Statistical Tests Using GraphPad Prism

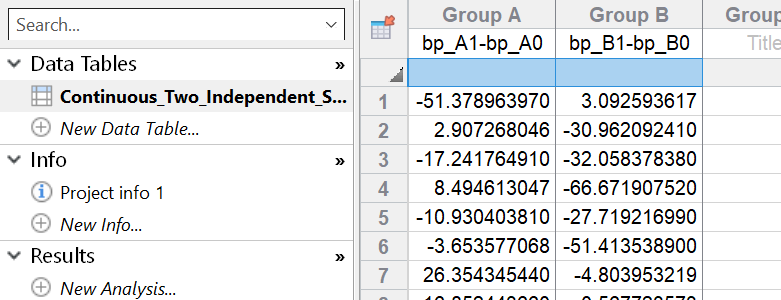
This document shows how to perform the tests that were presented in R during the course with Prism.

# 1. Continuous data: t-test

## 1.1 Two independent samples: Unpaired/two-sample Welch’s t-test

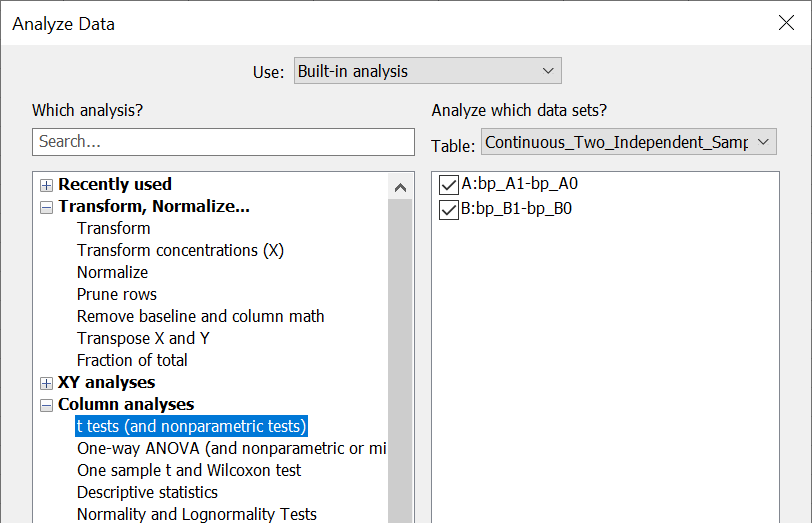
Table type:

Column, first two columns contain outcomes for the two groups, respectively

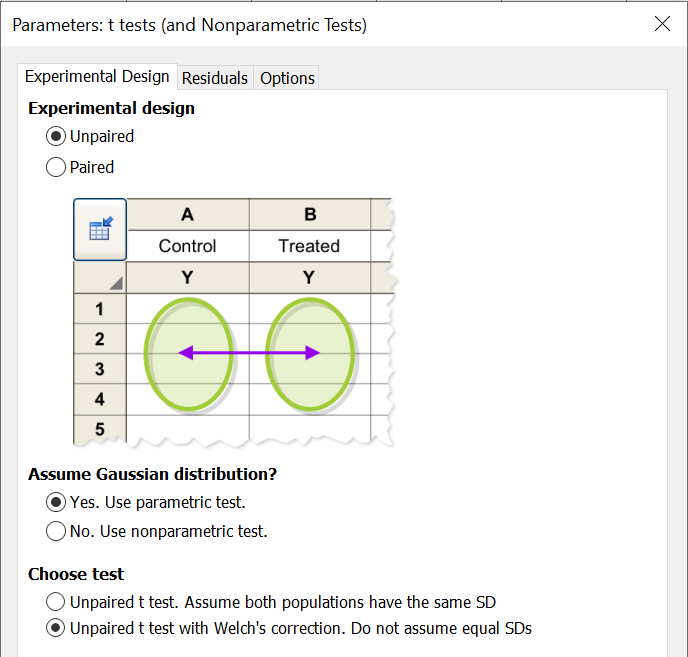


Analysis:

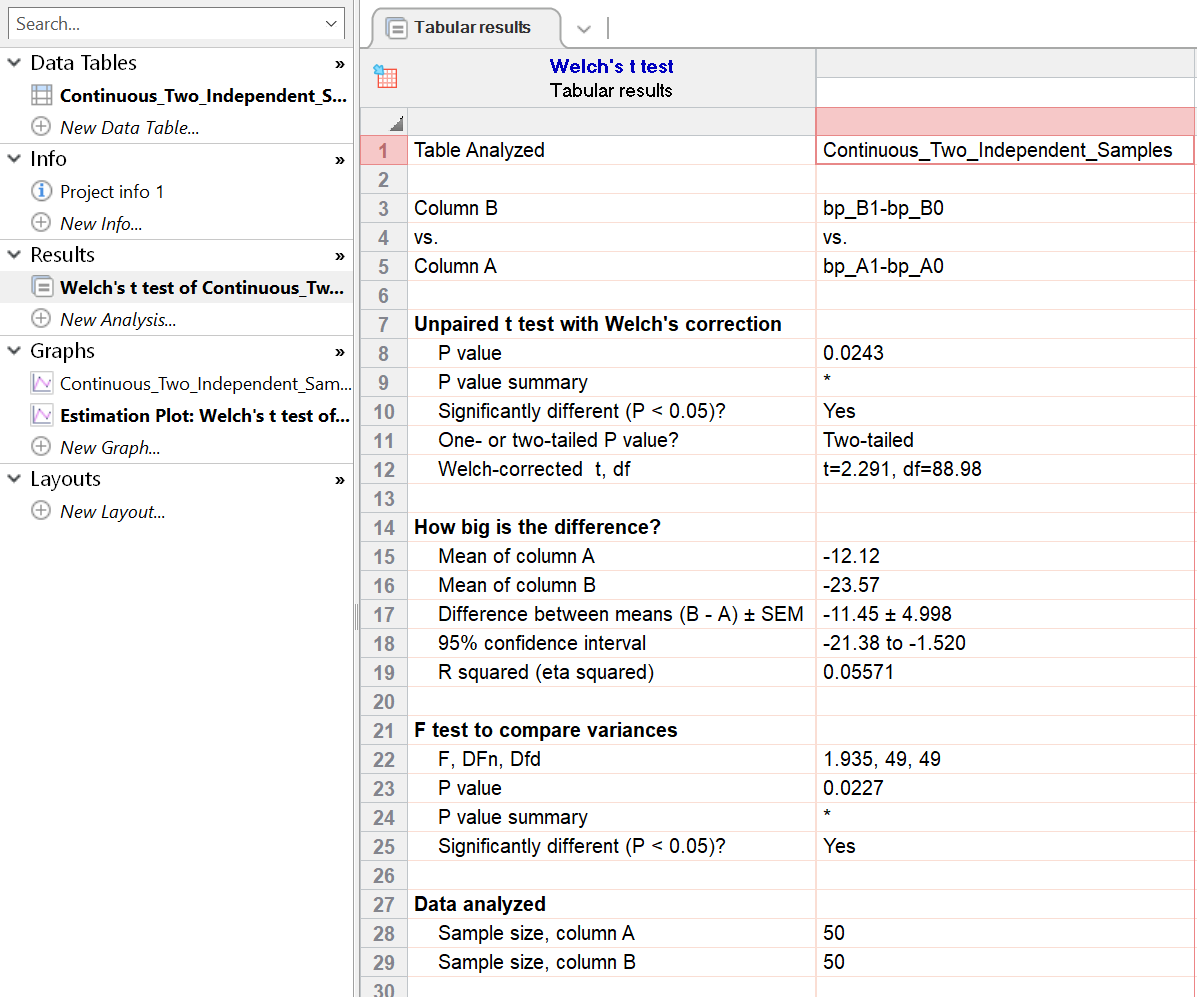
Test family: t tests (and nonparametric tests)



