



Epidemiological study designs and statistical methods for clinical research in oral and maxillofacial surgery: a narrative review

Tim T. Wang^{1,2^}, Sung-Kiang Chuang^{3,4,5,6}

¹School of Dental Medicine, Perelman School of Medicine, Philadelphia, PA, USA; ²Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA; ³Department of Oral and Maxillofacial Surgery, University of Pennsylvania, School of Dental Medicine, Philadelphia, PA, USA; ⁴Brockton Oral and Maxillofacial Surgery Inc., Brockton, MA, USA; ⁵Department of Oral and Maxillofacial Surgery, Good Samaritan Medical Center, Brockton, MA, USA; ⁶Department of Oral and Maxillofacial Surgery, Kaohsiung Medical University, School of Dentistry, Kaohsiung

Contributions: (I) Conception and design: Both authors; (II) Administrative support: SK Chuang; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Sung-Kiang Chuang, DMD, MD, DMSc. Department of Oral and Maxillofacial Surgery, University of Pennsylvania, School of Dental Medicine, Philadelphia, PA, USA. Email: sungkiangchuang@gmail.com.

Objective: This paper provides an overview of epidemiology, common types of clinical studies, statistical analysis, and common pitfalls in clinical research.

Background: As health care leaders at the nexus of dentistry and medicine, the field of oral and maxillofacial surgery (OMS) is relentless in its pursuit of innovation and improvements to enable surgeons to provide the best possible care for patients. To this end, it is of paramount importance to prioritize improving trainees' and attending surgeons' grasp on clinical research study design and statistical analysis. Only with robust scientific evidence can OMSs develop dependable guidelines and sustain shifts in standard of practice. An essential component of such clinical research requires a baseline understanding of epidemiology and biostatistics.

Methods: A review and synthesis of relevant literature was conducted using PubMed database.

Conclusions: This is by no means a comprehensive guide, as entire disciplines are devoted to each of these areas. Moreover, there is a difference between topic experts, who have an intimate knowledge of the clinical aspect of the study, and methodological experts who specialize on study design and biostatistical analysis. It is up to the surgeon researcher to determine whether they can fulfil both roles or should seek support for either or both roles. Ultimately, rigorous research methods are critical to continuing to improve evidence-based standard of practice in order to sustain clinical innovations, protect the scope of OMS, and most importantly enhance patient outcomes in OMS.

Keywords: Oral and maxillofacial surgery research (OMS research); surgical epidemiology; clinical research methods; biostatistical analysis

Received: 26 February 2021; Accepted: 29 June 2021; Published: 10 September 2022.

doi: 10.21037/fomm-21-16

View this article at: <https://dx.doi.org/10.21037/fomm-21-16>

[^] ORCID: 0000-0001-8279-3144.

The field of oral and maxillofacial surgery (OMS) is relentless in its pursuit of innovation and improvements to enable surgeons to provide the best possible care for patients. A prominent example of this is the development of three-dimensional virtual surgical planning, which has transformed the way surgeons approach surgeries and has improved the standard of care (1). Perhaps equally important as innovation is rigorously evaluating those innovations to ensure that they offer true improvements in the quality of care and patient outcomes. Only with robust scientific evidence can OMSs develop dependable guidelines and sustain shifts in standard of practice.

An essential component of such clinical research requires a baseline understanding of epidemiology and biostatistics. Unfortunately, evidence suggests that current OMS trainees on average may require more training in these areas (2). Given the rigorous and time-intensive training of OMS residency programs, however, the solution to the deficiencies in biostatistical skills may not be straight forward. In fact, many programs lack protected time for resident research endeavors (3).

Nevertheless, as OMS continues to lead health care at the nexus of dentistry and medicine, it is of paramount importance to devote resources to improve trainees' and attending surgeons' grasp on clinical research study design and statistical analysis (4). Beyond the aforementioned need to evaluate innovations, these skills are essential in understanding new scientific evidence, which can improve clinical decision making. The following paper provides an overview of epidemiology, common types of clinical studies, statistical analysis, and common pitfalls in clinical research. We present the following article in accordance with the Narrative Review reporting checklist (available at <https://fomm.amegroups.com/article/view/10.21037/fomm-21-16/rc>).

Methods

A literature review was completed utilizing the PubMed online database between June 2020 and December 2020 using search terms relevant to each section relevant to biostatistical study design and epidemiology. English language articles were screened by their titles and abstracts for their relevance and then synthesized to develop an overview relevant to surgeon-researchers in OMS. We summarized the evidence into the following sections: epidemiology and study design, clinical study designs, common statistical analyses, and types of errors and pitfalls.

Epidemiology and study designs

Epidemiology is defined as “the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems” (5,6). To dissect this statement, “distribution” refers to the frequency and pattern (e.g., temporal, geographic, etc.) of how health events occur in a population, while “determinants” refers to the search for factors that contribute to disease (6). “Health-related states or events” covers the gamut of phenomena that can affect the well-being of a population, ranging from communicable infectious disease to exercise and seat-belt use (6). “Specified populations” identifies an important distinction, especially relevant for OMSs, that epidemiological studies concern outcomes for entire populations, in contrast to patient care, which focuses on individuals. Finally, the “application” of epidemiological analyses can be incorporated and referenced when making policy and clinical decisions to improve the health outcomes in that population (6).

