TPM®

THE STATUS OF STATISTICAL ASSUMPTIONS IN PUBLISHED NURSING RESEARCH

MUAYYAD AHMAD
THE UNIVERSITY OF JORDAN

ENAAM AL-ANANBEH
MUTAH UNIVERSITY

HUTHAIFAH KHRAIS
ZARQA UNIVERSITY

MAYSOUN ATOUM
THE HASHEMITE UNIVERSITY, ZARQA

RULA AL-RIMAWI
AL-BALQA' APPLIED UNIVERSITY, AL-SALT

MOHAMMAD ALSADI
ELECTRONIC HEALTH SOLUTIONS, AMMAN

MOHAMMAD ALZAATREH
UNIVERSITY OF TABUK

Purpose: To examine the presence and appropriate usage of the statistical assumptions in nursing journals that have an impact factor. Methods: Using a retrospective exploratory design, 90 studies were examined. All the selected papers are classified by the Institute for Scientific Information and have an impact factor. The reviewed papers were distributed equally over the years 2005, 2010, and 2015. Results: More than half of the examined studies did not report any testing for the statistical assumptions. Even more, the percentage of the journals that did not report the statistical assumptions were the ones with the highest impact factor. Conclusion: The findings of this study will contribute in the direction of professional development by increasing the awareness of the importance of explicit reporting the statistical assumptions appropriately and adequately in the published nursing research.

Key words: Statistical assumptions; Nursing research; ISI; Impact factor; Methods.

Correspondence concerning this article should be addressed to Muayyad Ahmad, School of Nursing, The University of Jordan, Amman 11942, Jordan. Email: mma4jo@yahoo.com

The American Psychological Association (APA, 2010) and American Educational Research Association (AERA, 2016) suggested a guideline for reporting the statistical analysis in research. The emphasis was placed on reporting the effect size, sampling, and assumptions for the selected statistical analysis (Peng, Chen, Chiang, & Chiang, 2013). All parametric tests assume accomplishment of certain characteris-

tics of the data, known as assumptions such as normal distribution and linear relationship (Adams & Lawrence, 2015; Nimon, 2012). Violation of these assumptions changes the conclusion of the research and interpretation of the results. Therefore, all research, whether for a journal article, thesis, or dissertation, must test and report these assumptions for accurate analyses and results' interpretations.

Since each inferential statistical analysis has a set of assumptions, it is crucial to mention them in the research report. For example, Onwuegbuzie and Leech (2006) summarize the assumptions for correlation analysis and comparative analysis (such as *t*-test, ANOVA, etc.). Many researchers emphasize the importance of discussing the statistical assumptions in analyzing and interpreting the results (Chen & Zhu, 2001; Garson, 2012; Hoekstra, Kiers, & Johnson, 2012; Lumley, Diehr, Emerson, & Chen, 2002; Nimon, 2012; Osborne & Waters, 2002). Furthermore, Palmer and Sesé (2013) deemed that testing statistical assumptions is a prerequisite for producing appropriate analysis, and any failure in this testing will generate a misrepresentative result.

Although many researchers may underestimate the importance of examining and reporting statistical assumption, Chen and Zhu (2001) reported that the trustworthiness of results relies mainly on the accomplishment of the assumptions. The violation of statistical assumptions of any statistical test may increase proportions of Type I and Type II error, and the likelihood of reporting imprecise results (Osborne & Waters, 2002). On the other hand, if the validity evidence of the results is not offered, this will require the readers to create their own conclusion (Adams & Lawrence, 2015).

Previous studies found that researchers were inattentive in reporting the examination of assumptions for the used statistical analyses (Chen & Zhu, 2001; Hoekstra et al., 2012). This study is essential to the field of nursing, as the concern nowadays is on global health and the contribution of nursing science is vital to global health issues. Furthermore, the quantification of the value of scientific literature such as nursing is used to evaluate the significance of a discipline's contributions to the body of knowledge. Recently, the recognized international scientific publishing has given more emphasis to bibliometrics, particularly journals' impact factors. The IF has become widely used as a measure of the quality of professional journals (Johnstone, 2007).

Quality assessment approach of scientific literature is provided in a number of ways, such as international or national peer review, external expert review, and internal self -assessment by institutional experts. Both impact factors and citation rates are the dominant approaches to analyze quantitative quality journals. The impact factor provided by the Institute for Scientific Information (ISI) is a measure used to rank and compare journals in different fields. Impact factor is defined as ''a measure of the frequency with which the 'average article' in a journal has been cited in a particular year or period'' (Garfield, 1994 as cited in Johnstone, 2007, p. 36). The numerical impact factor represents only those journals selected for the ISI databases according to criteria defined by ISI (Thomson Reuters, 2015). International use of the impact factor supports the demand that journals ranked by the ISI embrace articles of high-level quality. In nursing, there are 110 journals that have an impact factor with a range from .015 to 2.901 (Thomson Reuters, 2015).