Experimental design: Unpaired  
Assume Gaussian distribution?: Yes. Use parametric test.  
Choose test: Unpaired t test with Welch’s correction. Do not assume equal SDs



Interpretation of results:

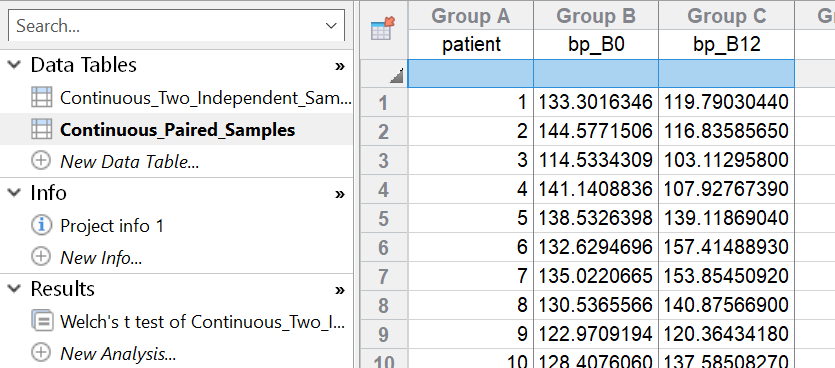


The important element here is the p-value of 0.0243 along with the 95% confidence interval of   
[-21.38,-1.52] which does not contain zero, indicating a significant difference between the groups. As the column means (-12.12 and -23.57, respectively) show, the blood pressure reduction in group B is significantly higher.

## 1.2 Paired samples: Paired/matched t-test

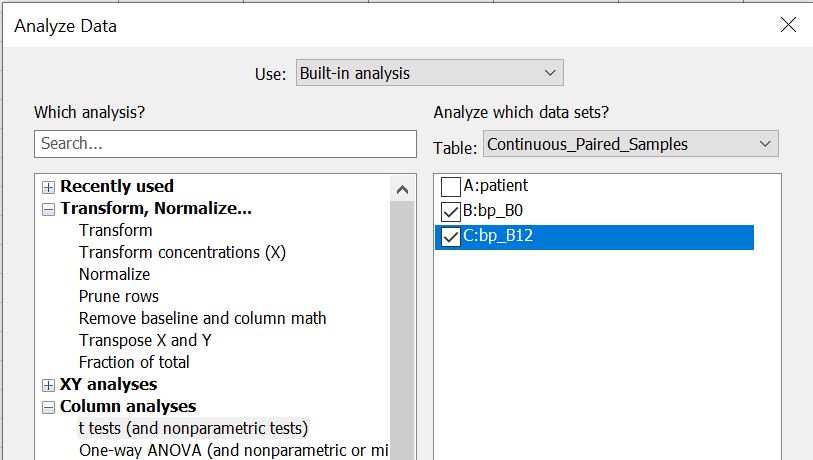
Table type:

Column, first column can contain pairing ID (but is not mandatory), next two columns contain outcomes for the data pairs, each in a row

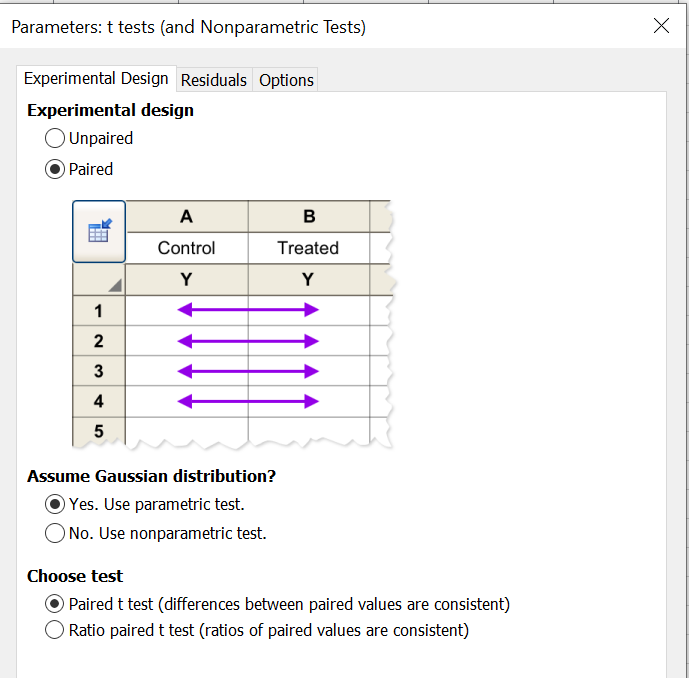


Analysis:

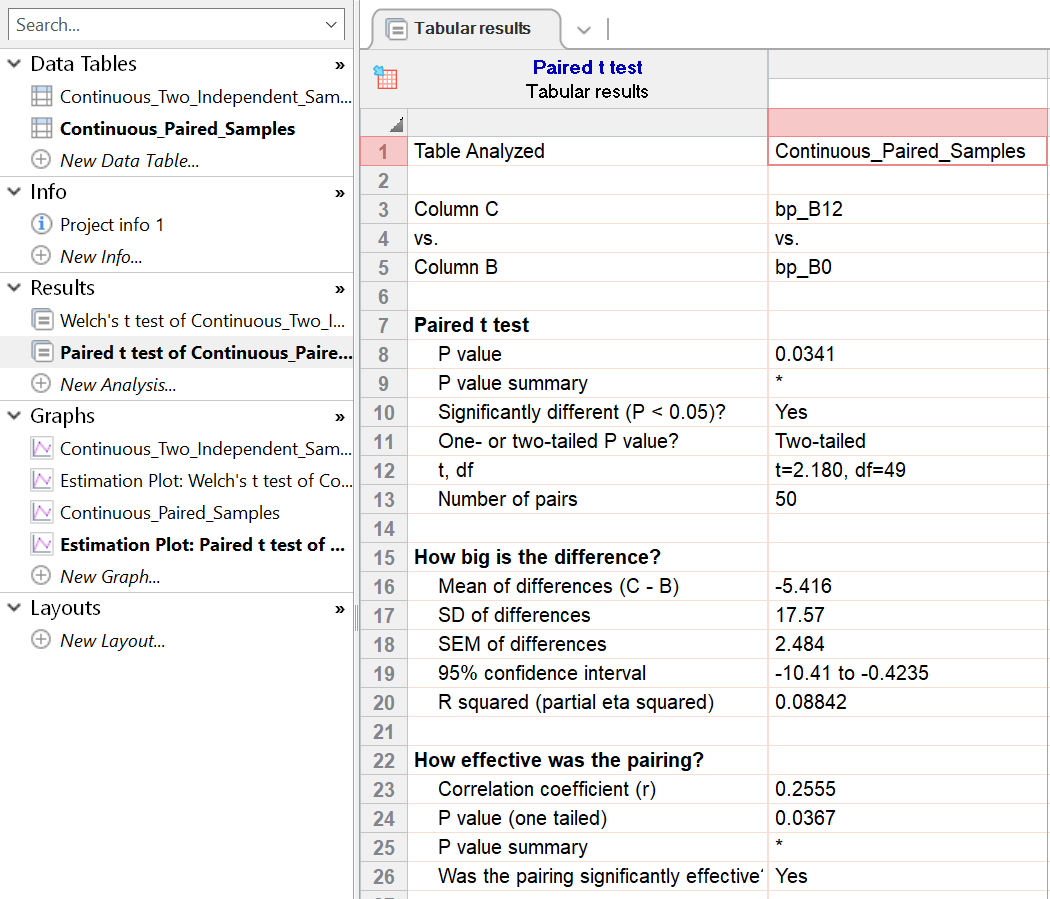
Test family: t tests (and nonparametric tests)



