

The panacea statistical toolbox of a biomedical peer reviewer

Younis Skaik

ABSTRACT

The main role of a peer reviewer is to make judgments on the research articles by asking a number of questions to evaluate the quality of the research article. Statistics is a major part of any biomedical research article, and most reviewers gain their experiences in manuscript reviewing by undertaking it but not through an educational process. Therefore, reviewers of the biomedical journals normally do not have enough knowledge and skills to evaluate the validity of statistical methods used in biomedical research articles submitted for consideration. Hence, inappropriate statistical analysis in medical journals can lead to misleading conclusions and incorrect results. In this paper, the most common basic statistical guidelines are described that might be a road map to the biomedical reviewers. It is not meant for statisticians or medical editors who have special interest and expertise in statistical analysis.

KEY WORDS: Biostatistics, Peer Reviewers.

doi: <http://dx.doi.org/10.12669/pjms.314.8611>

How to cite this:

Skaik Y. The panacea statistical toolbox of a biomedical peer reviewer. *Pak J Med Sci* 2015;31(4):999-1001.

doi: <http://dx.doi.org/10.12669/pjms.314.8611>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

'Statistics is the grammar of science'
by Karl Pearson, 1982.

After the passage of the research articles through the editorial quality tests, the articles are then forwarded to the gatekeepers "Peer reviewers" in a process called peer review. The reviewers are also called "authors' peers" as they work in the same research area. The main role of a peer reviewer is to make judgments on the research articles. Each research article is evaluated by the reviewer by asking a number of questions to evaluate the quality of the research article. Major questions like what is research about?, is it interesting and important?, is the methodology sound?, are the findings original and of considerable value?, are the conclusions appropriate?. Based on the answers of those questions, the reviewers judge on the quality of the research article, and they make a recommendation to the editor, who finally decides whether to publish the research article or not. Since 1960s there has been a dramatic increase

in the application of statistical methods in the biomedical sciences.¹ Adequate designed studies, representative samples (sample size), and appropriate statistical tests are all vital to display data in a concise and representative way and to estimate the probability (P value) of making an error.^{2,3} Most reviewers gain their experiences in manuscript reviewing by undertaking it but not through an educational process.⁴ Therefore, reviewers of the biomedical journals normally do not have enough knowledge and skills to evaluate the validity of statistical methods used in all kinds of research articles submitted for consideration.⁵ It is then difficult to have a reviewer armed with basic and/or advanced statistical knowledge.⁵ This has been confirmed by several studies which showed that half of the published articles in the biomedical sciences have incorrect statistical methods.⁶⁻⁹ Hence, inappropriate statistical analysis in medical journals can lead to misleading conclusions and incorrect results. In this paper, the most common basic statistical guidelines are described that may help the biomedical reviewers.

The panacea statistical toolbox of biomedical reviewers: The main topics of the basic statistical knowledge that may be useful for the biomedical reviewers are illustrated in Table-I.

1. Types of data: The target of most studies is to collect data about a particular topic, present them in a meaningful way, and to extract information from

Correspondence:

Dr. Younis Skaik,
Dept. of Laboratory Medicine, Faculty of Applied Medical Sciences,
AL-Azhar University, Gaza, Palestine.
E-mail: y_skaik@hotmail.com

- * Received for Publication: May 28, 2015
- * Revision Received: July 13, 2015
- * Revision Accepted: July 20, 2015

Table-I: The panacea statistical toolbox of a biomedical reviewer.

- Types of data.
- Which statistical test is the most appropriate to analyze the data?
- What is the P value?
- Numerical and graphical description of data.

them. Data are often discussed as variables; a variable is any characteristic that varies from one to another in a defined population. For example, weight in kilogram varies from person to person. There are two major types of variables: categorical or numerical. The importance of knowing the data type is to determine the most appropriate statistical test that can be applied to analyze the data.¹

Categorical (qualitative) data: Defined by the classes or categories into which an individual can belong.

Nominal data: Name only. Examples include gender, ethnicity, blood group, or marital status.

Ordinal data: Variable is ordered. Examples like "Yes/No" or "Low/Medium/High".

Numerical (quantitative) data: The variable has a numerical value.

1.2.1 Discrete data: A number obtained by counting. Examples include the number of biomedical reviewers who have statistical knowledge.

