

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)	Depression and associated factors among HIV-positive smokers receiving care at HIV outpatient clinics in Vietnam: a cross-sectional analysis
AUTHORS	Nguyen, Nam; Nguyen, Trang; Vu, Giap; Truong, Nga; Pham, Yen; Guevara Alvarez, Gloria; Armstrong-Hough, Mari; R Shelley, Donna

VERSION 1 – REVIEW

REVIEWER	Okafor, Chukwuemeka N The University of Texas Health Science Center at San Antonio
REVIEW RETURNED	05-Sep-2023

GENERAL COMMENTS	<p>The article describes factors associated with depression symptoms among persons with HIV receiving care at a HIV outpatient clinic in Vietnam. The analysis is fairly straightforward, and the results are for the most part expected. The following are a few suggestions for the authors:</p> <p>Methods</p> <ol style="list-style-type: none">1. It would help of the authors describe what constitutes "drug use"?2. Since the data is from a smoking cessation RCT, and assuming not everyone eligible agreed to participate, if possible, it would help to describe the characteristics of individuals who agreed and did not participate, to assess whether this group studied is representative of persons with HIV attending OPCs in Hanoi?3. This reviewer is curious on whether a specific behavioral theory guided the select of variables that were studied? Or was it based on what was assessed in the parent RCT?4. Is it depressive or depression symptoms? Also, for clarity and consistency, might suggest referring to the outcome as clinically significant depression symptoms as defined in the measures section. <p>Data Analysis</p> <ol style="list-style-type: none">1. Could you clarify why a p-value of <0.2 was used in selecting covariates to include in the multivariable model?2. Consider presenting the measures of associations as prevalence ratios rather than odds ratios, which overestimate associations when the outcome is common. <p>Results</p> <ol style="list-style-type: none">1. With n=22 females, could it be that these results are representative of males?2. Some variables in Table 2 are missing from Table 1. It would be helpful to describe these variables in Table 1 also.
-------------------------	--

	<p>3. Most the results described in the result section are repetitive because they are in the Tables.</p> <p>Discussion</p> <p>1. Some sentences in the discussion section require citations. For example, the 4th sentence in the second paragraph needs a citation.</p> <p>2. The reference to HIV-associated stigma, without a citation, appears speculative. Moreover, HIV-associated stigma or discrimination was not assessed in this study. Consider supporting this section with appropriate citations.</p> <p>3. I believe the authors should also discuss other variables that could impact depression symptoms that were not assessed in this study such as stigma, discrimination etc.</p>
--	--

REVIEWER	Peng, Lei Peking University
REVIEW RETURNED	21-Oct-2023

GENERAL COMMENTS	<p>This research includes 540 smokers receiving HIV treatment at HIV outpatient clinics. And then, the aims of this research were to assess the prevalence of depressive symptoms and associated factors among people living with HIV (PLWH) who are current smokers receiving treatment at HIV outpatient clinics (OPCs) in Ha Noi, Vietnam. From my point of view, This study has positive public health implications and I agree to publish it. However, The author of this article may also need to address the following issues:</p> <p>1.Lines 3-5 on page 5: The significance level the authors defined is 0.05, but the variables ultimately included in the regression model have a P-value less than 0.20. Is there any reference support for this? It is recommended to cite relevant literature here.</p> <p>2.I suggest that the author of this article change Table 1 to present the results of T-test and Chi square test.</p> <p>3.I don't know if the variables in Table 2 are included in binary logistic regression at the same time, or if each variable is regressed separately to obtain the results. The variables considered by the author are too many. Please explain to that if all of the variables in Table 2 are included in the regression model at the same time, because it may cause the model to not converge or the results to be inaccurate. At this point, FDR (false discovery rate) correction for P-value is also necessary.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Chukwuemeka N Okafor, The University of Texas Health Science Center at San Antonio

Comments to the Author:

The article describes factors associated with depression symptoms among persons with HIV receiving care at a HIV outpatient clinic in Vietnam. The analysis is fairly straightforward, and the results are for the most part expected. The following are a few suggestions for the authors:

Methods

1. It would help if the authors describe what constitutes "drug use"?

Response: We have added a drug use definition and description of questions asking about drug use to the method section – independent variables.

