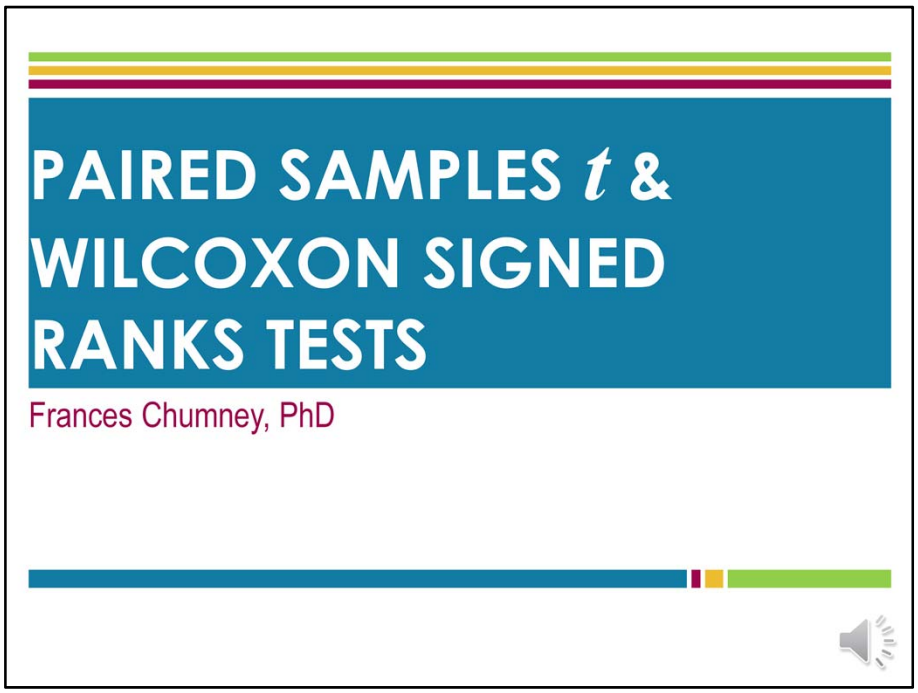
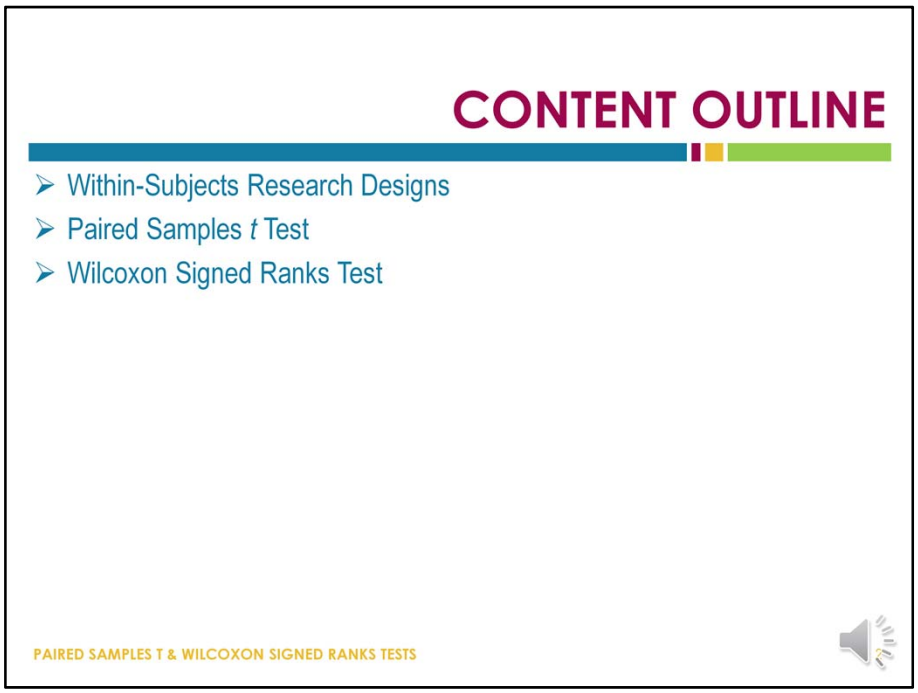


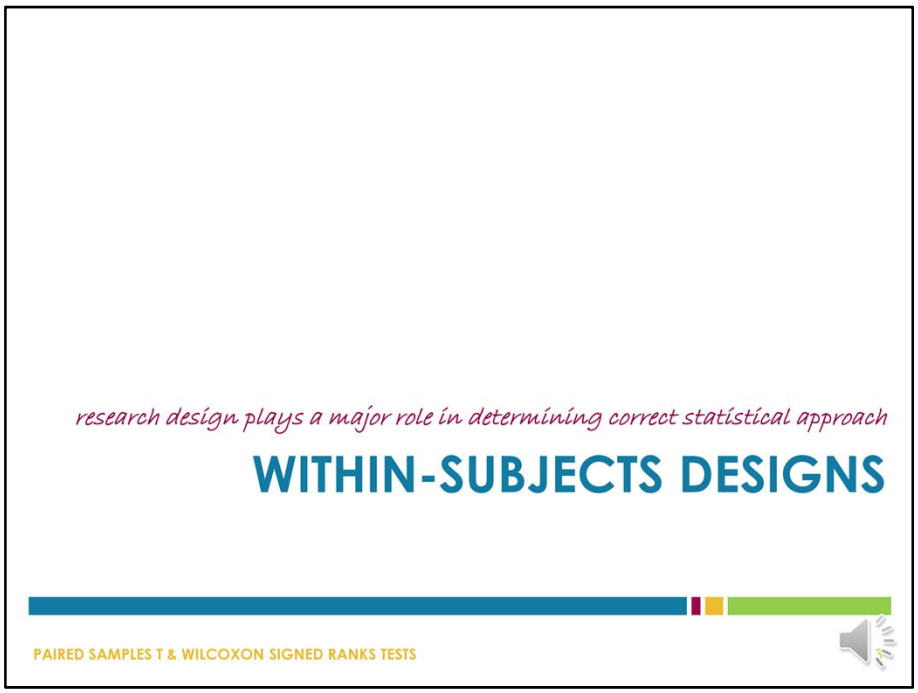
TRANSCRIPT:

This is Dr. Chumney with a brief overview of the paired samples *t* test and the Wilcoxon signed ranks test.



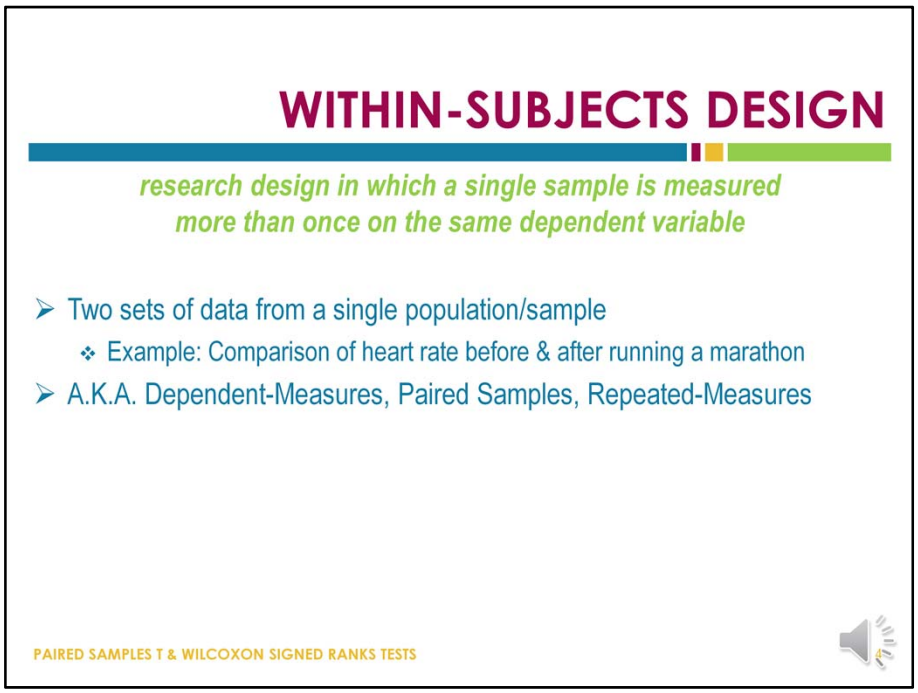
TRANSCRIPT:

In this brief lecture, I will first provide a quick reminder of what within-subjects research is, and then explain how and when to use paired samples *t* tests and Wilcoxon tests.



TRANSCRIPT:

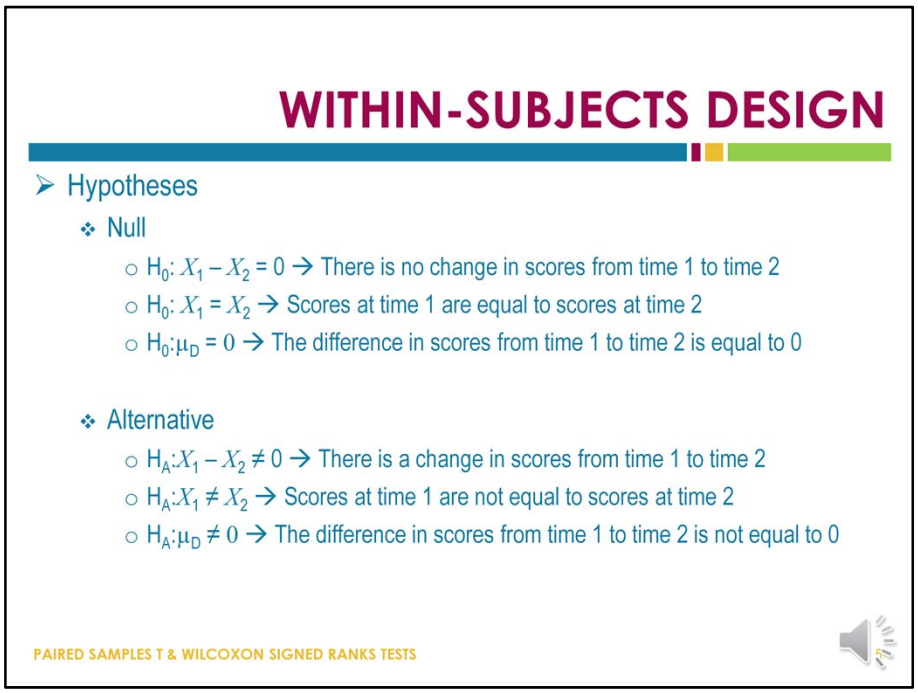
Paired samples *t* and Wilcoxon tests are analytic methods that are appropriate for within-subjects research designs, so let's start with a quick reminder of what a within-subjects design looks like.



TRANSCRIPT:

A within-subjects design is an approach to research in which one sample is measured more than once on the same dependent variable. In other words, we have two sets of responses just like with between-subjects designs, but now those data come from a single sample instead of from two separate samples that we want to compare.

Within-subjects is a type of research design – which is also sometimes called dependent-measures design - is sort of an umbrella term under which design choices such as paired samples, and repeated-measures fall. In a repeated-measures analysis, we have measured each person in the sample twice, as in a pre-/post- design. In a paired samples analysis, data are collected just once on each participant, but the participants come from the same sample and are matched in pairs based on key characteristics. For example, if a paired samples approach were used to examine the effects of a new drug, a sample would be selected and then members of that sample would be assigned to pairs based on sex, age, ethnicity, and other variables that the researcher believes are important for consideration. One member of each pair is then assigned to one condition, while the other member of each pair is assigned to another condition. When the data are analyzed, the data from one group is compared to the data from the other group, but the analysis accounts for the fact that the data are matched or paired.



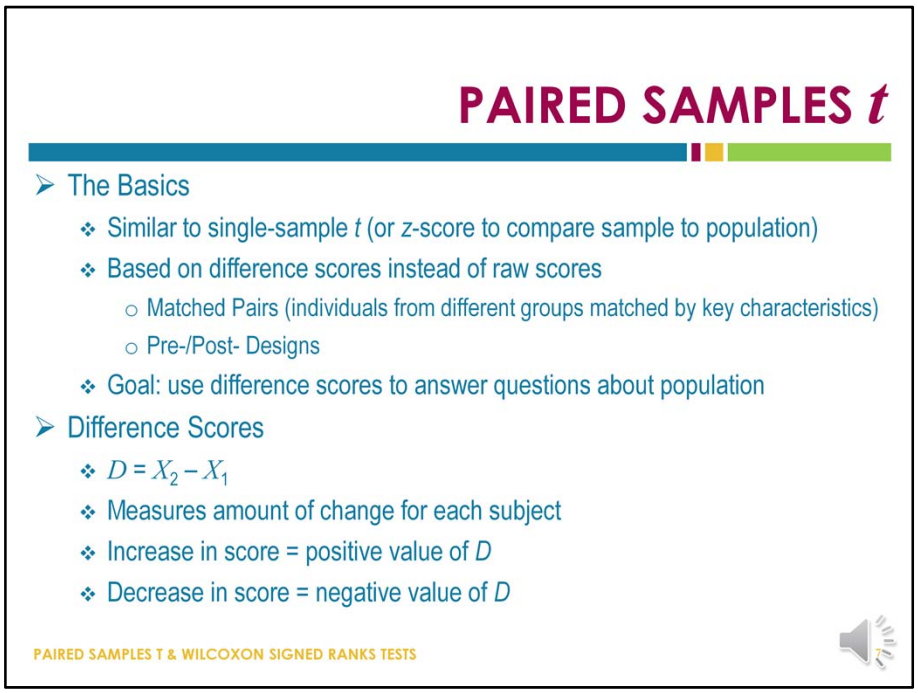
TRANSCRIPT:

In terms of hypothesis, the null hypothesis for a within-subjects design is always that there is no change in scores from time 1 to time 2, or that scores at time 1 are equal to scores at time 2, or that the difference in scores from time 1 to time 2 is equal to 0. The alternative hypotheses, then, are the opposites of those statement.

In the case of paired samples data, time 1 refers to the data from the half of each match assigned to group 1, and time 2 refers to the data from the half of each match assigned to group 2.

TRANSCRIPT:

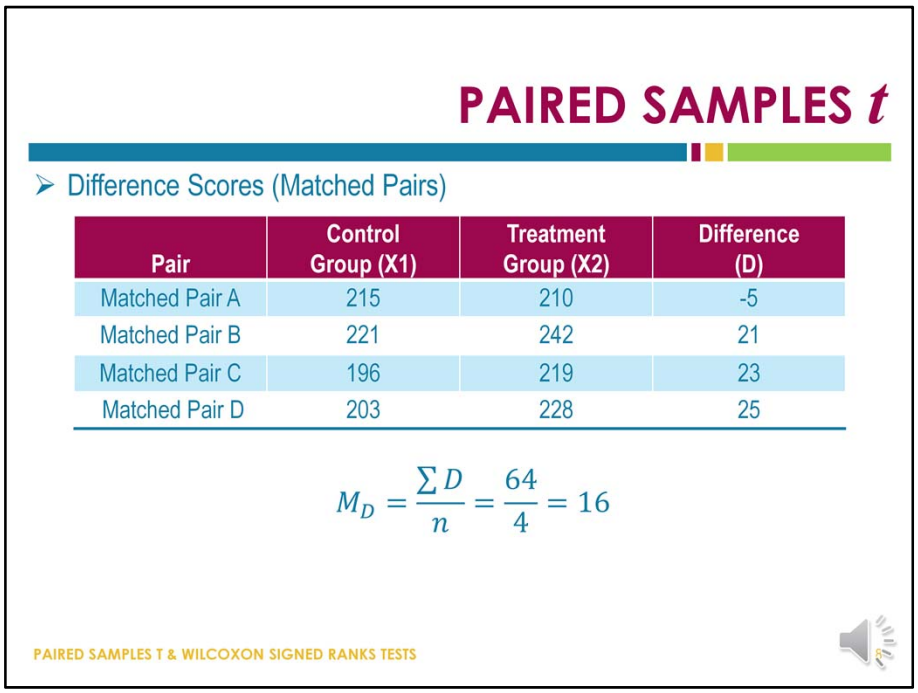
When we have data from a repeated measures or paired samples design, and those data consist of a single outcome, the most appropriate statistical methods are the paired samples *t* test and its nonparametric equivalent, the Wilcoxon test.



TRANSCRIPT:

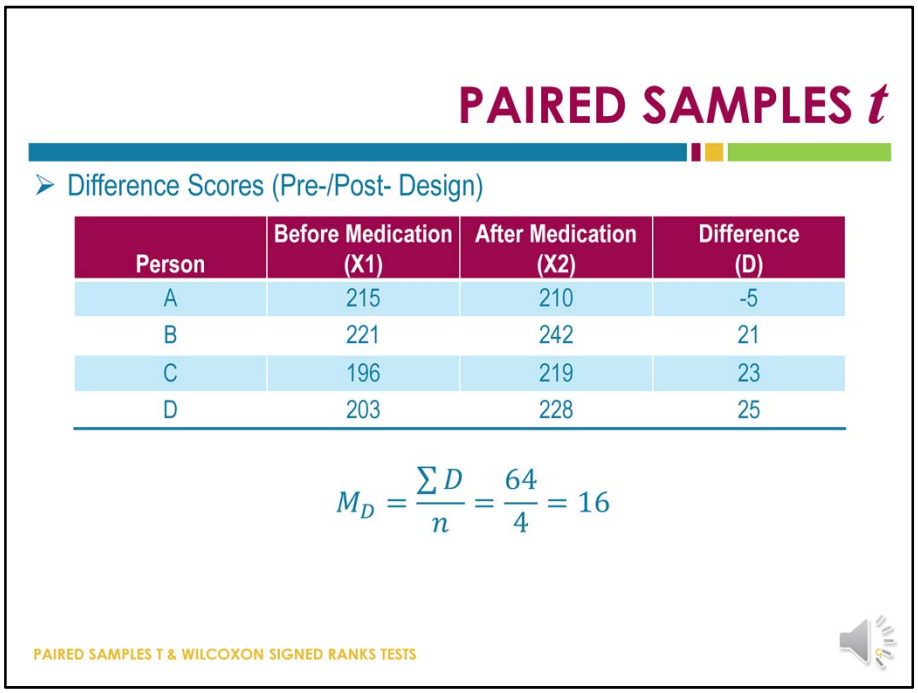
The paired samples *t* test is similar to a single-sample *t* test or z-score test. The difference is that a paired samples *t* test calculates difference scores between the two sets of observations, and the analysis is based on those difference scores. The goal of paired samples *t* test – like other inferential statistical approaches – is to use the difference scores of the observed data to make generalizations about the population.

Difference scores are calculated the same way as for other statistics, where one set of scores is subtracted from the other set of scores. Ultimately, it does not matter which variable serves as X1 or X2, but that decision does impact interpretation so it is important to know which variable is being used. It is also important to assign one variable to each conceptual label, and to not switch them back and forth within a single analysis.



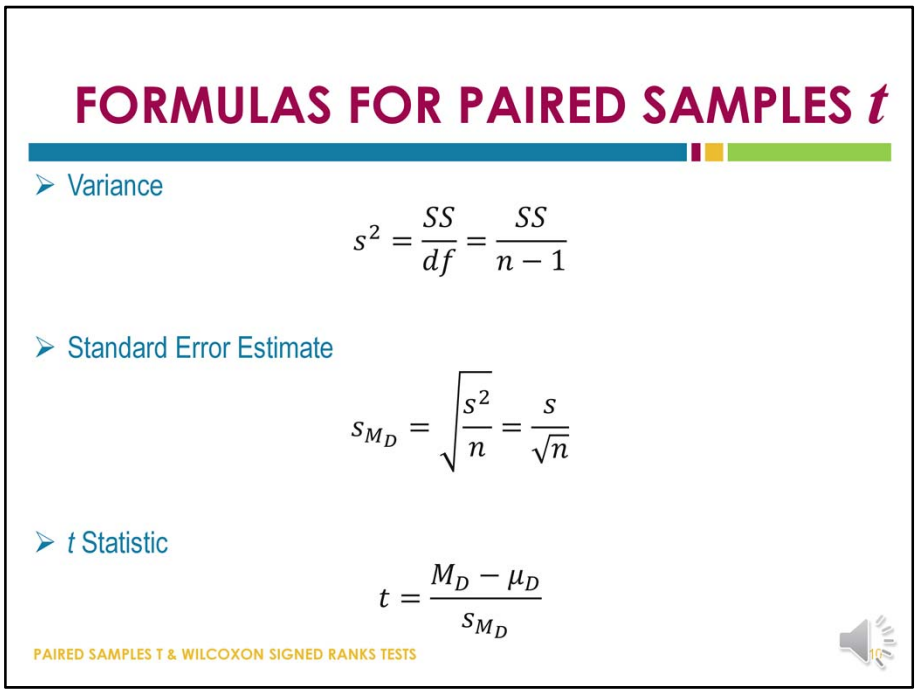
TRANSCRIPT:

Here are some example data that might be analyzed using a paired samples *t* test. These data are for matched pairs, where one variable is observations for the control group and the other variable is observations for the treatment group. The data appear on the same row of the data file because the participants have been matched on some relevant characteristics.



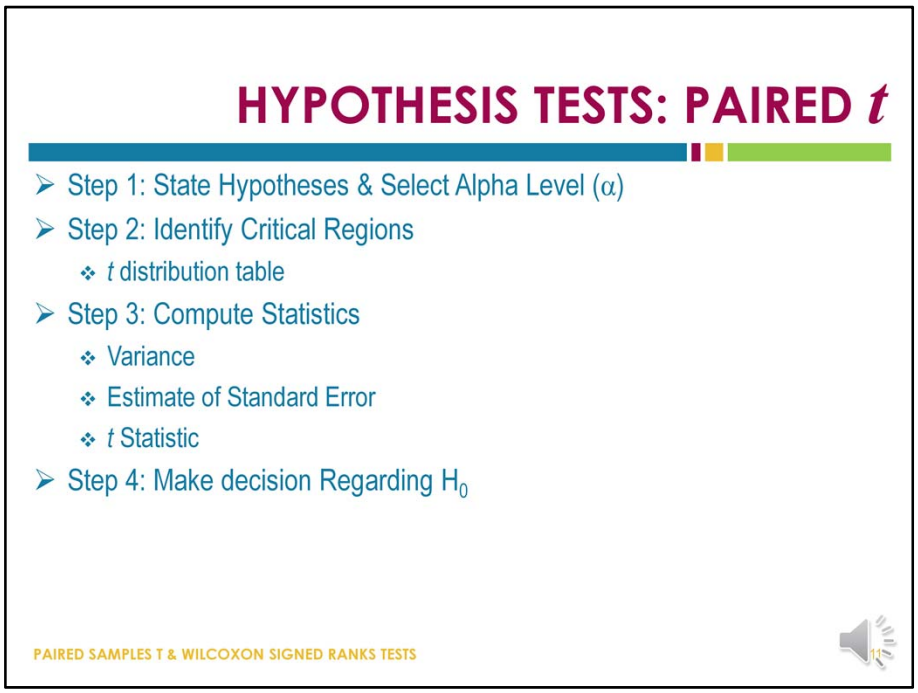
TRANSCRIPT:

This is the same set of data as on the previous slide, but now it is organized to reflect the data that might be collected in a repeated measures study. One variable is our pre-test or time 1 data, and the other variable is the post-test or time 2 data. In this example, the multiple observations appear on the same row because they belong to the same participant.



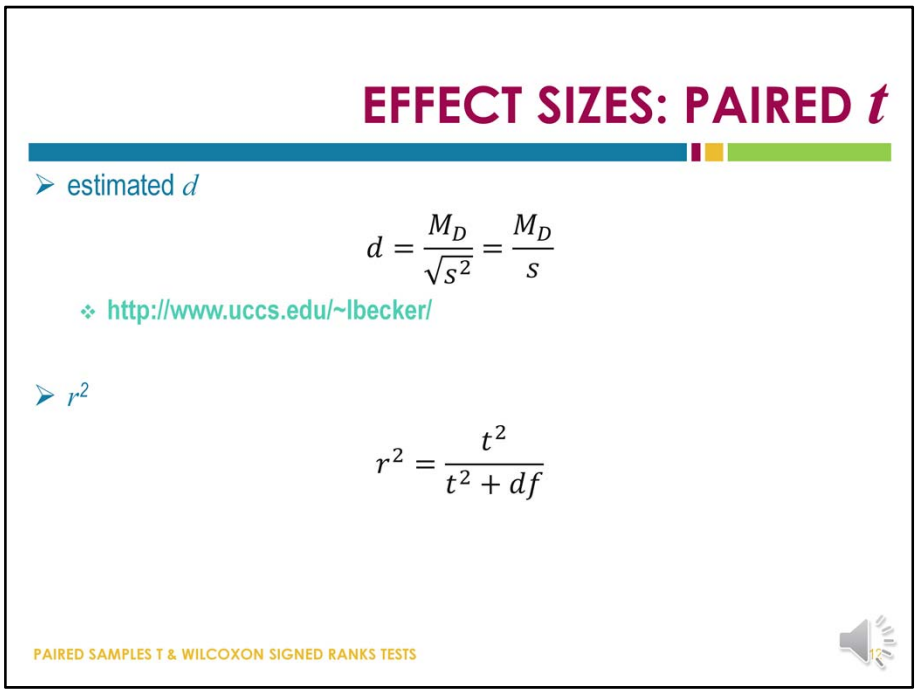
TRANSCRIPT:

Here are the key formulas that are used to calculate a paired samples *t* statistic. Calculate variance, then an estimate of the standard error for the mean difference scores, and then the *t* statistic.



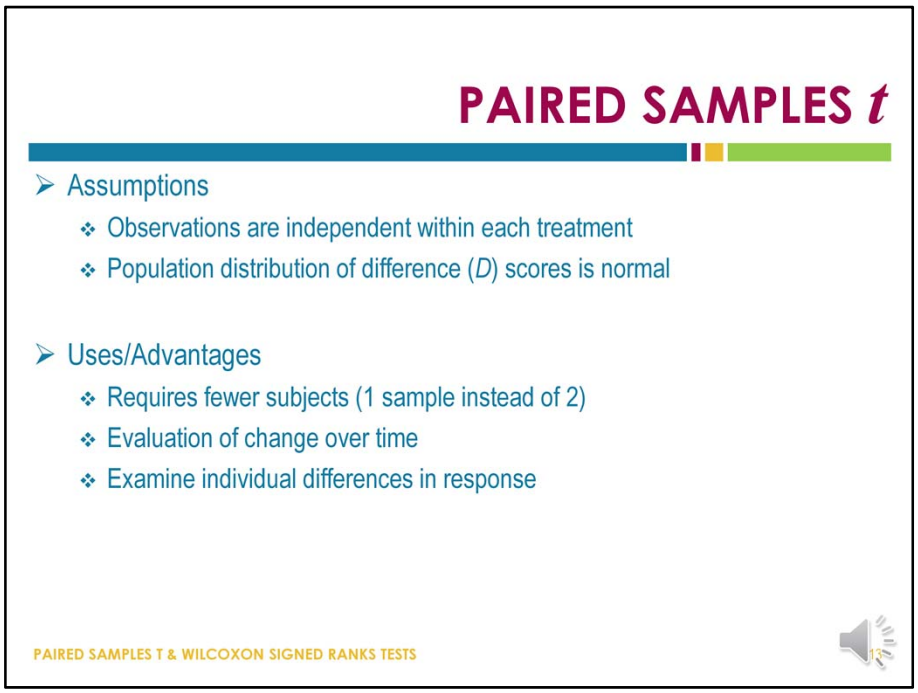
TRANSCRIPT:

This slide outlines the steps that are followed in a hypothesis test for a paired samples *t* test. This process is no different than the process for independent *t* and most other inferential statistics.



TRANSCRIPT:

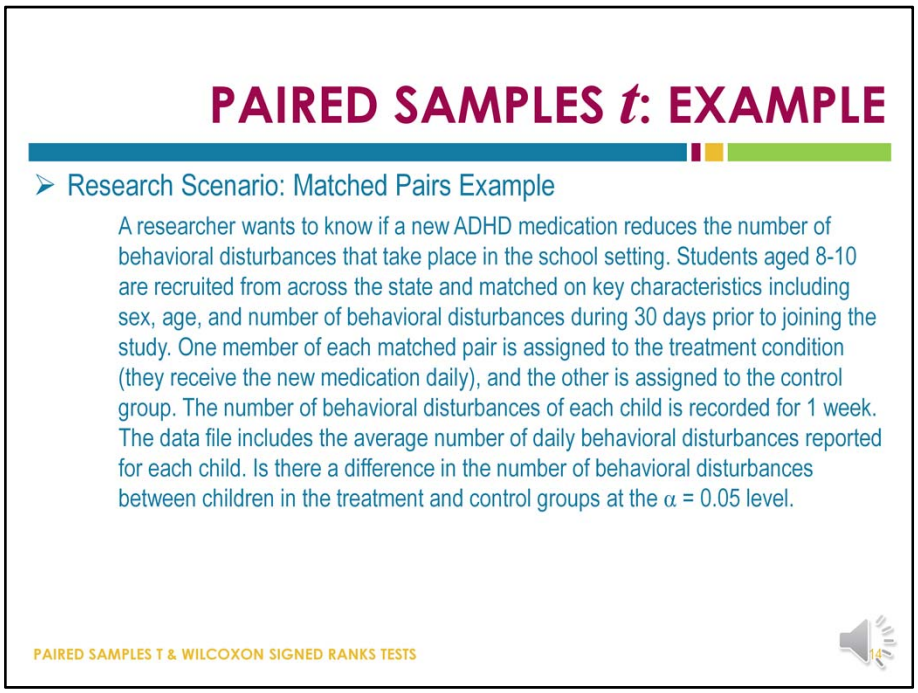
As we have discussed before for other statistics, we should calculate an effect size for paired samples *t* tests. To calculate Cohen's *d* effect size, use the online effect size calculator, and enter the means and standard deviations for each variable in the analysis.



TRANSCRIPT:

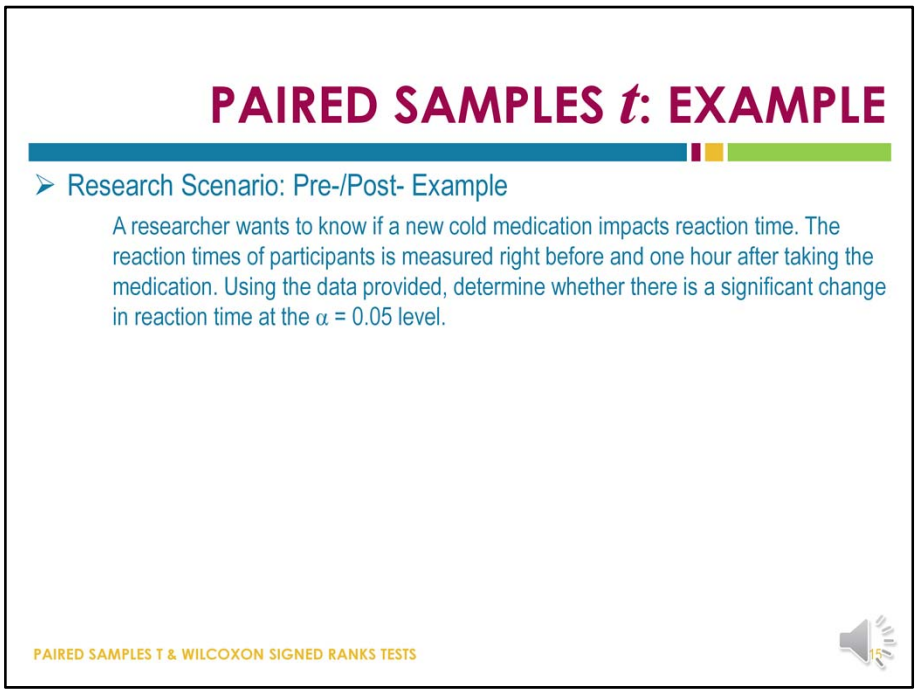
The advantages of paired samples *t* over other analytic approaches is that the research design requires fewer subjects and allows an examination of change over time.

Like other methods, though, there are some assumptions that should be met when using the paired samples *t* test. The most important of these assumptions is that observations should be independent within each time point or treatment group. This means that the value of the dependent variable for each participant should depend only on that participant, and should not be influenced by the data collected for the other participants in the sample. It is also assumed that the difference scores are normally distributed. When this assumption is not met, the nonparametric alternative to the paired samples *t* test should be used instead.



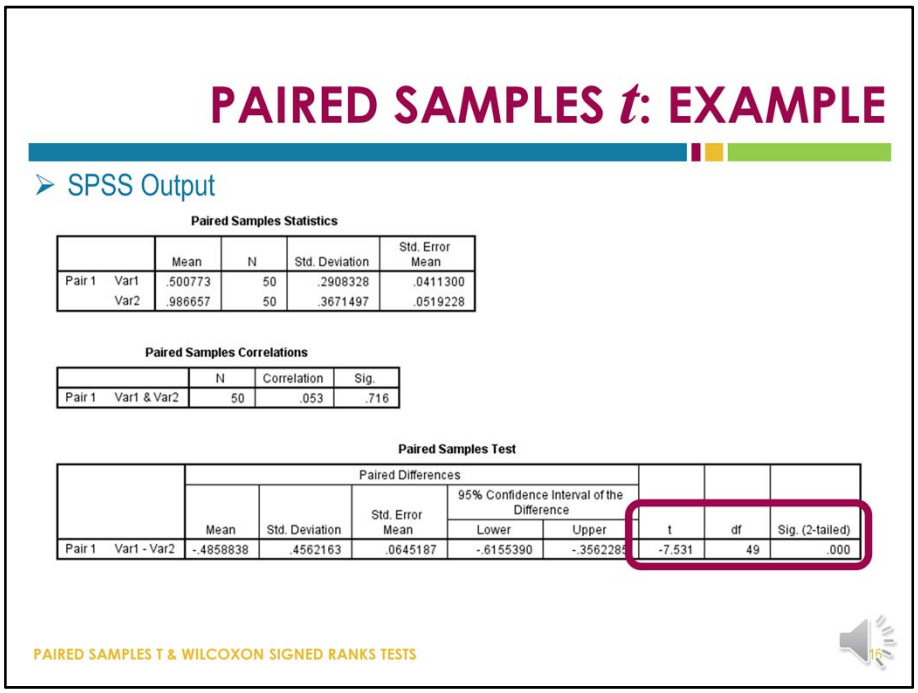
TRANSCRIPT:

Here we have an example research scenario laid out as if we are going to examine data from a matched pairs design. Notice that the sample is matched into pairs, and one member of each pair is assigned to each treatment group.



TRANSCRIPT:

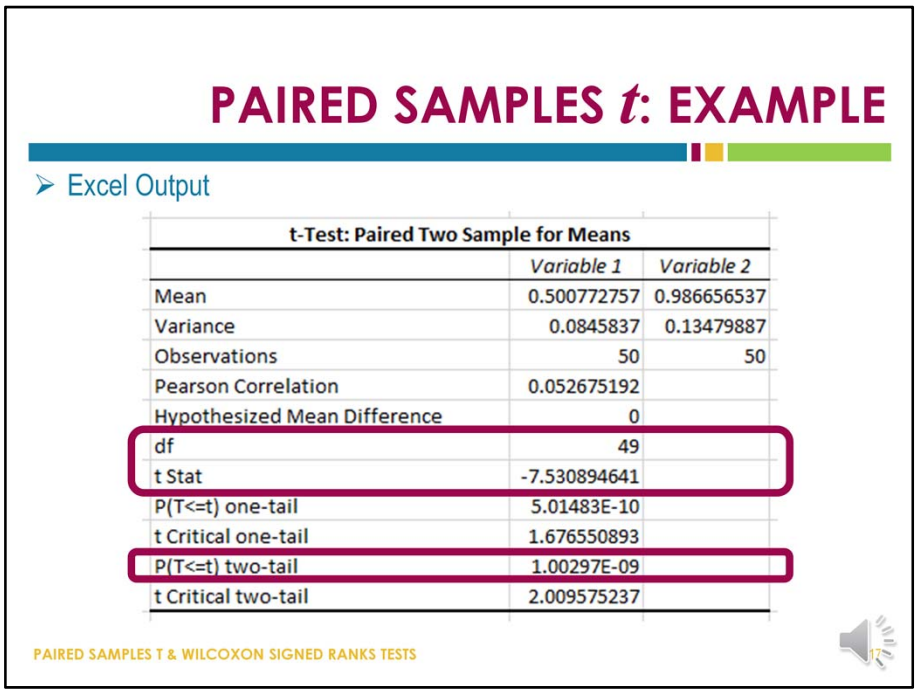
Here we have an example research scenario laid out as if we are going to examine data from a repeated measures design with pre-/post- data. Notice that each participant is measured twice.



