Reviewer’s report

**Title:** Knowledge of Pelvic Floor Disorders in Women Seeking Primary Care: A Cross-sectional Study

**Version:** 0  **Date:** 29 Jun 2018

**Reviewer:** Terry C.K. Lee

**Reviewer's report:**

The paper looked at the PFD knowledge in women presenting to three primary care clinics and also tried to identify factors that were associated with poor PFD knowledge. There are some major statistical issues which the author will need to address. Specific suggestions and questions are listed below.

1. There were two modes for enrolling participants. How many women were enrolled through each of the two modes? From a study design perspective, were there any systematic differences between these two modes that could cause a bias/difference in PFD knowledge? For example, did those pick up the questionnaire at advertised locations have more time to complete the questionnaires than those approached by a research team member (e.g., having more time might allow the participant to search answers on the internet).

2. The author defined lack of knowledge using the cutoffs of 50% and 80%. It would be useful to include more descriptive summaries using knowledge score as continuous variable. For example, what was the median score by education level?

3. Stepwise regression was used to derive the final logistic regression model. Please present the odds ratio from the full model together with those from the final model.

4. Income was said to be forced into the model, but the OR for income was not presented in Table 3.

5. Revisions are needed to the description of statistical methods in the method section. The sentence "Variance inflation factor (VIF) was performed to address the potential of collinearity inflating the standard error when performing multiple linear regression rather than logistic regression." needs to be revised as it does not accurately describe what VIF is. The sentence "Logistic regression diagnostics run on the UI and POP multivariate regression models led to the exclusion of certain variables." is unclear. Is it referring to large VIF?

6. VIF was used to assess collinearity between covariates. However, some variables are by definition or known to be highly correlated and so one should carefully select the variables to include in the regression analysis rather than purely relying on the VIF. E.g., those with a diagnosis of PFD would imply they were aware of PFD as medical conditions and so re-categorization (or combining) of these variables would be required to properly incorporate them into the multivariate model.
7. What variables were included in the starting model of stepwise regression? Putting covariates that were highly correlated into the starting model could produce incorrect results.

8. In the method section, the stated purpose of the link test seemed redundant as the model was derived from stepwise regression that already assessed which variables should be included.

9. It was surprising that age was statistically insignificant in the univariate analysis for PIKQ-POP, but became significant in the multivariate analysis. In addition, the ORs for age were very different between these two analyses. Further discussions/investigations on this are needed. Was this due to confounding? Or was it because variables highly correlated with age were included in the model and distorted the results.

10. The authors stated that "the mean VIF for both models were close to 1". Did any covariate had VIF that was substantially larger than 1? This would be more informative to report than the mean VIF.

11. For the last 10 lines of the results section which talked about regression diagnostics, I would suggest shorten the methodological descriptions as these were already presented in the method section. A simple sentence saying that regression diagnostics has been performed and no problems were identified would generally be sufficient.

12. In the discussion, the authors compared the % of respondents that were non-proficient to estimates from other studies. It would be useful to descriptively compare the mean/median score across studies as well. For example, in reference #19, the mean for POP scale was only 30.6 compared to 42 in the current study.

13. For the tables, please report the actual p-value instead of "NS".

14. Table 1
   a. For continuous variables that are presented as median, please provide the IQR so that readers can have a better understanding on the distribution of the data.
   b. Footnote (and in the methods section of the text) stated that "Continuous data was compared between proficient and non-proficient groups", was the comparison between clinic types?
   c. Please consider adding the summary for age using the same categories as in the logistic regression analysis. Without it, one cannot judge whether the insignificant result in Table 2 for age was due to insufficient sample size in some of the levels or was it a lack of signal.

15. Table 2
   a. In the text, the range of age was stated as 21-97. Please change the first category of age from "18-29" to "21-29".
   b. For weight, is the OR per 1 lb increase? Please consider changing the reporting unit to 10 lbs so that the OR is not so close to 1 after rounding to two decimal places.
   c. For "Information", the 95% CI excluded 1 and thus statistically significant (not "NS").

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.

Yes

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

Yes

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.
I declare that I have no competing interests

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