



Biostatistics Series – 1: Preparing for collection of data

Gangaboraiah Bilagumba

¹Professor Emeritus and Research Head,
Dayananda Sagar College of Dental Sciences,
Bengaluru, Karnataka, India

Correspondence

Gangaboraiah Bilagumba, Dayananda
Sagar College of Dental Sciences,
Bangalore.
Email: gangaboriah@dscds.edu.in

Abstract

Statistics is defined as the science of collecting, compiling, presenting, analyzing, and interpreting data to ensure research findings are implementable in curative or public health. This editorial, the first in a biostatistics series, outlines the critical framework for planning data collection based on study design, categorized into experimental and observational approaches. Experimental designs range from pre-experimental and quasi-experimental to true experimental studies, such as dental in-vitro research, which utilizes the three principles of randomization, local control, and replication to establish cause-and-effect relationships. For observational quantitative studies, the focus is on developing validated questionnaires that must undergo rigorous validation—including content, construct, face, and criterion validity—and demonstrate internal consistency, typically measured by a Cronbach's Alpha score greater than 0.7. Data collection methods vary from direct oral interviews to digital Google forms, each presenting unique limitations such as high dropout rates or Berksonian bias. Furthermore, the editorial details the importance of variables and scales of measurement—Nominal, Ordinal, Interval, and Ratio—which dictate the appropriate first-line statistical analysis, such as mean and standard deviation for quantitative data or proportions for qualitative data. Establishing these foundational statistical principles is vital for researchers to arrive at valid inferences in dental science.

KEYWORDS

Biostatistics, Data Collection, Experimental Design, Validity and Reliability, Scales of Measurement

1 | INTRODUCTION

Statistics being a science of collection, compilation, presentation, analysis and interpretation of data^{1,2} – the role of each of these terms should be properly understood and due care must be given to each of them to make the end result implementable in curative or public health.

2 | STUDY DESIGN

Firstly, planning of data collection is based primarily on design of the study which can be classified as Experimental and Observational. Experimental studies are further classified as pre-experimental, quasi experimental and true experimental. The invitro studies of dental research belongs to true experimental studies in which they offer high control, allowing for manipulation of independent variables, random assignment and the use of control groups to establish cause-and-effect relationship. A one-shot case study or a one group pre-and-post-test or a statistic studies form pre-experimental designs. In quasi experimental design –

it lacks in randomization.¹³ The experimental design typically adopts the three principles of randomization, local control and replication. The possibilities of bias will be zero or grossly minimized.

3 | IN-VITRO STUDIES

An in-vitro study¹⁴ is also a part of experimental design wherein the experiments are performed on cells, tissues, or molecules under a controlled environment outside the living organism. Though, it benefits in many cases cost-effectiveness and ethical advantages, can't fully replicate the dynamic interactions within a living system.

4 | OBSERVATIONAL QUANTITATIVE STUDIES

The observational quantitative studies mainly focuses on collection of data using a questionnaire developed for the purpose. The questionnaire so developed should be subjected to validation in terms of content

validity, construct validity, face validity and criterion validity. Check for internal consistency using Cronbach's Alpha (preferably which should be > 0.7). Once the required level of validity is established, then the following methods can be employed to collect the data.^{5,6,7,8,9,10,11}

1. Oral interview technique (either direct or indirect method)
2. Self-administered questionnaire
3. Telephonic conversation (disadvantages are more)
4. Through postal (dropout rates are high)
5. Google form (High dropout rate, Berksonian bias¹⁵)

If validated and established tools are used check for their appropriateness, verify the validity and reliability scores. In case, the already validated scale is modified to meet the researcher's requirement, the modified scale has to be revalidated.

5 | QUALITATIVE STUDIES

The qualitative studies use data collection through Delphi method, focus group discussion, analyzing pictures, videos, etc. In this method no serious statistical methods are employed to analyze data to draw conclusion. Usually, what comes out of qualitative research subjective results derived directly and therefore no sophisticated statistical are made use.

6 | MEASUREMENT OF VARIABLES

Statistics deals with measurement of data based the type of variable used. Basically, there are two types of variables viz., quantitative and qualitative variables. Quantitative variables are measured directly by numbers whereas the qualitative variables measured by attributes or certain criteria. However, on a group of individuals they can be quantified. The identification of suitable variables measured decides the type of statistical analysis to be done. For instance, if the variable measured is quantitative, then the first line of analysis will be computation of mean and standard deviation. The rest will be decided based on study protocol. If the variable measured is qualitative, commonly computed measure will be proportion as first line.