“Drug use was defined as the use of substances for psychotropic rather than medical purposes and was assessed using 2 questions that asked if respondents ever used and used in the past 3 months any of substances including Opium, Cocaine, Heroin, Amphetamine/Methamphetamine, Marijuana, Ecstasy, MDMA, Ketamine...”

2. Since the data is from a smoking cessation RCT, and assuming not everyone eligible agreed to participate, if possible, it would help to describe the characteristics of individuals who agreed and did not participate, to assess whether this group studied is representative of persons with HIV attending OPCs in Hanoi?

Response: We collected some information about gender, age, and smoking status from the patients who were invited to participate in the RCT, but did not agree to participate. We ran the analyses to compare the two groups and found no significant differences between the two groups in terms of gender, age, and smoking status.

We have added this information to the methods section – study design

“An analysis of patients who declined to participate demonstrated no significant differences in gender, age, and smoking status compared with those who enrolled”.

3. This reviewer is curious on whether a specific behavioral theory guided the select of variables that were studied? Or was it based on what was assessed in the parent RCT?

Response: We analyzed data from baseline surveys for the parent RCT.

The dependent variable - depressive symptoms was created based on CE-D 8 included in the baseline survey.

Independent variables were selected from variables that were potentially associated with depressive symptoms based on prior literature. We first included potential independent variables in the bivariate analyses. We then selected independent variables with a $p < 0.2$ in the bivariate analyses to include in the Poisson regression model.

4. Is it depressive or depression symptoms? Also, for clarity and consistency, might suggest referring to the outcome as clinically significant depression symptoms as defined in the measures section.

Response: The outcome of this paper was depressive symptoms.

We edited the paper to consistently use “depressive symptoms” rather than depression symptoms.

Data Analysis

1. Could you clarify why a p-value of < 0.2 was used in selecting covariates to include in the multivariable model?

Response: We used results from bivariate analyses to select variables for regression models.

Independent variables with a p-value < 0.05 (significantly associated with the dependent variable - depressive symptoms) in the bivariate analyses were included in the multivariate regression model.

Apart from including the variables with a p-value < 0.05 (from bivariate analyses), we included other variables with a p-value < 0.2 in the multivariate regression models based on two reasons:

First, including independent variables with a p-value < 0.05 (from bivariate analyses) in a multivariate regression is likely to omit important predictors whose association with the dependent variable only becomes significant when other independent variables are taken into account.

We used a higher p-value cutoff than 0.05 to reduce the omitted-variable bias. We used $p < 0.2$ as this cut-off is commonly used by researchers.

Second, the independent variables with p-value < 0.2 (from the bivariate analyses) included in the multivariate regression were also control variables in the models. The Prevalence Ratios (PR) in the multivariate model (Poisson Regression) were calculated for one independent variable by controlling/adjusting for other independent variables.

We have added a citation/reference for the use of p-value <0.2 to select variables from bivariate analyses to include in the multivariate regression models in the data analysis section.

2. Consider presenting the measures of associations as prevalence ratios rather than odds ratios, which overestimate associations when the outcome is common.

Response: We agree with your suggestion and have used Poisson regression for multivariate analyses. Poisson regression (estimating Prevalence Ratios) is a better alternative for the analysis of cross-sectional studies with binary outcomes (especially when the prevalence is high) than logistic regression (reporting Odd Ratios).

We have rerun the multivariate analyses using Poisson regression to estimate Prevalence Ratios. The results with Prevalence Ratios (PR) were presented in Table 2.

We have revised the data analyses section to reflect the change to use Poisson Regression and estimating Prevalence Ratios.

Results

1. With n=22 females, could it be that these results are representative of males?

Response: The lower number of females in this study reflects the low prevalence of females who smoke cigarettes (the national data on cigarette smoking by sex shows only 1.1% of females smoked cigarettes).

Although the number of females (n=22) was low, the bivariate and multivariate analyses showed significant differences in the association between gender and depressive symptoms. Females had a higher prevalence of depressive symptoms than males.

We think the results could still be representative for both males and females.

2. Some variables in Table 2 are missing from Table 1. It would be helpful to describe these variables in Table 1 also.

Response: We have changed Table 1 to present both participants' characteristics and results from bivariate analyses. Table 1 now presents all variables.

Table 2 now presents results from multivariate regression only.