Experimental design: Paired  
Assume Gaussian distribution?: Yes. Use parametric test.  
Choose test: Paired t test (differences between paired values are consistent)



Interpretation of results:



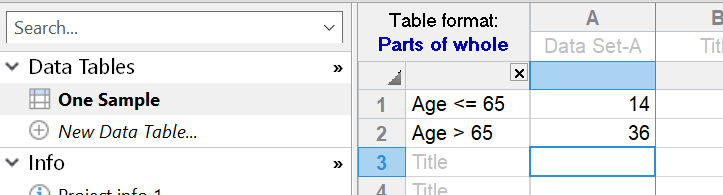
The p-value of 0.0341 and the 95% confidence interval of [-10.42,-0.4235] which does not contain zero indicate a significant difference between the groups.

# 2. Tests for categorical outcomes

## 2.1 Binary data from one population: Exact binomial test

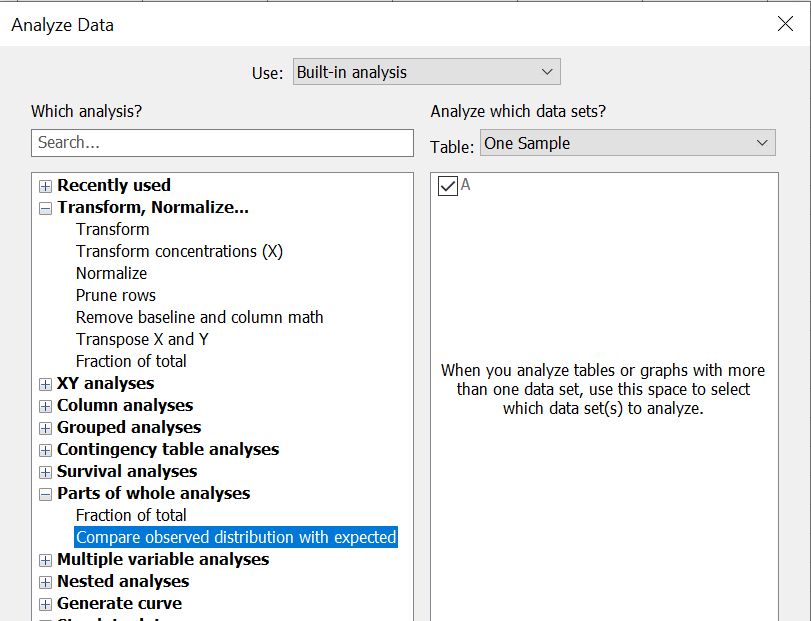
Table type:

Parts of whole, first two rows in the first column contain counts for the two possible outcomes, respectively

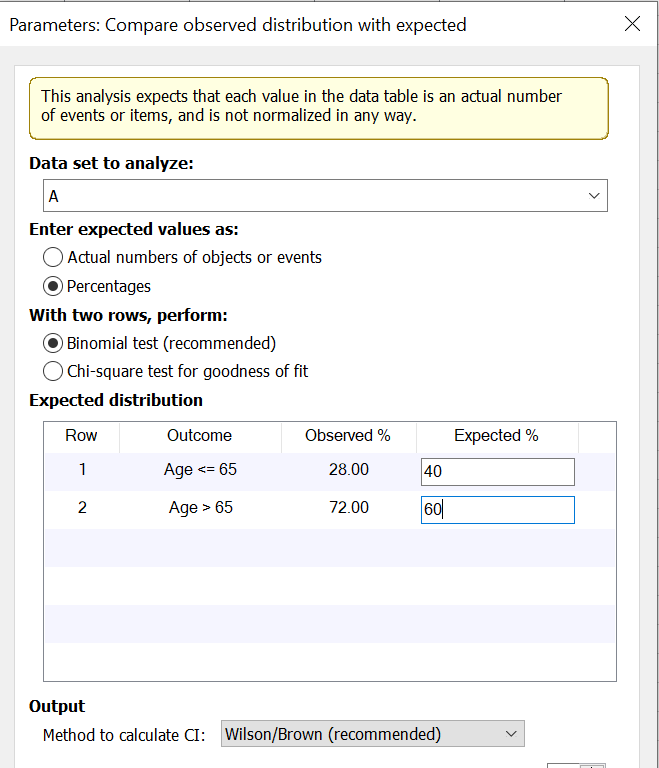


Analysis:

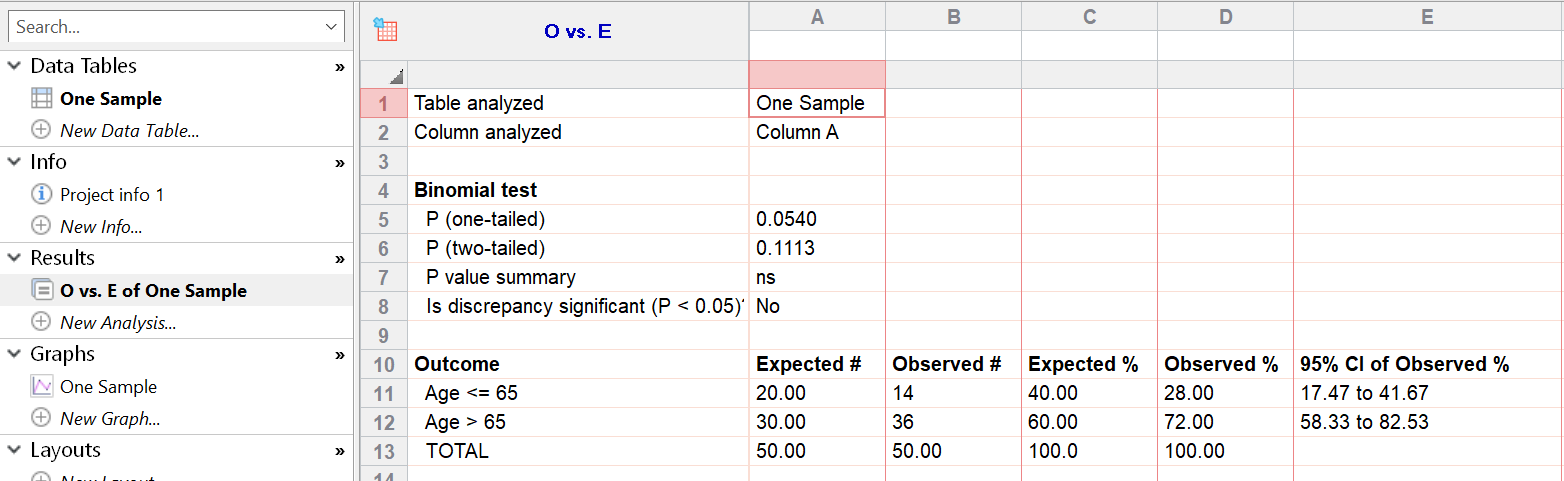
Test family: Compare observed distribution with expected



Enter expected values as: Percentages. I. e. if your null hypothesis is that the proportion of patients aged > 65 years is 60%, your expected percentages are 40% / 60% for Age <= 65 / Age > 65.  
With two rows, perform: Binomial test (recommended)  
Method to calculate CI: Wilson/Brown (recommended). This is definitely preferable to Wald CI (not implemented in Prism) which tends to be too liberal, and to the Pearson-Clopper CI used by the R function binom.test() which tends to be too conservative.



Interpretation of results:

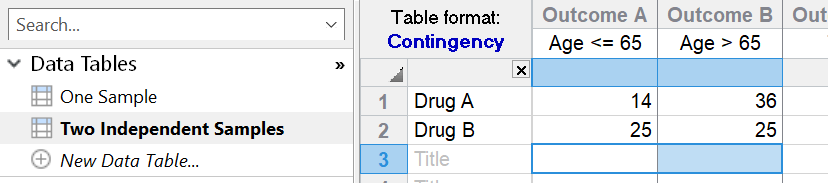


Since we test for difference and not for superiority, the two-tailed p-value of 0.1113 is the relevant one. Along with the 95% confidence interval of [58.33%,82.53%] which contains the value of 60%, it shows that the proportion of patients over age 65 years is not significantly different from 60%.

## 2.2 Binary data from two independent samples: Fisher’s exact test

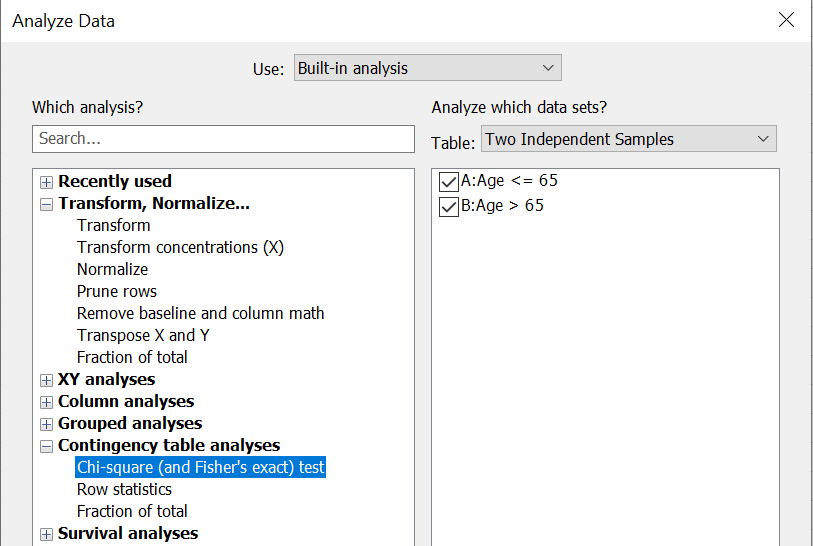
Table type:

Contingency, groups to be compared in the first two rows and possible outcomes in the first two columns

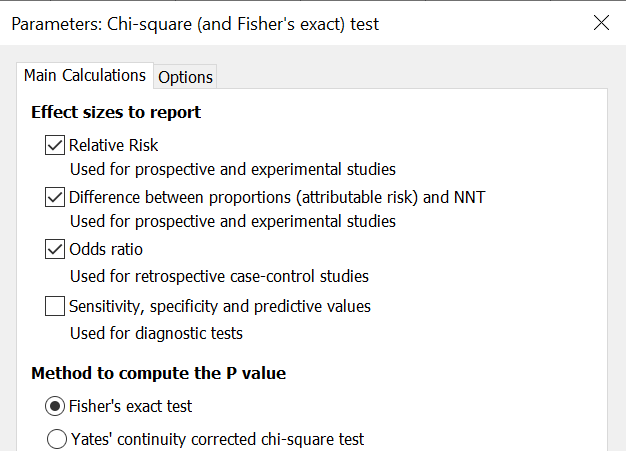


Analysis:

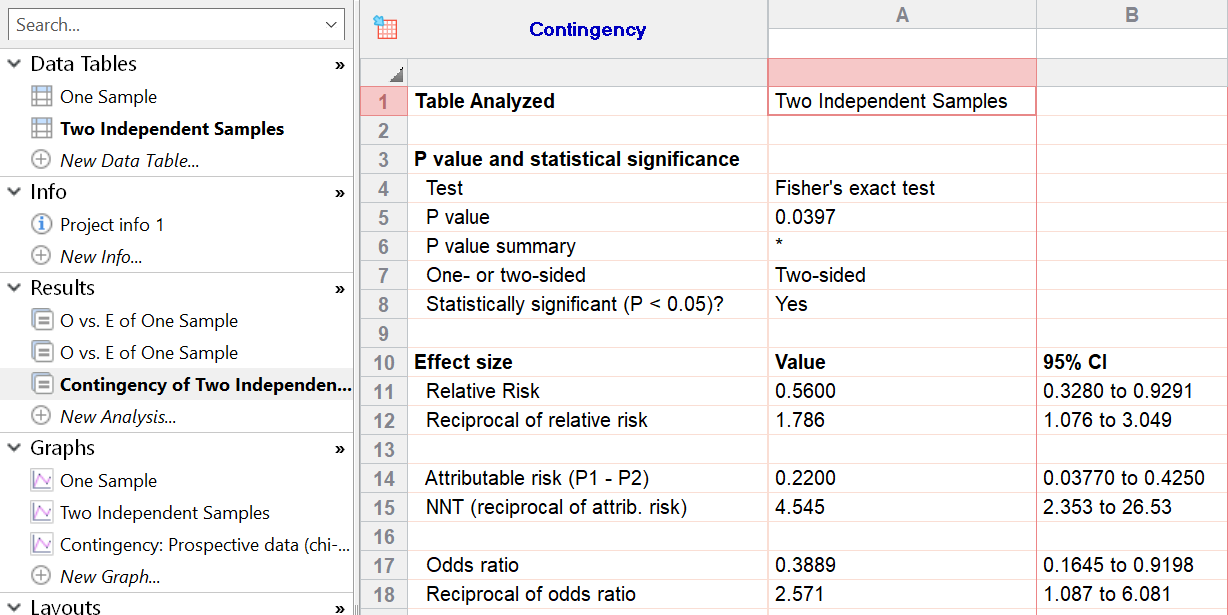
Test family: Chi-square (and Fisher’s exact test)



