Assessment of Submitted Manuscripts in Medical Sciences According to Statistical Errors

Tıp Bilimlerinde Gönderilen Çalışmaların İstatistiksel Hatalara Göre Değerlendirilmesi

İlker ERCAN,[a]
Gökhan OCAKOĞLU,[a]
Deniz SIĞIRLI,[a]
Güven ÖZKAYA[a]

[a]Department of Biostatistics,
Uludağ University Faculty of Medicine,
Bursa

Yazışma Adresi/*Correspondence:*
İlker ERCAN
Uludağ University Faculty of Medicine,
Department of Biostatistics, Bursa,
TÜRKİYE/TURKEY
ercan@uludag.edu.tr

ABSTRACT Objective: The aim of the present study was to list errors in statistical analysis and presentation detected during the first review of submitted manuscripts, compared to the literature reviewing published articles. **Material and Methods:** Statistical errors determined during the first review by the biostatistics reviewer in manuscripts included in the study were presented. All studies were assessed by the same biostatistics specialist. The defined statistical errors were classified by 4 biostatistics specialist respectively. **Results:** No statistical errors were found in 4.42% of 181 original manuscripts; 73.99% of 173 original manuscripts with statistical errors were accepted following the correction of the statistical errors. The most common statistical errors were errors related to test in 34.59%, p-values in 17.36%, summarizing data in 15.60% and incorrect and insufficient demonstration of descriptive statistics in 15.60%. **Conclusion:** Considering that the errors in scientific papers are reviewed and corrected by researchers according to the suggestions of the reviewer, researchers with lack of statistical solid background are found to make errors in the process of utilizing statistics in their studies. In some cases, although researchers get statistical counseling, they make errors in the presentation while writing the papers. In conclusion, it is of major importance that the submitted manuscripts be reviewed in terms of biostatistical analyses to prevent statistical and interpretation errors.

Key Words: Statistics; biostatistics; manuscripts, medical

ÖZET Amaç: Bu çalışmada, yayımlanmış makaleler üzerinde yapılmış çalışmalardan farklı olarak, makalenin ilk değerlendirmesi aşamasında, istatistiksel uygulamalarda ve sunuşlarda yapıldığı belirlenen hataların listelenmesi amaçlanmıştır. **Gereç ve Yöntemler:** Çalışmaya alınan orijinal makalelerde, biyoistatistik hakemi tarafından ilk incelemede belirlenen istatistiksel hatalar raporlanmıştır. Tüm çalışmalar aynı biyoistatistik uzmanı tarafından değerlendirilmiştir. Belirlenen hatalar 4 biyoistatistik uzmanı tarafından sınıflandırılmıştır. **Bulgular:** Yüz seksen bir çalışmanın %4,42'sinde istatistiksel olarak bir hata bulunmamıştır. İstatistiksel hata bulunan 173 çalışmanın %73,99'u istatistiksel hataların düzeltilmesi sonucunda kabul edilmiştir. En çok görülen istatistiksel hatalar sıklık sırasına göre şu şekilde sıralanabilir: testlerle ilgili hatalar %34,59, p değeri ile ilgili hatalar %17,36, verilerin sunulmasındaki hatalar %15,60, betimleyici istatistiklerin hatalı ve eksik gösterimi %15,60. **Sonuç:** Yapılan hatalar incelendiğinde ve yazarların istenen revizyonlarda hataları düzeltme gayretleri dikkate alındığında, özellikle yeterli istatistik bilgisine sahip olmayan araştırmacıların istatistikten faydalanma sürecinde çalışmalarda hatalar yaptıkları anlaşılmaktadır. Bazı durumlarda ise araştırmacıların istatistiksel danışmanlık almalarına rağmen, çalışmaların yazım aşamasında sunumuyla ilgili hatalar yaptıkları görülmektedir. Bu nedenle, istatistiksel analiz ve yorum hatalarının engellenmesi için dergilere gönderilen çalışmaların biyoistatistiksel açıdan incelenmesi önemli ve zorunludur.

Anahtar Kelimeler: İstatistikler; biyoistatistik; makale, tıbbi

Observation, hypothesizing based on the observation, testing the hypotheses regarding the new observations, and generalizing by suggesting theories and laws are the main culprits of modern sci-

entific approach.[1] Statistics take place in every moment of this scientific process. While statistical analyses are an essential part of a scientific paper, scientist without a solid statistical background may use statistical methods or present statistical data incorrectly.

