Editorial: C-value

Ferrás-Tarragó J

The sample size (n) is one of the most important parameters of a study, and as a general rule, It is considered that the higher n a study has, the higher its quality and the more likely the The results obtained are accurate, but if the meaning of the p value and the size of the sample, it will be seen that, when there is evidence with statistical significance, the lowest n is the most clinically significant.

The type of study that is most frequently carried out is the comparison of one technique with respect to another, to see if the new technique is better or not than a previous one. That happens in surgical techniques, diagnostics, medicines... in many fields. To do this, we compare the results obtained by the two techniques and a significance level a of the comparison is established, that is, it is assessed what probability we have that the differences found between both techniques are due random. If the probability that chance is responsible for the differences is less than 0.05 (value of significance a), it is usually affirmed that the differences are not due to chance, and therefore, that a technique is superior to the other, it which is a idea widely assumed in he field scientist.

The p value takes into account the n. A small n, often gives a large p-value. This is logical, yes. that the smaller the group that we study, the more certainty one has to affirm that these differences are "real" in the world, and not the reason for the error when selecting the sample to which we study (n). For this reason, it is normal that a small n does not allow us to affirm the differences between groups, but when p is less than the significance level a (and the study is methodologically correct), regardless of the number of n, it can be assumed that the differences found are not by random, with a probability of mistake assumable. Increase the n in these cases, allows to reduce the risk of error in the statement, but should not change the direction of the indication, or in other words: will reduce the p by 0.02 at 0.003, which means that the probability that these differences be by chance they will go from 2% to 0.3% which, although it is a statistical improvement, is not taken into account when accepting the statistical significance of the results. Furthermore, and to do so correct, the p that is assumed to be sufficient in the hypothesis test is also defined in the sample size formula, so reducing the p value should be done with the size calculation sample, defining of beforehand that worth p come on to estimate enough.

Another element contributed by the increase in sample size is the probability that small differences between both groups show differences mathematically significant, but No clinically significant, so may actually reduce the clinical utility of the study. said in In other words: if you take the entire world population for a study, any difference, for tiny that out, I would be statistically significant, although No out of relevance clinic.

And here comes the big key: can to plan in the design of the study that the differences statistically significant be also clinically significant with a worth mathematical?

The calculation of the sample size is one of the most important steps in a study, and the one that many sometimes it is not done correctly. It is the most important because it is the one that allows connecting the significance math and the significance clinic, thank you to the formulas of size sample.

When calculating the sample size, what you are actually doing is defining what number of patients is needed to see a difference that is defined prior to beginning the study of according to prior knowledge of the behavior of said variable. In other words, it what is done is establish how many people need study for that in the case that there is a difference X, definite in the methodology of the project, it could be detected.

To calculate the sample size in studies that analyze a quantitative variable (which are the most frequent), it is only necessary to know the physiological variance of the variable that we are going to analyze, the p-value that is assumed and the precision with which we want to determine is variable. If we see this from From a hypothesis testing point of view, to calculate the n we would need to know the variability of the difference of the main variable of the study between both techniques, the level of error (p) that is assumed (generally 0.05) and the precision with which we want to detect the differences (that is, the value of the difference that I want to analyze). Variability is an element that can be controlled, since it depends of the behavior of the organism before a technique, and it must be based on the bibliography to know as it behaves. The p-value can be selected by the researcher prior to the start of study, forks the level of error that the author assumes as acceptable for the definition of his hypothesis test, being generally 0.05, although it could be defined that the author considered, depending on the risk of mistake that HE want assume.

We can control the precision, and it is a fundamental element in the sample calculation, since is the only parameter of the denominator and is raised to the second power, so Small changes in precision will greatly alter the size of the sample that we need, the n being necessary elderly how much further precision we want.

Thus, if one analyzes two rotator cuff reconstruction techniques (for example), using the CONSTANT scale, it can be defined before the study what precision we want to use in depending on its clinical relevance. In other words, if a difference is established in the CONSTANT of 10 points between two techniques, the statistical test is probably being provided for the analysis of contrast of hypothesis of the power enough for detect differences of 10 points either further.

If with the sample size obtained it is not possible to obtain statistical differences differences between both techniques, it can probably be said that the differences we observed are less than 10, so they are not clinically relevant, although it cannot be said that there is no differences, because if I know will increase the n, surely the differences would significant, but lower at the level of clinical significance that has been defined before starting the study (in our example, 10 points).

Thus, one must be critical of the p value, and understand that statistical significance is not the clinical significance, but that there is a useful and simple way to equate in the preparation phase of the study clinical significance and statistical significance, being able to obtain the c-value, or "clinical" value prior to the study, defined as the precision value that we consider clinically relevant before of the study, to use n sufficient to detect likely clinically significant differences if there are, and rule out studies that find differences statistically significant without significance clinic.

This reasoning is not prepared for the demonstration of equality between groups, since for this, non-inferiority tests should be used, for example, since assuming absence of differences with a level of significance $p > a$ can always incur in a mistake type II. Nevertheless, the definition previous to the study of the level of clinical significance when determining the sample size, allows an intersection between clinical and statistical significance, being more critical when detecting differences clinics and statistically significant.