

ORIGINAL ARTICLE



Statistical and Practical Significance of Articles at Sports Biomechanics Conferences

Uday CH. Hasan *

Department of Physical Education and Sports Sciences, Al-Kitab University, Altun Kupri, Iraq.

Submitted 03 November 2020; Accepted in final form 23 January 2021.

ABSTRACT

Background. The importance of using statistical approaches has increased and became necessary for researchers and specialists in sports biomechanics because they need more objective and accurate methods to increase knowledge. **Objectives.** Evaluate the reality of using practical significance in the articles published in scientific conferences in the biomechanical sport. **Methods.** One hundred twenty-four articles were analyzed of 134 in terms of statistical approaches to calculate practical significance. These results were then compared with those of statistical significance to reveal the extent of similarities or differences between the results. **Results.** The mean test, which was the most commonly used descriptive statistical test, was applied in 114 articles (i.e., 92%); the T-test of paired samples, which was the most used difference measurement tests, was involved in 45 papers (i.e., 36%), statistical tests that measure the relationship between variables were used in 46 articles (i.e., 37%). Likewise, no items used advanced statistical tests except for six articles (i.e., 5%), which used regression and factor analysis. T-test independent samples are the most used statistical tests in sports biomechanics articles in which the results of practical significance matched those of statistical significance (88%). **Conclusion.** The use of practical significance was almost non-existent. Also, it was observed that there was a large percentage of practical significance mismatch with the statistical significance of many statistical tests, which was a considerable negative indicator that affected the quality of sports biomechanics articles.

KEYWORDS: *Biomechanics, Statistics, Sport, Effect Size, Significance.*

INTRODUCTION

Statistics is essential in sport biomechanics articles. After researchers analyze biomechanical variables in the kinetic analysis, the results are statistically processed, presented, and described concisely and accurately. In general, statistics is a mathematical science that deals with the collection, analysis, explanation, and presentation of large amounts of numerical data to extract relevant information. These observations-based analyses provide athletes with sufficient feedback; however, athletes lack data to achieve optimal improvement. For athletes to realize their full potential, higher data accuracy is required,

which can be attained using a perfect statistical model (1). Therefore, statistical methods are essential regardless of the type of study and its purpose. Specifically, statistical methods are relevant to articles related to imitation and modeling of mechanics of muscle work, mechanics of injury and rehabilitation, mechanics of equipment for athletes, or papers related to sports training, events, and sports. Hence, statistics is the main pillar in the field of sports biomechanics research. We can identify two types of applied statistics in sport biomechanics articles that are descriptive and inferential. There are no

*. Corresponding Author:

Uday CH. Hasan, Ph.D.

E-mail: dr_odayhasan@yahoo.com

studies that do not use descriptive statistical or inferential statistics methods, especially in quantitative research; also, there are parametric and nonparametric statistics (2, 3). The importance of using statistical approaches has increased and became necessary for researchers and specialists in sports biomechanics because they need more objective and accurate methods to increase knowledge; thus, research design quality and statistical analysis directly affect this area of science (4). There are concerns that the vast majority of published research findings are false. The main problem is that both scientists and the audience confuse statistics with reality. However, statistical inference is an intellectual experiment, which shows the predictive performance of models concerning reality (2). Because many sport biomechanics articles are published in scientific journals, the sports biomechanics literature requires continuous analysis and evaluation, carried out by some researchers (4-6). To add to the already published research, we aspire to study and evaluate the obtained results in the articles published at scientific conferences on biomechanical sport held in Iraq. Scientific research in mathematical biomechanics extensively depends on statistical significance tests as a scientific method to determine whether differences or correlations between variables are statistically significant. Therefore, the statistical significance of hypothesis testing has become a traditional method owing to the researchers' belief in the statistical ritual (7). The misuse of statistics using 'statistical significance' as a license to claim a scientific finding leads to a considerable falsification of the scientific process (8). Many experts in the field suggested abolishing using the term "statistically significant" altogether (9, 10).

