Reviewer's report

Title: Nursing students' perception of clinical learning environment and supervision in relation to two different supervision model - a comparative cross-sectional study

Version: 1 Date: 30 Apr 2019

Reviewer: Research Square

Reviewer's report:

"STATISTICAL REVIEWER ASSESSMENT:

Is the study design appropriate for the research question (considering whether the analyzed population accurately reflects the design and whether you see any problems with control/comparison groups, e.g., likely confounders)?

No - there are major issues

Are methodologies adequate and well implemented (considering whether assumptions are addressed and whether analyses are robust)?

No - there are major issues

Are the analyses adequately communicated (considering whether reporting details are adequate and whether figures and tables are well labeled and described)?

No - there are minor issues

Does the interpretation accurately reflect the analyses without overstatement (considering whether limitations/bias are acknowledged and whether accurate descriptors, e.g., 'significant', are used)?

No - there are major issues

Could an appropriately REVISED version of this work represent a statistically sound contribution?
STATISTICAL REVIEWER COMMENTS:

- This study compares one model to another but the authors don't articulate why, from a theoretical perspective, they expect to find differences in the measures they used to assess the outcomes.

Also, if learning is the outcome of interest, why were there no more objective measures of outcomes used?

- The authors describe the study as a "cross-sectional study." In some ways this is a correct description, but it's a bit general. From my perspective, this is more than just cross-sectional; it's a quasi-experimental design. The authors also refer to the survey data collected as "qualitative." But, by my reading, they have actually collected quantitative scale data (i.e., scores on several factors). In both cases, the authors might consider being more explicit in this regard.

- What's not clear in the study design is who was exposed to supervision models A and B? In other words, was this a within or between subjects design? Furthermore, were model A and B only used at certain hospitals, or in a given hospital, were both methods used? Based on the description under "procedures and sampling characteristics," it seems as though it's a between-subjects design, but this could be clarified.

- In the Method, when the authors describe the survey scales employed, they note: "Reliability measured with 16 scales' Cronbach's alpha ranging between 0.72-0.96." It's not clear, however, whether or not these reliability estimates were what they authors observed for the present study or for the scales when previously validated in previous studies? This needs clarification, as readers care most about the reliability scores for students in the current study.

- For the content validity evidence of the author-created survey, how many colleagues were used? As written, the authors only note that "The content validity is based on 21 interviews with a group of colleagues." This is a bit too vague. In addition, I have the same question as above about the Cronbach's alpha reported for these new subscales. In fact, more needs to be said about all the surveys used, even the "previously validated surveys," as we collect validity evidence for survey in certain contexts. Therefore, just because a survey 'worked' previously in a different context doesn't mean it will work the same in the authors' context.

- The data presented on pp. 8-9 and in Table 1 seem more like results than information to include in the Method: e.g., "Out of 381 eligible students, 244 filled out 22 questionnaires, giving a response rate of 64%. Of these, 170 students (90% women) got their 23 supervision in model A
and 74 students (88% women) in model B. The students' mean age 24 was 28 years and the range was between 21 and 50/51 years in both groups." These results belong in the Results section. In addition, why were the numbers of students so much higher in model A (170) versus model B (only 74)? Such an unequal cell size can have detrimental effects on the results of the statistical tests.

- The authors need a better rationale for the following decision: "In the analysis of questionnaire 2 (table 3) the negative response alternatives 'not at all' and 'rather small degree' were merged, as was the positive response alternatives 'fairly high' and 'very high degree', as the scorings on the negative scale was low." What is meant by "the scorings on the negative scale was low"? Generally speaking, it's bad statistical practice to consolidate response options as it results in a lose of data.

- The authors note in the Results "there were significant differences in their ratings on single items within the sub-dimension 'pedagogical atmosphere' in favor of model A. Why are the authors analyzing single items within a subscale that's designed to function together? This makes little sense and smells of p-hacking (i.e., running a bunch of statistical tests until "'statistically significant differences'" are found). If those items are part of a subscale, they shouldn't be analyzed independently. The same can be said for ANY of the individual items analyzed. The authors should instead calculate mean scores for the subscales and look for differences between those subscale scores (not differences between individual survey item scores).

- For all statistical analysis, it's important to not only report the p-values but also the means and SDs, as well as the effect sizes, as it's the size of the effects for any differences found that matter more than the specific p-values. In particular, the effect sizes should be included in the tables (but only for the subscale scores, not for individual survey items).

REQUESTED REVISIONS:

Please see my detailed comments above. All of these issues would need to be addressed to improve the quality and soundness of the work.

ADDITIONAL REQUESTS/SUGGESTIONS:

No. Please see above. "
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

This reviewer has been recruited by a partner organization, Research Square. Reviewers with declared or apparent competing interests are not utilized for these reviews. This reviewer has agreed to publication of their comments online under a Creative Commons Attribution License attributed to Research Square and was paid a small honorarium for completing the review within a specified timeframe. Honoraria for reviews such as this are paid regardless of the reviewer recommendation.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal