

STATISTICAL REPORTING IN THE “CLUJUL MEDICAL” JOURNAL

DANIEL-CORNELIU LEUCUȚA, TUDOR DRUGAN, ANDREI ACHIMAȘ

Department of Medical Informatics and Biostatistics, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania

Abstract

Background and aim. Medical research needs statistical analyses to understand the reality of variable phenomena. There are numerous studies showing poor statistical reporting in many journals with different rankings, in different countries. Our aim was to assess the reporting of statistical analyses in original papers published in Clujul Medical journal in the year 2014.

Methods. All original articles published in Clujul Medical in the year 2014 were assessed using mainly Statistical Analyses and Methods in the Published Literature guidelines.

Results. The most important issues found in reporting statistical analyses were reduced reporting of: assumptions checking, difference between groups or measures of associations, confidence intervals for the primary outcomes, and errors in the statistical test choice or the descriptive statistic choice for several analyses. These results are similar with other studies assessing different journals worldwide.

Conclusion. Statistical reporting in Clujul Medical, like in other journals, have to be improved.

Keywords: statistical reporting, SAMPL guidelines, biomedical research, Clujul Medical journal

Background

Medical research needs statistical analyses to understand the reality of variable phenomena. Readers of published medical research should be provided with all the necessary information in order to be able to understand what statistical analyses were used, why they were chosen that way, while results should carry enough information to make what was observed clearly understandable. This way a knowledgeable reader can identify the limits or even possible errors in the published literature. Also they will have all the information to get the precise image of the study data so judgments can be appropriately made. Poor statistical reporting impedes these to be achieved. In the end future research and even the health of patients managed on the basis of articles with statistical errors will suffer.

There are many books on how to perform statistical

analyses, and some of them inform on how to present the results of the analyses. There are also websites and guidelines to help authors to better write their articles, focusing [1–3] or including information about statistical analyses [4–6]. “Statistical Analyses and Methods in the Published Literature” (the SAMPL Guidelines) [3] is such a guideline.

There are numerous studies showing such bad practices in many journals with different rankings from different countries [7–10]. Clujul Medical is a journal that has published peer-reviewed articles in the medical field ever since 1920 almost without interruption [11]. We could not find a study assessing statistical reporting practices in articles published in Clujul Medical journal. In order to know the current situation and what is needed to improve the quality of published papers such a study is warranted. Thus our **aim** was to assess the reporting of statistical analyses in original papers published in Clujul Medical journal in the year 2014.

Manuscript received: 15.09.2015

Accepted: 01.10.2015

Address for correspondence: tdrugan@umfcluj.ro

Methods

All original articles published in Clujul Medical in the year 2014 were downloaded in full text from the journal website.

Each article was read and checked for reporting of statistical analyses. A list of items for reporting statistical analyses was created starting from the SAMPL guidelines [3]. Some items were identical to information found in the guidelines (further denoted in the article by the * sign), some items were modified by the authors (further denoted in the article by the \$ sign), and some new items were added by the authors (further denoted in the article by the + sign). We used almost the same list of items as in a previously published article assessing pharmaceutical articles [12].

The reporting of different items was presented by counts and percentages. All calculations were made using the package R environment for statistical computing and graphics version 3.2.0 [13].

Results

We found 29 original articles out of which 28 (96.6%) used statistics; the article that did not have statistics was a qualitative research.

Preliminary analyses were made in one article (1.6%), a log transformation of non-normally distributed variables.

Reporting of primary statistical analyses

The reporting of primary statistical analyses is presented in Table I. The most frequent problems found were the lack of reporting: - that data conformed to the assumptions; - if tests were one or two-tailed; - the statistical and the name of the statistical program used. We could identify only one paper with a clear error of analyzing skewed data with a parametric test.

