

Review of: "There is high prevalence of overnutrition among married and cohabiting women in Nigeria: Findings from the 2018 Nigeria Demographic and Health Survey"

Sumonkanti Das¹

¹ Australian National University

Potential competing interests: No potential competing interests to declare.

The authors conducted an interesting study on risk factors of overweight for married women in Nigeria using 2018 Nigeria DHS data. The authors follow the Socioecological model (SEM) as a framework for developing their models to investigate intra- and inter-personal along with community variables' impacts on the married women overnutrition status. In terms of SEM structure, the authors developed three logistic regression models using the covariates at three different levels - individually first and then with the presence of others. The developed models are statistically fine but not in terms of the data collected in the DHS, where the data are clustered in nature as per the survey sampling design. So, I wish to recommend the authors to update their models using clusters as the second level of the models. The models can be called two-level logistic regression models. Also, a logistic model with a single predictor variable is mentioned as bivariate and more than one predictor is mentioned as multivariate. Instead of these, the authors can denote them as unadjusted and adjusted logistic models respectively. When they use single and multiple covariates in the model, they can refer as simple and multiple logistic models respectively. When the authors updated these models by incorporating cluster as the level-two of the multilevel model, the models can be denoted as unadjusted and adjusted mixed effects logistic models. There are other options available, I suggest to use with which the authors feel comfortable.

The authors tried to show whether the inclusion of inter-personal and community variables increase the prediction power of the models. The paper of Nakagawa, and Schielzeth (2013) will help the authors to do this and also they can show how much variability are due to the cluster-level variability and how much due to the predictors.

To support the hypothesis of "the Southern region will have higher burden of overnutrition compared to the Northern region", the authors should not produce the Figure one, since the sample size is an issue for the prevalence estimation in this case. The authors can also provide the standard error of these estimates and I believe that the estimates with higher prevalence may have higher standard error not only for proportion but also for the sample size. A small area estimation method (Rao and Molina, 2015) will be required for getting better prevalence estimates at this detailed level. I would suggest they can make the partition of the country as the Region used for the models and then examine the hypothesis.

In terms of study design, the paper is nicely structured. My suggestion is to update the models and interpret according. Also, there were some discrepancies in the interpretation in the text relative to the Tables.

The term predictive power should be used with care, since the R-squared values particularly for logistic models cannot be

interpreted as the R-squared value as the linear regression models.

References:

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133-142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>

Rao, J. N., & Molina, I. (2015). *Small area estimation*. John Wiley & Sons.