TRANSCRIPT:

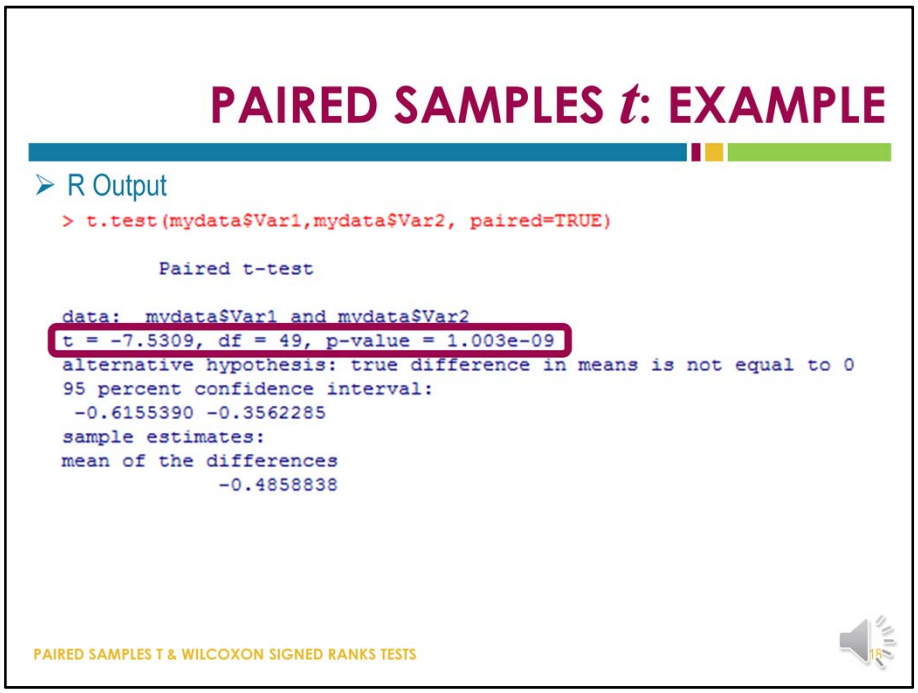
For these output examples, the data used would have been appropriate for either of the two scenarios you just viewed – either the matched pairs or repeated measures examples.

This excerpt from SPSS output includes the descriptive information for each variable, the bivariate correlation between the two variables, the t statistic, degrees of freedom, and p value. When the p value is less than the alpha level selected for the test, we determine that the test is significant and we reject the null hypothesis.



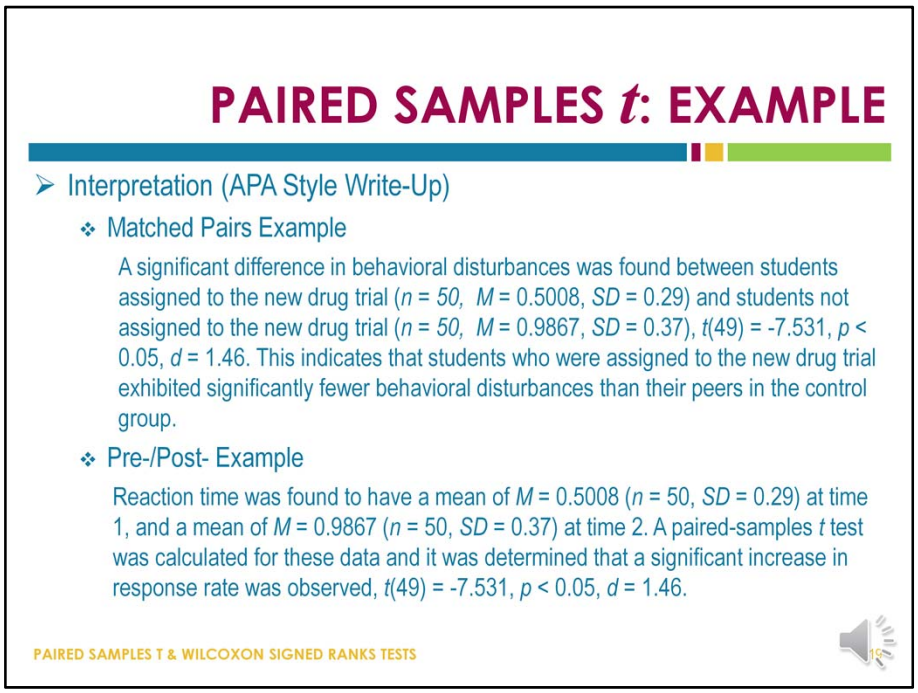
TRANSCRIPT:

Here is an excerpt of the same output when the data have been analyzed in Excel.



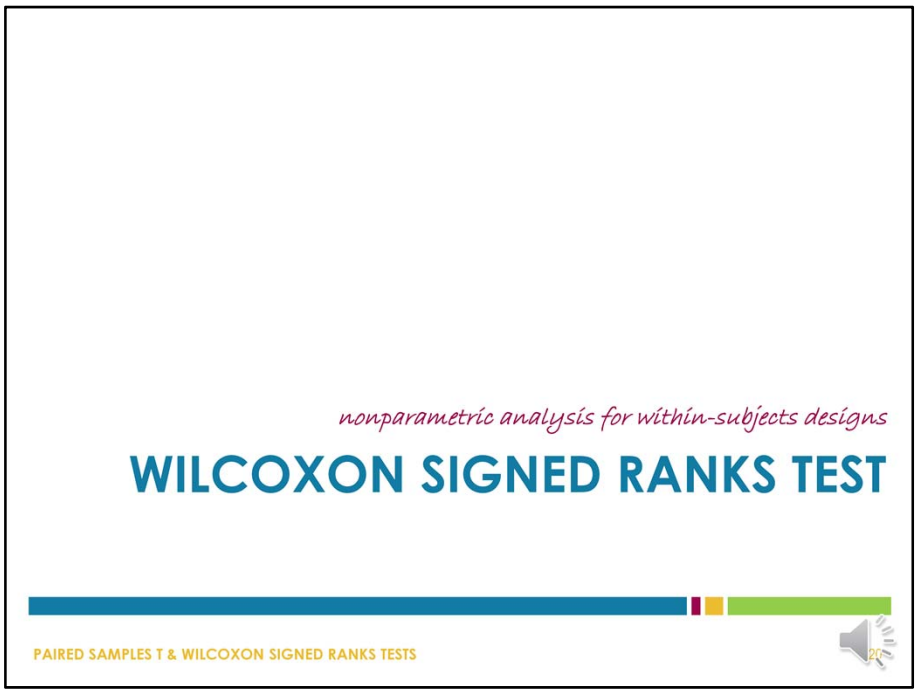
TRANSCRIPT:

Here is an excerpt of the same output when the data have been analyzed in R.



