> The following resources are associated:
> Paired t-test, Friedman, SPSS dataset 'Video'

Wilcoxon Signed-rank test in SPSS
(Non-parametric equivalent to paired t-test)

Used when two measurements of the same dependent variable are taken at different time points or under different conditions for each subject and the assumptions of the paired t-test have not been met.

Dependent: Continuous (scale) or ordinal

Independent: Time/condition

Data: The dataset 'Video' contains results from a study by a charity comparing the effectiveness of videos made to aid understanding of a particular medical condition. Participants watched three videos (A, B, C) and one product demonstration (D) and were asked several Likert style questions about each. These were summed to give an overall score for each e.g. TotalCold is the total score of the ordinal questions for the old video C. Ordinal questions for understanding the condition (Strongly disagree – strongly agree) for video C and the product demonstration are in the 2nd and 3rd columns of the data in the image below and total preference scores are in the 4th and 5th columns of this example. The 5th column contains the paired differences between the scores of video C and the demonstration (TotalCold – TotalDDEMO).

	Person	VideoCOldUnderstanding CONDITION	DEMOUnderstanding CONDITION	TotalCOld	TotalDDEMO	DiffCD
1	1	Neither agree or disagree	Strongly Agree	13	22	-9
2	2	Strongly Agree	Strongly Agree	20	22	-2
3	3	Disagree	Strongly Agree	14	23	-9
4	4	Strongly Agree	Strongly Agree	23	25	-2

The Wilcoxon test can be used to compare individual ordinal questions such as those asking for opinions on understanding the condition, or overall scores if the paired differences are not normally distributed.

Research question: Do members of the public prefer the demonstration technique to the old video C?

Reviewer: Chris Knox
University of Sheffield

The total scores for C and the demonstration are numeric so the paired t-test would be used if the paired differences are normally distributed. Paired differences can be calculated and added to the dataset using **Transform → Compute variable**. Give the new variable a name in the '*Target variable'* box and move the variables of interest to the '*Numeric Expression'* box separated by a '-' sign.
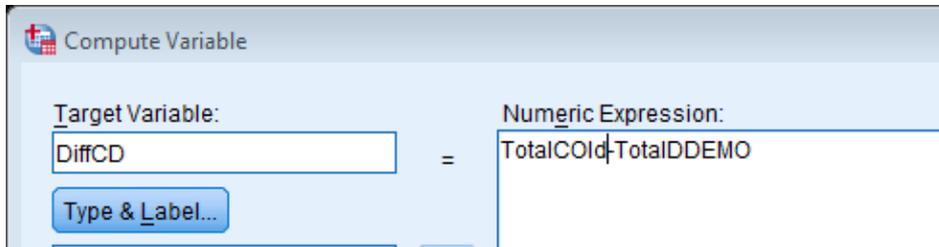
Produce a histogram using **Graphs → Legacy Dialogs → Histogram** and move the new variable 'DiffCD' to the Variables box.
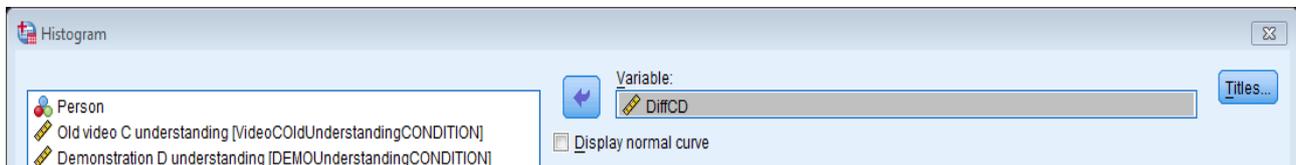
The paired differences are not normally distributed in this dataset so the paired t-test is not reliable and the Wilcoxon signed rank test is more appropriate.

