letters

RE: Perceptions of confidence among Saudi nursing interns during internship program: a cross-sectional study

The use of nonparametric statistical tests

To The Editor: I have read with care the paper of Aboshaiqah et at. entitled "Perceptions of confidence among Saudi nursing interns during internship program: a cross-sectional study".[1] I acknowledge the output of the authors, but found statistical errors which I would kindly like to bring to their notice.

Despite that the indicators in their research showed evidence of normal distribution of their data and linearity of variables, the authors used nonparametric statistical tests (Mann Whitney U test and Spearman Rank correlation). Actually, their sample size is big enough (n=301 interns). Also, the mean value of the overall perceived confidence score equals the median and has a standard deviation (SD) less than one third of this mean (both the mean and median=3.7, SD 0.9). Additionally, the mean value of the overall perceived impotence score almost equals the median and has a SD less than one third of the mean (the mean=4.1, SD 0.9, median=4.4). The median should appear as a integer and not as a decimal. There are two variables in Table 5 which were not used in Table 4.

Besides, the figures of perceived confidence and impotence were presented in the abstract in mean values and the authors wrote only the P of significance for Spearmen correlation and only for the confidence part of the SECS. I think the authors should clarify in their methods why they used the nonparametric statistics especially when para-

metric tests are more robust in this case.

Kitchen stated in her paper that the assumptions for non-parametric tests are weaker than those of parametric and researchers are better only resort to the first if normality of data was not met or the underlying distributions are heavily tailed or extremely skewed and/or the sample size is small (n<25).[2]

Mustafa Afifi

From the College of Health Sciences, Jumeira University, Dubai, United Arab Emirates

REFERENCES

1. Ahmad E. Aboshaiqah, Regie B. Tumala, Joel G. Patalagsa, Ahmed A. Al-Khaibary, Haya Al Fozan, John Paul Ben T. Silang. Perceptions of confidence among Saudi nursing interns during internship program: a cross-sectional study. Ann Saudi Med. 2018 Jul-Aug; 38(4): 288–298.
2. Kitchen, Christina M. R. Nonparametric versus parametric tests of location in biomedical research. Am J Ophthalmol. 2009 Apr; 147(4): 571–572.

Reply

Thank you very much Sir for providing your valued feedback to our published manuscript. We greatly value and appreciate learning from your wisdom and expertise. We open-heartedly take notice of your points.

1. On why we opted to use nonparametric statistics (Mann-Whitney U test and Spearman rank correlation coefficient: Classifying data based on the level of measurement is the basis for the statistical test that we used. When the dependent variable data is an ordinal level of measurement and assumption of normality of distribution is violated, nonparametric statistics may be used.[1] The reason we used nonparametric statistics (Mann-Whitney U Test and Spearman rank correlation coefficient) is that we treated the dependent variable (competence and importance) as an ordinal level of measurement since participants subjectively rate their competence or importance of a particular item. This is briefly described in the manuscript under METHODS section from line 21 to line 25 at the right column on page 289. The aggregated average therefore is a representation of the subjective ratings of the participants from the 30 items in the scales. One of the premises in treating our data as ordinal level of measurement is based on Field's suggestion[2] that in situations where people are asked to rate something subjectively the data should probably be regarded as ordinal. There are contrasting schools of thoughts on how to consider the level of measurement of a Likert-type scale data. One school of thought suggests that Likert-type scale is an ordinal level of measurement but can be considered an interval level of measurement if there are five response points in

the scale (Johnson & Creech, 1993 as cited by Langdridge & Hagger-Johnson, 2009, p47).[3] On the other hand, Miles and Banyard, 2007, (as cited by Langdridge & Hagger-Johnson, 2009, p47)[3] suggested that Likert-type scales with at least seven points may be considered an interval level of data.[2] This is supported by Tabachnick and Fidell[4] who stated that variables are often treated by statisticians or researchers as if they are continuous when the scale is thought to be continuous, but apparently the measured scale is ordinal, the number of scale is large, that is, seven or more and the data meet other assumptions of the analysis like normality in distribution.[4] Based on the apparent positions of Miles and Banyard (2007) and Tabachnick and Fidell,[4] Likert-type scales with less than seven response points are apparently considered an ordinal level of measurement and represent our second premise for considering competence and importance scales as ordinal level of measurement since our items have five response points. In summary, the suggestions of Field[2] that subjective ratings may be considered as ordinal level of measurement and the suggestions of Miles and Barnyard (2007) and Tabachnick and Fidell[4] that Likert-type scales with less than seven response points are considered an ordinal level of measurement support our premise that competence and importance scales have ordinal level of measurement. Since our dependent variables competence and importance are ordinal levels of measurement, we used nonparametric statistics.

2. We acknowledge that our data may be considered normally distributed based on a large sample (n=301) as per the central limit theorem.[1] This may warrant the use

of parametric statistics. However, we put weight on the level of measurement of our data as ordinal which called for the use of non-parametric statistics.