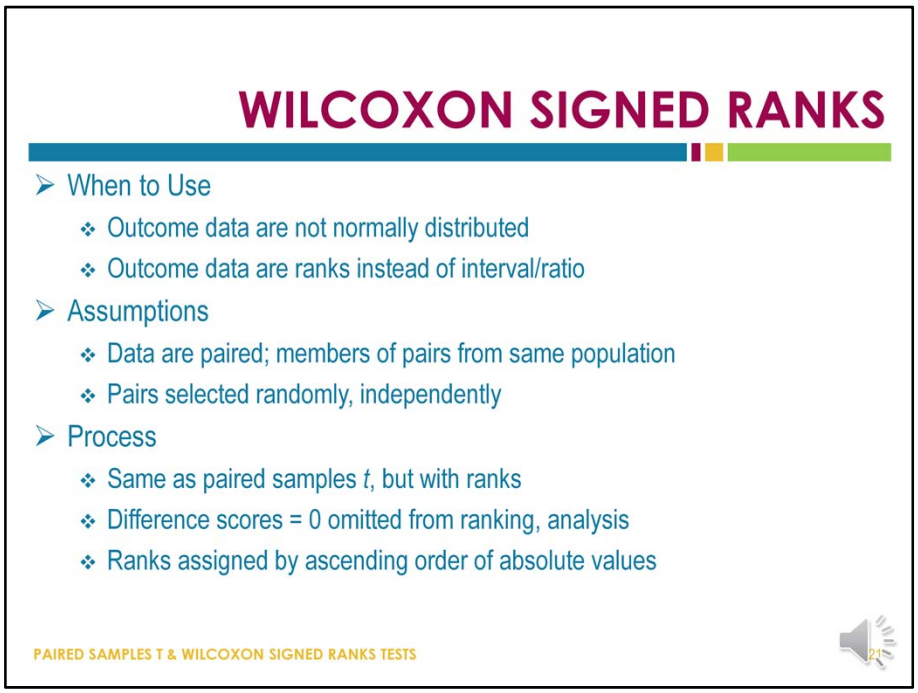
TRANSCRIPT:

Here is an example interpretation of these results for each of the research scenarios presented. Notice that the sample size, mean, and standard deviation for each variable is included, as well as the t statistic and effect size.



TRANSCRIPT:

In instances when the data are ordinal or the assumptions of the paired samples *t* test are not met, it is appropriate to use the nonparametric alternative to the paired samples *t* test, and that is the Wilcoxon signed ranks test, which we often just refer to as the Wilcoxon test.

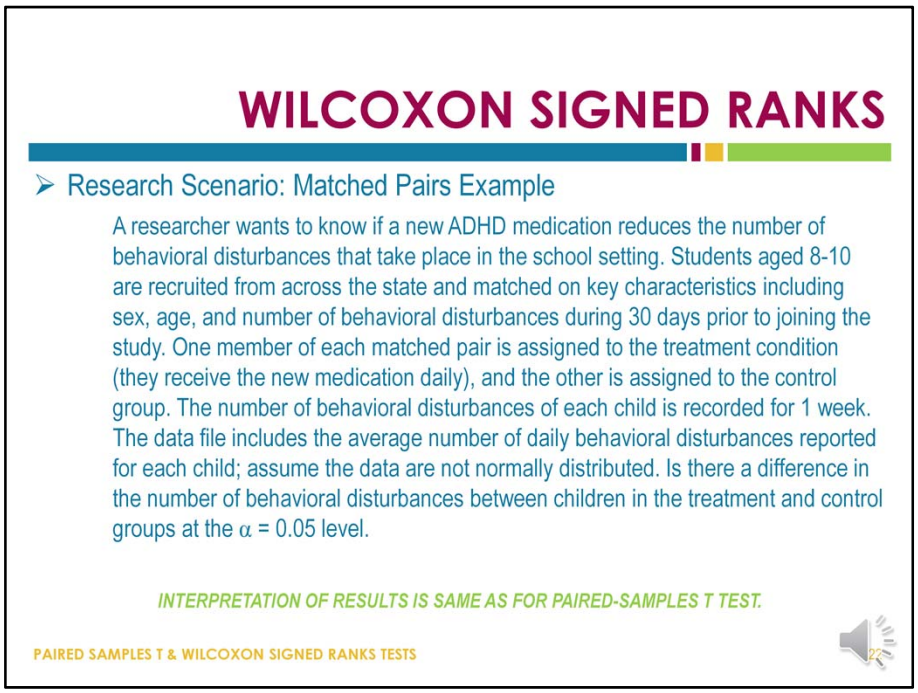


TRANSCRIPT:

As I mentioned, the Wilcoxon test is the nonparametric equivalent to the paired samples *t* test, and should be used when the dependent variable consists of ranks, ordinal data, or is not normally distributed.

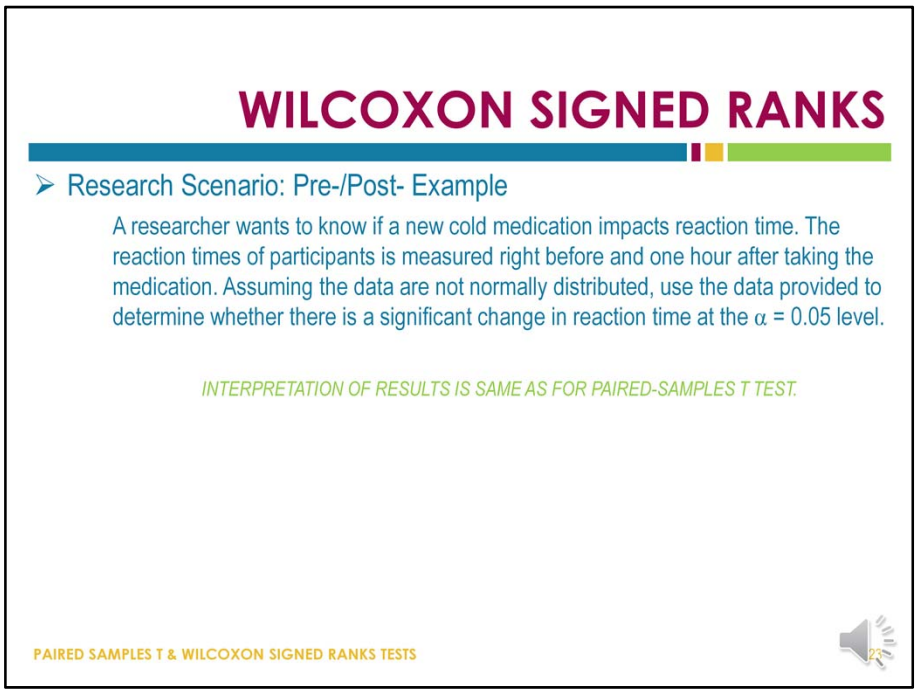
The Wilcoxon test does impose its own assumptions, and they are that the data are paired and that the members of each pair are members of a single population.

The process for calculating the test statistic and *p* value for the Wilcoxon test are the same as the process used for the paired samples *t* test. The only difference is that the Wilcoxon test conducts the analysis based on ranks assigned to the data instead of the raw data values themselves. First, the difference scores are calculated, and all difference scores equal to the value specified by the null hypothesis (which is typically 0) are omitted. The remaining difference scores are ranked in ascending order based on their absolute values.