Each journal has unique criteria to accept the paper for publication. There is a scarcity of published studies on the methodological issues that cover the statistical guidelines in the higher-education level. One content analysis study of three journals by Hutchinson and Lovell (2004) found that half of the reviewed articles would need more attention in reporting the statistical analysis and research methodology. The purpose of this paper was to examine the presence and appropriate usage of statistical assumptions in published nursing research. The specific aims were to: 1) explore a variety of nursing journals with different impact factors over three periods 2005, 2010, and 2015; and 2) examine the explicit reporting of statistical assumptions in published nursing research.

METHOD

Procedure of Data Collection

The technique of journal selection was performed with an objective method using a retrospective design. The number of nursing journals as classified by the ISI is 110 (Thomson Reuters, 2015). For this paper, and based on the impact factor, first, we selected 30 nursing journals. All the 110 journals were ranked in a descending order based on the impact factor, then the highest 10, the middle 10, and the lowest 10 were selected. Then, 10 journals, with the highest impact factor (1.741 to 2.901) were classified in level one, 10 journals in level two with the moderate impact factor (1.000 to 1.106), and 10 journals in level three with the lowest impact factor (0.115 to 0.404). Then, from each of these journals, three random papers published in 2005, 2010, and 2015 were selected (Thomson Reuters, 2015). The reason for choosing a 5-year interval was to reduce the possibility of having the influence of the same editor in the journal over two occasions of data collection. In this way the effect of the journal guideline rather than the editor style in each phase of data collection will be presented.

The eligibility criteria for selected papers were to: 1) be published in nursing journals under ISI classification; 2) have inferential statistics such as *t*-test, analysis of variance, and regression; 3) be published in English. In the designated years for the purpose of this study, the number of issues for each journal was determined, then one issue was randomly selected. Each of the chosen papers from the selected issue was also randomly selected with the condition of meeting the eligibility criteria; if the paper did not meet these criteria, the next paper in the issue was selected. The total sample size for the examined articles was 90. The selected articles were summarized in one matrix.

Measures

After selecting the journal articles as per the study eligibility criteria, the researchers critically reviewed each article regarding the necessary assumptions in relation to the statistical test used in the paper. Then, the necessary information for each article was entered into the matrix that was developed by the researchers using Microsoft Excel sheet. The entered information about each article included the journal name, impact factor of the journal, year of publication, authors' names, article title, study design, sampling design, sample size, inferential statistics used, and reporting status of the assumptions. Finally, frequencies and percentages were calculated and entered into the three tables presented in this study.

RESULTS

The 90 nursing studies reviewed in this paper are summarized in Table 1. It was found that the articles were conducted in 20 different countries. A high percentage (74.45%) of these studies used a convenience sampling design, while the cluster random design was the least utilized one. The randomization in the sampling design was utilized only in 10 studies (11.12%). The design of 71 studies (78.88%) was non-experimental. The inferential statistics used in the reviewed articles were: ANOVA (n = 25 times, 83.33%); ANCOVA (n = 4, 13.33%); *t*-test (n = 16, 53.33%); multiple linear regression (n = 27, 90.0%); binary logistic regression (n = 3; 10.0%); correlation (n = 26, 86.66%), considering that studies usually have more than one statistical test analyses.

TABLE 1
Characteristics of the journal articles in the study (N = 90)

Characteristic	IF^ = 1.741 - 2.901 Count (%)*	IF^ = 1.0 - 1.106 Count (%)*	IF^ = 0.115 - 0.519 Count (%)*	Total (%)*
Study Design				
Descriptive	20 (66.67)	18 (60.00)	20 (66.67)	58 (64.45)
Descriptive Correlational	3 (10.00)	4 (13.33)	4 (13.33)	11 (12.22)
Descriptive Observational		1 (3.33)	1 (3.33)	2 (2.22)
Experimental	3 (10.00)	5 (16.67)	2 (6.67)	10 (11.12)
Quasi-experimental	3 (10.00)	2 (6.67)	3 (10.00)	8 (8.88)
Pre-experimental	1 (3.33)			1 (1.11)
Total	30	30	30	90
Sample Design				
Convenience	22 (73.34)	26 (86.66)	19 (63.33)	67 (74.45)
Purposive		2 (6.67)	9 (30.01)	11 (12.22)
Consecutive	1 (3.33)		1 (3.33)	2 (2.22)
Simple random	4 (13.33)	2 (6.67)		6 (6.67)
Stratified random	2 (6.67)		1 (3.33)	3 (3.33)
Cluster random	1 (3.33)			1 (1.11)
Total	30	30	30	90

Note: IF^ = Impact Factor; *Percentage is calculated out of the number of articles per group/cell.

