EXPERT OPINION

Statistics, Causality and Dentistry: End and Means!

Vini Mehta[1], Visha Shailesh Pandya[2]

Received on: 05 March 2023; *Accepted on:* 25 March 2023; *Published on:* 15 July 2023

Journal of Scientific Dentistry (2023): 10.5005/jp-journals-10083-1032

[1]Department of Public Health Dentistry, Dr. D.Y. Patil Dental College and Hospital, Dr. D.Y. Patil Vidyapeeth, Pune, Maharashtra, India

[2]Department of Public Health Dentistry, Vaidik Dental College and Research Centre, Dadra and Nagar Haveli and Daman and Diu, Gujarat, India

Corresponding Author: Vini Mehta, Department of Public Health Dentistry, Dr. D.Y. Patil Dental College and Hospital, Dr. D.Y. Patil Vidyapeeth, Pune, Maharashtra, India, Phone: +91 9021887668, e-mail: vini.mehta@statsense.in

How to cite this article: Mehta V, Pandya VS. Statistics, Causality and Dentistry: End and Means! J Sci Den 2023;13(1):31–32.

Source of support: Nil

Conflict of interest: None

Research publications in medicine and dentistry heavily rely on this statistical reporting and data presentation. Statistical analysis is one of the most important and exciting steps in any research process, which deals with gathering, organizing, analyzing, and extrapolating data from samples to the entire population. Inadequate statistical techniques may produce false conclusions that may result in unethical behavior.[1]

Evidence-based dentistry (EBD) covers various activities, including developing research questions, planning studies, gathering and analyzing data, and interpreting and reporting data. With developments in each discipline of dentistry and medicine, some characteristics enhance the decision-making process, whether for the diagnosis of a condition or the formulation of public health policies.[2] Epidemiologists develop hypotheses, and the anticipated outcome may be a response to the question of whether there is a statistical relationship between an exposure and a consequence, such as morbidity or mortality. On the ladder of evidence, descriptive studies rank relatively low. Case reports, case series, and cross-sectional research are examples. Individual meta-analyses of randomized clinical trials diminish the role of chance, but they may introduce bias and confounding variables. While randomized controlled trials (RCTs) have long been regarded as the "gold standard" in causal inference, and dental research has historically been rooted in biomedical fields, other scientific fields, such as epidemiology, statistics, and empirical economics, have always faced difficulties in conducting RCTs.

Before a valid statistical association can be drawn using observational studies, the clinician must answer some fundamental questions about the study, such as whether it is a random occurrence or if there is a bias in the subject selection, and whether there are confounders or variables that have a negative effect on the outcome by masking independent risk factors. The completion of a statistical correlation is contingent upon the clinician's responses to these questions.[3] Even if a statistically valid link can be drawn from observational research after removing the effects of chance, bias, and confounders, this association may or may not be causal. Based on the totality of data from basic, clinical, epidemiological, and statistical research, causation is seen as having a high degree of credibility. As doctors, we ought to strive for this degree of quality in translational research.[4]

The availability of statistical software packages with new multivariate algorithms has contributed to a trend toward employing more advanced techniques, which can also be challenging to newcomers. As a result, the appropriate design of an epidemiological study or clinical trial requires adequate statistical understanding. When the research aims to examine the association between an outcome and (more than) one covariate, standard regression methods should not be used due to the hierarchical structure of such data. Alternately, multilevel analysis can be employed to account for the interrelated nature of dental data clusters. The zero-inflated models have also proven to be appealing methodological techniques for researching the excess number of zero DMFT/DMFT index. The reliability of diagnostic tests on various components of data collecting and strong agreement between observers were proven to be crucial procedures in data quality management. Since the number of units evaluated in dentistry is typically bigger than in other types of clinical investigations (teeth or surfaces instead of individuals), even the confidence interval of Pearson's statistic is typically narrow due to a large number of analyzed units, which may be misleading. Other statistical techniques for quantitative variables have been proposed. Traditionally, a fixed effect analysis of variance might be used. In cases with a large likelihood of chance agreement, it might not enhance meaningful differences between raters. The intraclass correlation, the concordance correlation coefficient, or the Bland and Altman approach can be utilized when quantitative indices like the DMFT/S-DMFT/S are used. Otherwise, a tetrachoric or polychoric correlation should be utilized if it is assumed that continuous latent variables underlie the contingency table. Other approaches have been proposed, and fresh, sophisticated statistical models seem promising for the near future.

Earlier observational study designs were increasingly prevalent, and determining statistical significance was the most frequently used for research. Also, according to a study by Nieminen P et al. in 2020.[5] The *t*-tests, ANOVA, Wilcoxon test, and McNemar's test were commonly used by readers of dental publications in 2017. For the study of longitudinal or intraclustered dental data, the various

methods for calculating intracluster correlations as additional extensions of fundamental regression models have not gained much traction. Thus, novel, highly computational, complicated data analysis techniques are still lacking. In dentistry, we need to increase reporting of multivariable regression, time-to-event models, and intracluster correlation approaches like multilevel modeling or generalized estimation equation analysis. Also, dental researchers did not use data mining, machine learning, and Bayesian methodologies promoted in the computer-oriented methodological literature. Therefore, a model that can account for data correlation must be utilized along with an appropriate statistical test.

Given the widespread inability to describe the fundamental methodologies employed, greater rigour is essential in reporting the methods in dental research articles. It is crucial that researchers, readers, and editors increase their statistical expertise to comprehend the correctness and validity of statistical results due to the widespread use of various statistical approaches.

References

1. Navarro P, Aleman I, Sandoval C, Matamala C, Corsini G. Statistical testing methods for data analysis in dental morphology. Int J Morpho 2020;38(5):1317–1324. DOI: http://dx.doi.org/10.4067/S0717-95022020000501317.
2. Nieminen P, Vähänikkilä H. Use of data analysis methods in dental publications: Is there evidence of a methodological change? Publications 2020;8(1):9. DOI: https://doi.org/10.3390/publications8010009.
3. Ahern J. Start with the "C-Word," follow the roadmap for causal inference. Am J Public Health 2018;108(5):621. DOI: 10.2105/AJPH.2018.304358.
4. Hernán MA. Methods of public health research – Strengthening causal inference from observational data. N Engl J Med 2021;385(15):1345–1348. DOI: 10.1056/NEJMp2113319.
5. Nieminen P, Uribe SE. The quality of statistical reporting and data presentation in predatory dental journals was lower than in non-predatory journals. Entropy (Basel) 2021;23(4):468. DOI: 10.3390/e23040468.