Effect sizes to report: Choose whichever you find useful. Odds ratio is not strictly limited to retrospective designs, but can also be ed in prospective designs.  
Method to compute the P value: Fisher’s exact test



Interpretation of results:

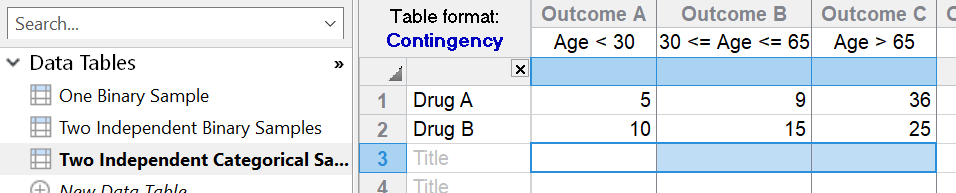


The important element here is the p-value of 0.0397, indicating that the distribution of the outcome is significantly different in the two treatment groups. This result can be seen from the confidence intervals, too: The CIs for relative risk and odds ratio do not contain the value of 1 which would be expected under the null hypothesis, and the CI for risk difference (attributable risk) does not contain zero.

## 2.3 Non-binary categorical data from two independent samples: Chi-squared test

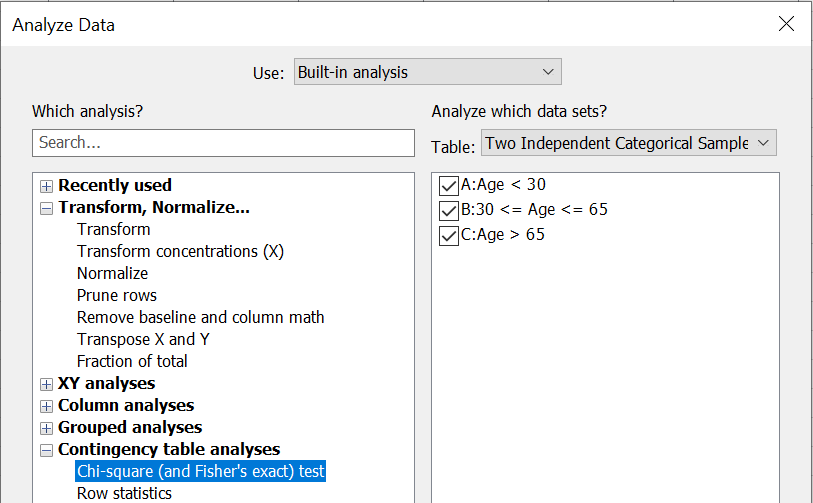
Table type:

Contingency, groups to be compared in the first two rows and possible outcomes in the first columns

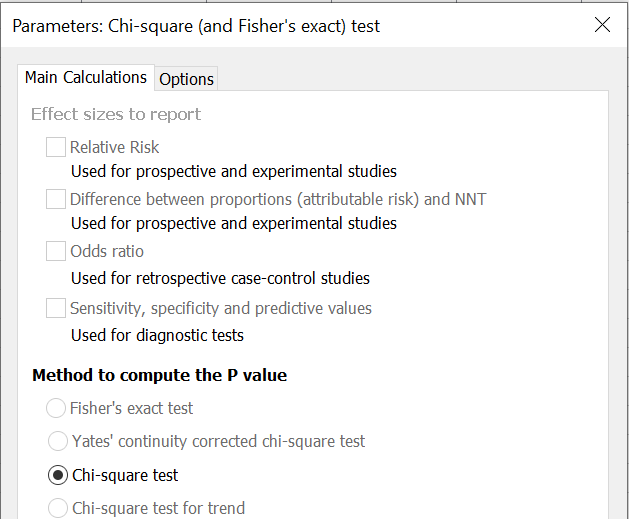


Analysis:

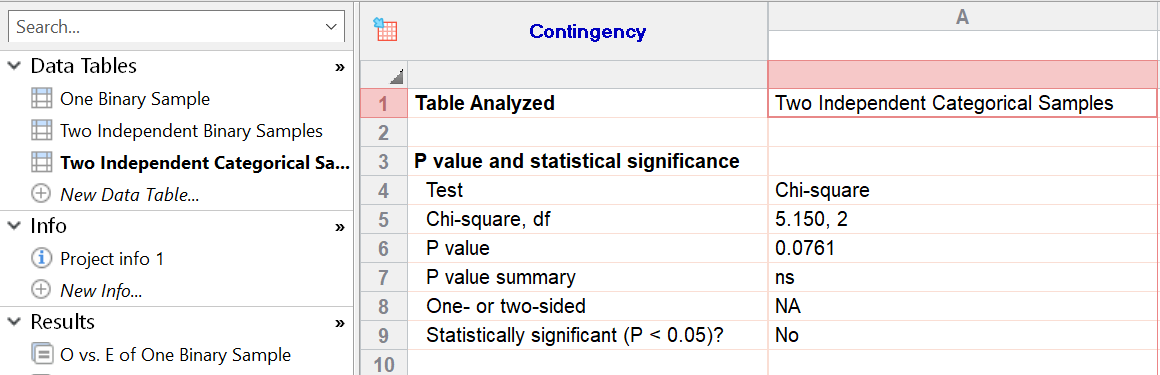
Test family: Chi-square (and Fisher’s exact test)



Method to compute the P value: Chi-square test



Interpretation of results:



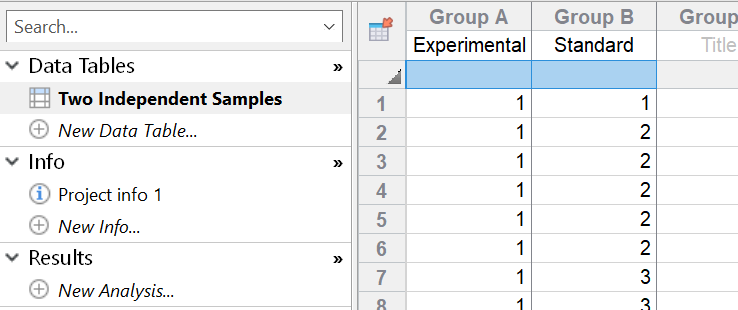
Here, the p-value of 0.0761 indicates that there is no evidence of different distribution of outcomes between the treatment groups.

