In the Crumlish and Magel (2011) study, a t-test was used to answer one of the hypotheses in the study. As t-tests often are used in research, I will discuss this statistical test in this column. A t-test is also known as Student's t-test. Student was the pen name of chemist William Sealy Gosset, who devised the t-test in 1908. The test was developed to monitor the quality of stout by the Guinness brewery in Dublin, Ireland (Raju, 2005). A t-test is a bivariate statistical test, comparing two groups or two variables at a time (Polit & Beck, 2010).

What Is the Purpose of a t-Test?

The t-test is used to compare the means of two different groups. This usually involves comparing the mean of a control or comparison group to the mean of an experimental or intervention group. In studies that test interventions, whether an education program or a clinical intervention, investigators want to determine if an intervention made a difference. In that case, they need to compare the people who received the intervention with a group of similar people who did not receive the intervention (Williams & Monge, 2001).

The mean is calculated by summing all the values and dividing by the number of participants (in other words, an average). Mean is one measure of central tendency and is the most stable; in multiple samples from a population, the means will be very similar. Thus the t-test compares the means of the two groups and the results indicate if the differences (they are rarely exactly the same) are greater than chance alone. If the difference is greater than by chance, the difference between the two groups is considered to be statistically significant.

When statistics such as a t-test are done, a null hypothesis is being tested. That is, researchers are testing the hypothesis that there is no difference between the two groups against the alternative or research hypothesis (Williams & Monge, 2001). This is usually not stated in the article, but it is part of the statistic and assumed. The result of the test either will support or reject the null. The null hypothesis is rejected when the research hypothesis is true and therefore the intervention worked; in scientific terms, a hypothesis can be supported, but never proved. When the t-test is significant, we therefore reject the null hypothesis of no difference, and support the research hypothesis that there is a difference.

Because a t-test compares means, the data should meet some important criteria or assumptions. The data should represent a normal distribution, meaning that in the normal population (e.g., people with heart attacks) under a null hypothesis, the data would fall (or vary from person to person) in the shape of a bell-shaped curve, with very few scores (data points) at either extreme of the curve and most data points around the mean in the center. If data are not normally distributed or are skewed, other tests can be performed. In addition, the data also should be interval or ratio level of measurement, meaning the difference between one score and the next is the same as with other scores and they can be added and subtracted. In addition, an independent t-test assumes the two groups are not related in any way. If they are related, such as the case when the same people appear in each group (pre-post test), the statistic that is used is a paired t-test. A paired t-test mathematically takes into consideration the use of the same group (Polit & Beck, 2010).

What about Non-Significant Results?

Just because a researcher gets non-significant results, meaning the null hypothesis was supported, does not mean the intervention did not work. There may have been flaws or errors in the conduct of the research or, in many cases, the sample was not large enough to see a statistically significant difference between the groups. In the Crumlish and Magel (2011) study, although the first hypothesis was supported, the second hypothesis involving follow-up telephone surveys was not supported. Attrition also was a factor in the study from the first test to the second, with only 62 of the original 113 people participating in the telephone surveys.

If researchers had a bigger sample, results might have been sufficient to reject the null given results that did occur were in the hypothesized direction. A study that does not have a large enough sample is said to have deficient power. It cannot be known if results would have been different with a bigger sample unless the study is replicated. This is an example of what is called a type II error, an error in which the null was accepted when it should have been rejected (Williams & Monge, 2001). A type I error is the opposite, with the null rejected when it should have been accepted.

The t-test is an important statistical test to understand, including when it is used and what the results of the test mean. The t-test examines the means and distributions of two samples. The results of the test indicate if the null hypothesis of no difference between the means should be accepted or rejected in favor of the research hypothesis. 

REFERENCES

Lynne M. Connelly, PhD, RN, is an Associate Professor and Director of Nursing, Benedictine College, Atchison, KS. She is Research Editor for MEDSURG Nursing.