The classical definition of statistical significance is $p \leq 0.05$, which means a 1/20 opportunity that the test statistic is due to the null hypothesis's normal variation (11). Statistical significance indicates whether the research result is due to chance or sampling variability; practical significance is concerned with whether the result is useful in real life (12). This means that the p-values estimate the error rate in rejecting the null hypothesis, which indicates that the p-value of

0.05 may give rise to the probability of a type I error by 5% (13). Many biomechanics studies contain wrong statistical analyses; therefore, inaccurate inferences and amplification of the effects are standard in the field (14). Studies, which depended on the p-value to infer whether a specific value has statistical significance or is not essential, are confusing or maybe misleading (15). Therefore, p-value does not provide evidence for a particular model or hypothesis (16). Knudson (2009) has specified that when observing statistical significance differences, research reports must provide confidence limits or effect sizes to document the size of the effects (5). Cohen (2013) mentioned that the effect size refers to a specific phenomenon in the population (differences between the means or a relationship between the variables). This means that when the null hypothesis is wrong, then, consequently, there is a considerable difference between the means (17). It can be calculated according to the standard deviation of the differences between the means. Standard deviation can be used in a way similar to using the T-test of paired samples. As for the control and experimental groups, the support point is the control group's mean, while the effect size is the standard deviation value for the arithmetic mean of the experimental group performance compared with the control group. When analysing the variance, the effect size refers to the relation between independent variable or variables and the dependent variable or variables. This can be done by calculating the variance value of the dependent variable(s), which can be duly interpreted according to the dependent variable(s) (18).

The abovementioned reasoning clearly shows that statistical significance is insufficient to make a correct decision and cannot be used as a statistical indicator to show a difference or relationship that is not possible; also, statistical significance cannot be considered for making decisions because it is affected by the sample size; therefore, the result does not have practical importance (19). The sufficiency is achieved only by calculating the test's functional significance, which is a statistical indicator of the ability to use the interpretative or applied results, mainly because they are not affected by the sample size.

Therefore, mistakes in the statistical analysis of sports biomechanics and the documentation of practical significance are common in the sports biomechanics literature. In this study, articles from specialized scientific conferences in the field of sports biomechanics that use statistical methods will be studied, and the extent of practical significance is used as a test to determine statistical relevance to examine hypotheses, detect the authenticity of results, and identify differences between the effects of statistical meaning and practical significance. This study's main objective is to find solutions that can improve the quality of sports biomechanics articles.

MATERIALS AND METHODS

Three scientific conferences specializing in sports biomechanics were held in Iraq. All published articles at these conferences were examined and analyzed. The results are published in "Al. Qadisiyah Journal for the Sciences of Physical Education" in one volume as different parts for each conference. Some of these articles were excluded (e.g., qualitative lectures and descriptive studies that do not use statistical means). Thus, the total number of articles subject to analysis was 124 articles of a total of 134 articles. The variables that were analyzed included the number of papers at each conference, specialization of articles, number of authors in each article, the research method used, the statistical program used, method of statistical significance, level of significance, type of hypothesis, details of the statistics used (descriptive or inferential). Also, each article was analyzed using statistical approaches to calculate practical relevance. Then, these results were compared with those of statistical significance to reveal the extent of similarities or differences between the results.

In the articles, the practical significance of sport biomechanical variables was determined by calculating the effect size (ES) (17, 18). The researcher calculated the effect size only for T-tests (one sample, independent samples, and paired samples), as well as one-way ANOVA and the Pearson correlation coefficient; the rest of the statistical tests could not be used to calculate the

effect size owing to the lack of data necessary to do so; the following explanation of inference size was used (15): 0.2-0.5 = small, 0.5-0.8 = moderate, > 0.8 = large for T-tests; 0.01-0.06 = small, 0.06-0.14 = moderate, and > 0.14 = large for one-way ANOVA; 0.1-0.3 = small, 0.3-0.5 = moderate, > 0.5 = large for the correlation coefficient.

