COVID-19 research: quality of biostatistics

Michal Ordak

Department of Pharmacodynamics, Centre for Preclinical Research and Technology (CePT), Medical University of Warsaw, Warsaw, Poland

Corresponding author:
Michal Ordak
Department
of Pharmacodynamics
Centre for Preclinical
Research and Technology
(CePT)
Medical University
of Warsaw
1B Banacha St
02-097 Warsaw, Poland
E-mail: mordak@wum.edu.pl

Abstract

Introduction: In recent years, unfortunately, low quality of statistical analyses in medicine has been observed. As it turns out, this also applies to COVID-19 subject matter.
Methods: The study included 2600 medical articles published between the beginning of 2020 and June 2021, in which the authors described the obtained results, i.e. related to COVID-19.
Results: Of the analysed articles, 39% were correct in terms of the statistical analysis performed.
Conclusions: There should be more emphasis on conducting statistical reviews in authors' contributions on various aspects of COVID-19.

Key words: biostatistics, COVID-19.

The global COVID-19 pandemic has generated health problems worldwide [1]. In recent years, a misuse of statistical analyses leading to flawed or inaccurate conclusions has been observed around the world [2]. At the International Statistical Congress held in Malaysia in 2019, it was mentioned that only 40% of accepted medical articles are statistically correct [3]. It transpires that this also applies to articles about COVID-19.

Methods. The analysis included 2600 articles on the medical aspects of COVID-19 (overall quality of life, medical/pharmacological approach) published between the beginning of 2020 and June 2021. International databases such as PubMed and Scopus have been searched for the medical aspects of COVID-19. For the analysis, articles published in journals with an impact factor of up to 15 were considered, of which over 95% had a maximum impact factor of 10. The statistical correctness of each article was analysed, namely in terms of all aspects related to the analysis performed. It includes the selection of appropriate statistical tests (including checking their assumptions), as well as the correct interpretation and recording of the obtained results (including the use of appropriate descriptive statistics). During the evaluation, additional factors were also considered, such as a description of the analysed variables, statistical software used, appropriate sample size, and considering possible data gaps. Incorrect analysis of the results, and referring to them in the discussion, resulted in treating the article as one in which the statistical analysis was conducted incorrectly. This applies, for example, to incorrect selection of statistical tests, failure to meet the assumptions of many statistical tests, visible and not described by the authors, etc.

AMS

Michal Ordak

Results. Firstly, of the 2600 analysed publications on the medical aspects of COVID-19, only 39% (n = 1014; unpublished) were statistically correct. This can even cause ambiguous results regarding various aspects of COVID-19, and this is not what we all expect. The most common mistakes include the use of inadequate statistical tests (including parametric equivalents despite unfulfilled assumptions), as well as incorrect estimation or underestimation of the correct size of

Table I. Basic statistical errors made by authors in published articles on COVID-19 (n = 2600)

Basic statistical errors in published articles made by authors COVID-19 (n = 2600)	Example	Recommended for reviewers/ editors to consider guidance when reviewing the manuscript
No information on the software used for statistical analysis is available	Including only 2 sentences on the statistical tests used, with no information on the software used	Providing detailed information on the statistical software used
A cursory, questionable description of the statistical tests used	A single description like: "The manuscript uses the student's t-test and Pearson's correlation analysis"	Including a detailed description of the statistical tests used: specific purpose and an adequate explanation of the selection of individual statistical tests
Incorrect selection of statistical tests and interpretation of results	The use of parametric equivalents of statistical tests despite many unfulfilled assumptions (very low group size, disturbances in normal distribution, heterogeneity of variance, etc.) that are visible to the naked eye. The use of regression analysis despite the strong correlation of predictors. Interpreting correlations as causation	Paying attention to factors such as the normality of the distribution (Gaussian distribution), the equality of groups (χ^2), the type of variables analysed, homogeneity of variance (Levene's test), etc.
There is no information in the results about where and which statistical test was applied. The effect – frequent doubts related to the quality of the presented results	General description of the various statistical tests used, and the results obtained include many tables, of which it is not known how the authors analysed the individual parameters	In-depth checking of whether the authors indicate where and what they used the for statistical testing in the obtained results
Putting p-value differently	Writing in the manuscript in the style of $p < 0.05$ on time and another time for example $p = 0.03$	Applying a uniform system of obtained results
Failure to record the results of statistical tests in accordance with scientific standards	Notation the result of the analysis of variance in the $p = 0.04$ style	Record statistical test results by standards, not just p-value
No valid explanation of outliers	No including outliers in regression	Paying attention to outliers/extreme cases when there is such need/ doubt, e.g. in a scatter plot
No explanation of the changes in the number of subjects	Missing data	Checking whether the authors describe any possible missing data
Failure to consider in the analysis various factors that may affect the obtained results	Including several hundred women and men suffering from COVID-19 in the analysis, while the analysis does not reflect the gender factor	Viewing in the reviewed article whether the authors consider the necessary factors such as sex, age, comorbidities, etc.
Too superficial use of descriptive statistics	Placing in the text only the mean value without other descriptive statistics relevant for a given study	Paying attention to descriptive statistics included in the manuscript, i.e. statistics tailored to the specific study conducted (median, standard deviation, etc.)
Others	No data on the recruitment of participants, inadequate sample size, unreadable list of variables, no clear baseline demographic and clinical parameters	Verifying whether the authors describe aspects such as the method of selecting the sample size, description of the test / control group, list of analysed variables, the size of the effect (e.g. eta-squared, Cohen's), etc.

