12	3	Yes
7	7	7	14	-7	No
8	8	5	8	-3	No
9	9	8	13	-5	No
10	10	12	3	9	Yes
11	11	4	9	-5	No
12	12	13	3	10	Yes
13	13	8	10	-2	No
14	14	10	2	8	Yes
15	15	11	4	7	Yes
16	16	13	7	6	Yes
17	17	6	1	5	Yes
18	18	6	11	-5	No
19	19	9	3	6	Yes
20	21	10	2	8	Yes
21	22	9	8	1	Yes
22	23	11	5	6	Yes

4. The characteristics of the graph

In T-test, according to the graph in the part "t-test to the test the null hypothesis the difference in overall Rose-Bengal staining is 0", it is not a symmetric graph because it is right skew. In the boxplot, the median is on right side far from the mean. Hence, it not a normal distribution. In the QQ plot, we can see that the data is not closet to the regression line. In Paired test, according to the graph in the part 'another way to perform t-test on paired-difference'. We can see that the graph is not symmetric. In the boxplot, the median is on right side far from the mean. In the QQ plot, we can see that the data is far from to the regression line. So it is not a normal distribution. In Sign test, from the table "test of H0:proportion=0.5", we can see that one side p-value(Pr <Z)=P(Pr<-1.0660)= 0.1432,which is equivalent to p-value(Pr <=p)=p-value(pr<=0.5)= 0.1431. In Wilcoxon Signed Rank test, the tables include all the results we obtained from the other three methods. Specifically, p-value(Pr > |t|)= 0.0595, p-value (Pr >= |M|)=0.2863, p-value (Pr >= |S|)=0.0405.

Section III Statistical Analyses

1. T-test

T-test method is to test if the null hypothesis the difference is 0. count are variables, data for 24 patients is the observation value. As can be seen from the T-test table, the Mean is 2.5, Standard Deviation is 1.2549, Minimum is -8, Maximum is 10. We know that t=(xbar-μ0)/s/ $\sqrt{n}$ has the t-distribution with n-1 degrees of freedom. P-value=2P(T>|t|)=2*P(T>|1.99|)=0.0595 which is greater than the confidence level $\boldsymbol{\alpha}$ (0.05), so p-value is not significant. Moreover, as can be seen from the T-test graphs, the distribution of count with 95% confidence interval for mean is not symmetric, and the median in the boxplot is on the right side far from the mean. Also, from the QQ-plot of count, in the polar coordinate system, x is quantile of the conditional distribution, x is expressed as variance count. We can see that the variance count is far from the linear regression line.

2. Paired t-test

In paired t-test method, count are variables, data for 24 patients is the observation value. Moreover, P-value=P(T>|t|)=P(T>|1.99|)=0.0595 which is greater than the confidence level $\boldsymbol{\alpha}$ (0.05), so p-value is not significant. In the graph of distribution of difference between hypotears and okerrinse, the distribution is not symmetric, it is not a normal distribution, and the agreement of hypotears, okerrinse are mostly far from the linear regression line.

3. Sign test

In sign test method, count are variables, data for 24 patients is the observation value. The sign test is a distribution free test because it use probability calculations that are correct for a wide range of population distributions. The sign test in matched pairs counts the number of positive differences. The p-value is computed from the B(N,1/2) distribution, where n is the number of nonzero differences. The sign test is less powerful than the t-test in cases where use of the t-test is justified. Let P be the probability of that

X=NO, which means the difference between the scores for patientnum and scores for hypotears. We want to test H0: p=1/2, p>1/2.

As can be seen from table "sign test, using count variable X", the percent of X=NO is 36.36%, the asymptotic confidence limits and test t. Specifically,95% lower confidence limit is 0.1399, and 0.5874. In the table 'Test of H0: Proportion = 0.5', Two-sided P-value(Z > |z|)=P(Z>|-1.0660|)=0.2864, P-value(P<p)=(P<0.5)=0.1431>$\boldsymbol{\alpha}$=0.05. Ignore pairs with difference 0, the number of trial n is the count of the remaining pairs. The test statistics is the count X of pairs with a positive difference. P-value for X are based on the binomial B(n,1/2) distribution.

4. Wilconxon Signed Rank Test

The Wilcoxon signed-rank test determines whether there is a statistically significant difference in the median of a dependent variable between two related groups. This method include all the result obtained from the previous three methods. As can be seen from the table 'Tests for Location: Mu0=0', the p value for T-test is: P(Pr> |t|)= P(Pr>|1.992253|)=0.0595>$\boldsymbol{\alpha}$=0.05; the p value for Sign test is : P(Pr >= |M|)=P(Pr>=|3|)=0.2863 >$\boldsymbol{\alpha}$=0.05; the p value for the Wilconxon Signed Rank test is P(Pr >=|62|)=0.0405<$\boldsymbol{\alpha}$=0.05, so only in Wilconxon Signed Rank test, the p-value is significant at confidence level $\boldsymbol{\alpha}$ 0.05.

5. Recommended statistical method

From the analysis for the method 1-4, Wilconxon Signed Rank Test is recommended for this study, because p value for T-test is 0.0595, p value for Sign test is 0.2863, only p value for the Wilconxon Signed Rank test is P(Pr >=|62|)=0.0405<$\boldsymbol{\alpha}$=0.05, which is significant at confidence level $\boldsymbol{\alpha}$=0.05. Even without comparing with $\boldsymbol{\alpha}$=0.05, p value for the Wilconxon Signed Rank test is also the smallest in the tree methods. Therefore, we have enough evidence so reject the null hypothesis.

Section IV. Conclusion

From the analysis of the data of 24 observation patients, and the variable count which means difference between overall scores between patients instructed to Oker-Rinse and Hypotears, we found the null hypothesis H0:μ=0 is rejected. We recommend Wilcoxon signed-rank test for this study because the p value is0.0405 which is the smallest in the four methods. Therefore, we have enough evidence to claim that there is difference between the two test preparation in the overall Rose-Bengal staining scores.