Misuse of Statistics in Medical Research – A Narrative Review

Aji Gopakumar, Quality Assurance and Institutional Effectiveness Unit

Jayakumary Muttappallymyalil, College of Medicine

Gulf Medical University, Ajman, UAE

GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

THUMBAY
Growth Through Innovation

Introduction

Appropriate use of statistical methods results in valid research inferences. A reasonable conclusion depends on the selection of representative sample, choice of proper sampling technique, identification of relevant variables, quality data and its proper classification, right use of statistical software, appropriate choice of statistical methods, and correct interpretation of results.

Different types of data require different statistical approach and its choice depends on the purpose of the study. All types of statistical methods are available in various statistical software which makes the process of data cleaning, descriptive/inferential and other advanced analysis easier. Analytical perspectives vary according to the study design and aim of the research. Objective of this review is to identify the presence of wrong usage of statistical tools and its probable reasons that are poorly addressed in the literature so far.

Methods

- A narrative review was conducted to identify the presence of misuse of statistics in medical research and the statistical methods which are commonly misused or misinterpreted.
- Published articles in the research area was collected from various electronic databases such as MEDLINE, ERIC, Scopus, PubMed, Science direct and EBSCO.
- As per the review inclusion criteria, the articles published from the year 1990-2021 were searched for the current study. Finally, 120 articles were included in the current review. (77% review papers, 6% original research & 16% books/chapter/ handbooks/conference proceedings)

Results

- Detailed review of published research articles revealed the misuse of statistical tools in different ways, such as incorrect application of descriptive statistics, wrong choice of statistical testing procedures, misinterpretation of results and inappropriate reporting.
- Literature shows, 50% of articles in medical and dental research contain one or more statistical errors. 78% of the articles reported with inappropriate use of basic statistical techniques. 17% of the study conclusions are not in alignment with the results and different analysis to be chosen for 39% of the research works. Almost 34% of studies had incorrectly used parametric methods instead of non-parametric methods.
- Research question and study design were not clearly defined and therefore the research outcome deviated from the study objective. A testable hypothesis was not identified statistically in few original research works
- Reliability of the items were not appropriately calculated in some of the studies. Alpha coefficients to be used for calculating the internal consistency of a scale as a reliability index and it is not about the items.
- Normality of the variables was not ensured to make a decision on the choice of inferential statistics.
- Missing information was one of the major issues that affects the validity of research
- Lack of efforts noticed at the preliminary stage of planning of data collection to ensure the minimum response
- Lack of reliable data and inappropriate management of outliers; quality of the data seems to be not guaranteed during the stage of data management.
- Researchers found more interested in extreme large samples which leads to significant effect size, but those are not clinically significant
- As statistical methods are based on a number of assumptions which are seldom fully met, which lead to incorrect choice of statistical procedures in few studies. Often it is failed to state the statistical assumptions.
- Categorization of quantitative data to categorical form reduces the precision of the measurements
- Mean/median, SD/range are the commonly used descriptive statistics; most common inferential statistics used were t-test/ANOVA/chi-square/Fishers-test and Wilcoxon/Mann-Whitney-tests, which were found to be widely misused.
- Arithmetic means are overused in an over dispersed continuous data.
- Misuse of association test was another common mistake observed when the basic test assumptions are violated (expected frequencies <5).
- Multiple t-tests were used for multiple group comparisons which led to higher type I error.
- Use of unpaired t-tests for paired data and the choice of ANOVA instead of repeated measure ANOVA when the same subjects were used to assess change over time.
- ANOVA is not always robust for groups with different variances, but have been applied in many scenarios.
- Wilcoxon rank sum tests are used instead of Wilcoxon signed rank test when the measurements are paired.
- Some of the studies reveals regression models are generated even if the explanatory variables and response variable are uncorrelated. Often the terms "correlation" and "regression" are synonymous for many non--statisticians.
- Deriving wrong conclusion about R^2 without looking into the residual pattern; R^2 are misused on the data which has no linearity.

Results

- In general, parametric tests were overused on nonparametric tests when its basic assumptions are not satisfied.
- Inappropriate presentation of research results in tabular/graphical form when the cell values are less and more suitable for description.
- Other mistakes identified in deriving inferences from small sample and generalization from small sample based on inferential statistics
- Presenting statistically significant difference in unimportant situations. Presenting significant p values without considering the sample size, power of the study, type I & II errors, confounding factors, biases which detect the real significance of effect size.
- Statistical errors in the interpretation of results are common and often unrecognized in medical research literature.

Conclusion

- Descriptive and inferential statistical methods are indispensable for valid research conclusions. Most commonly used statistical methods were found to be widely misused.
- Unscientific sample size calculation, violation of assumptions, overuse of parametric tests, regression models for uncorrelated variables, misinterpretation/ incorrect reporting of results were the areas significantly misused. Software-misuse was higher due to its increasing accessibility and ability to generate results on a click of button.
- Most kind of such misuses are mainly due to lack of knowledge about the purpose of using statistical methods and inexperience with the statistical software. Software misuse is higher due to its increasing accessibility and ability to generate results on a click of button
- Appropriate application of statistical techniques and correct interpretation of results ultimately leads to effective research conclusions. Each step in the research data analysis is viable when the data is in right hands compared to non-statisticians

Thiese MS, Ronna B, Robbins RB. Misuse of statistics in surgical literature. J Thorac Dis. 2016;8(8):E726-E730.
Thiese MS, Arnold ZC, Walker SD. The misuse and abuse of statistics in biomedical research. Biochem Med (Zagreb). 2015;25(1):5-11.
Yim KH, Nahm FS, Han KA, Park SY. Analysis of statistical methods and errors in the articles published in the korean journal of pain. Korean J Pain. 2010;23(1):35-41.

www.gmu.ac.ae