We found that the most frequently reported statistical assumptions were homogeneity of variance, the normality distribution, linearity, and the lack of multicollinearity. Out of the 90 journal articles included in this study, 62.2% (n = 55) did not report any examinations of statistical tests' assumptions, while 32.2% (n = 29) partially reported some aspects of the statistical assumptions. On the other hand, only 6.6% (n = 6) of the articles reported properly and completely the examinations of statistical tests' assumptions (Table 2).

TABLE 2
Assumptions reported per Impact Factor

Impact Factor	Assumptions testing not reported (%)	Assumptions testing partially reported (%)	Assumptions testing fully reported (%)
1.741 - 2.901	21 (23.3)	8 (8.9)	1 (1.1)
1.000 - 1.106	20 (22.3)	8 (8.9)	2 (2.2)
0.115 - 0.519	14 (16.6)	13 (14.4)	3 (3.3)
Total	55 (62.2)	29 (32.2)	6 (6.6)

The assumptions were considered as completely reported only if all the relevant assumptions to the statistical test were mentioned; and partially reported if incompletely mentioned. Surprisingly, the percentage of the journals which did not report the statistical assumptions were the ones with the highest impact factor (23.3%), then the moderate (22.3%), followed by the ones with the lowest impact factor

(16.6%). Moreover, it was found that the journals of the lowest impact factor have the highest percentage of articles that reported the statistical tests' assumptions (3.3%, n = 3).

Table 3 shows that the highest percentage of journal articles that did not report the statistical assumptions was in the year 2005 (22%, n = 20), followed by articles published in 2010 (20%, n = 18), and then articles published in 2015 (19%, n = 17). Whereas the percentage of completely reported assumptions slightly increased from 2005 (1%, n = 1) to 2015 (5%, n = 4).

TABLE 3
Assumptions reporting per year of publication

Year of publication	Assumptions testing not reported (%)	Assumptions testing partially reported (%)	Assumptions testing fully reported (%)
2005	20 (22)	9 (10)	1 (1)
2010	18 (20)	11 (12)	1 (1)
2015	17 (19)	9 (10)	4 (5)
Total	55 (61)	29 (32)	6 (7)

DISCUSSION

The uniqueness and the credibility of any research findings depend on many factors. One of the most important factors is the appropriateness of the statistical tests used to analyze the research questions or hypotheses (Adams & Lawrence, 2015; Hoekstra et al., 2012). All inferential statistical analyses have underlying assumptions that should be met. The violation of these assumptions might have no effect on the research conclusion. Conversely, it can produce a profound effect that ultimately undermines the trustworthiness of the study (Garson, 2012; Onwuegbuzie & Leech, 2006). In this paper, we assessed the explicit reporting and appropriate testing of statistical assumptions in 90 published articles in ISI journals selected equally over the years 2005, 2010, and 2015.

Surprisingly, our findings revealed that almost two-thirds (61%) of the examined studies did not report any examination of the assumptions required for the selected statistical analyses. This result is congruent with the findings of other researchers who indicated that most of the reviewed articles did not report the examination of assumptions for the used statistical analyses (Chen & Zhu, 2001; Hoekstra et al., 2012).

Another important finding was that about one-third of the reviewed studies (32%) partially examined the required assumptions. This result is consistent with an old meta-analysis conducted by Hutchinson and Lovell (2004) which concluded that researchers need more advanced training in reporting statistical analyses appropriately. Furthermore, our findings indicated that the reporting of statistical assumptions examination appropriately is very low (7%). Similarly, previous articles have reported that only 9% of the researchers reported statistical assumptions examination clearly and appropriately (Chatoupis & Vagenas, 2011; Chen & Zhu, 2001). Poor adherence to statistical assumptions testing is attributed to the fact that researchers lack the necessary knowledge about assumptions, how they are tested, and the robustness of the analytical procedure used (Hoekstra et al., 2012; Nimon, 2012).

Moreover, the current study unexpectedly showed that the journals with the highest and moderate impact factor have less reporting of the assumptions' testing. Conversely, the journals with the lowest impact factor more frequently report assumptions testing. This result coincides with Starbuck's (2005) study

which found that the majority of the articles published in prestigious journals contain either improper reporting or missing statistical assumptions.

Implications and Recommendations

One of the main contributions of this study is that it encourages researchers and educators to seek and possess knowledge and skills necessary to understand and practice testing statistical assumptions. Among the reviewed journals for this study, only three of them focused on statistical assumptions in its authors' guidelines. These journals clearly stated that they expected authors to report the analysis plan and testing of statistical assumptions such as normality of distribution, lack of multicollinearity, level of measurement, heteroscedasticity in addition to consulting a statistician.