# 3. Nonparametric tests

## 3.1 Two independent samples: WMW test

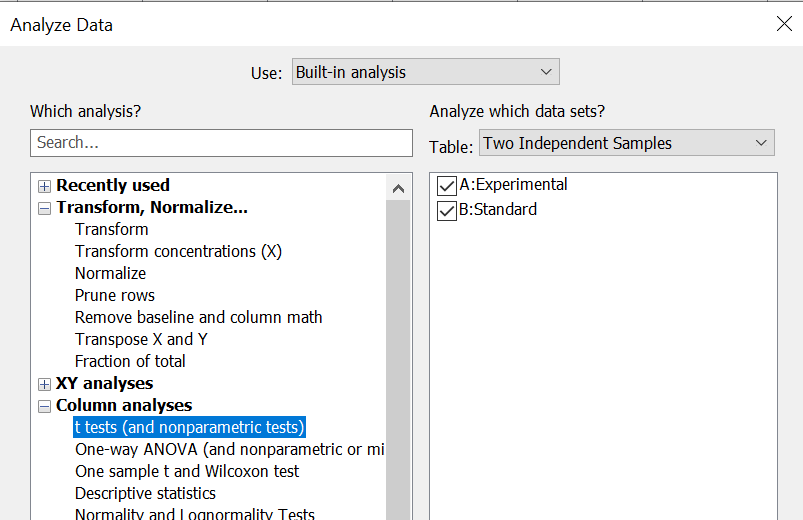
Table type:

Column, first two columns contain outcomes for the two groups, respectively

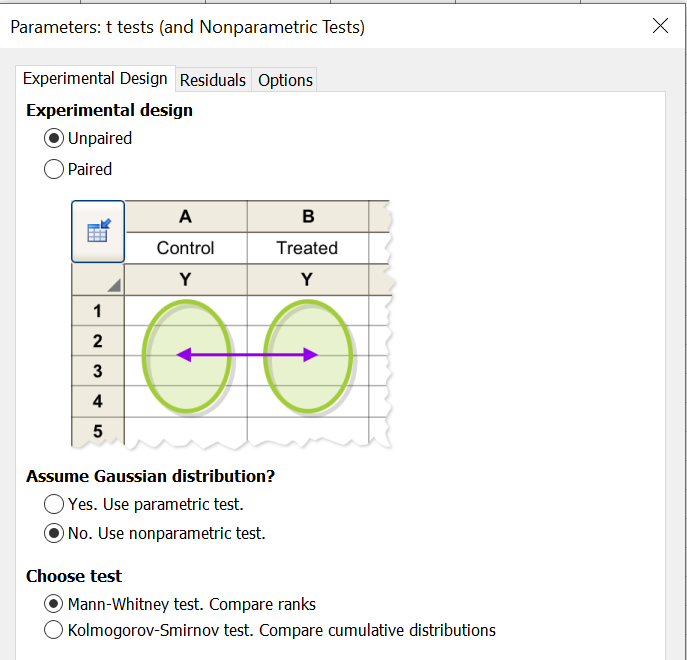


Analysis:

Test family: t tests (and nonparametric tests)

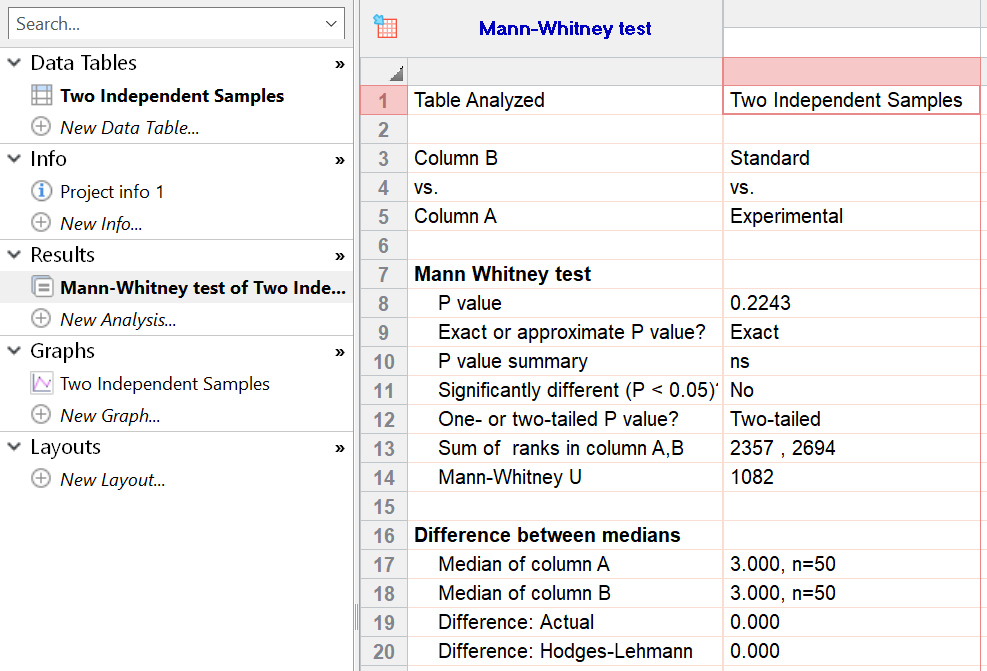


Experimental design: Unpaired  
Assume Gaussian distribution?: No. Use nonparametric test.  
Choose test: Mann-Whitney test. Compare ranks



It is not possible to choose between exact and asymptotic tests – Prism switches as soon as the smaller sample size exceeds 100.

Interpretation of results:

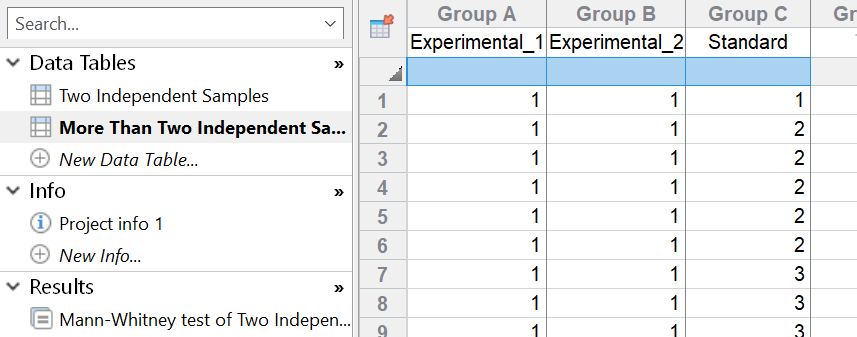


The important element here is the p-value of 0.2243, indicating that there is no evidence of different distribution of outcomes between the groups.