However, the definition of surgical epidemiology is more ambiguous (7). While this discussion can quickly get technical, it is important for OMSs to keep in mind the general principles of epidemiology and to design studies in a thought manner to appropriately answer research questions. Study design is of critical importance, as it dictates the subsequent analyses and conclusions.

Clinical study designs

Randomized controlled trials (RCTs) are considered the gold-standard in clinical research. Each subject is randomly assigned to a different treatment arm, often including a control group, and evaluated for certain outcomes in a prospective manner. Alternative methods to assign subjects to groups include randomization between match pairs or using a crossover method where subjects sequentially experience one treatment and then another (8). The specific treatment can be single-blinded (to the subject only), or double-blinded (to both the subject and researcher). Both randomization and blinding allow RCTs to minimize known and unknown biases that can skew data. It is also important to note that the sample size of the study must be large enough to enable statistical analyses to detect a difference between treatments, a determination that depends on the frequency and nature of the effect being studied (9).

Despite their positive qualities, RCTs are not always feasible, such as due to limited resources, so there are

other ways to investigate research questions. The closest alternative, quasi experimental research, are studies where subjects are not randomly allocated to study arms, which can introduce biases to the results. The next alternative to RCTs are prospective cohort studies, in which subjects are followed over a period of time and evaluated in two group based on whether or not they developed the effect being studied (e.g., disease).

In situations where there is limited time and resources, researchers can also conduct retrospective studies to identify clinical trends. Case-control studies involve comparing two groups of subjects, based on whether or not they developed a certain outcome, and then analyzing their attributes and looking for their statistical associations with the outcome. Finally, observational studies can provide a snapshot overview of the study population and identify statistical associations as well. These can be used to identify broad trends and potential areas of further reinvestigation using RCTs. In terms of data source for retrospective studies, researchers can use a range of databases, from the hospital level to large publicly available ones at the national level (10). Finally, researchers can use case-series to study trends in very rare pathologies and events.

Despite these different types of study designs, there are many nuances within each. For example, clustered observations are commonly used in OMS research, and can be collected as part of various clinical studies, ranging from RCTs to observational studies. To illustrate, multiple implants placed on one individual requires a different design and analyses than single implants placed on multiple individuals. To conclude this section, it is critical to note that only RCTs can be used to prove causation between factors and the outcome, whereas all of the other aforementioned studies can only be used to prove correlation.

Common statistical analyses

OMS researchers must be mindful to select the appropriate statistical test to analyze their data. To this end, the statistical test used must be based on the hypotheses of the study. Support from statisticians should be obtained when necessary.

When choosing a test, it is important to determine the type of data being analyzed. Data can be broadly categorized into qualitative and quantitative. Within qualitative data, there are nominal and ordinal data. Nominal data have no inherent order, while ordinal data does, such as a scale of low, medium, and high. Within quantitative data, there are

interval and ratio data. Interval and ratio data have ordered differences between measurements, with the difference between the two being that ratio data have a true zero (e.g., age, height, weight) while interval does not (e.g., IQ, year, temperature not in Kelvin). The type of data that constitute the independent (input) and dependent variables (outcomes) dictate the type of statistical test that should be used.

It is also important to consider whether the data is normally distributed, as this can affect the test being used. For example, the paired t-test can be used to compare two normally distributed quantitative datasets, while the Wilcoxon signed-rank test is used if those two datasets are not normally distributed, nor non-parametric (8).

There are many nuances in statistical analyses and choosing the right test for the right situation. The *British Medical Journal* provides an excellent summary table for this (8). However, it is worth reiterating that if there is any uncertainty on the part of the researcher about statistical analysis, it is always better to consult a methodologic expert to ensure quality.

Types of errors & pitfalls

Scientific hypotheses are tested by using the data to either reject or fail to reject a null hypothesis, which suggests that the differences or associations in data are due to chance, and there is no true difference or association. Statistical tests will provide researchers with a P value, which signifies the probability that the differences in data observed occurred by chance. The P value helps researchers decide whether to reject the null hypothesis and conclude that there is a statistically significant difference or fail to reject the null hypothesis and conclude that there is no statistically significant difference.

Rejecting the null hypothesis to conclude that there is a statistical difference when in reality there is none constitutes a type 1 error (alpha), or a false positive. Meanwhile, failing to reject the null hypothesis and concluding that there is no difference when in reality there is constitutes a type 2 error (beta), or a false negative. It is generally accepted alpha and beta should be less than or equal to 0.05 and 0.2, respectively. Relatedly, the power of a study, or the ability of a study to successfully avoid false negatives, is equal to 1-beta. The power of a study can be calculated based on the sample size and the expected magnitude in difference being studied (9,11).

