Author’s response to reviews

Title: Perceived Medical School stress of undergraduate medical students predicts academic performance. An observational study

Authors:

Thomas Kötter (thomas.koetter@uni-luebeck.de)
Josefin Wagner (josefin.wagner@uni-luebeck.de)
Linda Brüheim (linda.brueheim@uni-luebeck.de)
Edgar Voltmer (edgar.voltmer@thh-friedensau.de)

Version: 1 Date: 19 Jun 2017

Author’s response to reviews:

We thank the editor and reviewers for their time and thoughtful criticism of our manuscript. We have reviewed each comment carefully and our responses are directly below the original reviewer’s comment.

Lisa Lampe (Reviewer 1):

1. My main query regards the interpretation of the statistical tests and therefore the conclusions that have been drawn. From Table 1 and the accompanying text it is evident that female students have a significantly higher mean M1 grade than males (I am assuming that means a better performance, but for readers not familiar with the M1 and its scoring, this could be clarified). The text reports that students who sat the M1 after more time in the course had a higher score – this would also make their mean age higher, presumably. Thus, at this point we are expecting to find that regressions will confirm that older, female students do better.

However, from the linear regression analysis the authors concluded that older and female students did more poorly, in contrast to what the t tests seemed to be suggesting. It seems that the Bs in the linear regression have been interpreted as Odds Ratios and I wonder whether this is correct?
My thinking was that the linear regressions were predicting score on the M1, not a categorical outcome. My interpretations for time1 were thus:

- For each year of increase in age the M1 grade increases by 0.05 units
- As gender increases by one unit (i.e. becomes female), the M1 grade increases by 0.33 units
- As PMSS increases by 1 unit, M1 grade increases by 0.03 units

This latter finding is quite interesting – it may suggest that although stress increases across time in the course, it isn’t necessarily detrimental to performance. I wonder whether an increase (or a decrease) of 0.03 units is meaningful anyway?

We thank the reviewer for these important remarks. However, there might have been misunderstandings leading to her raised concern. Firstly, the lower the M1 score, the better (we have clarified this in the revised version of the manuscript: see page 6 lines 1 et seq.). Secondly, we took the age from a fixed time point after matriculation, independently from when the actual exam took place. So, taking the exam after having spent a longer time on the course didn’t make the mean age used for the statistical analyses higher. We agree with the reviewer on how to interpret the results of our linear regression analysis. In the revised version of the manuscript, we have reworded our conclusions more cautiously (page 14, line 22) and state that especially those students with all three of the predictive characteristics might be considered as a potential risk group.

2. More information is needed about the M1 for readers outside of Germany. Is there discretion allowed in terms of when the examination is taken? This is implied (“The M1 exam can be taken after two years of Medical School”) but not clearly stated. What is the possible range of scores? On page three, L5: “… our aim was to examine … academic performance (first medical examination [M1] grade) for the first time” is unclear. Do the authors mean that this is the first
time this has been studied or do they mean that what is being studied is the first assessment in medical school?

We have added further information about the M1 to the Methods section of the revised version of the manuscript, with regard to prerequisites, passing rates, duration of pre-clinical studies (page 6, line 3), possible range of scores (page 6, lines 1 et seq.) and whether we studied this for the first time or studied the first assessment in medical school (page 5, line 4).

3. I would suggest that the titles for tables 2 and 3 should be “Linear regression analysis at t1/t2” or “Linear regression analysis M1 at t1/t2” without reference to PMSS since the latter is merely one of the predictors for M1. And perhaps capitalise the T so it is not confused with a t test.

Following the reviewer’s suggestion, we have reworded the titles for tables 2 and 3 (page 10, line 16 and page 10, line 18) and capitalized the T throughout the whole manuscript.

4. I agree with the authors’ suggestion that more consistent use of a structured instrument would more easily allow comparison across studies. However, the stresses being measured by this instrument (financial worries, workload and social isolation) are by no means specific to medical school, and therefore I question exclusion of consideration of the findings of studies that used other measures. In terms of wording, I would suggest that whilst the PMSS may be an instrument specific for use in medical school, it is not correct to say that what is being measured are “specific medical school stresses”.

The reviewer is right. We have reworded the sentence, stating that the PMSS is a “stress questionnaire designed for use in Medical School” (page 2, lines 6 et seq. and page 4, line 16).