Reporting of numbers and descriptive statistics

Positive findings in reporting numbers and descriptive statistics (Table II) were the reporting of the total or group size for analyses, not exaggerating with the decimals, and using the correct descriptive statistics when the distribution

was stated. We found a lack of reporting variability of the data with standard deviations or quartiles or ranges, and a lack of reporting numerators and denominators for percentages. A problematic finding was the reporting of the mean along with a non-parametric test without reporting the median.

Reporting hypothesis tests

Regarding the reporting of hypothesis tests (Table III), the poorest reported items were the differences between compared groups, the use of confidence intervals, the checking of assumptions for the test and the type of test (for paired or independent samples). Summarizing data with descriptive statistics fares better.

Reporting association analyses

Association analyses were found in 4/28 (12.3%) of the studies. Three of them (75.0%) were reporting them as a primary analysis. The reporting was appropriate for: - identifying the variables used * 4/4 (100.0%); - summarizing each with descriptive statistics * 4/4 (100.0%); - identifying the test of association used * 4/4 (100.0%); - for tests of association (e.g., a Chi-square test), reporting the P value of the test * 4/4 (100.0%); - for primary comparisons, including the full contingency table for the analysis * 3/3 (100.0%). Presenting a measure of association was less frequent * 1/3 (33.3%);

Reporting correlation analyses

The reporting of correlation analyses (Table IV), was suffering especially on the side of confirming that the assumptions of the analysis was met, and reporting confidence intervals.

Reporting regression analyses

Regarding the reporting of regression analyses (Table V), we observed the same lack of reporting the checking of assumptions, and confidence intervals. On the other hand, reporting the variable selection process by which the final model was developed and providing a measure of the model's "goodness-of-fit" to the data were well done.

Table I. Reporting of primary statistical analyses.

Item	n/total (%)
summarize each variable with descriptive statistics \$ partially	25/28 (89.3) 2/28 (7.1)
verify that that data conformed to the assumptions *	7/20 (35.0)
skewed data were analyzed with non-parametric tests *	4/6 (66.7%)
paired data were not analyzed with unpaired tests \$	0/3 (0.0)
the study had multiple comparisons +	1/19 (5.3)
indicate whether and how any allowance or adjustments were made for multiple comparisons *	1/1 (100.0)
say whether tests were one- or two-tailed	1/19 (5.3)
how many were reported as one-tailed +	0/1 (0.0)
report the alpha level (e.g. 0.05) *	12/19 (63.2)
alpha levels used: 0.05 +	12/12 (100.0)
name the statistical package or program used *	15/28 (53.6)

Table II. Reporting of numbers and descriptive statistics.

Item	n/total (%)
round to a reasonable extent *	24/28 (85.7)
report total or group sample size for analyses *	28/28 (100.0)
report numerators and denominators for all percentages *	14/23 (60.9)
partially	4/23 (17.4)
the normality of the data was reported +	1/26 (3.8)
summarize data that were stated as normally distributed with means \$	1/1 (100.0)
summarize data that are were stated as normally distributed with standard deviations \$	1/1 (100.0)
they use means? (regardless of the distribution) +	21/23 (91.3)
they use SD? (regardless of the distribution) +	15/23 (65.2)
use the form: mean (SD), not mean \pm SD +	0/14 (0.0)
summarize data that was stated as not normally distributed with medians \$	1/1 (100)
summarize data that was stated as not normally distributed with inter-percentile ranges, ranges, or both \$	1/1 (100)
reporting medians? (regardless of the distribution) +	4/23 (17.4)
report the mean along with a non-parametric test (no median) +	4/8 (50.0)
report the upper and lower boundaries of inter-percentile ranges and the minimum and maximum values of ranges *	11/14 (78.6)
do not use the standard error of the mean (SE) to indicate the variability of a data set *	22/22 (100.0)
use standard deviations, inter-percentile ranges, or ranges *	17/23 (73.9)

Table III. Reporting hypothesis tests.