RESULTS

Table 1 shows that the smallest number of articles was in the first sports biomechanical conference (31 articles), which accounted for 25% of the total number of articles; the number of articles in the second and third conferences was similar (i.e., 47 and 46), which accounted for 37.9% and 37.1%, respectively; the smallest number of articles was in the basketball and fencing specialization (i.e., three articles), which accounted for 2.4% of the total number of articles. The most number of articles was in the track and field specialization (i.e., 45 articles; 36.3% of the total number of articles). The number of articles published by individual authors was 26 (i.e., 21%). Authors published most articles with co-authors (i.e., 98 articles; 79%); the descriptive method was used in 93 articles (i.e., 75%); experimental methods were used in 31 articles (i.e., 25%); most articles used the statistical package for social sciences (SPSS) (i.e., 73 articles; 58.9%); the number of articles that did not use this program was 18 (i.e., 14.5%). The majority of papers used a tabular value to detect the moral significance of the number (i.e., 61 articles; 49.2%); the number of articles, in which two methods were used to tabulate values and to calculate the level of significance to determine significance, was 10 (i.e., 8.1%). Most articles at sports biomechanical conferences used the significance level of 0.05 (i.e., 83 articles; 66.9%). However, the number of articles in which the authors used the significance level of 0.01 was 2 (i.e., 1.6%). The number of the hypotheses that are used by the researchers in their articles are 167; 117 times used an alternative two-tailed hypothesis (i.e., 63.9%), while the number of times that used the null hypothesis was 5 (i.e., 2.7%).

Table 1. Distribution of the Number of Sports Biomechanics Articles According to the Study Variables

Variables	No
Conference	
First	31
Second	47
Third	46
Specialization	
Track and Field	45
Gymnastic	15
Weight lifting	14
Volleyball	8
Racquet games	7
Football	8
Swimming	4
Basketball	3
Fencing	3
Handball	10
General	7
Authors	
Single	26
Co-authorship	98
Research methodology	
Descriptive	93
Experimental	31
Statistical program used	
SPSS	73
Not used	18
Not mentioned	33
Method of statistical significance	
Tabular value	61
Sig	27
Tabular value and sig	10
Not used	25
Not mentioned	1
Significance level	
0.05	83
0.01	2
0.05 and 0.01	6
Not used	20
Not mentioned	13
Type of hypothesis	
Ho	5
Two-tailed	117
One-tailed	28
No hypothesis	26

Table 2 clearly shows that the mean test, which was the most commonly used descriptive statistical test, was applied in 114 articles (i.e., 92%); the median test was used less (i.e., 7 articles; 6%). Table 3 shows that the T-test of paired samples, which was the most used difference measurement test, was applied in 45 articles (i.e., 36%), and chi-square was used less (i.e., 1 article; 1%). Statistical tests that measure the relationship between variables were used in

46 articles (i.e., 37%). Likewise, no articles used advanced statistical tests except for six articles (i.e., 5%), which used regression and factor analysis.

Figure 1 shows that T-test independent samples are the most used statistical tests in sports biomechanics articles. The practical significance results matched those of statistical significance (88%); the T-test of paired samples was the least used test (54%).

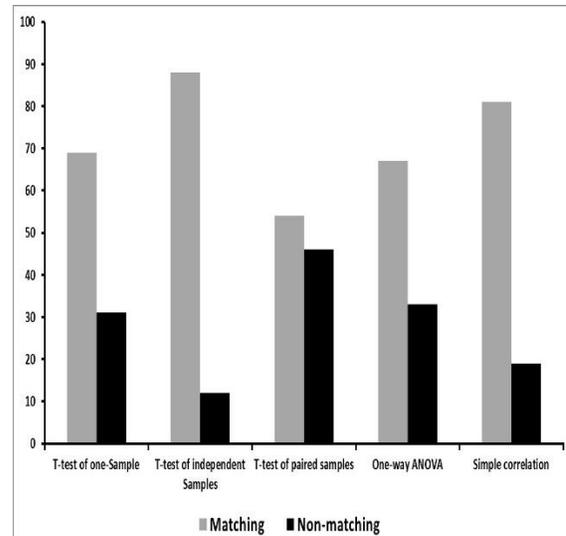


Figure 1. Percentages for Matching and Non-Matching of the Results of Practical Significance with the Results of Statistical Significance

