FRESENIUS KABI
caring for life

Dummy Guide to Statistics

Author: Conor Kerly (PhD, BSc, H.Dip, MINDI)

June 2017

This article is intended to assist dietitians in understanding basic statistics and performing their own statistical tests. I confess that I am no statistical expert but perhaps that is a good thing for the purposes of this article. Additionally, there can be no substitute for actually performing statistical tests (or at least observing them being performed). Therefore this article will serve as a basic introduction.

Most people don't understand and hate statistics (me too…to a point!). However a basic understanding of statistics can help you read a paper/series of papers and pass judgement on whether the observations are trustworthy or not.

Why do we need statistics?

In science we make observations based on either observing life as it happens or based on asking people to adjust their life in some manner. Statistics help us determine if our observations are likely due to chance or not.

P-values

Generally in science, a p-value of <0.05 is used. If a p-value of 0.05 is reported, this means that the probability or 'p' of an event occurring due to chance is 5%. i.e. 95% not due to chance. It is important to understand that even with a p-value of <0.05, the event may still be due to chance (i.e. 5% chance) and that's a major reason why in science we look for consistency i.e. multiple studies under similar and difference conclusions demonstrating the same thing. In science we also look for a mechanism(s) to explain results. So if we have consistent statistical significance in multiple studies as well as at least one mechanism, we can be more confident in that the observations are real. However, nutritional science is complicated and when we have inconsistent results and statistics, we can't be so confident the observations are real.

Statistical significance vs. clinical significance

It is important to understand that statistical significance is not the same as clinical significance! Statistical significance merely means that there is less than 5% probability that an observation is due to chance. Clinical significance means that the effect is clinically meaningful. The p-value is important, but sometimes there is an overemphasis on the p-value. Ideally, we want both statistical and clinical significance!

For example, if a new expensive, monotonous diet was proven in multiple trials to statistically significantly decrease body weight, would you recommend it to your obese patients? We have statistical significance, so maybe it would be a good recommendation. However, what about if the weight loss was 50g over 6 months? Not very clinically significant. Vice versa: if a new diet was studied and did result in weight loss, but the results were not statistically significant, would you recommend it? We don't have statistical significance. But what about if the weight loss was 10kg in 6 months?

In this scenario, it is likely that the result would in fact be statistically significant but not always, for example if the study included only a few participants.

Sample sizes, effect sizes and power calculations

An important consideration when reading a paper or designing your own research is what is/should be the sample size. If the sample size is too small, you may observe a clinically significant effect but non-statistically significant. However, if a study recruits more subjects than needed, it is a waste of resources, including the researchers and the participants.

In general smaller studies are more detailed while larger studies are less detailed. Some people are seduced by the size of a study or the length of a study but if the methods are not very detailed, the observations are dubious at best. For example: lets say there is a study of 1 billion people. At the start they were asked to weight themselves and asked what types of fruit they eat. One year later they are asked to weight themselves again and the researchers observe that those eating berries gained the most weight...can we believe that? On the other hand, let's say there was a lab based study performed with 20 adults living in a research lab for 4 weeks. Everything they ate and drank was measured and detailed assessments of body composition were conducted at the start and end of the 4 week period. The people who were fed the most berries lost weight…which study can you believe? It would be impossible to do a lab study with 1 billion subjects! Therefore clinicians and policy makers must make sense of small, detailed studies and larger, less detailed studies!

The effect size of an intervention will have a major determinant on what the sample size should be. Let's say diet A is thought to result in large weight loss (a large effect size), we would need a very large sample to determine if the diet worked or not i.e. we think diet A can reduce weight by 10kg in 6months. We do a power calculation and find that we need 40 subjects in a randomized, controlled trial i.e. 20 to diet A and 20 to control diet…after 6 months we compare the 2 groups and either:

1. the diet worked and there was statistically significantly greater weight loss with diet A (i.e. there is less than 5% probability that the weight difference between groups is due to chance)
2. the diet didn't work and the weight difference between groups is not statistically significant (i.e. there is greater than 5% probability that the weight difference between groups is due to chance).

In this type of study, if less than 40 participants were included the result could be statistically non-significant but only due to an inappropriately small sample. Conversely, if we recruited more than 40 subjects the weight loss could be modest but still be statistically significant.

If a trial has a small sample but reports a statistically significant effect, this means that the intervention was powerful i.e. had a large effect size. In contrast, if a trial has a very large sample and the result is barely statistically significant, this means the intervention does seem to have an effect but that effect is modest. The good news is that there are some dietary trials showing large differences with statistical significance. The bad news is many people criticise these studies because the samples are small!

Absolute risk vs. relative risk

These common terms are used widely when reporting scientific studies, usually large scale, observational studies. However, these terms are not interchangeable and can be used inappropriately. To help illustrate, I will use an example: Among a given cohort, adherence to the DASH diet may be associated with a cardiovascular event in 1% while the Atkins diet may be associated with a cardiovascular event in 2%. In this scenario, the absolute increase in risk with the Atkins diet compared to the DASH diet is 1%. However, the relative increase in risk is 50%.

Do you have access to a statistics program?

There are many statistics programs (e.g. SPSS, SAS, STATA). Most third level institutes and some hospitals will provide access to at least one of these packages. Further, one of your colleagues may have personal access either through their affiliation or through their own personal account…ask! Some basic statistics can be done on Microsoft Excel and this can be a great starting point.

Simple descriptive statistics are an easy way to start i.e. if you're looking at cholesterol levels:

- How many people do you have results on
- What are their characteristic e.g. age, gender, BMI etc.
- What is the average (mean) cholesterol level
- What is the standard deviation
- What is the highest (maximum) and lowest (minimum) cholesterol value? If there is a large difference between the maximum and minimum this means there is a large spread in the data.

Is your data normal (or parametric)?

In statistics, one of the best places to start is simply to assess the normality of the data i.e. is the data parametric (normal) or non-parametric (non-normal)? This can be done on most statistical packages or by either eye-balling or plotting your data. Mini-cheat, if the standard deviation value is close or greater than the mean value, it's very likely that your data is nonparametric. On the other hand, if the standard deviation is much smaller than the mean value, it is likely that the data is normally distributed. Whether your data is normal or not will determine what types of statistical tests you should do.

Further resources

Colleagues

Speak to your colleagues, particularly anyone who has research experience. Some long-time clinicians have great knowledge of statistics!

Youtube!

Youtube actually has some great statistics videos. If you have a specific problem, Youtube can be a great starting point As an example, type in: how to perform a t-test on excel and you'll find many videos showing you step by step how to set up and perform the test!

CSTAR

Professor Leslie Daly has set up the Centre for Support and Training in Analysis and Research (C-STAR). C-STAR run courses all many universities throughout Ireland. Further, C-STAR has been used as a consultancy (for a fee) which can be handy if you're applying for a grant or writing a paper.

Post-graduate courses

If you're so inclined, several universities offer post-graduate courses in statistics.

Online courses

Similarly, there are several reputable online courses (some free) which will help guide you regarding statistics.
A combination of Youtube and colleagues is a great starting point for most.

This article was written by Dr. Conor Kerley (PhD, BSc, H. Dip, MINDI) as requested by Fresenius Kabi in June 2017.