5. Regarding participants (p 5). There could be important differences between those who passed the M1 and those who failed: one might reasonably wonder if those who failed had particularly high levels of stress. Therefore I think that more consideration of the excluded students is needed. Could the authors test for even basic differences (gender, age) between those who passed
and those who failed? And similarly, are there any defining characteristics of students who did not participate in the LUST study? There is also reference to “incomplete datasets” – was there an analysis of any trends in what data was missing from whom? It is possible that this final sample ended up rather highly selected, since to be included they had to pass the M1, participate in the LUST study and have complete data.

Since about 95% of the students pass the M1 exam, the number of those who failed is quite small. However, we have added to the revised version of the manuscript a descriptive analysis with respect to the gender, age and PMSS scores of those who failed (page 8, lines 16 et seq.). A comparison of the age and gender distribution of our sample with the whole classes is also provided to enable the reader to make a judgement about the representativeness of the results (page 8, lines 3 et seq.). Since less than 20% of all students from the included classes had to be excluded for not having participated in the LUST study and not having passed the M1, and the sociodemographic characteristics of our sample match those of the whole classes quite well, the risk for selection bias should be rather small. However, we have added information on students with missing / incomplete data to the Results section of the revised manuscript (page 10, lines 5 et seq.).

6. Results p7, L22: suggest rewording as: “After the exclusion of incomplete data-sets, 386 PMSS T1 scores and 352 PMSS T2 scores could be matched to 456 M1 grades (84.7 and 77.2% respectively)”

Following the reviewer’s suggestion, we have reworded the sentence (page 8, line 2).

7. Results p8, PMSS: although there was a statistically significant increase, do the authors think it was of a clinically meaningful magnitude? This could be considered in the discussion.

We have added a paragraph discussing the increase in the mean PMSS score between T1 and T2 in the context of changes achieved in interventional studies aiming to reduce Medical School related stress (page 12, lines 20 et seq.).
8. The authors suggest that older students may have more stress. Two points need to be considered here:

1. Were “older” students meaningfully older?

2. Did the authors examine for a relationship between age and stress? The data would have allowed this and it might be quite instructive rather than just speculating!

In the revised version of the manuscript, we have added information about the age distribution of the sample (page 8, lines 7 et seq.). We further examine the relationship between age, stress and the M1 grade by calculating bivariate correlations and comparing older students with younger students using t-tests (page 8, lines 12 et seq., page 9, lines 8 et seq. and lines 23 et seq.).

9. There are a few typos here and there.

We have checked the manuscript for typos and corrected them.

10. Regarding "necessary controls" I am aware you could not include a control group, so my recommendation is based on the fact that I think more consideration of excluded students and missing data is required.

We understand the reviewer’s concerns. As mentioned above, we have added information on excluded students and those with missing / incomplete data to the revised version of the manuscript (page 10, lines 5 et seq.).

Harlina Halizah Siraj (Reviewer 2):
1. This study attempted to explore the relationship between perceived medical student's stress scores and academic performances using a purely quantitative approach. With a large sample size of 456 students, the findings might be credible enough. However, it is more appropriate to indicate that an additional qualitative approach should be applied to explore the actual links and relationship between the stress score and academic performance of medical students. In-depth interviews of the two extreme groups: high achievers and poor performers could provide some insights into the subject matter. It is quite improper to come to such a strong conclusion and generalization of the relationship between age and gender with academic performance simply based on a quantitative study.

We thank the reviewer for this important suggestion. We have now addressed the need for qualitative studies to explore the links between stress scores and academic performance of medical students in the Discussion and reworded the Conclusion in the revised version of the manuscript (page 14, lines 16 et seq. and line 24).

2. Can the PMSS scores of the medical students be further analysed according to the possible stressors of workload, competition, social isolation and financial worries? This can provide some insights as to which stressor is more dominant during this stage of medical education.

Following the reviewer’s advice, we further analysed the PMSS results and added the findings to the Results and Discussion sections (page 9, lines 16 et seq. and page 12, lines 23 et seq.).

3. Current trend in studying the impact of stress to student's achievement is to measure their resilience or ability to bounce back. It would be good to include some comments on the resilient factor as one of the important confounders.

We have added a sentence on coping abilities / the ability to bounce back as potential confounders (page 12, lines 17 et seq.).
4. About 14 references cited are from more than five years ago. Appreciate if newer references are included.

Following the reviewer’s suggestion, we have added some newer references. However, there are some older “landmark” studies that we must cite to provide a framework for our study and to support our claims. The median year of publication of our references is now 2012.