TRANSCRIPT:

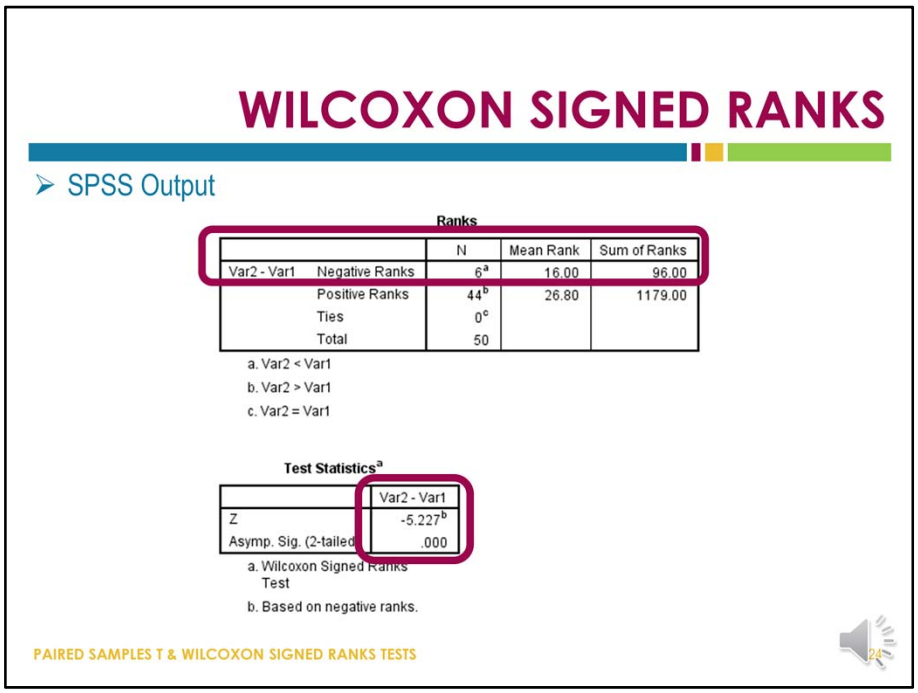
Here is an example research scenario for a matched pairs design for which the Wilcoxon test is an appropriate data analytic technique.



TRANSCRIPT:

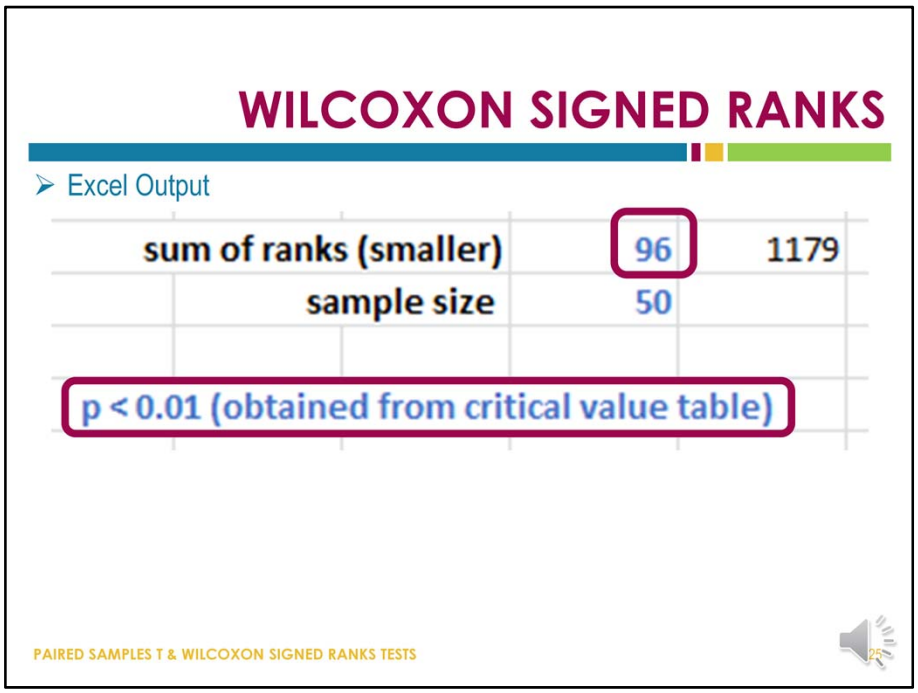
Here is an example research scenario for a repeated measures design for which the Wilcoxon test is an appropriate data analytic technique.

Once the analysis is complete, the interpretation of the Wilcoxon test results is identical to the interpretation of the results of a paired samples *t* test.



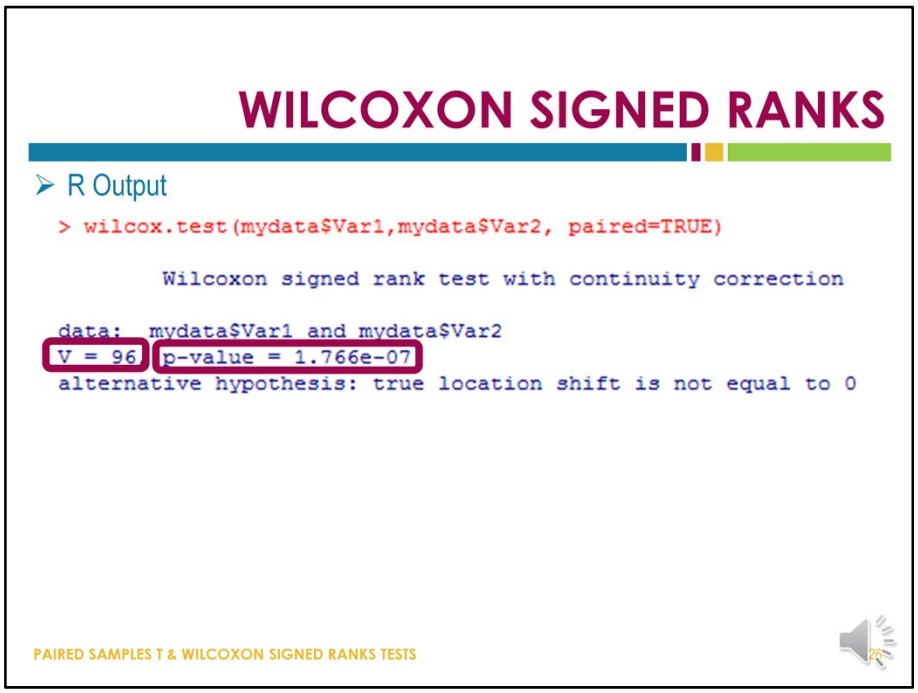
TRANSCRIPT:

This slide includes an excerpt of SPSS output from a Wilcoxon test. The test statistic and p value are calculated based on the sum of the negative ranks. If the p value is smaller than the critical value or the selected alpha value, then we conclude the test is significant and we reject the null hypothesis.



TRANSCRIPT:

Here is an example of some Excel output. The only catch with Excel is that once you have the sum of the ranks, you use the smaller of the two sum of ranks values along with sample size and then compare those numbers to a table of critical values similar to the unit normal table to determine the approximate p value.



TRANSCRIPT:

Finally, here is an excerpt of output for a Wilcoxon test computed in R.