3. The use of mean in our paper is for the purpose of interpreting the responses of the participants per item and overall confidence and importance. Since we used Mann-Whitney U test, the median is used in the comparative data presented in Table 4.

4. The following table shows

egories since they are all public tertiary training hospitals and we could not make distinctions. We could not make credible comparisons of school graduated since the participants in private schools are relatively few (n=22; 7%) (Table 1, p.291) compared with the number of participants in public schools. Besides, the figures for perceived confidence and impotence were presented in the abstract in mean values and the authors wrote only the *P* value of significance for the

Mean range interpretation table

Mean range	Confidence Interpretation	Importance Interpretation
5.00-4.21	Quite a lot of confidence	Very important
4.20-3.41	Confident	Important
3.40-2.61	Fairly confident	Fairly important
2.60-1.81	Little confidence	Little importance
1.80-1.00	Very Little confidence	Very little importance

the mean range interpretation that we used for each item and competence and importance as a whole. The purpose of Table 2 and Table 3 is to present how the participants perceived their competence and importance of the learning objectives included in their training. We believe that plainly presenting the mean without interpreting it based on the range of responses would leave users of published studies hanging.

5. The value used in the overall competence and importance is the aggregated average of all ratings. This is the reason why the median has a decimal point in Tables 2 (pp.292-293) and Table 3 (pp 293-294).

6. We did not include training hospital and school graduated in the comparison tests due to the representativeness of participants per category of data. We could not merge hospitals into two cat-

Spearmen correlation and only for the confidence part of the SECS. I think the authors should clarify in their methods why they used nonparametric statistics especially when parametric tests are more robust in this case. Thank you very much for raising your observations.

We believe that the brief presentation at the methods section on page 289 of the manuscript can be interpreted such that the scales used in the study are an ordinal level of measurement. We also believe that the presentation of the statistical test used on page 290 is also brief and specific on what tests were used and variables included in the analysis. We are sorry for missing the rs-values in the abstract. This is attributed to the word count limitation. If the Editor-in-Chief so needed and allow us to have an addendum, we would gladly add the missing rs-values. We are sorry for missing the r-value

for the importance correlation in the abstract. If the Editor-in-Chief would allow as have an addendum, we would gladly add the missed result. Kitchen stated in her paper that the assumptions for nonparametric tests are weaker than those of parametric and researchers are better to only resort to the first if normality of data was not met or the underlying distributions are heavily tailed or extremely skewed and/or the sample size is small (n<25).[2] We acknowledge that parametric statistics are more robust than non-parametric tests if all assumptions for using parametric statistics are met. However, we still stand by our choice of using nonparametric test since our dependent variable data is at ordinal level of measurement. Nonparametric tests can still be considered robust. To establish additional robustness to our results, the effect sizes of variables with significant findings were computed and presented in the comparison tests. Effect sizes represent an objective and standardized measure of the magnitude of observed effect.[1] Furthermore, since rs-value is considered as the same as the effect size, we computed the coefficient of determination (rs2) to include the variance explained by the independent variable to the dependent variable.

In closing, we wholeheartedly respected and valued the scholarly views of our esteemed reader. In behalf of my co-authors, we are one in declaring that we observe all best practices and ethically approved methods that we know and within our means in conducting our study and in reporting the results of our study. We owe it to the research users all over the world to provide robust, credible, and sound study and manuscript based on ethically approved standards and best practices.

We thank you very much for providing us an opportunity to provide clarifications on the points raised in our study. This is evidence that our research users and readers are critically appraising not only our study, but all studies published in the esteemed *Annals of Saudi Medicine* (ASM). This exercise strengthens the integrity of articles published in ASM.

Ahmad E. Aboshaiqah, Joel G. Patalagsa

From the College of Nursing, King Saud University, Riyadh, Saudi Arabia

Correspondence: Ahmad E. Aboshaiqah RN, PhD · Associate Professor, College of Nursing, King Saud University, Riyadh 11421, Saudi Arabia · T: +966 50 315 4993 · aa.boshaiqahksu@gmail.com · ORCID: http://orcid.org/0000-0002-9763-2658

REFERENCES

1. Weinberg SL, Abramowitz SK. Data analysis for the behavioral sciences using SPSS. New York: Cambridge University Press. (Reprinted 2007). 2002:257.
2. Field A. Discovering statistics using IBM SPSS Statistics. (5th ed.). Thousand Oaks, California: SAGE Publications Inc. 2018.
3. Langdridge D, Hagger-Johnson G. Introduction to research methods and data analysis in psychology. (2nd ed) England: Pearson Education Limited. 2009:47.
4. Tabachnick BG, Fidell LS. Using multivariate statistics. (6th ed.). Upper Saddle River, New Jersey: Pearson Education Inc. 2013;7.