7 | DERIVED VARIABLES AND ASSOCIATION

Nevertheless, with quantitative and qualitative variables as basic ones, for further analysis based on these variables' others are derived like categorical variable, independent and dependent variables. For example, tobacco chewing can be categorized as yes/ no and with another variable oral cancer categorized as present/ absent, it is possible to find out the odds ratio and the strength association between them.

8 | SCALES OF MEASUREMENT

Measuring of statistical data is based on scales of measurement viz., Nominal, Ordinal, Interval, and Ratio scales. The nominal scale (categorical without any order) and ordinal scale (categorical with order) are used for qualitative variables whereas quantitative variables are measured using interval and ratio scales, all by numbers. These scales decide the suitable statistical methods applied to analyze the data to arrive at valid inference. Apart from these four scales, Knowledge (0, 1), Attitude (Likert 5-point) and Practice (0, 1) are also used to measure the data. There are also certain developed and validated with good reliability scores which are to be verified properly and used based on the topic selection.

9 | NOTE

Biostatistics series 2: Description of data by descriptive statistics will be in the next issue of the World Journal of Dental Science, published from Dayananda Sagar College of Dental Sciences, Bangalore.

REFERENCES

1. Daniel W W , Cross C L. Biostatistics: A Foundation for Analysis in the Health Sciences, 10th Edition, John Wiley & Sons, 2013.
2. Triola M F. Elementary Statistics, 13/e, Pearson, 2017.
3. Weiss, N. A., Introductory Statistics, 10th Edition, Pearson, 2017
4. Mohan Malhotra V, Kapoor A, Kaur R. Questioning the questionnaire: Testing Validity and Reliability of Questionnaires. International Journal of Public Health Research: 2022, 9(3) DOI: <https://doi.org/10.17511/ijphr.2022.i03.02>.
5. Bolarinwa, Oladimeji Akeem. Principles and Methods of Validity and Reliability Testing of Questionnaires Used in Social and Health Science Researches. Nigerian Postgraduate Medical Journal 22(4):p 195-201, Oct-Dec 2015.
6. Cohen, J. Statistical power analysis for the behavioural sciences. 1988. 2/e. Hillsdale, NJ, Erlbaum.
7. Cronbach L J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16(3), 297-334.
8. Sylvester Chibueze Izah, Ligeiaziba Sylva, Milan Hait3. Cronbach's Alpha: A Cornerstone in Ensuring Reliability and Validity in Environmental Health Assessment. ES Energy and Environment. ES Energy Environ., 2024, 23, 1057, <https://dx.doi.org/10.30919/eseee1057>
9. Nunnally J C, Bernstein I H. Psychometric theory, 1994, 3/3 McGraw Hill. (Known for defining the commonly used 0.70 cutoff; < 0.5 is low and closer to 1 is better)
10. Cortina J M. What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 1993; 78(1), 98-104.
11. Schmitt N. Uses and abuses of coefficient alpha. Psychological Assessment, 1996; 8(4), 350-353.

12. Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*. 2009; 74(1), 107-120. (Often cited for limitations)
13. Field, A P. *Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll)*. 1920, (5th ed.). Los Angeles: SAGE Publications.
14. George D, Mallery P. *SPSS for Windows step by step: A simple guide and Reference*. 2003, 11.0 update, 4/e. Boston: Allyn & Bacon.
15. Bernadette Capili, Joyce K Anastasi. An Introduction to Quasi-experimental Design (Non-randomized Design). *Am. J Nurs*. 2024, Oct; 124 (11): 50-52. DOI: 10.1097/01.NAJ.0001081740.74815.20
16. Clovis Mariano Faggion Jr, Med Dent. Guidelines for reporting Pre-clinical In-vitro Studies on Dental Materials. *Journal of Evidence Based Dental Practice*. 2012; December, 12 (4), 182-189.
17. Daniel Westreich. Berkson's bias, Selection bias, and Missing data. *Epidemiology*. 2012, Jan 23 (1), 159-164

How to cite this article:

Biligumba G. Biostatistics Series – 1: Preparing for collection of data. *World J Dent Excell*. 2025;1(2):1-3.