DISCUSSION

Based on the abovementioned results, it is clear that many researchers participated in such specialized scientific conferences. Therefore, the success achieved by the first scientific conference in Iraq and the benefits provided to the researchers led to increased participation in the next panels. Therefore, these meetings are essential for the development of science; however, these meetings require the proper organization to improve the benefits that researchers receive after attending these scientific conferences (20). Many articles on track and field are attributed to the inclusion of many sports competitions (e.g., running, throwing, and jumping). Likewise, the large number of articles specializing in track and field and individual games may be due to the ease of controlling the search variables, movements, and skills that characterize these games. They are closed skills, not open, which facilitates the performance evaluation process. Biomechanics is particularly

suited for evaluating sports skills; the technical performance of these skills mainly determines its success. Because many sports with high technology components are also individual sports, biomechanics can be applied to these sports, including track and field games (21). Teamwork results in better research in terms of quantity and quality; therefore, the partnership between scientific disciplines related to biomechanics is advantageous. Besides, work-related partnerships provide an interactive environment between researchers and enhance bonds between them to serve the community. Recently, the number of co-authored articles in

applied biomechanics and kinesiology has considerably increased (22). The descriptive approach is frequently used in sports biomechanics articles because the experimental approach is more difficult compared to the descriptive system and requires to control external variables that affect the dependent variable (23). Generally speaking, in the biomechanical articles, no specified statistical methods for the test-retest reliability are available (24); also, biomechanics tools are expensive, and the processes used for data extraction is usually time-consuming (21). If the experimental method is used, a researcher will use these tools again.

Table 2. Number of Descriptive Statistical Tests Used and Not Used in Sport Biomechanical Articles

Statistical Tests	Mean	Median	Mode	Std. Deviation	Skewness	Coefficient of Variation	Simple Correlation
Used (number)	114	7	9	110	30	24	46
Not used (number)	10	117	115	14	94	100	78

Table 3. Number of Inferential Statistical Tests Used and Not Used in Sport Biomechanical Articles

Statistical Tests	T-Test of One-Sample	T-Test of Independent Samples	T-Test of Paired Samples	One-Way ANOVA	Chi-Square	Regression	Factor Analysis	Effect Size
Used (number)	2	21	45	14	1	6	6	2
Not used (number)	122	103	79	110	123	118	118	122

Regarding the statistical program used, the researcher believes that SPSS is easy to use because it works under the Windows system; therefore, a user of this program will have prior knowledge of the operating system. It will be easier to use the features of SPSS (25). Regarding the method of statistical significance, the author believes that many researchers are not able to benefit from the P-value of the calculated significance level that can be obtained as part of the output of statistical analysis results when using SPSS, as shown by the results in Table 1 (i.e., approximately 60% of the articles used it in SPSS); therefore, the users of the tabulated P-value method, when they use SPSS, adhere to the outdated ways. Also, the use of both methods simultaneously (i.e., tabulated P-value and the level of significance) is unnecessary, has no meaning, and does not change statistical decisions; thus, it is sufficient to calculate the significance level determine the significant results. The author attributes the frequent use of the level of statistical significance of 0.05 to the linking of Types I and II errors with a reciprocal relationship, and neither can be reduced at the same time; if we seek to reduce the Type II error, the probability of the Type I error will increase

and vice versa. Therefore, the strategy is to achieve an acceptable balance between the two in advance. Traditionally, this is done by specifying a sufficient (0.05) value that does not increase a lot for the Type II error (26, 27). Statistical hypothesis tests provide a framework to decide whether the observed values differ from what would be expected by coincidence (28). Related to the type of hypotheses, the two-tailed hypotheses are most frequently used because there is no supported opinion for a specific trend or the researchers have doubts about determining the direction of the relationship or differences. There is a misconception among many researchers that studies that do not reject the null hypothesis are not published (29). Many researchers have indicated that one-tailed-related hypotheses are not always suitable, and they should not be used (30).

Statistics involves the collection, description, and analysis of data that are subject to random variation. Descriptive statistics summarizes collected data according to the specialized features of their distribution. Inferential statistics aims to analyze sample data to obtain an estimate or predict characteristics of the larger population from which the sample is obtained (26).

Descriptive statistics can be considered a first step for selecting and using statistical hypothesis tests (31). The mean is the most important statistical method because it is regarded as the main pillar for conducting and understanding all other complex statistics. Therefore, it is essential to identify the most suitable statistics for different measurement levels (32). When the measured biomechanics variables are the ratio scale levels, the most appropriate method to describe these variables is the mean; this is why it is extensively used in sport biomechanics articles. The T-test is considered to be one of the inferential statistics that are widely used in sports biomechanics articles to determine whether there is a significant difference between the means of two independent or paired groups (33); this observation matches that of Yim et al. (2010) who determined that T-test is one of the most widely used statistical hypothesis tests (34). The author did not identify a considerable use of advanced inferential statistics, except in rare cases. This observation indicates a weakness of sports biomechanical articles that use factor analysis and regression analysis to reveal factors or variables that contribute the most to the achievement of results; the author did not identify any indication of the availability of conditions and assumptions for the use of inferential statistical methods because the reliability of the results of statistical analyses depends on achieving a set of beliefs that are made about the data in the published studies in physical education (6). Therefore, studies that evaluate variance in the reliability between tests or tools require designs and analyses that researchers rarely correctly apply (35).