Statistical methods may be inappropriately used in every stage of a medical research related to data analysis such as the design of the experiment, data collection and pre-processing, analysis method and implementation, and interpretation. Using inappropriate statistical methods, techniques and analyses is a waste of time and financial resources and most importantly, considering scientific ethics, it is detrimental for the scientific concepts and humanity. Even when the study is carefully planned, the use of incorrect statistical approaches may end up with misleading results, which will be cited by other researchers.[2]

Unfortunately, a great number of published medical research contains statistical errors.[3] A number of studies have assessed the presence of incorrect statistical analyses and presentations in published articles.[3-10] Statistical review of articles is of great importance for editorial boards of journals to avoid incorrect and misleading results. Biostatisticians have started to take place in many editorial boards of the journals.

In published articles, the main reasons for investigating the statistical errors are as follows: (i) for authors/researchers to draw their attention to the necessity of not using statistical methods, techniques, and analyses without a solid statistical background, and to ask for biostatistical help form statisticians and (ii) for editors to draw their attention to the necessity of asking for statistical review of submitted manuscripts. The aim of the present study was to list errors in statistical analysis and presentation detected during the first review of submitted manuscripts, compared to the literature reviewing published articles. Accordingly, the results of the first biostatistical review of 181 manuscripts submitted to the Türkiye Klinikleri Journal of Medical Sciences (TKJMS) indexed by SCI expanded and other scientific indexes were reported.

MATERIAL AND METHODS

One hundred and eighty-one original articles submitted to the TKJMS between 2009 and 2011 were included in the study. Statistical errors determined during the first review by the biostatistics reviewer in manuscripts included in the study were presented. All studies were reviewed by the same biostatistics specialist (İ.E.). The defined statistical errors were classified by 4 biostatistics specialists respectively.

In the first review, statistical errors were classified as follows; "errors related to p values" (p-values expressed in closed form, p-values not included after the statistical test result, incorrect p-values, incorrect demonstration of p-values), "errors related to the statistical tests" (statistical technique used but not defined, insufficient data presented for the statistical test, incorrect name for the statistical test, statistical technique defined but not used, use of incorrect test, statistical analysis required but lacking), "errors in summarizing data", "mathematical demonstration errors", "statistical symbol errors", "incomprehensible statistical terms", "inappropriate interpretation", "errors in (statistical) terminology", "incorrect and insufficient demonstration of descriptive statistics" and "presentation of statistical method-analysis and results in the incorrect section of the manuscript".

The percentage of statistical errors were calculated first according to the number of manuscripts and second according to the total number of errors (respectively Table 1, Table 2).

In the present study, the results regarding the statistical errors in submitted manuscripts and published articles were investigated in a similar manner to that in other studies. Data of the present study were presented as median (minimum-maximum) or counts and the related percentage values.

RESULTS

Statistical errors were found in 95.58% (n=173) of the 181 manuscripts investigated and the total number of statistical errors was 795.

TABLE 1: Distribution of statistical errors in first submitted manuscripts and results of other similar studies.									
Source of Errors	First Review (Present study) Number of total reviewed manscripts=181	Welch and Gabbe[5]	Welch and Gabbe[4]	Hanif and Ajmal[9]	McGuigan[6]	Harris et al.[8]	Šimundić and Nikolac[3]	Glantz[10]	Lukiće and Marušiœe[7]
	% (n)	%	%	%	%	%	%	%	%
p-values given in closed form	37.57(68)			16.25	51.22				
p-values lacking after statistical test	9.94(18)								
Incorrect p-values	15.47(28)								
Incorrect demonstration of p-values	2.21(4)								
Statistical technique used but not defined	14.92(27)	6.21	47	26.25	13				
Insufficient data presented for the statistical test	18.78(34)			47.50					
Incorrect name for the statistical test	11.60(21)			12.50					
Statistical technique defined but not used	6.08(11)			21.25					
Use of incorrect test	28.18(51)	31.70		28.75			62	57	[ta]27 [tb]35
Statistical analysis required but not performed	26.52(48)								
Errors in summarizing data	43.65(79)				25.8				
Mathematical demonstration errors	19.34(35)								
Statistical symbol errors	1.66(3)								
Incomprehensible statistical terms	1.66(3)								
Inappropriate interpretation	10.50(19)	52.60		13.75 *32.5	17 #2	[$a]35 [$b]10 [$c]10 [$d]5	[£a]55 [£b]22		4
Errors in (statistical) terminology	20.44(37)								
Incorrect and insufficient demonstration of descriptive statistics	59.67(108)			16.25	27		34.55		[ta]16 [tb]11
Presentation of statistical method-analysis and results in the incorrect section of the manuscript	4.97(9)								

