Guest editorial

Overreliance on P-value interpretation in dental scientific reports

Unquestionably, statistics is regarded as the fundamental part of the vast majority of scientific reports in various medical fields. It is the basic tool used to differentiate materials, procedures and techniques subjected to countless modalities of scientific evaluations. It is unusual to read a scientific report in the medical or dental fields without a statistical description of the behavior of the variables under test. Most of those studies, when not relying sole in descriptive statistics, use the P-value as the cut-off point in which a decision to accept or not the previously set null hypothesis can be made, based on the so-called **statistical significance**. This is rather largely considered a sound way to interpret the raw data and provide unbiased conclusions regarding the performance of the groups under test [4].

Nevertheless, a P-value interpretation may be of little or no **clinical significance** [1, 5, 6]. For instance, a 10-day difference in the duration of an orthodontic treatment between two techniques may imply in statistical significance, but certainly has no clinical relevance. However, the given P-value-based results are usually explored by the authors in such a (statistical) manner that influences clinicians' decision-make for treatment A or B. This attitude could lead a given treatment to be prematurely abandoned or to suffer from a severe reduction in interest by the scientific community. In fact, the sole P-value interpretation might be an inefficient method to drive the data analysis towards the conclusions, especially when dependent variables of clinical impact are under analysis. In such cases, the inclusion of Confidence Intervals (CI's) would be of great importance to highlight how the variation for both groups is highly coincident [1, 5, 6]. This can give more precise information on how these results shall impact in the clinical reality.

In biology, assumed a sample size of enough power has been selected, it is unusual that two given conditions are really equal. Therefore, there is a real probability that a difference is found while comparing biological events [1]. However, this statistical find, may be of neglect or very little clinical importance. Thus, many studies may result in a treatment to be rejected by clinicians based on the statistical results only. The scientific community should carefully rethink this attitude.

In the dental field, several typically studied variables are very prone to drive people into limited interpretations, as long as P-values are the sole-performed evaluation. Some remarkable examples are leakage data, materials penetration into tubules, time required to finish a dental movement, push-out strength, teeth resistance to fracture, instruments resistance to torsion, etc... Here, a perceived statistical significant difference between materials usually results in material's rejection or lower acceptance in daily practice, especially if new, untested materials are on the focus. A simple observation of P-value difference, innumerous times, hide a neglectable clinical difference.

For instance, if a restoration A de-adheres from the tooth at 250N which is statistically different from restoration B at 260N, it must be questioned whether this 10N has any impact on the clinical use of restoration B, specially if those mean values are above the range of chewing forces. The report of 95% CI's would help authors to demonstrate that the interval of variance between the materials is not in a range of clinical importance, although null hypothesis has been rejected [5, 6].

I strongly believe that it is time for Peer-Reviewed Journals to increase attention on drawing more strict guidelines on how to report and interpret statistical data. Establishing more specific rules for statistical reports is a current tendency by some dental journals [2, 3], but indubitably much more should be done in this direction, such as, for instance, the set of an editorial board exclusively to deal with statistics of manuscripts. While we wait for this movement to increase, I believe that raw data table and CI's are to be required routinely from any accepted study, especially those when clinical parameters are the dependent variables. A very simple way to comprise with CI's description is to include diagram bars. Discussion and conclusions should also be drawn from this data interpretation.

This obviously might result in more space to be required for a single manuscript, and more accurate review from the Editorial team. However, for the sake of knowledge, would help to keep dental science in a taught-rising direction.

References

1. Dancey CP, Reiddy J. Estatística sem matemática para psicologia. Tradução de Lorí Viali. Porte Alegre: Artmed; 2006. 608 p.

2. De-Deus G. Research that matters – root canal filling and leakage studies. Int Endod J. 2012 Dec;45(12):1063-4.

3. Fouad AF. Journal of endodontics: new guidelines, new directions. J Endod. 2013 Feb;39(2):159.

4. Krithikadatta J, Valarmathi SJ. Research methodology in dentistry: part II – the relevance of statistics in research. Conserv Dent. 2012 Jul;15(3):206-13.

5. Lucena C, Lopez JM, Pulgar R, Abalos C, Valderrama MJ. Potential errors and misuse of statistics in studies on leakage in endodontics. Int Endod J. 2012 Jul;24.

6. Polychronopoulou A, Pandis N, Eliades T. Appropriateness of reporting statistical results in orthodontics: the dominance of P values over confidence intervals. Eur J Orthod. 2011 Feb;33(1):22-5.

Erick Souza
DDS, MSc, PhD in Endodontics
Professor of Florence Institute, São Luís, MA, Brazil