When using a non-parametric test, it is better to summarise using medians rather than means. Examination of the medians suggests that people

Median

Overall score (video C)	Overall score (demo D)
14	22

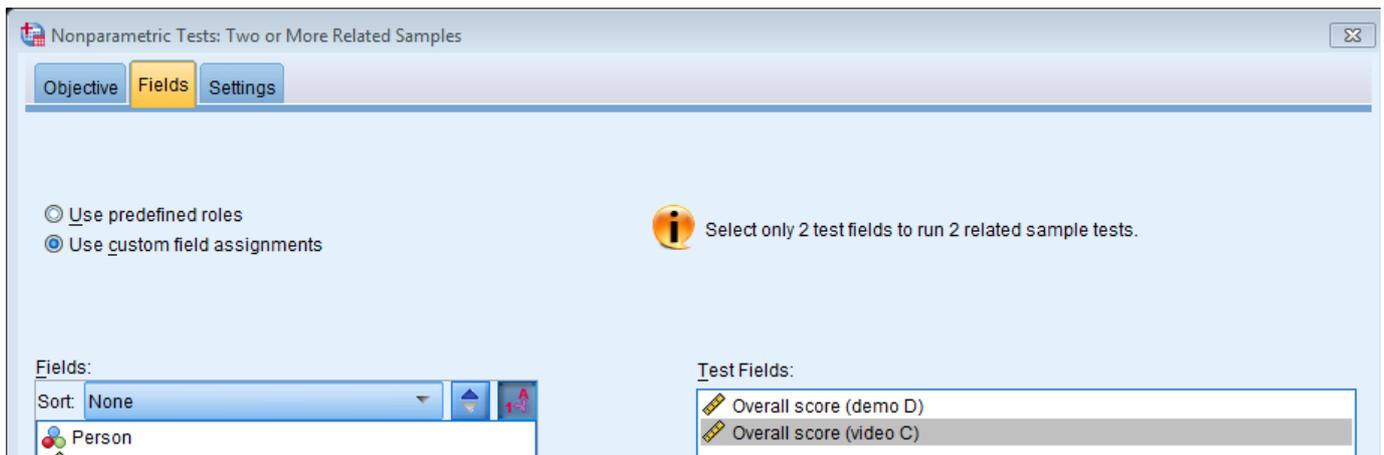generally preferred the new method for explaining the condition as the median is higher.

Conducting the Wilcoxon signed rank test in SPSS (new method)

There are two methods for carrying out a Wilcoxon test in SPSS, the legacy dialogs method and the 'New' procedure that we will use here.

Go to *Analyze → Nonparametric → Related samples* and move the two variables of interest, TotalDEMO and then TotalCold, to the '*Test Fields'* section on the **Fields** tab.

Important note: SPSS will only allow variables classified as scale **Scale** to be analysed using Wilcoxon. Even if you are using ordinal variables, you must classify them as scale in **variable view** before running the analysis.

Nonparametric Tests: Two or More Related Samples

Objective | Fields | Settings

◯ Use predefined roles
◉ Use custom field assignments

ⓘ Select only 2 test fields to run 2 related sample tests.

Fields:
Sort: None
👥 Person

Test Fields:
Overall score (demo D)
Overall score (video C)

Clicking on **Run** will now run the standard Wilcoxon signed rank test and give the output below.

The output shows that there is significant evidence of a difference between the scores for the old video and the new demonstration method.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Overall score (demo D) and Overall score (video C) equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

If you double click on the results in the Output window, a new **Model Viewer** window with more information opens.
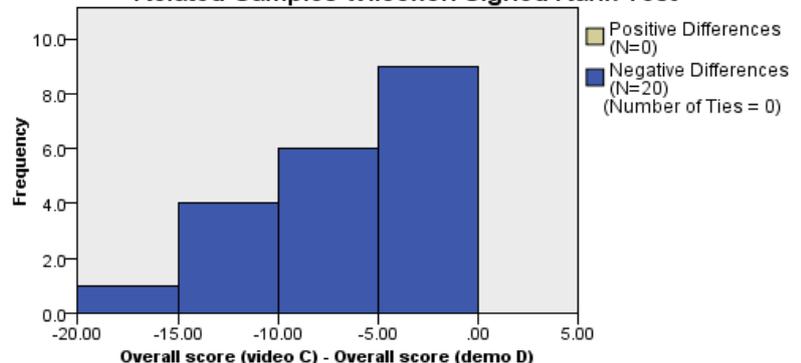
The test involves ranking the absolute paired differences of each subject e.g. a rank of 1 would indicate the smallest difference between the scores of one person for the old and new methods. The ranks are then classified as positive (Old Video > Demo) or negative (Demo > Old video). The histogram shows everyone ranked the Demo higher than the Old Video as there are only negative ranks. Positive ranks would appear in brown. The sums of the positive and negative ranks are compared and SPSS uses the sum of the positive ranks as the **Test Statistic** (W). As there are no positive ranks to add, the test statistic is 0.