* Errors in interpreting the p-values;

[#] Misinterpreted p values, "no significant difference" interpreted as "no difference";

[$a] Lack of understanding of the limitations of the analysis, the need for replication, and sensitivity analysis;

[$b] Drawing inferences that go beyond the data, e.g., causal claims for cross-sectional data;

[$c] Comparing p values in separate tests (e.g., in paired t test) to assess group differences;

[$d] Too much made of "marginally significant" results;

[£a] Incorrect interpretation of correlation analysis;

[£b] Incorrect interpretation of p value;

[ta] Before the statistical editor assigned;

[tb] After the statistical editor assigned.

The distribution of statistical errors in 181 manuscripts and results of similar studies were listed in detail in Table 1. Table 2 presents the distribution of 795 statistical errors determined in our study.

No statistical errors were found in 4.42% (n=8) of 181 original manuscripts. Of the 173 manuscripts with statistical errors, 73.99% (n=128) were accepted following the correction of the statistical errors, 19.65% (n=34) were rejected due to inadequate correction, 5.20% (n=9) is still in the process of correction, and 1.16% (n=2) was withdrawn by the authors. Following the suggestions of

Source of Error		Number of total errors=795	
		% (n)	
Errors related to p-values	p-values given in closed form	10.19 (81)	17.36 (138)
	p-values lacking after statistical test	2.52 (20)	
	Incorrect p-values	4.15 (33)	
	Incorrect demonstration of p-values	0.50 (4)	
Errors related to tests	Statistical technique used but not defined	4.03 (32)	34.59 (275)
	Insufficient data presented for the statistical test	6.92 (55)	
	Incorrect name of statistical test	3.27 (26)	
	Statistical technique defined but not used	1.64 (13)	
	Use of incorrect test	8.30 (66)	
	Statistical analysis required but not performed	10.44 (83)	
Errors in summarizing data		15.60 (124)	
Mathematical demonstration errors		5.41(43)	
Statistical symbol errors		0.38 (3)	
Incomprehensible statistical terms		0.50 (4)	
Inappropriate interpretation		2.89 (23)	
Errors in (statistical) terminology		6.42 (51)	
Incorrect and insufficient demonstration of descriptive statistics		15.60 (124)	
Presentation of statistical method-analysis and results in the incorrect section of the manuscript		1.26 (10)	
TOTAL		100.00 (795)	

TABLE 2: Distribution of statistical errors in first submitted manuscripts according to the number of total errors.

the reviewers, the median number of revisions in accepted and rejected manuscripts was 3 (1-11) and 3 (1-6), respectively. The overall median number of revisions was 3 (1-11).

DISCUSSION

Stages of a scientific study consist of the design of the experiment, data collection and pre-processing, analysis method, implementation, and interpretation. Statistics is required at every stage of the research for scientific significance and to obtain reliable results.[2] Since these scientific processes require statistical procedures, errors in the analysis and presentation of data may occur if the researcher is not equipped with adequate statistics knowledge. Therefore, biostatistics reviewers are asked to assess the submitted manuscripts in the review process.

In some journals, following the pre-review by the editor, the submitted manuscripts are first investigated by the biostatistics reviewer and are for-

warded to reviewers of the related discipline if confirmed; on the other hand, other journals ask the reviewers of the related discipline to assess and forward the manuscripts to the biostatistics reviewer if appropriate. Reviewing both algorithms revealed that submitting the manuscripts first to the biostatistics reviewer rather than the reviewers of the related discipline would be time saving. Although reviewers of the related discipline accept the submitted manuscripts, errors can be found by the biostatistics reviewer in the analysis and presentation of data, which leads to the revision of the manuscripts. This may result with alterations in the interpretation of data and the writing of the discussion, which ultimately causes the reviewers of the related discipline to re-review the manuscript, which ends up with delays in the review process.

Biostatistics specialists publish studies in which they analyze and report the statistical errors in published articles and the first-review submitted manuscripts to draw attention of editors and re-

searchers. Occasionally, editors may ask biostatistics specialists to review the articles in terms of statistical errors so that editors can be aware of them. In conclusion, editors can revise the review process owing to the feedback given by the biostatistics specialist.

