

How to measure the precision of clinical study results

Como mensurar a precisão dos resultados de estudos clínicos

When one makes up a clinical study, after posing a question to be answered by means of the data obtained by research, one can define the population to be studied. As it is not possible in general to evaluate all the population studied, one can select a sample. Sampling is when the analysis is performed based in a (representative) part of a population. In this stage, it is important to bear in mind that the only form of guarantee the representativity of that sample will be doing the selection randomly, so each element of the population will have exactly the same probability of being selected, without suffering influence on the part of the researcher ⁽¹⁻³⁾.

In randomly selected samples of a population it is possible to measure the veracity of the obtained results, by means of tools like the calculation of confidence interval (CI). We generally define 95% CI as ideal. It means that with a precision of 95%, the percentage value of each studied variable corresponds to the real value of the population percentage. Nevertheless, we have to take into consideration that even that interval is not an assurance, because there is a confidence of 95%, or better, there is a 5% probability of the actual percentage in the studied population to and out of those limits. The bigger the sample, the shortest will be the confidence interval and, thus, the greater will be the probability of having precise extrapolations of the real percentages of the population. Small samples generate imprecise data.

That is why it is important to calculate the dimension of the sample, so one can give the precision he wants to the results. The dimension of the sample is related to the precision of the desired confidence intervals, in order to extrapolate the sample results for the population. But, in case the sample is not randomly selected, it does not make any sense to apply the confidence intervals, for the sample in that situation will not be representative of the population.

Thus, the CI helps verify if the size of the sample was sufficient and if obtained results came from the intervention or randomly (the narrowest the CI the greater will be the certainty that the sample represents the population). Besides the size of the sample and the CI, the “p” value (significance level) helps define the precision of the results of the study. “P” represents the probability of the differences between the results found having occurred randomly and not because of the intervention. In general, one considers the value of “p” < 5% or 0.05, so the chance of the results not being real is minimum, less than a chance among 20.

One can consider that “p” indicates the possibility of detecting a difference that does not exists (positive false). This type of error (type I) is not good for the patients, because in the case of studies dealing with new medications, it may happen that a drug which is not efficient can do good, injuring the users who will not be adequately treated. False negative errors (type II), which correspond to the possibility of not recognizing a difference that really exists, are largely more tolerated up to 20%, because, in this case, the injured are the pharmaceutical industries. Those industries develop an efficient drug, but it is considered inefficient by the studies. Therefore, this drug will not be used.

Consequently, sampling is a statistical technique that means extracting a part (sample) from the whole (population), with the objective of evaluating certain characteristics of that population. The validity of the study is directly linked to randomization, to the size of the sample, to the indicators of data precision (CI and “p”) and to the magnitude of follow-up loss. It is important that readers analyze those items in order to evaluate the possibility of extrapolation of sample results for the population to be studied.

Researchers and readers should also observe the fact that in very big samples few effects can be considered statistically significant, having “p” < 5%, though differences are clinically irrelevant. To such an extent, readers

should also consider the relevance of clinic characteristics of the results, taking into account the critical analysis of the articles, besides statistic data.

Another factor of confusion is the fact that, depending of the quantity of statistic tests for data analysis, some may have positive significance while the great majority is negative. Because the level of significance utilized is generally $< 5\%$, one expects that one case in each 20 is falsely positive. If the researcher decides to force a positive result, he can, for instance, perform several statistical tests and consider the only that was positive.

In this way, researches that incur in methodological selection errors, those of gauging or data interpreting, tend to create divergent results in similar studies ^(4,5). That is why it is important that the reader consider possible interest conflicts registered by the authors and also evaluate several similar studies before developing an opinion about a certain subject. That is the ideal way of making a critical analysis of literature.

Newton Kara-Junior

Editor-Chefe da Revista Brasileira de Oftalmologia

Professor Colaborador, Livre-docente e Professor do Programa de Pós-graduação,
Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil.

REFERÊNCIAS

1. Siqueira RC. Pesquisa translacional na oftalmologia: o caminho para a medicina personalizada. Rev Bras Oftalmol. 2012; 71(5):338-42.
2. Kara-Junior N, Espindola RF, Gomes BA, Ventura B, Smadja D, Santhiago MR. Effects of blue light-filtering intraocular lenses on the macula, contrast sensitivity, and color vision after a long-term follow-up. J Cataract Refract Surg. 2011;37(12):2115-9
3. Kara-Junior N, Jardim JL, de Oliveira Leme E, Dall'Col M, Susanna R Jr. Effect of the AcrySof Natural intraocular lens on blue-yellow perimetry. J Cataract Refract Surg. 2006;32(8):1328-30.
4. Chamon W. [Plagiarism and misconduct in research: where we are and what we can do]. Arq Bras Oftalmol. 2013;76(6):V-VIII. Portuguese.
5. Chamon W. [Fine prints at the bottom of the page]. Arq Bras Oftalmol. 2013;76(3):v-vi. Portuguese.