PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Predictors of Older Adults' Health Behaviors to Prevent COVID-19 Transmission: A Multilevel Analysis
AUTHORS	Lee, Jiwon; Seok, Jo Woon

VERSION 1 – REVIEW

REVIEWER	Shelley, Mack Iowa State University, Department of Political Science, and Department of Statistics
REVIEW RETURNED	05-Feb-2024

GENERAL COMMENTS	This is a very strong study with generally excellent content. One major need is to provide more thorough discussion of data-driven recommendations for public policy and for practice by public health entities. Recommedations for practice by patients and for the general population are well articulated, but more needs to be said about appropriate initiatives for the public health sector policymakers and for practice by professional health practitioners. A comparatively minor but non-trivial point is to be careful about reporting statistical significance. A p-value of 0.000 is possible only if the test statistic achieves its maximum (often infinity) and reporting "p<0.000)" should be corrected to "p<0.001" as the incorrect statement literally says that the probability is negative--which is totally impossible.

REVIEWER	タカダ, ミドリ Osaka Center for Cancer and Cardiovascular Disease Prevention
REVIEW RETURNED	08-Apr-2024

GENERAL COMMENTS	The authors have examined factors that may influence the presence or absence of knowledge and health behaviors to cope with COVID 19. It seems interesting that many factors are being considered at the same time. However, the descriptions of Methods, Results were vague in many parts, and the appropriateness of the statistical methods used and the validity of their interpretation did not seem plausible. The manuscript seems to have much room for improvement. Abstract Primary and Secondary Outcomes ---Although the authors regarded "COVID-19 coping-related knowledge and poor preventive behaviors" as primary and secondary outcomes, the primary outcome of this study is

"individual and community level variables influencing deficient knowledge regarding coping COVID-19 and not-practicing COVID-19 preventive behaviors", is it not?
In addition, the authors should refer to "regional difference in individual and community level variables influencing deficient knowledge regarding coping COVID-19 and not-practicing COVID-19 preventive behaviors" as secondary outcome in the Abstract.

Results
---The statement of " Older rural resident adults were more likely to not practice preventive behaviors compared to urban resident adults." seems to be different from the results of Table 2.

Methods
Data and subjects:
Figure 1:
---The authors should also explain the exclusion criteria in the box of "Data excluded   (n=156, 457)".
---The proportion of missing in Income is larger than the proportion of missing in other variables (around 15%). Unless the mechanism of missing all variables is missing completely at random (MCAR), the complete case analysis currently employed by the authors may lead to biased results.

Variables 1) Independent variables
Individual-level variables
---The authors should explain what "Yes/No" of economic activity means.

2) Dependent variables
--- The authors confuse terms such as minor/mild and aggravated/worsening, which seem to represent the same thing. It would be better to unify these terms into one of them
--- How was "Mild or aggravated symptoms" defined? It is also helpful to describe how respondents to the questionnaire recognized and distinguished between them.
--- The authors describe in the main text that both "knowledge about COVID-19-related coping" and "preventive behaviour practice" were categorised as "practice-health behaviour group (0)" or "not practice-health behaviour group (1)". However, the presentation in Table 2 categorises "knowledge about COVID-19-related coping" into "Deficient knowledge regarding coping to symptoms" and others, and "preventive behaviour practice" is classified as "Not-practicing preventive behaviours on COVID-19 infection" and others. The terminology in the text and in the Table should be aligned.

Statistical analysis
--- What is analysed with "the null model, meaning that it had no variables"? And what is the meaning for this analysis?
--- The authors describe null model, second model and final model in the main text, but the models in the Table are described as model 0, model 1 and model 2. It would be better to either align the notation of the terms in the text and the Table, or to clearly explain which model in the main text corresponds to which model in the Table.
--- It is unclear what is meant by the sentence "Intra-class correlation (ICC) was used to measure the effect of variables at individual and community levels." Although I am not an expert in

| | statistics, Isn't the intraclass correlation coefficient (ICC) a reliability measure used in intra-rater and inter-rater reliability analysis? Is the use of ICC suitable in this situation?
--- It would be better to state in the Statistical analysis how p-values were calculated.

Results
Demographic and general characteristics
---With regard to the sentence "those who were economically active had household incomes ranging from 0 to 2500,000 won ($2011.05).", I got the impression that this text does not explain the results presented by the authors in Table 1.
---I would like you to provide a citation for the sentence "Suggesting that the collected data reflected the initial response pattern of older adults to COVID-19."
---Please also explain the characteristics of the Enabling factors. As the authors focus on the results on the presence or absence of a living family member in the Abstract, it is necessary to explain the characteristics of the Enabling factors in Results in the main texts.
---The authors state that "Model 2 was also the most explanatory (ICC:0.064)", but for an ICC that ranges from 0 to 1, 0 indicates unreliability and 1 indicates complete reliability. Therefore, it may be that "Model 2 was the least explanatory".
---There is no description at all on individual and community factors influencing deficient knowledge regarding coping to symptoms.
---Why have the authors not analysed "the differences in influencing factors between living in urban and rural areas regarding deficient knowledge regarding coping to symptoms?"

