



Statistical recommendations for the authors of manuscripts submitted to the Journal of Cancer Research and Clinical Oncology

Michał Ordak¹

Received: 8 February 2022 / Accepted: 9 February 2022 / Published online: 3 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In recent years, a negative picture of statistical analyses carried out in medicine has been observed around the world. Unfortunately, as it turns out, this also applies to COVID-19. The most important guidelines for the members of the readers and authors of articles submitted to the Journal of Cancer Research and Clinical Oncology, i.e., on numerous factors related to the statistical analysis, are presented.

Keywords Biostatistics · Medicine

Dear readers and the authors of articles submitted to the Journal of Cancer Research and Clinical Oncology.

In recent years, a deteriorating quality of statistical analyses in medicine has been observed (Diong et al. 2018). For example, the most recent data indicate that only 39% of the 2600 accepted articles related to various aspects of COVID-19 meet the requirements of statistical correctness (Ordak 2022). The prevailing pandemic should increase the involvement of biostatisticians in the development of new methods and indicators for modeling and preventing COVID-19 (O'Neill 2021). According to 2020 data published in PLOS ONE, 34% (36/107) of journal editors stated that they rarely or never use a specialized statistical review. This percentage has not changed since 1998, despite greater care that is placed on the credibility of research (Hardwicke and Goodman 2020). For this reason, it seems right to educate members of the editorial board of journals on the most common mistakes made by authors related to the statistical analysis conducted, as well as on possible ways to reduce this problem. It is recommended that authors of articles submitted to the Journal of Cancer Research and Clinical Oncology take into account several principles.

First, before authors submit an article to the journal, their statistical analysis should be reviewed by an expert in

biostatistics. This applies to all aspects related to the statistical analysis. Second, authors are required to note this fact in their cover letter as well as in the submitted manuscript, i.e., by providing a reference to this editorial. Third, authors are advised to take into account a number of factors related to their statistical analysis (Table 1).

Statistical guidance for the readers and authors of articles submitted to the Journal of Cancer Research and Clinical Oncology is provided. First, it is not recommended that the authors describe a few statistical tests in just one sentence. To better illustrate to the reader the correctness of the selected statistical tests, the authors should describe in more detail the sense of their application (e.g., comparing three groups of patients and the relationship between the level of anti-SARS-CoV-2 antibodies and the severity of the disease). The same applies to the extension of the description of the abbreviations of the more advanced statistical analyses used.

Second, when using non-parametric or parametric equivalents of the statistical tests used, the authors should pay attention to the use of appropriate descriptive statistics. It should also be explained why in the case of these specific analysis, for example, the non-parametric equivalent of the statistical test was used (type of variable, normality of distribution, group equivalence, etc.). Unfortunately, in many journals, the authors very often use the wrong counterparts of the statistical tests used, which may result in the incorrect interpretation of the obtained results, and thus, incorrect drawing of conclusions. The end result of such a situation may be the ambiguity of the obtained research results on the

✉ Michał Ordak
mordak@wum.edu.pl

¹ Department of Pharmacodynamics, Centre for Preclinical, Research and Technology (CePT), Medical University of Warsaw, 1B Banacha Street, 02-097 Warsaw, Poland

Table 1 Statistical guidance to readers and authors of articles submitted to the Journal of Cancer Research and Clinical Oncology

Variable	Statistical factors
1. Description of the applied statistical tests	The authors should describe in which cases they used specific tests, i.e., not only in general (type of variables, normality of distribution, etc.)
2. Statistical package used	The most important type of information should be included, e.g., GraphPad Software, Inc., San Diego, CA
3. Statistical test result	To increase the reliability of the obtained results, they could be written according to scientific standards, e.g., in the case of the Mann–Whitney U test: $U=60.5$; $p=0.03$
4. Designation under the table of statistical tests used in it	Next to the p value in the table, there could be a symbol, e.g., “*”, and under the table an explanation of what it means the applied statistical test
5. Assumptions of statistical tests	The authors should describe on the basis of exactly what they use parametric/non-parametric equivalents of various statistical tests
6. Results in the abstract	In the abstract, the authors should include p values to increase the clarity of the results obtained
7. Size effect	To increase the value of the work, it is advisable to study the size of the effect (<i>Cohen’s d</i> , <i>Eta-squared</i> , <i>Hedges’ g</i> , <i>Glass’s δ</i> , etc.)
8. Reliability and validity of survey information	It is recommended that the integrity of these new surveys be tested (<i>Cronbach’s alpha</i> , etc.)
9. Use of appropriate descriptive statistics	With different type of statistical tests, the authors should include in their article adequate descriptive statistics such as the median, etc.
10. Clarity of the described advanced parameters/statistical tests for the reader	The reader could be briefly explained, i.e., with simple words, what advanced parameter/statistical test is. This would allow many people who are not familiar with biostatistics to understand the meaning of this type of parameters
11. Detection of outliers	Influence of cases of outliers on the obtained results could be evaluated by means, e.g., Cook’s distance, Leverage—LEV, Mahalanobis Distances, Adjusted Predicted Value—ADJ, etc.
12. Post hoc tests	The authors should clearly describe in future papers which post hoc test they chose and why (homogeneity of variance, sample size, etc.)