Related-Samples Wilcoxon Signed Rank Test

Positive Differences (N=0)
Negative Differences (N=20)
(Number of Ties = 0)

Overall score (video C) - Overall score (demo D)

Total N	20
Test Statistic	.000
Standard Error	26.744
Standardized Test Statistic	-3.926
Asymptotic Sig. (2-sided test)	.000

The Wilcoxon W **Test Statistic** is simply the sum of the positive ranks but in order to calculate the p-value (Asymp. Sig), SPSS uses an approximation to the standard normal distribution to give the **Standardized (Z) test statistic** (-3.296) and resulting p-value in the **Asymptopic Sig. (2- sided test)** row (p < 0.001). This approximation is not as reliable for small samples so running an Exact test (see end of sheet) is preferable.

Reporting a Wilcoxon test

A Wilcoxon signed rank test showed that there was a significant difference (Z = -3.926, p< 0.001) between scores given for the old video compared to the new demonstration method. The median score for the new demonstration method was 22 compared to 14 for the old video. Therefore, the charity should start using the new method for explaining the condition to the general public.

Effect size

An effect size can be calculated by dividing the absolute (positive) Standardised test statistic z by the square root of the number of pairs. $\dfrac{Z}{\sqrt{n}} = \dfrac{3.926}{\sqrt{20}} = 0.88$

Here the effect size is 0.88 which is very large according to Cohen's classification of effect sizes which is 0.1 (small effect), 0.3 (moderate effect) and 0.5 and above (large effect).

Exact test using legacy dialogs procedure

If you have a very small sample, you will need to use the Exact test using the Legacy Dialogs method for the Wilcoxon. Go to *Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples* and move the two variables of interest to the 'Test Pairs' section. Then click on the **Exact** button and select the 'Exact' option.
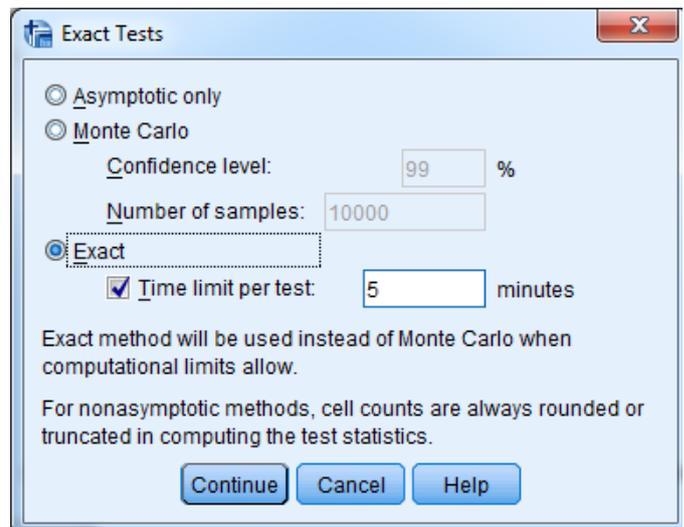
Test Statistics[a]

	Overall score (video C) - Overall score (demo D)
Z	-3.926[b]
Asymp. Sig. (2-tailed)	.000
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Use the 'Exact Sig. (2-tailed)' p-value (p<0.001) and conclude that there is a significant difference between the conditions.