It is important to prescribe caution regarding the use of P values. While it conveys important information, researchers

must avoid what is colloquially known as p-hacking (12). P-hacking describes situations where researchers use a P value of less than or equal to 0.05 to make conclusions without interpreting the P value in context of the study. For example, while surgeon researchers should be encouraged to use large publicly available databases, the large sample size in these databases allow statistical tests to detect statistical differences of very small magnitude (10,12). For example, a study using a large national database comparing two different surgical techniques for mandibular advancement may find statistically significant differences with a P value <0.05 between an advancement of 7 mm in one group and 8 mm in the other. However, even though there is a statistical difference, the magnitude of the 1 mm difference has little meaningful clinical impact in terms of outcomes for the patient. Another area prone to p-hacking involve using a dataset with many variables to look for associations (12). In these datasets, there will be many variables that have statistically significant associations with P value <0.05. Continuing the example above, researchers may be able to find statistically significant associations (P value <0.05) between the magnitude of mandibular advancement and a factor that is obviously not clinically relevant, such as the color of the shoes that the patient wore on the day of surgery. It is up to the researcher to interpret the results using their clinical expertise to rule out statistically significant but clinically meaningless findings (11). After all, the P value is merely meant to be a supplement, rather than a driver, in making conclusions.

In recent years, in fact, some surgical researchers have pushed to replace or augment P value s by requiring studies to also report confidence intervals, or the range of values that the “true” value will be in between 95% of the time (13).

Conclusions

This paper offers a brief review of epidemiology, study design, statistical methods, and common pitfalls. This is by no means a comprehensive guide, as entire disciplines are devoted to each of these areas. Moreover, there is a difference between topic experts, who have an intimate knowledge of the clinical aspect of the study, and methodological experts who specialize on study design and biostatistical analysis. It is up to the surgeon researcher to determine whether they can fulfil both roles or should seek support for either or both roles. Ultimately, rigorous research methods are critical to continuing to improve evidence-based standard of practice in order to sustain

clinical innovations, protect the scope of OMS, and most importantly enhance patient outcomes in OMS.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Frontiers of Oral and Maxillofacial Medicine* for the series “Clinical Outcomes and Innovations in Oral and Maxillofacial Surgery”. The article has undergone external peer review.

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at <https://fomm.amegroups.com/article/view/10.21037/fomm-21-16/rc>

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at <https://fomm.amegroups.com/article/view/10.21037/fomm-21-16/coif>). This series “Clinical Outcomes and Innovations in Oral and Maxillofacial Surgery” was commissioned by the editorial office without any funding or sponsorship. SKC served as the unpaid Guest Editor of the series and serves as an unpaid editorial board member of *Frontiers of Oral and Maxillofacial Medicine* from July 2022 to June 2024. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Miloro M, Markiewicz MR. Virtual Surgical Planning for

- Inferior Alveolar Nerve Reconstruction. *J Oral Maxillofac Surg* 2017;75:2442-8.
2. Best AM, Laskin DM. Oral and maxillofacial surgery residents have poor understanding of biostatistics. *J Oral Maxillofac Surg* 2013;71:227-34.
 3. Mohammad AE, Best AM, Laskin DM. Attitudes and opinions of residency directors and residents about the importance of research in oral and maxillofacial surgery residencies. *J Oral Maxillofac Surg* 2011;69:2064-9.
 4. Kyzas PA. Evidence-based oral and maxillofacial surgery. *J Oral Maxillofac Surg* 2008;66:973-86.
 5. Best J. editor. *Dictionary of epidemiology*. 4th ed. New York: Oxford University Press, 2001.
 6. *Principles of Epidemiology in Public Health Practice, Third Edition An Introduction to Applied Epidemiology and Biostatistics* [Internet]. Centers for Disease Control and Prevention. 2020 [cited 2021 Jan 31]. Available online: <https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section1.html>
 7. Thind A, Mock C, Gosselin RA, et al. Surgical epidemiology: a call for action. *Bull World Health Organ* 2012;90:239-40.
 8. Swinscow T. *Statistics at Square One* [Internet]. 9th ed. Campbell M, editor. BMJ Publishing Group; 1997. Available online: <https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one>
 9. Wang TT, Chuang SK. Power and Sample Size: An Opportunity to Optimize Randomized Controlled Trials in Oral and Maxillofacial Surgery Research. *J Oral Maxillofac Surg* 2020;78:1880-2.
 10. Wang TT, Lee KC, Chuang SK. Large Publicly Available Administrative Databases: An Opportunity to Improve Evidence-Based Care in Oral and Maxillofacial Surgery. *J Oral Maxillofac Surg* 2021;79:1195-6.
 11. Barkan H. Statistics in clinical research: Important considerations. *Ann Card Anaesth* 2015;18:74-82.
 12. Dodson TB. The Problem With P-Hacking. *J Oral Maxillofac Surg* 2019;77:459-60.
 13. Karadaghy OA, Hong H, Scott-Wittenborn N, et al. Reporting of Effect Size and Confidence Intervals in JAMA Otolaryngology-Head & Neck Surgery. *JAMA Otolaryngol Head Neck Surg* 2017;143:1075-80.

doi: 10.21037/fomm-21-16

Cite this article as: Wang TT, Chuang SK. Epidemiological study designs and statistical methods for clinical research in oral and maxillofacial surgery: a narrative review. *Front Oral Maxillofac Med* 2022;4:26.