

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

| | |
|----------------------------|--|
| TITLE (PROVISIONAL) | Determinants of COVID-19 related knowledge and preventive behaviors among students in re-opened secondary schools: cross-sectional study |
| AUTHORS | Handebo, Simegnaw; Adugnaw, Asmamaw; Kassie, Ayenew; Shitu, Kegnie |

VERSION 1 – REVIEW

| | |
|------------------------|--|
| REVIEWER | R. F. Alves University of Minho Institute of Education, CIEC – Research Centre Child Studies |
| REVIEW RETURNED | 23-Feb-2021 |

| | |
|-------------------------|---|
| GENERAL COMMENTS | <p>This is an interesting cross-sectional study exploring knowledge and preventive health behaviors toward COVID-19 among secondary school students.</p> <p>I just have a few recommendations to consider before publication:</p> <ul style="list-style-type: none">- Objective: Doing the results presented in manuscript, the objective need be reformulated. Adding, for example, factors associated to knowledge and preventive behaviours.- Explain in the section “Study Variables”, all the five sections including the in questionnaire: missing: socio demographic, social support, constructs of Health Belief Model.- The variable “Source of information about COVID-19” is not in the study variables.- Explain in the introduction and in the discussion the importance of Health Belief Model.- Why did you not correlate the main variables? Better knowledge contribute for better preventive behaviors is an issue. Why the variable knowledge is not in the regression analysis? The entire discussion refers to the influence of knowledge on preventive behaviors. However, given the analysis carried out, it was not (yet) possible to perceive this influence.-Why does table 2 show all items on the knowledge scale and does it not show a table with all items on the scale of preventive behaviours?- Why not show in Table 1, a test of differences between groups, for example, Chi-Square?- The title of manuscript is a question. Were you able to answer this question? <p>Minor:</p> <ul style="list-style-type: none">- p. 6 line 7 [16-18] should be [16, 18]?- p. 8 line 21-22. Should be in the next section: “Data processing and analysis”?- Abstract: line 30: add the maximum of preventive behaviours.- In the table 3, explain what is “cut 1” and “cut 2”.- Cronbach’s alpha usually doesn’t appear in percentages. |
|-------------------------|---|

| | |
|--|---|
| | <ul style="list-style-type: none"> - Missing reference for "Blooms cut-off point" - Doesn't table 2 lack the answer category "Don't know"? Or no answer was given in this category? - Are figures 1 and 2 important for the manuscript? - Reference [22] it's a study with University Students. |
|--|---|

| | |
|------------------------|---|
| REVIEWER | Rubén López-Bueno University of Zaragoza |
| REVIEW RETURNED | 07-Mar-2021 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | <p>This is an interesting study dealing with a hot topic. To improve the quality of the manuscript some points should be addressed. My main concern is that authors are using the regression prediction function instead of the regression estimation effect function, which hampers the study design and makes statements regarding findings difficult to follow.</p> <p>Abstract</p> <ul style="list-style-type: none"> -The main results of the regression analyses should be shown in the results. <p>Introduction</p> <ul style="list-style-type: none"> -It is generally well-conducted. This can be extended commenting some findings from prior research observing hand-washing practises in low countries as initial benchmark. Also, the closure of schools can have had detrimental effects on health-related behaviours in both children and adolescents. -Providing a study hypothesis would be fine since your background invites to think in potential percentages. <p>Methods</p> <ul style="list-style-type: none"> -Is it possible to show values for internal and external validation of your main assessment tool? -How were the questionnaires distributed? Were they written questionnaires? -How did you deal with the missing values? <p>Results</p> <ul style="list-style-type: none"> -My main concern is that you are mixing concepts since it seems that you are speaking about predictor variables (also in the discussion), as if you had built a model to predict your outcome, but this has not been stated in the objectives of your study. <p>Discussion</p> <ul style="list-style-type: none"> -This is well conducted. However, could be a bit more expanded trying to explain your main findings. For example, although those watching more Tv seemed to be more aware of preventive measures to avoid COVID-19 spread, there is still the possibility that misinformation on the internet can have produced the opposite. -One important limitation that should be included is that referred to a recall bias, since your data were self-reported by participants. |
|-------------------------|--|

| | |
|------------------------|------------------------------------|
| REVIEWER | Asif Imtiaz University of Dhaka |
| REVIEW RETURNED | 12-Mar-2021 |

| | |
|-------------------------|---|
| GENERAL COMMENTS | <p>Dear Authors,</p> <p>1. In the background section, you have mentioned the children can act as agents of change to adopt healthy behaviours to prevent COVID-19 and you are measuring the level and predictors of knowledge about COVID-19 and preventive behaviours. The missing link is why you need to measure them? Please explain why you want to know the level and predictors of the outcome variables to make</p> |
|-------------------------|---|