Regarding the effect size, unfortunately, it was used only in 2 articles out of 124; this observation is a negative indication for sports biomechanics because statistical significance is not the same as practical significance (36). However, statistical significance and practical significance (effect size) are related; they complement each other, but they do not replace each other. Therefore, good research practice is required to make the right research decision; therefore, both characteristics must be considered. The effect size is essential in the experimental studies; that is why most articles in this field highlight its importance to communicate the results' practical significance. Therefore, effect sizes are most useful for scientists because they facilitate cumulative science (18). Thus, the presented results show a

mismatch between statistical significance and practical significance in many of the reported statistical tests, which indicates that this is a major challenge and that there are many errors in statistical analysis in sports biomechanics articles. The authors of research articles in sports biomechanics have to document the practical significance of the effects that reach the standard of statistical significance. Therefore, we recommend using the effect size to support the results of the statistical significance test to address statistical problems and obtain accurate and reliable results that help improve knowledge of sports biomechanics.

CONCLUSION

This study showed that researchers wanted to participate as co-authors at biomechanical conferences; their articles were in the track and field specialization and used a descriptive approach to test two-tailed hypotheses. Although many researchers used SPSS, they heavily relied on the tabulated P-value to interpret their results. There was a diversity in the use of descriptive and inferential statistical methods; however, we did not observe the use of advanced inferential statistics, except in very few cases. Unfortunately, the use of practical significance was almost non-existent. Also, it was observed that there was a large percentage of practical significance mismatch with the statistical significance of many statistical tests, which was a considerable negative indicator that affected the quality of sports biomechanics articles.

APPLICABLE REMARKS

- We recommend using the practical significance tests to complement the statistical significance tests in the biomechanical articles.
- We believe that biomechanical researchers need to refer to the actual value of the functional significance in their studies to better comprehend such studies' results. At the same time, use these results by basing future reviews on suitable theoretical and practical bases.
- We also believe that researchers could do better by relying on advanced inferential statistics due to their significance in revealing the nested variables' relations.
- A further recommendation is that the research papers that are accepted for publication in the refereed journals need to be reviewed by statisticians and referees.