3. Most the results described in the result section are repetitive because they are in the Tables.

Response: We have revised the result sections.

Discussion

1. Some sentences in the discussion section require citations. For example, the 4th sentence in the second paragraph needs a citation.

Response: We have reviewed and added citations for the discussion when needed (including the 4th sentence in the second paragraph).

2. The reference to HIV-associated stigma, without a citation, appears speculative. Moreover, HIV-associated stigma or discrimination was not assessed in this study. Consider supporting this section with appropriate citations.

Response: We have added citations for that sentence about HIV-associated stigma and throughout that section.

3. I believe the authors should also discuss other variables that could impact depression symptoms that were not assessed in this study such as stigma, discrimination etc.

Response: We have added a paragraph discussing the impact of biological and psychosocial factors including HIV-associated stigma, and discrimination on depressive symptoms among people living with HIV.

Reviewer: 2

Dr. Lei Peng, Peking University

Comments to the Author:

From my point of view, This study has positive public health implications and I agree to publish it. However, The author of this article may also need to address the following issues:

1.Lines 3-5 on page 5: The significance level the authors defined is 0.05, but the variables ultimately included in the regression model have a P-value less than 0.20. Is there any reference support for this? It is recommended to cite relevant literature here.

Response: We used results from bivariate analyses to select variables for regression models. Independent variables with a p-value <0.05 (significantly associated with the dependent variable - depressive symptoms) in the bivariate analyses were included in the multivariate regression model. Apart from including the variables with a p-value <0.05 (from bivariate analyses), we included other variables with a p-value <0.2 in the multivariate regression models based on 2 reasons: First, including independent variables with a p-value < 0.05 (from bivariate analyses) in a multivariate regression is likely to omit important predictors whose association with the dependent variable only becomes significant when other independent variables are taken into account. We used a higher p-value cutoff than 0.05 to reduce the omitted-variable bias. We used p<0.2 as this cut-off is commonly used by researchers. Second, the independent variables with p-value <0.2 (from the bivariate analyses) included in the multivariate regression were also control variables in the models. The Prevalence Ratios (PR) in the multivariate model (Poisson Regression) were calculated for one independent variable by controlling/adjusting for other independent variables.

We have added a citation/reference for the use of p-value <0.2 to select variables from bivariate analyses to include in the multivariate regression models in the data analysis section.

2. I suggest that the author of this article change Table 1 to present the results of T-test and Chi square test.

Response: We have changed Table 1 to present both participants' characteristics and results from bivariate analyses with p-values from T-test and Chi-square test. Table 2 now presents results from multivariate regression only.

3.I don't know if the variables in Table 2 are included in binary logistic regression at the same time, or if each variable is regressed separately to obtain the results. The variables considered by the author are too many. Please explain to that if all of the variables in Table 2 are included in the regression model at the same time, because it may cause the model to not converge or the results to be inaccurate. At this point, FDR (false discovery rate) correction for P-value is also necessary.

Response: We selected only variables with p-value <0.2 in the bivariate analyses to include in the multivariate regression model. In total, 10 variables were included in the regression model. As a common rule, at least 10 participants (sample size) are needed for each independent variable

included in the regression model. Our sample size of 527 was enough for 10 variables to be included in the regression model.

In the original manuscript, Table 2 included all bivariate and multivariate analysis results so it included many variables. Now as we have revised Table 2 to present only the Poisson regression results, only 10 variables included in the regression model were presented.

Following the comment from reviewer 1, we have used Poisson Regression for multivariate analyses instead of logistic regression.

The Poisson Regression with robust variance is better than logistic regression in terms of convergence issues[1, 2].

1. Tamhane AR, Westfall AO, Burkholder GA, Cutter GR. Prevalence odds ratio versus prevalence ratio: choice comes with consequences. Stat Med. 2016;35(30):5730-5.
2. Vittinghoff E, Shiboski S, Glidden D, McCulloch C. Regression Methods in Biostatistics: Linear, Logistic, Survival and Repeated Measures Models. New York: Springer; 2011.

VERSION 2 – REVIEW

REVIEWER	Peng, Lei Peking University
REVIEW RETURNED	27-Nov-2023
GENERAL COMMENTS	All comments have been resolved and I agree to publish this paper.

VERSION 2 – AUTHOR RESPONSE