1.2.2 Continuous data: It reflects a measurement. There is no limitation on the values that the variable can take. Examples include blood pressure, height, or weight.

Transforming continuous data into two or more ordinal data is often observed in the biomedical literature to make distributions closer to normal distribution, and sometimes to make easier both interpretation of the results and data comparison.¹⁰ There are several methods to transform data, which all can reduce the precision of measurements and hence causes false negative or positive results.^{11,12} Therefore, the reviewers should pay an attention whether the authors mention how and why the transformation was done, if there is any. In addition, non-parametric tests that do not assume normal distribution are always an alternative to analyze the data rather than transforming data.

Which statistical test is the most appropriate to analyze the data?

According to the data type, it is of utmost importance to select the most appropriate statistical test to avoid the unsound conclusions and incorrect research results. Table-II shows the most common statistical tests for simple analysis of data, which can be used by the reviewers to check the appropriateness of statistical tests used in the submitted research articles. More statistical tests which are used for regression studies, longitudinal studies, and assessing evidence are more advanced and hence are not highlighted in this paper.

The research question, the data type, and the number of groups involved in the study are the main factors which play a role in selecting the most appropriate statistical test.

What is the P value?

P value is the estimated probability but not a biological importance of occurrence of observed effect if the null hypothesis (H_0) of a study question is true. It is used to assess the statistical significance, however, the strength of association and effect size are limitations.¹³ The reviewers should look for the exact P values in the submitted papers but not more or less the significance level (0.05 or 0.01). If the exact P-value of an observation (e.g., comparing two groups) was 0.049 with a sample size of 20, and the authors reported the P value as < 0.05 , then many researchers would not think to replicate the results with a larger sample size. It is also recommended to report in addition the 95% confidence interval for the difference.¹⁴

Numerical and graphical description of data

Numerical and graphical summaries of data would save the time of the readers and the data would be more informative. However, the question that should be raised by the reviewers is that which graphical methods and summary statistics would be more valuable in certain situations to avoid the incorrect usage by the authors?¹⁵

Fig.1 illustrates a wide spectrum of visualizing the data, which are all important for the reviewers to have an idea about the basic concepts of each item. The reviewers should be aware of the following common

Table-II: Most common statistical tests used for simple analysis of data.

Type of data	No. of comparing groups / Categories	Statistical test
Numerical	One group	t-test, Sign test
	Two paired (matched)	Paired t-test, Wilcoxon sign rank test, Sign test
	Two independent groups	Unpaired t-test, Wilcoxon rank sum test
	More than 2 groups	One-way ANOVA, Kruskal-Wallis test
Categorical	Two categories (1 group)	Z test
	Two categories (2 groups paired)	McNemar's test
	Two categories (2 independent groups)	Chi-squared test, Fisher's exact test
	Two categories (more than 2 groups)	Chi-squared test, Chi-squared trend test
	More than two categories	Chi-squared test

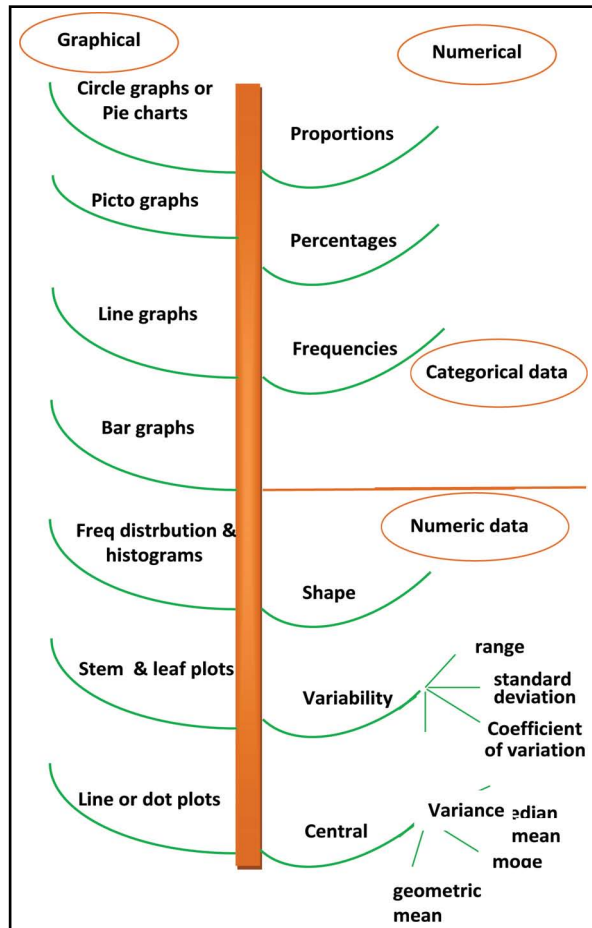


Fig.1: Representation of data.

mistakes in the graphical and numerical description of data and how to avoid them.