Throughout, there is a mixture of notations to two decimal places and three decimal places, or there is more than one term for one meaning (e.g. oldest-old and older-old). Unifying the notation would make it easier for readers to read the manuscript. |

VERSION 1 – AUTHOR RESPONSE

Reviewer1
1. provide more thorough discussion of data-driven recommendations for public policy and for practice by public health entities. (appropriate initiatives for the public health sector policymakers and for practice by professional health practitioners)

: We added initiatives for policymakers and health practitioners in the discussion.

2. modification of A p-value from 0.000 to p<0.000

: Thanks for your valuable comment. We modified the format of p-value as you suggested.

Reviewer2

1. Abstract

1.1. Although the authors regarded "COVID-19 coping-related knowledge and poor preventive behaviors" as primary and secondary outcomes, the primary outcome of this study is "individual and

community level variables influencing deficient knowledge regarding coping COVID-19 and not-practicing COVID-19 preventive behaviors"

: The sentence has been modified in the abstract to make it easier to understand.

1.2. the authors should refer to "regional difference in individual and community level variables influencing deficient knowledge regarding coping COVID-19 and not-practicing COVID-19 preventive behaviors" as secondary outcome in the Abstract.

: We added the following sentence to result part of the abstract: "Specifically, older adults living in rural areas were less likely to engage in COVID-19 preventive behaviors compared to those in urban areas, and the effectiveness of mass media as a motivating factor for adopting preventive measures was only significant in rural populations."

2. Methods

2.1. The authors should also explain the exclusion criteria in the box of "Data excluded (n=156, 457)
: To explain definition of the study population and reasons for exclusion, we inserted Figure1. Study population flow chart.

2.2. The proportion of missing in Income is larger than the proportion of missing in other variables (around 15%). Unless the mechanism of missing all variables is missing completely at random (MCAR), the complete case analysis currently employed by the authors may lead to biased results.

: Thank you for your observation about the missing income data. We have acknowledged this issue in the limitations section of our manuscript, noting that the non-random missing values could bias the results and necessitating cautious interpretation.

1) Independent variables Individual-level variables

2.3. The authors should explain what "Yes/No" of economic activity means.

: We added the following explanations in the Materials and Methods section: "Yes (engaged in paid employment)" or "No (not employed)."

2) Dependent variables

2.4. The authors confuse terms such as minor/mild and aggravated/worsening, which seem to represent the same thing. It would be better to unify these terms into one of them

Thank you for your critical comment. We have unified the terms into 'mild' and 'aggravated'.

2.5. How was "Mild or aggravated symptoms" defined? It is also helpful to describe how respondents to the questionnaire recognized and distinguished between them.

: We have added a brief description of the symptoms as follows: "First, regarding knowledge about COVID-19-related coping, questionnaires assessed coping involving mild (fever, cough, and fatigue without signs of severe pneumonia) or aggravated symptoms (severe respiratory distress, chest pain, and confusion indicating potential severe illness requiring immediate medical attention)."

4

2.6. The authors describe in the main text that both "knowledge about COVID-19-related coping" and "preventive behaviour practice" were categorised as "practice-health behaviour group (0)" or "not practice-health behaviour group (1)". However, the presentation in Table 2 categorises "knowledge about COVID-19-related coping" into "Deficient knowledge regarding coping to symptoms" and others, and "preventive behaviour practice" is classified as "Not-practicing preventive behaviours on COVID-19 infection" and others. The terminology in the text and in the Table should be aligned.

: We have unified the terms into "Deficient knowledge regarding coping with symptoms" and "Not-practicing preventive behaviors on COVID-19 infection."

3) Statistical analysis
2.7. The authors describe null model, second model and final model in the main text, but the models in the Table are described as model 0, model 1 and model 2. It would be better to either align the notation of the terms in the text and the Table, or to clearly explain which model in the main text corresponds to which model in the Table.
: We described Models 0, 1, and 2 in the statistical methods to avoid confusion. A null model containing no variables is used in the same sense as Model 0. We discussed statistical analysis and its interpretation in collaboration with a statistician.

2.8 What is analysed with "the null model, meaning that it had no variables"? And what is the meaning for this analysis?
: Thanks for your comment. Model 0 (i.e. the null model) in our study serves as a baseline to quantify variance in outcomes attributable solely to differences between groups, providing justification for further multilevel analysis.

2.9. It is unclear what is meant by the sentence "Intra-class correlation (ICC) was used to measure the effect of variables at individual and community levels." Although I am not an expert in statistics, Isn't the intraclass correlation coefficient (ICC) a reliability measure used in intra-rater and inter-rater reliability analysis? Is the use of ICC suitable in this situation?
: Intra-class correlation (ICC) is the ratio of the residual variance between groups (i.e., between communities) to the total residual variance, which is represented as the sum of the residual variance between groups and intra-group (i.e., within individuals). Here, the residual variance between groups accounts for 5 percent of the total residual variance, suggesting that multilevel analysis is suitable for explaining differences between groups. Additionally, the ICC decreased from Model 1, which included only individual variables, to Model 2, which incorporated both individual and community variables. This data implies that the variance between communities decreases when these variables are considered, suggesting improvements in the model's performance.
2.10 It would be better to state in the Statistical analysis how p-values were calculated.
: Thank you for your valuable comment. Here, the p-value is calculated from a t-value, which is derived by dividing the estimated coefficient by its standard error. To more clearly describe the statistical methodology, we have added the following explanation to the note accompanying the table: 'The p-value is calculated from a t-value, which is the estimated coefficient divided by its standard error.'