| | |
|--|---|
| | <p>the children agents of change on the verge of the schools' reopening.</p> <p>2. Using a 5% anticipated non-response rate, the final sample size to be estimated was 403 participants. But you have analysed a sample of 370 participants with 8.2% non-response rate. How do you justify analysing a lower sample size than the requirement to perform any inferential technique? Can we perform regressions in any sample that is lower than its required size? Please justify whether the sample has become a non-probabilistic one or not due to this lower than required sample size.</p> <p>3. It is not clear whether you have used ordinal logistic regression as you have mentioned parallel line tests, proportional odds model but wrote ordinary logistic regression. Please clarify.</p> <p>4. It is unclear that why students aged 20-30 are still in secondary level. Please provide background information of this attribute of the educational system of the country.</p> <p>5. Please take help from any native English speaker or language expert to improve the language quality.</p> |
|--|---|

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1. Objective: Doing the results presented in manuscript, the objective need be reformulated. Adding, for example, factors associated to knowledge and preventive behaviors.

Response: This section was revised accordingly (page 2 line 7 and page 5 line 29)

2. Explain in the section “Study Variables”, all the five sections including the in questionnaire: missing: socio demographic, social support, constructs of Health Belief Model. The variable “Source of information about COVID-19” is not in the study variables.

Response: This was due to typo error. We have corrected accordingly (page 7 line 14)

3. Explain in the introduction and in the discussion the importance of Health Belief Model.

Response: we have not included health belief model constructs in this study. Due multiple research work we had we mistakenly include it in the list of variables. So, we have corrected the introduced typo error in the revised manuscript. As a result, we haven't explained about health belief model in this manuscript.

4. Why did you not correlate the main variables? Better knowledge contribute for better preventive behaviors is an issue. Why the variable knowledge is not in the regression analysis? The entire discussion refers to the influence of knowledge on preventive behaviors. However, given the analysis carried out, it was not (yet) possible to perceive this influence.

Response: We have included knowledge as an independent variable in simple linear regression analysis. But knowledge is not significant in the analysis at p-value less than 0.25. As a result, it was not included in multiple linear regression. We have discussed this limitation in page 12 line 20-22.

5. Why does table 2 show all items on the knowledge scale and does it not show a table with all items on the scale of preventive behaviors?

Response: We have separately added table 3 with preventive behavior items (page 22)

6. Why not show in Table 1, a test of differences between groups, for example, Chi-Square?

Response: we have included Chi-square p-value in the last column of table 1 in the revised manuscript (page 18 and 19)

7. The title of manuscript is a question. Were you able to answer this question?

Response: we have changed the title of the manuscript by including study design and objective. Previously we stated in a question format.

Minor comments

1. p. 6 line 7 [16-18] should be [16, 18]?

Response: This is too mean reference from 16 to 18.

2. p. 8 line 21-22. Should be in the next section: "Data processing and analysis"?

Response: We have corrected accordingly

3. In the table 3, explain what is "cut 1" and "cut 2".

Response: The cut-points shown at the bottom of the output indicate where the latent variable is cut to make the three groups that we observe in our data. This means the estimated cut points on the latent variable used to differentiate low knowledge from middle and high knowledge when values of the predictor variables are evaluated at zero.

4. Cronbach's alpha usually doesn't appear in percentages.

Response: Sometimes it was stated as percentage. But we have revised accordingly in the revised manuscript.

5. Missing reference for "Blooms cut-off point"

Response: reference number 29 was cited in the methods section. Page 6 line 28

6. Doesn't table 2 lack the answer category "Don't know"? Or no answer was given in this category?

Response: This was operationalized in the methods section. Page 6 line 23-25

7. Are figures 1 and 2 important for the manuscript?

Response: We hope they are more informative.

8. Reference [22] it's a study with University Students.

Response: Yes, it was done among higher education students in Portugal. We have considered this as a related literature.

Reviewer: 2

Abstract

1. The main results of the regression analyses should be shown in the results.

Response: We didn't include the regression coefficients in the result part of the analysis due to the journal limited the number of word abstract section (less than 300 words).

2. Providing a study hypothesis would be fine since your background invites to think in potential percentages.

Response: we have included the study hypothesis in the revised manuscript page 5 line 25-29 and page 6 line 1-2

3. Is it possible to show values for internal and external validation of your main assessment tool?

Response: Content validity values are included in the revised manuscript. Page 7 line 16-19

4. How were the questionnaires distributed? Were they written questionnaires?

Response: we have used written questionnaires which were distributed for the selected participant. COVID-19 safety measures were followed.

5. How did you deal with the missing values?

Response: Since we checked the collected data at the field, we didn't have missing data.