1. Most biological data are not normally distributed.¹⁶ Hence, mean and standard deviation (SD) are not the correct tools to describe the skewed data. Alternatively, medians and inter quartile ranges can be used in such situations.
2. The mean alone is not enough to describe the variability of data.
3. The SD and standard error of the mean (SEM) are interchangeably used in the biomedical literature despite they are different.¹⁷ SD is used in the summarized descriptive data to describe that quantifies variability within the sample. However, SEM is limited to compute the confidence intervals (CI) and hence quantifies uncertainty in the estimate of the mean.
4. Figures and tables to summarize the data should be presented in a way to assist the readers and not to mislead them.
5. Graphs that do not start at zero are tricky and can mislead the readers. Therefore, starting the graphs and charts at zero would give accurate comparability of the columns.

6. Scales with equals intervals are highly recommended rather than compressing or lengthened one of the axis which can lead to distorted relationship between the two axes.

CONCLUSIONS

Notwithstanding the statistical software packages facilitate the task of data analysis for statistically unskilled researchers; major statistical problems are still determined in the biomedical research articles, due to the insufficient knowledge of researchers of the statistical ideas and mathematical concepts. Therefore, screening the submitted articles for the validity of statistical analysis of the data is an additional task of the biomedical reviewers. Thus, the reviewers should have at least basic statistical knowledge to be able to end up with sound results and correct conclusions. In addition, consulting a statistician or statistically skilled experts should be a choice for the editors and reviewers to enhance the statistical quality of the biomedical published articles.

Declaration of interest: None

REFERENCES

1. Reverberi R. The statistical analysis of immunohematological data. *Blood Transfus.* 2008;6:37-45.
2. Joel B. Damned lies and statistics: untangling numbers from the media, politicians, and activists. Berkeley-London: University of California Press, 2001.
3. Bailer JG. Medical uses of statistics. 2nd ed. Waltham (MA): CRC Press, 1992.
4. Benos DJ, Kirk KL, Hall JE. How to review a paper. *Advan Physiol Edu.* 2003;27:47-52.
5. Petrovecki M. The role of statistical reviewer in biomedical scientific journal. *Biochemia Medica.* 2009;19:223-230.
6. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circultaion.* 1980;61:1-7.
7. Altman DG. The scandal of poor medical research. *BMJ.* 1994;308:283-284.
8. Lang T. Twenty statistical errors even you can find in biomedical research articles. *Coratian Med J.* 2004;45:361-370.
9. Stasak A, Zaman Q, Pfeiffer K, Göbel G, Ulmer H. Statistical errors in medical research-a review of common pitfalls. *Swiss Med Wkly.* 2007;137:44-49.
10. Bland JM, Altman DG. The use of transformation when comparing two means. *BMJ.* 1996;312:1153.
11. Bland JM, Altman DG. Transforming data. *BMJ.* 1996;312:770.
12. Bland JM, Altman DG. Transformations, means and confidence intervals. *BMJ.* 1996;312:1079.
13. Pencina M, D'Agostino R, Vasan R. Statistical methods for assessment of added usefulness of new biomarkers. *Clin Chem Lab Med.* 2010;48:1-9.
14. Bailar JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals. *Ann Intern Med.* 1988;108:266-273.
15. Wullschlegler M, Aghlmandi S, Egger M, Zwahlen M. High incorrect use of standard error of the mean (SEM) in original articles in three cardiovascular journals evaluated for 2012. *PLoS One.* 2014;9:e110364.
16. Lang T. Common statistical errors even you can find. *AMWA.* 2003;18:67-71.
17. Nagele P. Misused of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. *Br J Anaeth* 2001;90:514-516.