3. Results
3.1. The statement of " Older rural resident adults were more likely to not practice preventive behaviors compared to urban resident adults." seems to be different from the results of Table 2.
: Thanks for your valuable comment. In table 2, we showed that odd ratio of non-practising preventive behaviors on COVID-19 infection on rural resident adults to urban resident adults were 1.247 (CI: 1.141-1.363). This means rural resident adults were tended not to practice preventive behaviors compared to urban resident adults.

3.2. With regard to the sentence "those who were economically active had household incomes ranging from 0 to 2500,000 won ($2011.05).", I got the impression that this text does not explain the results presented by the authors in Table 1.
: Thanks for your critical comment. We removed the sentence to avoid ambiguity.

3.3. I would like you to provide a citation for the sentence "Suggesting that the collected data reflected the initial response pattern of older adults to COVID-19."
: According to the report from The Korea Disease Control and Prevention Agency, the first national wide spread of SARS-CoV2 took place from 20, 1, 2020 to 6, 7, 2021. Data collection was conducted from 8, 2020 to 10, 2020 during the above period. To make the sentence more clearly, we modified the sentence as follows: "the collected data reflected the response pattern of older adults to COVID-19 during the first national-wide transmission (Lee et al., 2022)."

3.4. Please also explain the characteristics of the Enabling factors. As the authors focus on the results on the presence or absence of a living family member in the Abstract, it is necessary to explain the characteristics of the Enabling factors in Results in the main texts.
: Thanks for your suggestion. Regarding enable factors, we described significant predictors including living alone and number of supportive people in the abstract section. To describe more clearly, we modified our text as follows: "Older adults who reside alone, as opposed to with their spouse, and who lack the support of another individual."

3.5. The authors state that "Model 2 was also the most explanatory (ICC:0.064)", but for an ICC that ranges from 0 to 1, 0 indicates unreliability and 1 indicates complete reliability. Therefore, it may be that "Model 2 was the least explanatory".
: In this study, the intra-class correlation (ICC) represents the proportion of residual variation between groups relative to the total residual variation, with a higher ICC indicating that between-group variation exceeds intra-group variation. In Model 2, the ICC decreased compared to Model 0 or Model 1, indicating a reduction in between-group variation when considering various individual- and community-level variables. Therefore, we believe that Model 2 is the most appropriate model for explaining the between-group variation.

3.6. There is no description at all on individual and community factors influencing deficient knowledge regarding coping to symptoms.
: In our results section, we described results showed on table 2, specifically focusing on individual and community factors influencing deficient knowledge regarding coping to symptoms with the title "Individual and community factors influencing deficient knowledge regarding coping with COVID-19 symptoms". Please consider that we only highlighted the significantly affecting predictors on the dependent variable.

3.7. Why have the authors not analysed "the differences in influencing factors between living in urban and rural areas regarding deficient knowledge regarding coping to symptoms?"
: In table 2, we presented that odds ratio on deficient knowledge regarding coping to symptoms between rural residents and urban residents were not significantly different. So, we believe that it is not needed to be further analyzed by sub-group analysis.

3.8. Throughout, there is a mixture of notations to two decimal places and three decimal places, or there is more than one term for one meaning (e.g. oldest-old and older-old). Unifying the notation would make it easier for readers to read the manuscript.
: Thanks for your suggestions. We reviewed our manuscript and unified two words with same meaning to a word. In addition, we changed two decimal places into three decimal places.

VERSION 2 – REVIEW

REVIEWER	Shelley, Mack Iowa State University, Department of Political Science, and Department of Statistics
REVIEW RETURNED	26-May-2024

GENERAL COMMENTS	Apologies for an apparent typo that caused confusion regarding the reporting of p-values. For very small p-values, the document should state "p<0.001," rather than "p<0.000" (which would be impossible as it implies a negative probablity, which cannot exist). Also, please be sure that the "p" in statements of probability levels is always italicized.

REVIEWER	タカダ, ミドリ Osaka Center for Cancer and Cardiovascular Disease Prevention
REVIEW RETURNED	10-Jun-2024

GENERAL COMMENTS	Thank you for responding to my comments.

VERSION 2 – AUTHOR RESPONSE

Reviewer1

Apologies for an apparent typo that caused confusion regarding the reporting of p-values.

For very small p-values, the document should state "p<0.001," rather than "p<0.000" (which would be impossible as it implies a negative probablity, which cannot exist).

Also, please be sure that the "p" in statements of probability levels is always italicized.

: Thanks for your critical comment. We modified the format of p-value as you suggested. Additionally, we corrected our typos and error in notation of p.