same topic, i.e., conducted by independent authors (Nahm 2016).

Third, in the case of post hoc testing, it is not enough to write one general sentence. There are a number of different post hoc tests, each with strengths and weaknesses. Some tests are more liberal and others are conservative. The use of different statistical tests in the same study can produce quite different results (Lee and Lee 2020). For this reason, the authors using a specific post hoc test, or e.g., pointing to the failure to meet the assumption of sphericity of the variance, should describe why they chose this test/correction and not another. Thanks to this type of extended description, the credibility of the obtained research results would increase significantly.

Fourth, to increase the significance of the obtained research results, it is recommended to calculate the size of the effect, i.e., the statistics indicating the strength of a specific phenomenon, e.g., the difference. An example here is Cohen’s d measure, Hedges g , eta-square, F Cramer, Glass’s rank two-series correlation coefficient, etc. Contrary to p value, the strength of the effect makes it possible to assess the practical significance of the result, as well as to compare the results of many studies in meta-analysis (Sullivan and Feinn 2012; Ialongo 2016).

The penultimate recommendation relates to outliers, the presence of which may play a significant role in the results obtained. The presence of outliers can result in overstated or underestimated values. For this reason, it is recommended to use tests to detect this type of observation, allowing to answer a question like: what would happen if the particular observation were not present in the model? An example here is the Cook distance and the Mahalanobis distance (Kwak and Kim 2017).

The last suggestion has to do with the recording of statistical test results obtained according to scientific standards. The authors should include in the table appropriate symbols and their description, i.e., denoting the use of specific statistical tests for individual variables. It is also advisable to record the results of statistical tests according to accepted scientific standards, not just p value (Arifin et al. 2016). This is another factor that increases the transparency and credibility of published research results.

Improving the quality of statistical analyses allows to improve the transparency and credibility of published research results, which may be reflected in the improvement of the quality of life of medical patients.

Declarations

Conflict of interest The author have no financial conflicts of interest.

Ethical approval Not applicable.

References

- Arifin WN, Sarimah A, Norsa'adah B, Najib Majdi Y, Siti-Azrin AH, Kamarul Imran M, Aniza AA, Naing L (2016) Reporting statistical results in medical journals. *Malays J Med Sci* 23(5):1–7
- Diong J, Butler AA, Gandevia SC, Héroux ME (2018) Poor statistical reporting, inadequate data presentation and spin persist despite editorial advice. *PLoS ONE* 13(8):e0202121
- Hardwicke TE, Goodman SN (2020) How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey. *PLoS ONE* 15(10):e0239598
- Ialongo C (2016) Understanding the effect size and its measures. *Biochem Med (zagreb)* 26(2):150–163
- Kwak SK, Kim JH (2017) Statistical data preparation: management of missing values and outliers. *Korean J Anesthesiol* 70(4):407–411
- Lee S, Lee DK (2020) What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol* 73(6):572
- Nahm FS (2016) Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean J Anesthesiol* 69(1):8–14
- O'Neill RT (2021) Reacting to crises: The COVID-19 impact on biostatistics/epidemiology. *Contemp Clin Trials* 102:106214
- Ordak M (2022) COVID-19 research: quality of biostatistics. *Arch Med Sci* 18(1):257–259
- Sullivan GM, Feinn R (2012) Using effect size-or why the P value is not enough. *J Grad Med Educ* 4(3):279–282

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.