6. My main concern is that you are mixing concepts since it seems that you are speaking about predictor variables (also in the discussion), as if you had built a model to predict your outcome, but this has not been stated in the objectives of your study.

Response: we have incorporated this in the revised manuscript. Page 5 line 27-29 and page 10 line 19.

Discussion

7. This is well conducted. However, could be a bit more expanded trying to explain your main findings. For example, although those watching more Tv seemed to be more aware of preventive measures to avoid COVID-19 spread, there is still the possibility that misinformation on the internet can have produced the opposite.

Response: we have incorporated this in the revised manuscript. Page 12 line 9-11

8. One important limitation that should be included is that referred to a recall bias, since your data were self-reported by participants.

Response: we have incorporated this in the revised manuscript. Page 12 line 26

Reviewer: 3

1. In the background section, you have mentioned the children can act as agents of change to adopt healthy behaviors to prevent COVID-19 and you are measuring the level and predictors of knowledge about COVID-19 and preventive behaviors. The missing link is why you need to measure them? Please explain why you want to know the level and predictors of the outcome variables to make the children agents of change on the verge of the schools' reopening.

Response: The health message from student was accepted by their family members and neighbors. Educated students were well accepted by most community members. In addition, the student's engagement in such practice will facilitate the behaviors through observational learning. We have incorporated this in page 5 line 25-27.

2. Using a 5% anticipated non-response rate, the final sample size to be estimated was 403 participants. But you have analyzed a sample of 370 participants with 8.2% non-response rate. How do you justify analyzing a lower sample size than the requirement to perform any inferential technique? Can we perform regressions in any sample that is lower than its required size? Please justify whether the sample has become a non-probabilistic one or not due to this lower than required sample size.

Response: Post-hoc power analysis was conducted to ascertain statistical power for sample size adequacy and it indicates that our sample was adequate to run regression analysis to identify factors associated with COVID-19 related knowledge and preventive behaviors among the students. Page 6 line 14 -16

3. It is not clear whether you have used ordinal logistic regression as you have mentioned parallel line tests, proportional odds model but wrote ordinary logistic regression. Please clarify.

Response: This was due to typo error we have corrected in the revised manuscript.

4. It is unclear that why students aged 20-30 are still in secondary level. Please provide background information of this attribute of the educational system of the country.

Response: In rural areas where facilities are often thinly spread, significant number of children are out of school at their official school ages. Formal education in Ethiopia begins at age seven. However, late enrolment is pervasive. Only 57 percent of these were enrolling 'on time' in terms of formal age for grade. That is why these students are enrolled in secondary school.

VERSION 2 – REVIEW

| | |
|------------------------|--|
| REVIEWER | R. F. Alves University of Minho Institute of Education, CIEC – Research Centre Child Studies |
| REVIEW RETURNED | 09-Apr-2021 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | Congratulations! The authors had responded to all questions and comments. In my opinion, the manuscript is ready to publish. |
|-------------------------|--|

| | |
|------------------------|---|
| REVIEWER | Rubén López-Bueno University of Zaragoza |
| REVIEW RETURNED | 26-Mar-2021 |

| | |
|-------------------------|------------------------------|
| GENERAL COMMENTS | The manuscript has improved. |
|-------------------------|------------------------------|

| | |
|------------------------|------------------------------------|
| REVIEWER | Asif Imtiaz University of Dhaka |
| REVIEW RETURNED | 30-Mar-2021 |

| | |
|-------------------------|---|
| GENERAL COMMENTS | <ol style="list-style-type: none"> 1. Page 6 of 64, line 26: You wrote "thru". It should be "through" 2. Page 9 of 64, line 01: Again you wrote "ordinal" as "ordinary" 3. Please provide the calculation of determining the adequacy of the sample size using post-hoc power analysis using appropriate reference. 4. Please carefully check for any other typo. |
|-------------------------|---|

VERSION 2 – AUTHOR RESPONSE

Reviewer #3

1. Page 6 of 64, line 26: You wrote "thru". It should be "through"

Response: Corrected accordingly

2. Page 9 of 64, line 01: Again, you wrote "ordinal" as "ordinary"

Response: Corrected accordingly

3. Please provide the calculation of determining the adequacy of the sample size using post-hoc power analysis using appropriate reference.

Response: We conducted a post hoc power analysis to verify that the sample was adequate to detect the required estimates precisely based on the following assumption; the amount of explained variance in the endogenous variable (0.43) number of predictor (exogenous) variables (5), significance level (0.05), and sample size (370). Based on that, the observed power was calculated to be 1.0, which is acceptable (> 0.8) (1,2). In addition to this, according to previous literature, a response rate of 60% and above will not be problematic to infer results to the target population (3,4).