the tested sample. The more advanced statistical analyses used by the authors in the manuscript should be reviewed by experienced people or, for example, statistical reviewers [4, 5].

The first table below shows only the basic statistical errors made by the authors when publishing the obtained research results related to various aspects of COVID-19 (Table I).

Secondly, the retraction of articles on COVID-19 due to incoherent data also applies to prestigious international periodicals. Among the results of this situation is information chaos, which undermines trust in reliable sources of information and affects the approach to vaccination against COVID-19. In a study of 3480 non-medical students, 75% of them ($n = 2.610$; unpublished) stated that due to such situations, their confidence in the results of studies on COVID-19 published in prestigious journals significantly decreased.

Third, as can be seen for many years, young researchers who want to have their articles published in ranked journals commit statistical fraud frequently. For example, a study was carried out on a group of 14,000 people working in various fields of medicine (including physicians, graduate students, PhD students, PhDs, and professors) indicating, for example, that as many as 76% of respondents stated that they did not know what type 1 cumulative error is. Forty-six percent of people admitted that they often performed several or over a dozen t-tests instead of conducting an analysis of variance. While 10% of them did so due to a lack of knowledge, others wanted to increase the chance of obtaining a statistically significant result [3].

The same, unfortunately, applies to analysing the results obtained for COVID-19. Researchers deal with this issue because it is topical now, and so they think that even if they carry out their analysis incorrectly, they will have a chance to obtain points for the published article. Instead of improving the quality of life of people suffering from COVID-19, they can unfortunately work to their detriment. In an exemplary study that I carried out in 2020 on a group of 550 scientifically working psychiatrists, 48% of them ($n = 264$; unpublished) began research on COVID-19, explaining that it was a topical issue, and hence they would be able to publish the obtained results even faster. As we know, this plays a significant role in the functioning of research units. The speed in question may unfortunately result in the publication of scientific articles in which statistical analysis is incorrectly carried out. It may be related to the factors motivating employees to conduct scientific research. This applies, e.g., to competitive pressures, institutional, regional, and national recognition, and financial remuneration [6, 7].

Discussion. Optimizing protection against infection in wealthy nations (e.g. UE, USA) and not helping low-income countries with vaccinations will lead to a prolonged pandemic. The editors of medical journals should pay more attention to validating the authors' statistical analysis. Having statistical reviewers or statistical editors is essential for biomedical journals. The authors of an article published in PLoS One point out that there is a huge need to improve statistical education [8]. The substantive assessment of the submitted articles is clearly insufficient. We will have to face this pandemic for a long time, not (only) because of the non-observance of the statistical rigors by the researchers, but also because of the non-observance of the specific prevention measures by a large proportion of the general population.

In conclusion, the results on various aspects of COVID-19 sent to journals should be subjected to thorough statistical review.

Conflict of interest

The author declares no conflict of interest.

References

1. Burlacu A, Mavrichi I, Crisan-Dabija R, et al. "Celebrating old age": an obsolete expression during the COVID-19 pandemic? Medical, social, psychological, and religious consequences of home isolation and loneliness among the elderly. Arch Med Sci 2020; 17: 285-95.
2. Diong D, Butler AA, Gandevia SC, et al. Poor statistical reporting, inadequate data presentation and spin persist despite editorial advice. PLoS One 2018; 13: e0202121.
3. Ordak M. Critical image of statistical analyses in medicine between 2006 and 2018. Department of Statistics Malaysia (DOSM). Proceeding of the 62nd ISI World Statistics Congress 2019, Contributed Paper Session 1: 23-30.
4. https://2019.isiproceedings.org/Files/7.Contributed-Paper-Session(CPS)-Volume-1.pdf
5. Kyrgidis A, Triaridis S. Methods and Biostatistics: a concise guide for peer reviewers. Hippokratia 2010; 14: 13-22.
6. Greenwood DC, Freeman JV. How to spot a statistical problem: advice for a non-statistical reviewer. BMC Med 2015; 13: 270.
7. Marco CA, Larkin GL. Research ethics: ethical issues of data reporting and the quest for authenticity. Acad Emerg Med 2000; 7: 691-4.
8. Hardwicke TE, Goodman SN. How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey. PLoS One 2020; 15: e0239598.