Item	n/total (%)
summarize the data for each variable with the appropriate descriptive statistics *	15/19 (78.9)
report whether the test was for paired samples *	3/19 (15.8)
report whether the test was for independent samples *	4/19 (21.1)
confirm that the assumptions of the test were met by the data *	6/19 (31.6)
report the difference between groups +	2/16 (12.5)
precision for differences between groups \$	0/2 (0.0)
use confidence intervals \$	2/19 (10.1)

Table IV. Reporting correlation analyses.

Item	n/total (%)
reporting correlation is a primary analysis +	2/7 (28.6)
identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman) *	5/7 (71.4)
confirm that the assumptions of the analysis were met *	1/7 (14.3)
report the value of the correlation coefficient *	6/7 (85.7)
do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined	8/12 (66.7)
for primary comparisons, report the (95) confidence interval for the correlation coefficient, whether or not it is statistically significant *	0/2 (0.0)
for primary comparisons, consider reporting the results as a scatter plot *	2/2 (100.0)
the correlation coefficient on the chart *	1/2 (50.0)
the confidence interval on the chart *	1/2 (50.0)

Table V. Reporting regression analyses.

Item	n/total (%)
describe the purpose of the analysis *	3/3 (100)
confirm that the assumptions of the analysis were met *	0/3 (0.0)
for example, in linear regression indicate whether an analysis of residuals confirmed the assumptions of linearity *	0/3 (0.0)
report how any outlying values were treated in the analysis *	0/3 (0.0)
report how any missing data were treated in the analyses *	0/3 (0.0)
for either simple or multiple (multivariable) regression analyses, report the regression equation *	2/3 (66.7)
for multiple analyses: report whether the variables were assessed for co-linearity *	0/2 (0.0)
for multiple analyses: report whether the variables were assessed for interaction *	0/2 (0.0)
for multiple analyses: describe the variable selection process by which the final model was developed (e.g. forward-stepwise; best subset) *	2/2 (100.0)
report the regression coefficients (beta weights) of each explanatory variable *	2/3 (66.7)
report the associated confidence intervals *	1/3 (33.3)
report the p values *	2/3 (66.7)
provide a measure of the model's "goodness-of-fit" to the data *	3/3 (100.0)
provide the coefficient of determination, R ² , for simple regression *	2/3 (66.7)
provide the coefficient of multiple determination R ² , for multiple regression *	2/2 (100.0)
specify whether and how the model was validated *	0/3 (0.0)
for primary comparisons analyzed with simple linear regression analysis, consider reporting the results graphically, in a scatter plot showing the regression line *	1/2 (50.0)
show the regression line confidence bounds *	1/1 (100.0)
do not extend the regression line beyond the minimum and maximum values of the data *	0/1 (0.0)

Summing up all the proved errors regarding statistical tests (choosing the wrong statistical test, e.g.: skewed data analyzed with non-parametric tests, using independent sample tests for paired data), we identified 2/19 (10.5%) articles with problems. If we add the papers that report means along with non-parametric test, then we have 6/23 (26.1%) articles with problems.

Discussion

We fulfilled the aim of the study by identifying the situation of reporting statistics in original articles published in Clujul Medical journal, in 2014.

The most important problems found in reporting results were low level reporting of: checking whether assumptions are met, reporting the difference between groups or measures of associations, reporting confidence intervals for the primary outcomes, and some errors in choosing the statistical test or the descriptive statistic for several analyses. There are aspects that are appropriately reported, and some that are not. There is room for improvement in almost all the observed aspects that were checked. Not all reporting issues are equally important. Some are only about clarity (e.g. number of decimals used), while others are clear errors in analysis. Not all statistical errors can be identified while assessing a paper, since there

is no access to the data, in order to check if the assumptions hold. So these errors will always stay under the radar, and will be underestimated. Their prevalence in reality is higher. Given that the available publishing space for a paper is often limited, the reporting of all the statistical aspects that can be reported is not possible.