REFERENCES

1. Davis SG. Statistics and Biomechanics: An Interdisciplinary Evaluation of the Mathematical, Practical, and Athletic Applications of Principal Component Analysis. Gardner-Webb University 2018.
2. Amrhein V, Trafimow D, Greenland S. Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication. *Am Stat [Internet]*. 2019;**73**(1):262-270. doi: [10.1080/00031305.2018.1543137](https://doi.org/10.1080/00031305.2018.1543137)
3. Nataliia B, Lolita D, Oksana S, Serhiyenko K, Vitaly U, Olha S. Using the methods of mathematical statistics in sports and educational research of masters in physical education and sport. *J Phys Educ Sport [Internet]*. 2019;**19**(3):1030-1034.
4. Vagenas G, Palaiothodorou D, Knudson D. Thirty-year Trends of Study Design and Statistics in Applied Sports and Exercise Biomechanics Research. *Int J Exerc Sci*. 2018;**11**(1):239-259.
5. Knudson D. Significant and meaningful effects in sports biomechanics research. *Sports Biomech*. 2009;**8**(1):96-104. doi: [10.1080/14763140802629966](https://doi.org/10.1080/14763140802629966) pmid: [19391497](https://pubmed.ncbi.nlm.nih.gov/19391497/)
6. Chen A, Zhu W. Revisiting the Assumptions for Inferential Statistical Analyses: A Conceptual Guide. *Quest [Internet]*. 2001;**53**(4):418-439. doi: [10.1080/00336297.2001.10491756](https://doi.org/10.1080/00336297.2001.10491756)
7. Gigerenzer G. Statistical Rituals: The Replication Delusion and How We Got There. *Adv Methods Pract Psychol Sci [Internet]*. 2018;**1**(2):198-218. doi: [10.1177/2515245918771329](https://doi.org/10.1177/2515245918771329)
8. Amrhein V, Greenland S, McShane BB. Statistical significance gives bias a free pass. *Eur J Clin Invest*. 2019;**49**(12):e13176. doi: [10.1111/eci.13176](https://doi.org/10.1111/eci.13176) pmid: [31610012](https://pubmed.ncbi.nlm.nih.gov/31610012/)
9. Gibson EW. The Role of p -Values in Judging the Strength of Evidence and Realistic Replication Expectations. *Stat Biopharm Res [Internet]*. 2020:1-13. doi: [10.1080/19466315.2020.1724560](https://doi.org/10.1080/19466315.2020.1724560)
10. Wasserstein RL, Schirm AL, Lazar NA. Moving to a World Beyond " p < 0.05.". *Am Stat [Internet]*. 2019;**73**(1):1-19. doi: [10.1080/00031305.2019.1583913](https://doi.org/10.1080/00031305.2019.1583913)
11. Heston T. A new definition of statistical significance. *J Nucl Med*. 2013;**54**(2):1262.
12. Kirk RE. Practical Significance: A Concept Whose Time Has Come. *Educ Psychol Meas [Internet]*. 1996;**56**(5):746-759. doi: [10.1177/0013164496056005002](https://doi.org/10.1177/0013164496056005002)
13. Harrison AJ, McErlain-Naylor SA, Bradshaw EJ, Dai B, Nunome H, Hughes GTG, et al. Recommendations for statistical analysis involving null hypothesis significance testing. *Sports Biomech*. 2020;**19**(5):561-568. doi: [10.1080/14763141.2020.1782555](https://doi.org/10.1080/14763141.2020.1782555) pmid: [32672099](https://pubmed.ncbi.nlm.nih.gov/32672099/)
14. Knudson D. Confidence crisis of results in biomechanics research. *Sports Biomech*. 2017;**16**(4):425-433. doi: [10.1080/14763141.2016.1246603](https://doi.org/10.1080/14763141.2016.1246603) pmid: [28632059](https://pubmed.ncbi.nlm.nih.gov/28632059/)
15. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. *Int J Sports Physiol Perform*. 2006;**1**(1):50-57. pmid: [19114737](https://pubmed.ncbi.nlm.nih.gov/19114737/)
16. Wasserstein RL, Lazar NA. The ASA Statement on p -Values: Context, Process, and Purpose. *Am Stat [Internet]*. 2016;**70**(2):129-133. doi: [10.1080/00031305.2016.1154108](https://doi.org/10.1080/00031305.2016.1154108)
17. Cohen J. Statistical power analysis for the behavioral sciences. New York: Routledge; 2013.
18. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;**4**:863. doi: [10.3389/fpsyg.2013.00863](https://doi.org/10.3389/fpsyg.2013.00863) pmid: [24324449](https://pubmed.ncbi.nlm.nih.gov/24324449/)
19. Wilkinson L. Statistical methods in psychology journals: Guidelines and explanations. *Am Psychol [Internet]*. 1999;**54**(8):594-604. doi: [10.1037/0003-066X.54.8.594](https://doi.org/10.1037/0003-066X.54.8.594)
20. Pierce G. The dilemma of attending (or not) scientific conferences. *Can J Physiol Pharmacol [Internet]*. 2014;**92**(1). doi: [10.1139/cjpp-2013-0412](https://doi.org/10.1139/cjpp-2013-0412)
21. Lees A. Biochemical assessment of individual sports for improved performance. *Sports Med*. 1999;**28**(5):299-305. doi: [10.2165/00007256-199928050-00001](https://doi.org/10.2165/00007256-199928050-00001) pmid: [10593642](https://pubmed.ncbi.nlm.nih.gov/10593642/)
22. Knudson D. Twenty Years of Authorship, Sampling, and References in Kinesiology Research Reports. *Int J Kinesiol High Educ [Internet]*. 2017;**1**(2):44-52. doi: [10.1080/24711616.2017.1282760](https://doi.org/10.1080/24711616.2017.1282760)
23. Gravetter FJ, Forzano LB. Research methods for the behavioural sciences. 4th ed. Canada: Wadsworth Cengage Learning; 2012.
24. Pini A, Markstrom JL, Schelin L. Test-retest reliability measures for curve data: an overview with recommendations and supplementary code. *Sports Biomech*. 2019:1-22. doi: [10.1080/14763141.2019.1655089](https://doi.org/10.1080/14763141.2019.1655089) pmid: [31578129](https://pubmed.ncbi.nlm.nih.gov/31578129/)

