



Understanding Research: Those Scary Statistics

Most authors who publish research articles use statistics to make their conclusions. Hold on to your seats! Statistics have a way of losing even the best of readers. What we'll try to do here is give you a very simple, streamlined understanding of statistics.

In general, statistics are used to *describe* something or to examine *differences* among groups or *relationships* among characteristics. Statisticians will use terms like *mean*, *median*, and *standard deviation*.

Mean:

Mean is just a fancy word for average. It's the sum of *all* the values or scores, divided by the *number* of people in the study or group.

Median:

The **median**, on the other hand, is nothing more than the score or value that falls closest to the *middle*. Half of the individual scores are higher than the median while the other half are lower. For example, if you have five numbers: 0, 0, 5, 10, 30, the **mean** or average would be 9. ($0 + 0 + 5 + 10 + 30 = 45$; $45/5 = 9$). The **median**, however, would be 5, for there are two scores above and two scores below.

It's pretty obvious how a **mean** or an average might be used, but why do we care about **medians**? In fact, researchers use the median to give the reader more information about the mean. While the **mean** tells you what's average, the **median** tells you more about what's *typical*. Consider this example: The average price of a home in your city may be \$250,000. This is partly because the lowest price home is \$90,000 (there really are no homes out there that cost \$10, \$15, or even \$500 to figure into the average) — and partly because a small number of multi-million-dollar homes drive up the **mean** or average. In a case like this, the **median** can tell you that despite the high average, the “middle ground,” typical or median home is more like \$120,000.

The content in this document is intended for general informational purposes only and is not a substitute for professional medical advice or treatment for specific medical conditions. No professional relationship is implied or otherwise established by reading this document. You should not use this information to diagnose or treat a health problem or disease without consulting with a qualified healthcare provider. Many of the resources referenced are not affiliated with Craig Hospital. Craig Hospital assumes no liability for any third party material or for any action or inaction taken as a result of any content or any suggestions made in this document and should not be relied upon without independent investigation. The information on this page is a public service provided by Craig Hospital and in no way represents a recommendation or endorsement by Craig Hospital.

Standard Deviation:

Another statistic that you'll see in scientific research is the **standard deviation**. It tells you how spread out the data or information is. For example, imagine that you're going to have a relatively new spinal surgery. You've been told that only five people have had this surgery at your hospital, and that their **mean** (average) length of stay was 24 days. Is this all you need to know to tell your employer, or to plan for personal care needs? No way. Think about it: suppose the hospital stays for those five people were 22, 22, 24, 25 and 26 days. The **mean** length of stay is 24 days. Suppose instead their hospital stays looked like this: 6 days, 8 days, 10 days, 31 days, and 65 days. The **mean** is again 24 days. In the first example, the values are very close together, with only one day between each. In the second example, the scores are more spread out, with many days between each pair of scores. It is this spread between scores or values that the **standard deviation** describes.

Everything we've described so far falls into the realm of *descriptive* statistics. They are intended to give you good information on what the subjects and the data look like. Is the typical subject anything like you? Are the means and medians similar? Are the findings very spread out?

Showing Relationships: **Statistical Significance**

The next level of statistics attempts to show *differences* between two or more groups, or *relationships* between two or more different things. Suppose, for example, a researcher wants to show that a new medication is effective in reducing the number of bladder infections. Even if those who take the drug seem to have fewer infections than those who do not, there's always the possibility that any improvement that *does* occur is because of chance. Also, there's always the possibility that anything that happens does so because of some factor *other than* what the researcher was studying. To lessen the effects of these possibilities, statisticians do two things:

1. They test their findings – those means and averages we talked about above – for **statistical significance**.
2. They try to *control* for other factors outside of the ones they're studying.

When researchers test for statistical significance, they compare different sets of values – such as bladder infections before and bladder infections after using a medication – while taking into account how many people participated in the research, how dramatic their findings seem to be, and what some of the characteristics were of the people they compared. They then use complicated mathematical formulas to calculate *probability* values. For our bladder infection drug, this probability value will tell us how likely it is that people in the study got better because the drug did its job, or whether they got better simply because of chance or due to some other unknown factor other than the drug.

If the researcher finds that the probability value is low (usually less than 5%, 1% or even 1/10th of 1%), he or she can conclude that the drug really does work. These probability values – called **p** values – represent percentages, but are typically represented as $p < .05$, $p < .01$, or $p < .001$, or more specifically, as $p = .023$, and $p = .0067$. For example, $p < .01$ means that there is less than 1% chance that our bladder medication seemed to work because of chance alone. If probabilities are low, researchers describe them

Research: Statistics

as **statistically significant**. These are key words you should look for as you read. Remember: the *lower* the **p** value, the smaller the percentage, the *greater* the significance, and the *less* likely that something happened just because of chance.