An important advantage of this study over similar studies that assess the statistical reporting is the fact that we tried to present the percentage of reporting problems not from the total number of studies that were assessed, but from the total number of studies for which that reporting issue was relevant (e.g. the number of errors when choosing a statistical test out of the number of studies that used that statistical test). This way our results are closer to reality. A drawback is that for a naïve reader it would seem that the reporting is poorer than in other journals, since other studies that did not use this technique are biased to underreport these issues.

Another advantage of this study is its detailed nature that tries to show both important but also finer issues to be found in statistical reporting. Since there are so many aspects that were checked, there is no room to discuss all of them. Also our purpose is not to explain why the assessed issues represent errors, but only to show the observed reporting situation and point to the reader the issues that need careful attention. Literature abounds in statistical errors explanation,

their importance, and how one should report [1–3,7,14,15].

The SAMPL guidelines are not designed to assess a published paper, but to guide one writing a paper for publication. This is the reason we couldn't use all the ideas in the guideline, or why we modified certain ideas, or added new ones.

The preliminary analyses, that were used only once are not compulsory, they depend on the nature of the data, and sometimes they can be replaced by using nonparametric methods. Thus this does not represent a problem. By reading a paper it is not usually possible to assess if preliminary analyses are missing.

Regarding the reporting of primary statistical analyses, there are positive and negative findings. There was only one clear mistake. The other reporting problems cannot prove that mistakes were made due to their corresponding action. Nevertheless the most important thing to change is to report the conformity to the assumptions, followed by saying whether tests were one or two-tailed. Failure to check the assumptions or using a one-tailed test instead of a two-tailed one can cause big errors. The fact that adjustments were made for multiple comparisons is positive, but there was only one study with this situation in our assessment.

The identified mistake in reporting numbers and descriptive statistics of reporting mean along with a non-parametric test without reporting the median might be due to choosing either a wrong statistical test, or some wrong descriptive statistics, the former being the worst. Variability should be always described for variables, along with centrality statistics, and also the numerators and denominators for percentages. Although there are positive findings in reporting numbers and descriptive statistics, there is room for improvement.

As previously mentioned in reporting of primary statistical analyses chapter, the lack of checking assumptions for statistical analyses reporting is low – same is for the tests. Failure to do so is very problematic. In fact, checking assumptions for any statistical analysis tool is cornerstone. The majority of the studies that were comparing groups show only the statistics on each group, failing to help the reader with the computed difference between groups and its precision. Confidence intervals should be presented whenever possible for primary outcomes to help the reader with the generalizability of the findings.

Association analyses were well reported. Presenting a measure of association was less frequent. This is not necessarily bad practice, since sometimes one doesn't need such a measure to convey the information.

Correlation analyses reporting, as well as regression analyses reporting have the same problem as other analyses – the lack of confirming that the assumptions of the analysis were met. The trust of the reader when reading a paper without mentioning those is diminished. Also confidence intervals should be reported more for primary analyses.

Not specifying whether and how the model was validated is expected in the observed studies since the number of subjects wouldn't allow for such techniques.

We identified similar findings in other Romanian journals, but in the pharmaceutical papers [12]. Clujul Medical is a general medical journal, publishing papers in the field of medicine, dentistry and pharmacy. Since the case mix of analysis types found in Clujul Medical is different from other journals some comparisons with the literature are difficult to make. Nevertheless poor statistical reporting practices were found worldwide from high rated journals to smaller ones. Summing all the proved errors regarding statistical tests (choosing the wrong statistical test, e.g.: skewed data analyzed with non-parametric tests, using independent sample tests for paired data) we identified 2/19 (10.5%) articles with problems. If we add the papers that report means along with non-parametric test, then we have 6/23 (26.1%) articles with problems. Proved errors regarding the choice of statistical tests or errors linked to the assumptions of the analyses found in this paper (26.1%) are similar with findings from other studies: 41.6% (British Medical Journal in 1977 [16]), 31.7% (2 pharmacology journals in India - 2002-2010 [10]), 21.3%-67.9% (10 leading medical journals in China – 2008 [9]), 28.7% (80 papers in medical journals in Pakistan [7]). Difficulties in comparisons with the literature are present also due to the fact that our paper used many items that were not used in other studies assessments, and the fact that other papers identified other types of errors depending on the analyses types they found.