In the present study, statistical errors determined by a biostatistics specialist in the first review of submitted manuscripts pre-reviewed by the editor but not forwarded to the reviewers of the related discipline, were reported. It must be acknowledged that there is no unique definition of either "statistical error" or "statistical error rate" which makes the comparison of different statistical reviews difficult.[6]

When errors related to p-values are studied, it is clear that "p-values are not given in open forms" in 37.57% of manuscripts. This type of statistical error was reported in 16.25% of published articles by Hanif and Ajmal and in 51.22% by McGuigan.[6,9] Some authors do not consider closed forms of p-values an error. However, p-values given in open form enables published articles to be a source for meta-analysis. It also helps us determine statistical errors during the review of submitted manuscripts, if inappropriate statistical methods are used with inaccurate p-values as well as preventing unethical approaches. Additionally, readers can get more information with p-values given in open form.[11] Editors have started to ask for p-values in order to prevent generalizing based on small-scale studies. For instance, Dr. Franz J. Ingelfinger prohibited the use of the word "significant" without giving p-values in his journal during his career in "*The New England Journal of Medicine*" between 1967 and 1977.[11,12]

Other errors related to p-values in the present study were p-values lacking after the statistical test in 9.94% of manuscripts, incorrect p-values in 15.47% and incorrectly demonstrated p values in 2.21%. Šimundić and Nikolac reported incorrect demonstration of p-values in 66% in their similar study on submitted manuscripts.[3]

Regarding errors related to tests in the present study, statistical test was used but was not defined in 14.92% of manuscripts. In similar studies on published articles, this reate was reported as 6.21% by Welch and Gabbe in a former and 47% in a further study; 26.25% by Hanif and Ajmal and 13% by McGuigan.[4-6,9] In the present study, the rate of giving insufficient data for the statistical test performed was 18.78% whereas Hanif and Ajmal reported 47.50% in published articles.[9] In the present study, 11.60% of manuscripts included incorrect names of statistical tests. In their similar study on published articles, Hanif and Ajma[9] reported this rate as 12.50%.[9] The statistical technique was defined but was not used in 6.08% of manuscripts in the present study; on the other hand, Hanif and Ajmal had a higher rate of 21.25%.[9] In 28.18% of the manuscripts reviewed in our study, an incorrect statistical test was used. Welch and Gabbe, Hanif and Ajmal and Glantz reported the same rate 31.70%, 28.75% and 57% respectively in their studies on published articles.[5,9,10] Lukiæ and Marušiæ found this rate as 27% before statistical editor was assigned and 35% after assignment of statistical editor.[7] Šimundić and Nikolac reported this rate as 62% in their similar study of manuscripts in the process of submission.[3] In the present study, statistical analysis was required but not performed in 26.52% of manuscripts.

In the present study, 43.65% of manuscripts included errors in summarizing data. McGuigan found this rate as 25.8% in his study on published articles.[6] Mathematical demonstration errors were present in 19.34% of manuscripts, statistical symbol errors in 1.66% and incomprehensible statistical errors in 1.66% in our study.

The rate of inappropriate interpretation of statistical analysis results was 10.50% in our study, 52.60% in the study by Welch and Gabbe, and 4% in the report by Lukiæ and Marušiæ including published articles. Hanif and Ajmal and McGuigan found this rate as 13.75% and 17%, respectively in published articles.[5-7,9] Errors related to interpretation of p-values was reported as 32.5% by Hanif and Ajmal and 2% by McGuigan.[6,9] Harris et al. subcategorized the errors related to interpretation of statistical analysis results in published articles and reported 24% for "lack of understanding the limitations of the analysis, and the need for replication and sensitivity analysis"; 10% for "drawing infer-

ences that go beyond the data such as casual claims for cross-sectional data; 10% for "comparing p-values in separate tests (e.g. in paired t test) to assess group differences; and 5% for the "too much made from "marginally significant" results".[8] Šimundić and Nikolac found the rate of misinterpretation of correlation analyses as 55% and 22% for the misinterpretation of p values in their similar study on the manuscripts in submission process.[3]

The rate of errors in statistical terminology was 20.44% in our present study and 4.97% of manuscripts had presented their statistical methods, analyses and results in the incorrect section of the manuscript.