As the results revealed that most studies did not adequately or explicitly report the assumptions testing, an investigation into the causes for this issue is recommended as well as seeking the opinion of editors from different journals to assess the criteria for evaluating the statistical analyses' techniques. On the other hand, the publication manuals, such as the one by the American Psychological Association and journals' guidelines for authors, must include directions and rules on statistical assumptions' reporting and emphasize the importance of this condition for submitting research papers. Furthermore, specialized training programs may be needed for nursing researchers on how to examine and appropriately report the relevant statistical assumptions. Journal editors must ascertain the absence of errors related to statistical assumptions in order to improve the quality of nursing research and empower the advancement of the nursing profession. The question, whether authors of the examined studies really did not care about testing assumptions or whether they simply failed to report the results can be considered as a limitation for our study. Furthermore, it is recommended that future research follows the preferred reporting items for systematic reviews and meta-analyses (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2010) guidelines for systematic review and meta-analysis. A systematic review could contribute significantly more to the literature than a random selection of papers.

Conclusion

In contrary to expectations, the percentage of journals with a low-impact factor, which reported the statistical assumptions' testing, was higher than the moderate or high-impact factor journals. A considerable number of manuscripts are usually rejected because of inappropriate designs or flaws in reporting, but few are rejected as a result of a flaw in examining the statistical assumptions. If nursing journals continue to use the same evaluation criteria, their reputation might be jeopardized. Hopefully, this study may increase the awareness of researchers and journal editors towards testing and explicitly reporting the statistical assumptions appropriately and adequately in the published manuscripts.

Acknowledgements

The authors acknowledge the support from the University of Jordan.

REFERENCES

Adams, K., & Lawrence, E. (2015). *Research methods, statistics, and applications.* Thousand Oaks, CA: Sage.

American Educational Research Association (AERA). (2016). *Standards for Reporting on Empirical Social Science Research in AERA Publications*. Retrieved from http://www.aera.net/Portals/38/docs/12ERv35n6_Standard4Report%20.pdf

American Psychological Association (APA). (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: APA.

Chatoupis, C., & Vagenas, G. (2011). An analysis of published process-product research on physical education teaching methods. *International Journal of Applied Sports Sciences*, *23*(1), 271-289.

Chen, A., & Zhu, W. (2001). Revisiting the assumption for inferential statistical analysis: A conceptual guide. *Quest*, *53*, 418-439. doi:10.1080/00336297.2001.10491756

Garson, D. (2012). *Testing statistical assumption*. Asheboro, NC: Statistical Publishing Associate.

Hoekstra, R., Kiers, H., & Johnson, A. (2012). Are assumptions of well-known statistical techniques checked, and why (not)? *Frontiers in Psychology*, *3*, 137. doi:10.3389/fpsyg.2012.00137

Hutchinson, S., & Lovell, C. (2004). Review of methodological characteristics of research published in key journals in higher education: Implications for graduate research training. *Research in Higher Education*, *45*, 383-403. doi:10.1023/B:RIHE.0000027392.94172.d2

Johnstone, J. (2007). Journal impact factors: Implications for the nursing profession. *International Nursing Review*, 54, 35-40. doi:10.1111/j.1466-7657.2007.00527.x

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*, 151-169. doi:10.1146/annurev.publhealth.23.100901.140546

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. J. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Surgery of Surgery*, *8*, 336-341. doi:10.1016/j.ijsu.2010.02.007

Nimon, K. (2012). Statistical assumptions of substantive analyses across the general linear model: A mini-review. *Frontiers in Psychology*, *3*, 233. doi:10.3389/fpsyg.2012.00322

Onwuegbuzie, A., & Leech, N. (2006). Linking research questions to mixed methods data analysis procedures 1. *The Qualitative Report*, *11*(3), 474-498.

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, *8*(2). Retrieved from https://pareonline.net/getvn.asp?v=8&n=2

Palmer, A., & Sesé, A. (2013). Recommendations for the use of statistics in clinical and health psychology. *Clínica y Salud*, *24*, 47-54. doi:10.5093/cl2013a6

Peng, C.-Y., Chen, L.-T., Chiang, H.-M., & Chiang, Y.-C. (2013). The impact of APA and AERA guidelines on effect size reporting. *Education Psychological Review*, *25*, 157-209. doi:10.1007/s10648-013-9218-2

Starbuck, W. (2005). How much better are the most-prestigious journals? The statistics of academic publication. *Organization Science*, *16*, 180-200. doi:10.1287/orsc.1040.0107

Thomson Reuters, T. (2015, June 18). *Clarivate Analytics – Master Journal list*. Retrieved from http://ip-science.thomsonreuters.com/cgi-bin/jrnlst/jloptions.cgi?PC=master