The poor reporting practices of statistical analyses found in our study, as well as in other studies, sustain the body of evidence that calls for a change. There is a need to improve the way statistical analyses are done, as well as they are reported. Answers to this need are guidelines for improving statistical reporting like the one used in this paper [3]; guidelines for improving reporting of specific types of studies [4–6]; initiatives like the new one called STRATOS [17] (STRengthening Analytical Thinking for Observational Studies) that are thought to guide knowledgeable statisticians in their work; using statistical advisors for journals; always asking for a statistician's advice from the design phase of a study till the writing of the article; improving statistical literacy of researchers as well as of readers of medical papers.

Conclusion

The most important issues found in reporting statistical analyses were reduced reporting of: assumptions checking, difference between groups or measures of associations, confidence intervals for the primary outcomes, and errors in statistical test choice or the descriptive statistic choice for several analyses. Measures have to be implemented to improve statistical reporting.

Acknowledgment

This paper was published under the frame of European Social Found, Human Resources Development Operational Programme 2007-2013, project no. POSDRU/159/1.5/S/138776.

References

1. Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Ann Intern Med.* 1988;108:266–273.
2. Fukuda H, Ohashi Y. A guideline for reporting results of statistical analysis in Japanese Journal of Clinical Oncology. *Jpn J Clin Oncol.* 1997;27:121–127.
3. Lang TA, Altman DG. Basic Statistical Reporting for Articles Published in Biomedical Journals: The “ Statistical Analyses and Methods in the Published Literature ” or The SAMPL Guidelines”. *Sci Ed Handb.* 2013;29–32. Available from: <http://www.equator-network.org/wp-content/uploads/2013/07/SAMPL-Guidelines-6-27-13.pdf>.
4. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* 2007;335(7624):806–808.
5. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ.* 2003;326(7379):41–44.
6. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials.* 2010;11:32.
7. Hanif A, Ajmal T. Statistical Errors in Medical Journals (a Critical Appraisal). *Ann King Edward Med Univ.* 2011;17(2):178–182.
8. Gardner MJ, Altman DG, Jones DR, Machin D. Is the statistical assessment of papers submitted to the “British Medical Journal” effective?. *Br Med J (Clin Res Ed).* 1983;286:1485–1488.
9. Wu S, Jin Z, Wei X, Gao Q, Lu J, Ma X, et al. Misuse of statistical methods in 10 leading Chinese medical journals in 1998 and 2008. *Sci World J.* 2011;11:2106–2114.
10. Jaykaran, Yadav P. Quality of reporting statistics in two Indian pharmacology journals. *J Pharmacol Pharmacother.* 2011;2(2):85–89.
11. Clujul Medical. Available from: <http://www.clujulmedical.umfcluj.ro/index.php/cjmed/index>.
12. Leucuta DC, Drugan T, Farcas A, Achimas A. Statistical reporting in pharmaceutical papers from Romanian journals. *Farmacia.* 2015;63(3):394–401.
13. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2015. Available from: <http://www.r-project.org>.
14. Lang T. Twenty statistical errors even you can find in biomedical research articles. *Croat Med J.* 2004;45(4):361–370.
15. Polychronopoulou A, Pandis N, Eliades T. Appropriateness of reporting statistical results in orthodontics: The dominance of P values over confidence intervals. *Eur J Orthod.* 2011;33:22–25.
16. Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in *BMJ* from January to March 1976. *Br Med J.* 1977;1(6053):85–87.
17. Sauerbrei W, Abrahamowicz M, Altman DG, le Cessie S, Carpenter J; STRATOS initiative. STrengthening analytical thinking for observational studies: the STRATOS initiative. *Stat Med.* 2014;33(30):5413–5432.