Incorrect and insufficient demonstration of statistical results was present in 59.67% of manuscriptsin our study. Hanif and Ajmal and McGuigan reported the same rate as 16.25% and 27%, respectively.[6,9] In the study by Lukiæ and Marušiæ the rate of incorrect and insufficient demonstration of statistical results decreased from 16% before the assignment of statistics editor to 11% after the review of the statistics editor.[7] Šimundić and Nikolac reported this rate as 34.55% in submitted articles as in our study.[3]

In contrast to the results of other studies, in the present study, 795 errors of 181 manuscripts were studied as distribution of errors within error categories. Errors related to statistical tests with 34.59% ranks the highest place among all the error types in the first review. The subcategorization of errors related to statistical tests revelaed 4.03% for statistical test used but not defined, 6.92% for insufficient data given for the statistical test performed 3.27% for incorrect name for the statistical test, 1.64% for statistical technique defined but not used, 8.30% for use of incorrect test and 10.44% for statistical analysis required but not performed.

The overall rate of errors related to p-values was 17.36%. The subcategorization of errors related to p-values revealed 10.19% for p-values not given

in open form, 2.52% for test performed but p-values not given, 4.15% for incorrectly expressed p-values and 0.50% for incorrectly presented p-values.

The following errors in the present study rated 15.60% for errors in summarizing data, 15.60% for incorrect and insufficient demonstration of descriptive statistics, 6.42% for incorrect statistical terminology, 5.41% for mathematical demonstration errors, 2.89% for interpretation errors, 1.26% for the presentation of statistical methods/analyses and results in the incorrect section of the manuscript, 0.50% for incomprehensible statistical definitions, and 0.38% for statistical symbol errors.

In general, when the error types were sorted among themselves, the most common ones were errors related to statistical tests, errors related to p-values, errors in summarizing data and incorrect and insufficient demonstration of descriptive statistics, respectively.

In conclusion, 73.99% of the manuscripts evaluated as incorrect were accepted after their errors were corrected; 19.65% were rejected due to not being corrected. Authors seem to have a tendency to correct their errors.

Increasing knowledge with the improvement of tools used to obtain knowledge and the complex structure of the knowledge required, necessitate the analysis of data and that is only provided by statistics.[2] Considering that the errors in scientific papers are reviewed and corrected by researchers according to the suggestions of the reviewer, researchers who do not have a solid statistical background are found to make errors in the process of utilizing statistics in their studies. In some cases, although researchers get statistical counseling, they make errors in the presentation while writing the papers. In conclusion, it is of major importance that the submitted manuscripts be reviewed in terms of biostatistical analyses to prevent statistical and interpretation errors.

▌REFERENCES

1. Özdamar K. [Science and scientific method]. SPSS ile Biyoistatistik. 5. Baskı. Eskişehir: Kaan Kitabevi; 2003. p.1-6.

2. Ercan I, Yazici B, Yang Y, Ozkaya G, Cangur S, Ediz B, et al. Misusage of statistics in medical research. Eur J Gen Med 2007;4(3):128-34.

3. Šimundić AM, Nikolac N. Statistical errors in manuscripts submitted to Biochemia Medica Journal. Biochem Med 2009;19(3):294-300.

4. Welch GE 2nd, Gabbe SG. Statistics usage in the American Journal of Obstetrics and Gynecology: has anything changed? Am J Obstet Gynecol 2002;186(3):584-6.

5. Welch GE 2nd, Gabbe SG. Review of statistics usage in the American Journal of Obstetrics and Gynecology. Am J Obstet Gynecol 1996;175(5):1138-41.

6. McGuigan SM. The use of statistics in the British Journal of Psychiatry. Br J Psychiatry 1995;167(5):683-8.

7. Lukić IK, Marusić M. Appointment of statistical editor and quality of statistics in a small medical journal. Croat Med J 2001;42(5):500-3.

8. Harris A, Reeder R, Hyun J. Survey of editors and reviewers of high-impact psychology journals: statistical and research design problems in submitted manuscripts. J Psychol 2011;145(3):195-209.

9. Hanif A, Ajmal T. Statistical errors in medical journals (a critical appraisal). Annals of King Edward Medical University 2011;17(2):178-82.

10. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation 1980;61(1):1-7.

11. Ercan I. [should P-values be written in the open or closed form? Letter to the editor]. Journal of Pediatric Infection 2010;4(1):47.

12. Feinstein AR. Common faults in biostatistics in medical research. Scientific Writing, Editing and Auditing in Medicine Symposium. Ankara: Tubitak Health Sciences Research Group; 1994. p.49-55.