How significant is significant *enough*? It depends on what's being studied, the potential benefit or harm of the findings, and the author's own standards. If that bladder drug is relatively cheap and safe and your bladder infection isn't life threatening, the small chance that the drug doesn't really work as effectively as the researchers had indicated (in other words, a *significance* of $p < .05$, or a less than 1 in 20 chance that it works) may not be a big deal. If, instead, that drug is a treatment for cancer and causes severe side effects, you want to be very sure it's extremely likely to work before you take it, so maybe only a significance as low as $p < .001$ will be good enough.

Showing Relationships: *Control*

Then, there is the issue of *control*. Researchers can control for other factors that might affect their results. Keep thinking of our bladder infection study. Do the researchers know before they start if the person's age or gender has any effect on how they react to the drug? What about the type of programs the study participants use to manage their bladders? Or how much they drink or whether they take Vitamin C? To deal with issues like these, researchers collect extra information on each study participant. Then, they include this information into their statistical analyses and learn if their findings still hold in light of these possibly complicating factors. In their articles, researchers usually will tell you which factors they *controlled for*. If *you* can think of factors that might have had an impact that the researcher failed to control for, you probably should interpret his or her findings with a bit more caution.

"Proof"

Always keep in mind that **proof** is a very strong word in statistics. Researchers – especially in disability and health related research should not typically be telling you that they have *proven* that one thing causes or leads to another. Instead, the more likely message should be that one thing is *related* to another. Be wary. It's easy to unintentionally (or sometimes intentionally) mislead the reader – especially when only a small amount of information is given. An example will make this clearer. We found that among aging SCI survivors, people who were coffee drinkers at one point in time had more shoulder pain three years later. This sounds like a situation where one thing causes another; the two are even separated by time. But this still is only a *relationship* – coffee drinking was somehow *related* to shoulder pain. This finding could have appeared for a number of reasons – maybe people who drink coffee are more active, maybe their sensation of pain is heightened; maybe there is some other connection that we just haven't thought of. Or, this could simply illustrate that weird findings just happen. Remember when we talked about how $p < .05$ means that there is a 5% possibility that the finding occurred purely because of chance or a fluke? Well, for all we know, this relationship between coffee and pain could fall within that 5%! That's why with really important data – more like the cancer drug example we used earlier than with the coffee-drinking quirk we're describing here – good researchers *replicate* their findings. They conduct whole research projects a second time, with the same methodology, to see if they get the same results.

Research: Statistics

Validity

The last concept to talk about here is called **validity**. It's best explained with an example.

John Q Researcher, from the nation's top university, reported that his latest study showed a **significant relationship** between level of spinal cord injury and intelligence: People who have paraplegia, he said, are much smarter than people with quadriplegia. How did he come up with this finding? He used the best and most popular intelligence test available. All his subjects were of the same age and all had the same level of education. All, of course, spoke English and could read the test questions easily. He followed the test's rules for administration, and tested everybody together in the same classroom to make sure the testing situations were equal. He had helpers to make sure no one cheated. Everyone had two hours to complete the test. John Q. corrected and analyzed the tests, and sure enough, those quads just didn't cut the mustard. His conclusion that people with quadriplegia just aren't as smart made all the newspapers, and landed him on a few talk shows.

But – there's one little point here that he missed. Can you guess what it is? The test was *timed*. Can people who have quadriplegia write as fast as those with paraplegia? No. Some can't write at all. So, in truth, it's very likely that it wasn't their *brains* that accounted for their poor test scores, but their *arms and hands*. We're being flippant here, but this is a very important concept. *John Q. wasn't really measuring what he thought he was measuring*. He believed he was testing intelligence; instead, he was assessing writing speed. His research *wasn't valid* – it didn't measure what he said it did. How could he have made his research valid? He might have given all his subjects an oral test; he could have given all more time. Keep this concept in mind – as well as the ones we described above – and you'll be on your way to critiquing those research reports you hear and read about.

This brochure originally appeared as an article in the 2001 issue of SCI Life (pages 32-33), titled "Understanding Research: Part III: Don't let the statistics scare you!" It is reprinted here with the permission of the National Spinal Cord Injury Association. This brochure was prepared with funding from the U.S. Department of Education's National Institute on Disability and Rehabilitation Research. The opinions contained in this publication are those of the grantee and do not necessarily reflect those of the US Department of Education.