Dear editor

This letter is in response to the article by Williams et al titled “Empathy levels among first year Malaysian medical students: An observational study” published in Advances in Medical Education and Practice, May 2014.

This observational study addresses the issue of educating medical students about empathy as an attribute – a matter which is often not given enough importance – and including the topic in the academic curriculum, in view of the declining empathy levels observed in clinical practice. We would like to raise some concerns regarding the methods used for data analysis in the paper. The authors have analyzed the data using descriptive statistics such as means and standard deviation (SD) that are not appropriate parameters since the instrument used a seven-point Likert scoring scale. It would have been better if they had used median with minimum and maximum rather than the mean and SD.

Paired t-test was used to compare the scores before and after the intervention. As the data do not follow a normal distribution, it is better to use Wilcoxon signed rank test than the t-test. In the Results section, it is mentioned that mean and SD are 112.08±10.67 and 117.93±13.13 respectively before and after intervention and that the difference is statistically significant. The effect size calculated was 0.48, with no mention of how the value was derived. Additionally, as a Likert scale with 1 to 7 scale points has been used, the interpretation cannot be precise. The coefficient of variation within the two groups, according to our calculations, was 9.5% in the before-workshop group and 11.1% in the after-workshop group. The researchers have used the difference in mean scores whereas the coefficient of variation would have been more appropriate. Furthermore, the mean, standard deviation and coefficient of variation are not the recommended measures for ordinal scale values.

Our query is also applicable to the comparison of the empathy scores before and after among the males and the females.

The authors have used Cronbach alpha to measure the reliability of Jefferson Scale of Physician Empathy–Student Version (JSPE-S), as mentioned in the Results and Discussion sections (α=0.70, before workshop and α=0.83, after workshop). The Cronbach alpha is used to determine the reliability of the tool during its construction: Why the authors have calculated the same before and after the interaction is not clear.

In Table 2, the confidence interval (CI) and its P-value are given. Here, too, the normality assumption is used. Many of the P-values and CIs do not match. In some
places the CI includes the value zero and is indicated as being significant, whereas in some others the value zero is not included and is described as being not significant. Also, why the authors considered the negative value for mean difference is not clear.

Disclosure
The authors declare no conflicts of interest in this communication.

Reference
Authors’ reply

Brett Williams1
Sivalal Sadasivan2
Amudha Kadirvelu2
Alexander Olaussen1

1Department of Community Emergency Health and Paramedic Practice, Melbourne, Australia; 2Jeffrey Cheah School of Medicine and Health Sciences, Sunway Campus, Selangor, Malaysia

Correspondence: Brett Williams
Department of Community Emergency Health and Paramedic Practice, Faculty of Medicine, Nursing and Health Sciences, Monash University Peninsula Campus, PO Box 527, McMahon’s Road, Frankston, Melbourne, Victoria 3199, Australia
Tel +61 3 9904 4283
Fax +61 3 9904 4168
Email brett.williams@monash.edu

Dear editor

Firstly, we would like to thank Gopakumar et al for the statistical critique of our recent publication “Empathy levels among first year Malaysian medical students: An observational study” in Advances in Medical Education and Practice. It is good to have critical debate about published research and to have different perspectives reported by prospective readers. They have raised a number of important points that we would like to address.

While the issue of using parametric statistics with ordinal level data is something that stimulates ongoing debate and is passionately contested in the literature on an ongoing basis in research statistics circles, a number of scholars take a more practical approach1-5 to some of the statistical assumptions that have been raised by Gopakumar et al. These very points are something that many of us as researchers grapple with; do we simply not use parametric statistics when some (or all) of the assumptions are not met? Do we simply report all results (both mean, median, and interquartile ranges)? Do we perform ordinal regression? Do we strictly only use non-parametric analyses when Likert-type rating scales are used (putting aside tests for normality and sample sizes etc)?

A quick review of the latest issues of three of the top ranking medical and health education journals did not yield one paper using Likert-based rating scales that conformed to the strict (and correct) mathematical and statistical assumptions that Gopakumar et al identified. The broader argument is not whether studies such as ours use parametric or non-parametric statistics; the dispute and discourse, we believe, are the publishing and authorship guidelines that journals provide to authors. While clearly this is a very difficult challenge, but if journals and their peer-reviewers were to make a stance on the conventions of non-interval level data and statistical analyses then perhaps such discussions would not take place on this issue. In summary, these discussions are very important and again acknowledge the points raised by Gopakumar et al.

While Gopakumar et al have suggested that the Cronbach alpha is to be used to determine the reliability (or more correctly internal consistency) during an instrument’s construction, which of course is correct, we would suggest that this is not the only time it should be used.6 Our rationale for using Cronbach alpha at both pre-test and post-test times was an attempt at examining the internal consistency at both time periods, thus adding to the psychometric data about the tool. Regarding the confidence intervals, we would like to thank Gopakumar et al for identifying the error; we have re-run the statistics and will revise the table accordingly.

In closing, we thank Gopakumar et al for their interest and comments about our article.

Disclosure

The authors declare no conflicts of interest in this communication.

References