

## A Significant Confusion Over Significance

One of the observations I have made as the editor-in-chief of JOMI is to see and experience both sides of the editorial process. One is to be a clinician-scientist with an interest in asking relevant (clinical and scientific) questions that will impact patient care. The second is to frame those questions in a way that generates the ability to measure (qualitatively or quantitatively) a difference, and to then assess the observed differences from a statistical level of significance. We will come back to the important differences between these two concepts in future articles, but for now, let me outline a fundamental misunderstanding that many authors tend to make that I see as the editor.

Statistical approaches to data analysis generally use one of two approaches: first, the frequentist approach, where the event occurred because of an intervention, some other unknown intervention (or noncontrolled for intervention), or simply random chance. The second approach is Bayesian statistics, which consider the probability of the outcome before an intervention, then use a likelihood measurement to update the initial probability of the outcome by calculating the “posterior probability.” The former frequentist approach is where we typically use the statistical probability of error (Type I and II) and balance the assumption that if we measure enough events, we will measure a frequently occurring event enough to make a statement of “confidence” that the outcome may be related to the intervention (or not). As a reminder, statistical significance is evidence of correlation or association, not causation. Thus, we ask for a power analysis, where we assume an acceptable error rate (by convention, 5% for Type I and 20% for Type II) and we imply a critical difference that we consider important. The last point is where clinical relevance becomes important. Say you compare two implant systems; if one measures average marginal bone changes to determine a net change in mean marginal bone levels, and you repeat the same measurement (using a tool where you know the resolution with some measure of accuracy [low variation]) with precision (a measure of the accuracy relative to the gold standard, or “truth”), you could make that measure 10,000 times and find a statistical difference between the two systems of, say, 0.003 mm net difference in changes in MBL. From that, you could make a declarative statement this is “statistically significant” ( $P < .0001$ ), with one system being superior to the other! Thus, heaven becomes apparent and the clouds part! Obviously, this is not a clinically relevant difference, and there’s the rub. What is a clinically significant difference? Especially to a patient, let alone a busy provider of care? The key for a clinical investigator—and

later the editor of a paper under review—is to understand this nuance of assumption, power, and sample size, especially when it relates to the conclusions being drawn. And, by the way, the 5% rule you learned in stats class for “significance differences” is just a convention, a choice. You could certainly choose a different Type I error you are willing to live with...

To help the readership of IJOMI with many of these issues, we are starting a series of short tutorials written by biostatistical experts and data analysis teams to review basic and emerging trends in biostatistics, data analytics, and information management. These are intended as brief reviews, reminders, and supportive primers to assist dental education, residency programs, and fellowships, along with helping our reviewers and readers to better understand the complexities and assumptions baked into all statistical approaches.

IJOMI is a journal dedicated to expanding and improving the art and science of tooth replacement therapy through the use of oral implants. Core to our mission is providing the best evidence and validated clinical decision support systems we can to help all of the global providers who are impacting patients’ lives every day. The Academy of Osseointegration recently completed the 2022 Summit in Chicago and the papers from that meeting are being refined and will appear in early 2023. One of the key observations from an in-depth review of the literature in outcome measurements of flap vs nonflap surgical approaches, ceramic implants, and the implant-abutment connections was a confusion over what “significance” really means. Just remember the aphorism, “absence of evidence is not evidence of absence” (attributed to Drs Martin Rees or Carl Sagan). We need to ask the important questions, utilize the data analysis that is appropriate with an understanding of the limitations, and temper our enthusiasm of results until we see the holy grail of reproducibility.

Thank you



Clark M. Stanford, DDS, PhD, MHA  
Editor-in-Chief

Jin Xie, PhD  
Associate Dean for Research and  
Director of Biostatistics  
University of Iowa  
Iowa City, Iowa