Reviewer's report

Title: Comparative cross-sectional study of empathy among first year and final year medical students in Jimma University, Ethiopia: Steady state of the heart and opening of the eyes

Version: 1 Date: 31 March 2012

Reviewer: Jay Orlander

Reviewer's report:

Understanding the relationship between clinical training and the impact on empathy for medical students and physicians remains an important area of study. In the proposed manuscript the authors attempt to compare scores on two different empathy assessments of First year, compared to final year medical students in Ethiopia and assess any influence that social factors, such as living situation, career goals, presence and position among siblings, gender, and use of social media has on the measurement. The sample size and recruitment methods are reasonable as were the analytical methods. The writing is clear throughout. The conclusions are limited and they do not over-interpret their data.

The literature review seems to cite many and recent related studies. The literature review and discussion may be enhanced by including specific additional details. The instruments used as outcome measures are cited each in a single reference, and apparently not previously used in medical settings nor in Ethiopia. The paper would be improved if the authors would discuss why these instruments were selected, particularly the BEES – a self-administered instrument, over instruments used in prior published studies of medical students, where comparative data would be available. It seems that in many of the US based studies and at least 2 of the international studies, used the Jefferson Scale of Physician Empathy - JSPE, (although I am not familiar with all of the cited references). In addition, a brief understanding of the development of each instrument for the reader is essential when no other studies using these in medical students are cited. For readers who will be unfamiliar with the BEES, providing sample questions or the full instrument in a table or appendix would be helpful.

The authors found overall scores to be low compared to published norms for the instruments, but should describe what population those norms were developed on. The RME-R for example seemed to be tested on “adults” and it would not be surprising to this reader if we learned that older individuals, particular adults who are parents, may have developed better cognitive skills than this likely, much younger group of medical students who are subjects in the current study. Understanding such issues is critical in the reader ability to interpret the results. They found no change on their measure of emotional empathy, where as the literature has generally suggested a decline when using self-administered instruments. In their discussion they site a few contradictory studies, but do not
mention how empathy was measured in these studies. Conversely the authors found improvement in their measure of cognitive empathy from beginning to last year students, which could reflect a positive effect of clinical training. They found several demographic and social factors did impact empathy such as female gender (consistent with the existing literature) and presence of siblings and use of social media – each of seems to be a new finding and raises interesting hypotheses.

I like the fact that they chose to measure two different aspects of empathy. In this regard, this is similar to reference 3, where a self-administered score is compared to another type of assessment. I believe this is the strength of their study. Some additional discussion of this point, placed in the context of the existing literature would enhance the paper.

Major essential revisions

Enhance the description of the instruments selected as outcome measures including some additional discussion of how they were developed.

Enhance the discussion section to include description of how the instruments used in the studies cited in the literature can impact their results as this will put your own results in a better context for the reader.

Minor essential revisions

Tables 1, 2, 4 “Totals” are a meaningless statistic in these tables. P values for test of comparison should be used here instead.

Tables 3 and 5; p – values and 95% confidence intervals reflect the same information, choose only one value to put in your table. P < values preferred.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.