25. Šebjan U, Tominc P. Impact of support of teacher and compatibility with needs of study on usefulness of SPSS by students. *Comput Human Behav [Internet]*. 2015;**53**:354-365. doi: [10.1016/j.chb.2015.07.022](https://doi.org/10.1016/j.chb.2015.07.022)
26. Hazra A, Gogtay N. Biostatistics Series Module 2: Overview of Hypothesis Testing. *Indian J Dermatol*. 2016;**61**(2):137-145. doi: [10.4103/0019-5154.177775](https://doi.org/10.4103/0019-5154.177775) pmid: [27057011](https://pubmed.ncbi.nlm.nih.gov/27057011/)
27. Cramer AO, van Ravenzwaaij D, Matzke D, Steingroever H, Wetzels R, Grasman RP, et al. Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychon Bull Rev*. 2016;**23**(2):640-647. doi: [10.3758/s13423-015-0913-5](https://doi.org/10.3758/s13423-015-0913-5) pmid: [26374437](https://pubmed.ncbi.nlm.nih.gov/26374437/)
28. George BJ, Beasley TM, Brown AW, Dawson J, Dimova R, Divers J, et al. Common scientific and statistical errors in obesity research. *Obesity (Silver Spring)*. 2016;**24**(4):781-790. doi: [10.1002/oby.21449](https://doi.org/10.1002/oby.21449) pmid: [27028280](https://pubmed.ncbi.nlm.nih.gov/27028280/)
29. Figueiredo Filho DB, Paranhos R, Rocha EC, Batista M, Silva JrJA, Santos MLWD. When is statistical significance not significant? *Brazilian Polit Sci Rev [Internet]*. 2013;**7**(1):31-55. doi: [10.1590/S1981-38212013000100002](https://doi.org/10.1590/S1981-38212013000100002)
30. Banerjee A, Chitnis UB, Jadhav SL, Bhawalkar JS, Chaudhury S. Hypothesis testing, type I and type II errors. *Ind Psychiatry J*. 2009;**18**(2):127-131. doi: [10.4103/0972-6748.62274](https://doi.org/10.4103/0972-6748.62274) pmid: [21180491](https://pubmed.ncbi.nlm.nih.gov/21180491/)
31. Twycross A, Shields L. Statistics made simple Part 2 Standard deviation, variance and range. *Paediatr Nurs [Internet]*. 2004;**16**(5):24. doi: [10.7748/paed.16.5.24.s24](https://doi.org/10.7748/paed.16.5.24.s24)
32. McHugh ML. Descriptive statistics, Part II: Most commonly used descriptive statistics. *J Spec Pediatr Nurs*. 2003;**8**(3):111-116. doi: [10.1111/j.1088-145x.2003.00111.x](https://doi.org/10.1111/j.1088-145x.2003.00111.x) pmid: [12942890](https://pubmed.ncbi.nlm.nih.gov/12942890/)
33. Derrick B, White P. Comparing two samples from an individual Likert question. *Int J Math Stat*. 2017;**18**(3).
34. Yim KH, Nahm FS, Han KA, Park SY. Analysis of statistical methods and errors in the articles published in the Korean journal of pain. *Korean J Pain*. 2010;**23**(1):35-41. doi: [10.3344/kjp.2010.23.1.35](https://doi.org/10.3344/kjp.2010.23.1.35) pmid: [20552071](https://pubmed.ncbi.nlm.nih.gov/20552071/)
35. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;**30**(1):1-15. doi: [10.2165/00007256-200030010-00001](https://doi.org/10.2165/00007256-200030010-00001) pmid: [10907753](https://pubmed.ncbi.nlm.nih.gov/10907753/)
36. Gelman A, Stern H. The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant. *Am Stat [Internet]*. 2006;**60**(4):328-331. doi: [10.1198/000313006X152649](https://doi.org/10.1198/000313006X152649)