Role of biostatistics in medical research

Medical students, especially in their early years of study, always ask a question: why do we study biostatistics? We are going to be doctors, not statisticians. On the other hand, graduated doctors feel day by day that they are in need of better understanding of statistics, not only as researchers, but as readers of medical literature. This gap in knowledge is partly attributed to the curricula of some medical colleges since biostatics was not emphasized upon previously. With the introduction of “evidence based medicine” as a cornerstone for patient’s management, more importance was given to biostatistics, study designs and research methodology. Guidelines of evidence based medicine mainly depend on the results of Meta analysis which is based on evidences extracted from a number of randomized controlled trials. It provides the most powerful evidence which aids doctors in decision making process on the best management options for their patients. In spite of the mentioned advances in the statistical and methodological procedures, many researchers think that the ‘p’ value is the only method that shows whether or not there is significant difference between the study groups, and hence building inferences accordingly. In order to assure the validity of a study, readers must also be able to exclude role of chance, bias, and confounding variables. Bias and confounding variables are methodological issues which researchers must put in mind prior to conducting the study. It is highly recommended to write a detailed protocol as well as considering these issues. Once bias occurred, it is difficult to be dealt with in the stage of analysis hence it would be advisable to eliminate or minimize it at earlier stages. The introduction of statistical packages could partly reduce the effects of confounders during study analysis.

Most of articles submitted to our journal contain the probability of chance (the p value). In my observation, some of the researchers mention that ‘p’ is less or more than 0.05 or 0.01 without mentioning the actual value of ‘p’. A ‘p’ value of 0.051 and of 0.9 are considered not significant (more than 0.05), but it is evident that the first one could become significant (less than 0.05) by simply increasing the sample size as it is close to the significance level. The size of the ‘p’ value is a function of both the strength of the association and the sample size. The ‘p’ value should not be considered as a fast method for establishing the role of chance, but as a guide to the likelihood that chance could be an explanation of the findings. No ‘p’ value albeit being small, could exclude chance findings completely. The output of some statistical packages present the ‘p’ value as 0.000. In fact it is not zero, and should be written as < 0.001 when presented to the journal.

A related but more informative measure used to evaluate the role of chance is the confidence interval (CI) which is defined as “the range within which the true value or the population value is likely to lie, with certain degree of assurance”. The CI provides information about the level of significance and the sample size of the study. The larger the sample size, the narrower the C.I. Thus the P value and C.I together provide the best information about the role of chance. Nearly all the researches are conducted on samples (not the whole population) aimed at generalizing the results obtained from studying the sample to the population from which the sample was drawn from.
It is worth mentioning that the sample size is an important determinant of the power of the study to detect significant differences between the study groups if these differences truly exist.

Bigger sample size gives more power to the study. In conclusion, biostatistics has a paramount role in medical researches when performed in a sound and scientific manner.

Professor
Namir Ghanim Al-Tawil
College of Medicine, Hawler Medical University