## 3.2 More than two independent samples: Kruskal-Wallis test

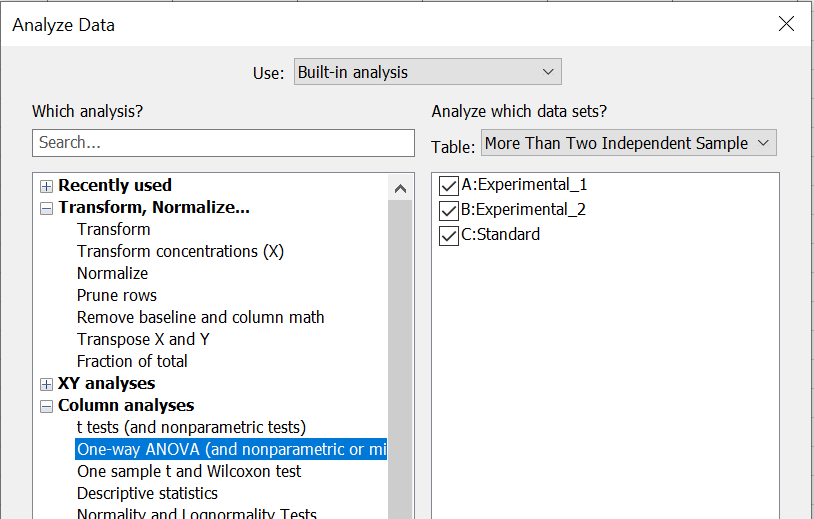
Table type:

Column, first k columns contain outcomes for the k groups, respectively



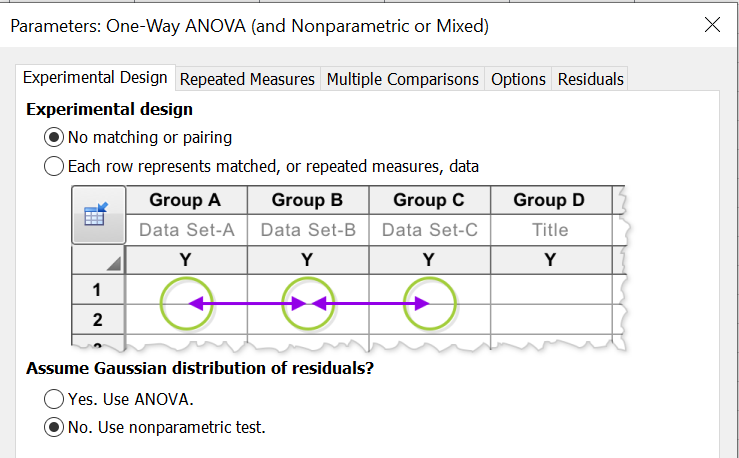
Analysis:

Test family: One-way ANOVA (and nonparametric or mixed)

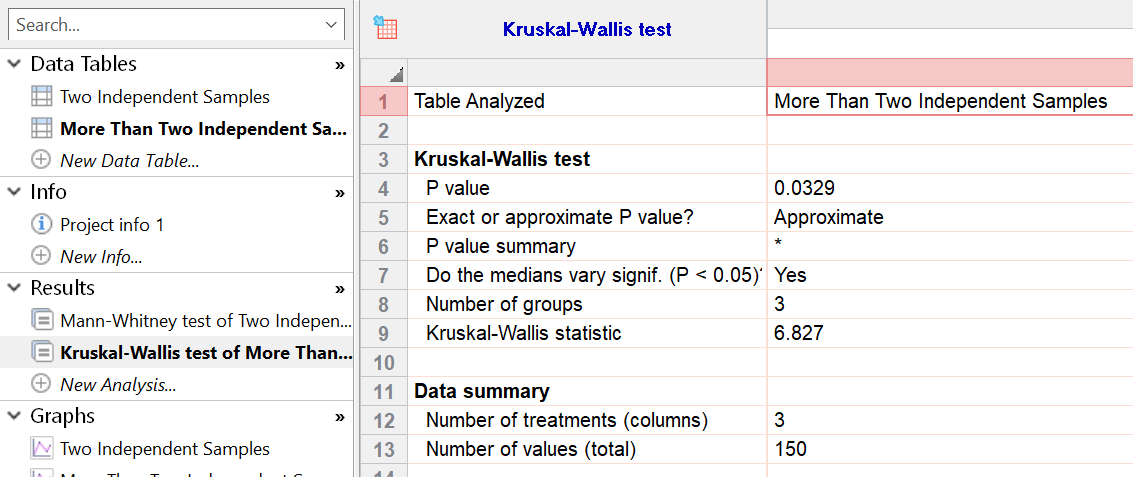


Experimental design: No matching or pairing  
Assume Gaussian distribution of residuals?: No. Use nonparametric test.

The test is always based on ranks, not on pseudoranks. It is not possible to choose between exact and asymptotic tests – Prism switches as soon as the sample size gets too large (the exact cutoff is not given in the documentation).



Interpretation of results:

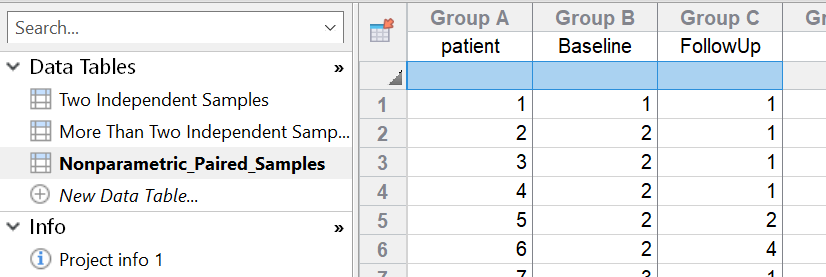


The important element here is the p-value of 0.0329, indicating that the distribution of the outcome is not the same in all three groups. In a next step, both experimental groups could be tested against the standard group (with α = 2.5% to allow for multiplicity) using the WMW test to find out where this global difference comes from.

## 3.3 Paired samples: Wilcoxon signed ranks test

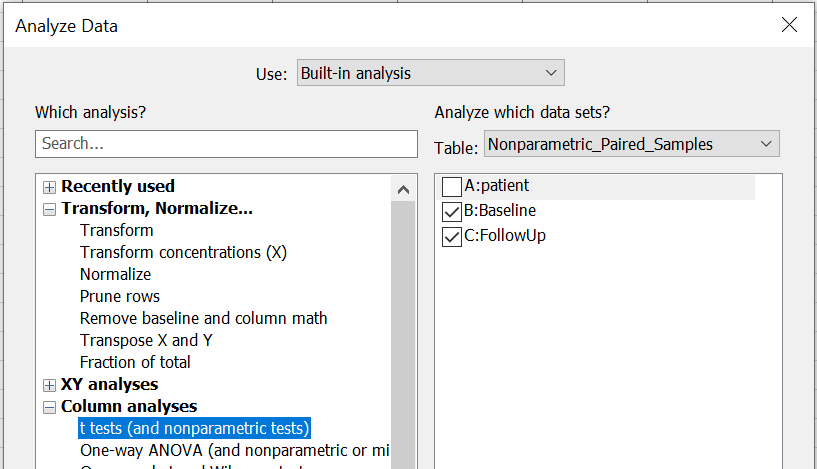
Table type:

Column, first column can contain pairing ID (but is not mandatory), next two columns contain outcomes for the data pairs, each in a row

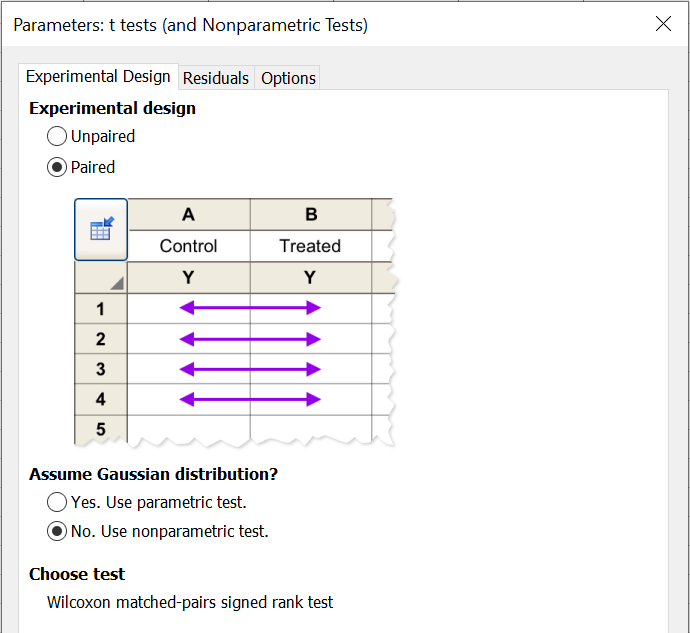


Analysis:

Test family: t tests (and nonparametric tests)

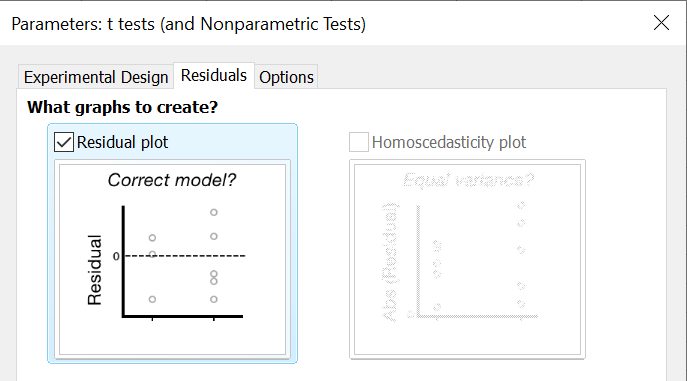


Experimental design: Paired  
Assume Gaussian distribution?: No. Use nonparametric test.



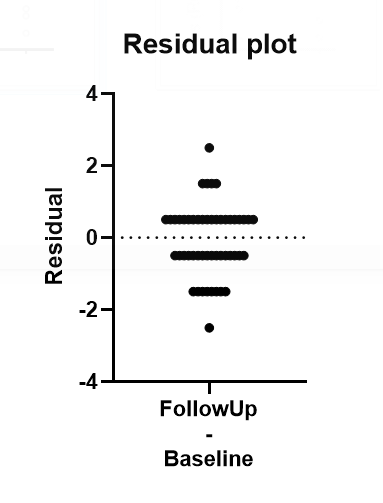
Paired ranks test and sign test are not implemented in Prism. Thus the only option in this scenario is to use the Wilcoxon signed ranks test (but of course you can always rank your outcomes by hand, import the ranks into Prism and perform a paired t-test on the ranks).

Create a residual plot to check for symmetric differences, which is a prerequisite for validity of the test:

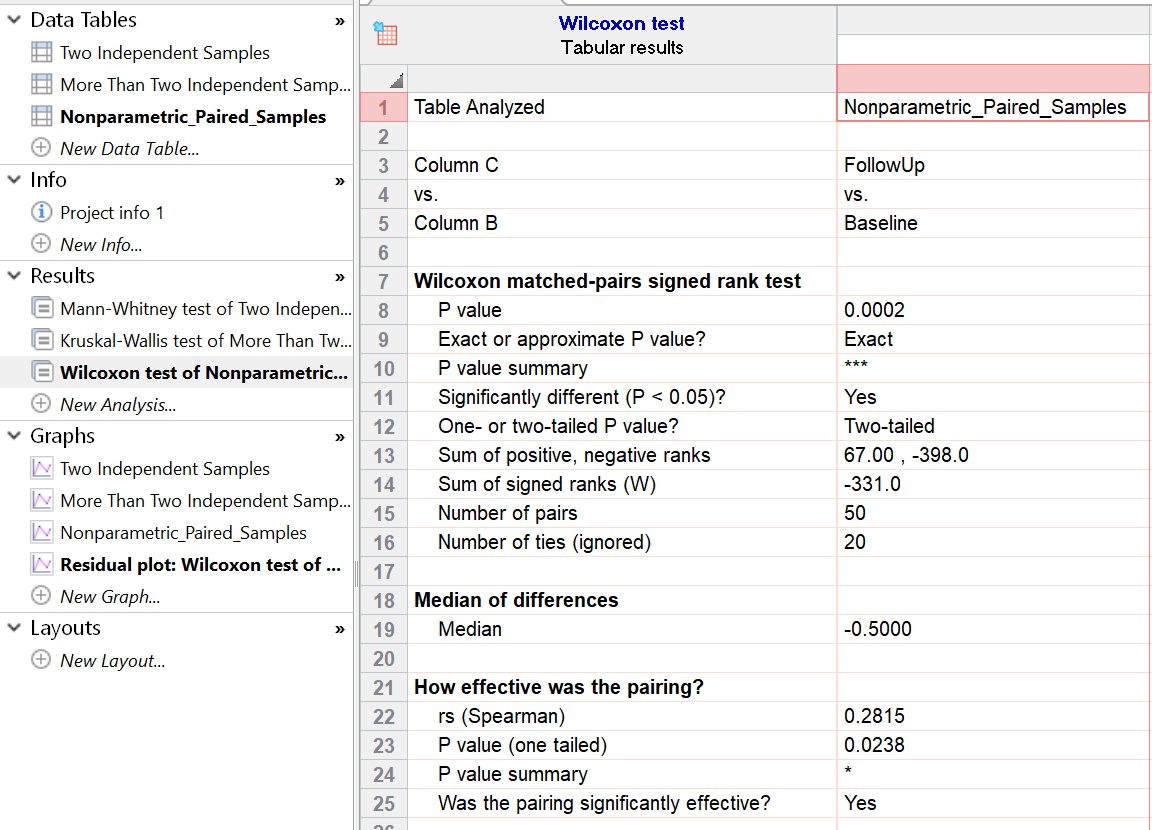


Again, it is not possible to choose between exact and asymptotic tests – Prism switches as soon as there are more than 200 pairs.

Interpretation of results:



The residuals (differences) show a rather symmetric distribution to the positive and negative side of zero. So we can go on with the test itself:



The p-value of 0.0002 indicates a highly significant difference between the groups.