

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

#### Title (Provisional)

Social support status and associated factors among methadone maintenance patients: a multicenter, cross-sectional study in Vietnam during the COVID-19 pandemic

#### Authors

Nguyen, Huong Thi Thanh; Quang, Dien Tran; Dinh, Dai Xuan

---

### VERSION 1 - REVIEW

---

<b>Reviewer</b>	<b>1</b>
<b>Name</b>	<b>Van, Hoang Thi Hai</b>
<b>Affiliation</b>	<b>Hanoi Medical University, Global Health</b>
<b>Date</b>	<b>05-Mar-2024</b>
<b>COI</b>	<b>n/a</b>

---

1. In the literature review section, I would encourage the authors to expand on this section by discussing the COVID-19 context of Vietnam in more detail. Furthermore, I would encourage the authors to ground their research in a theoretical framework.
  2. The methodology is adequately described. However, how were clinics in each province selected?
  3. Results: Certain findings are expected.
  4. The Discussion section is comprehensive and the authors have done well in comparing and contrasting their findings with the existing literature. However, the differences between the three provinces should be combined with their COVID-19 pandemic and the difference between pandemic time and normal time situation should be discussed.
  5. The implications are too general.
- 

<b>Reviewer</b>	<b>2</b>
<b>Name</b>	<b>Nguyen, Diep Bich</b>

**Affiliation**                    **Hanoi Medical University**  
**Date**                            **16-Apr-2024**  
**COI**                              **I have no conflict of interest**

---

This manuscript presents findings from a cross-sectional study aimed at exploring social support and its associated factors among Vietnamese patients undergoing methadone maintenance treatment amidst the COVID-19 pandemic. Regrettably, the manuscript fails to incorporate or discuss the pivotal context of the COVID-19 outbreak during the study period. This absence significantly diminishes the novelty and relevance of the study's outcomes.

Introduction:

The rationale for investigating social support within the framework of the COVID-19 pandemic needs to be clearly articulated. Why is this context essential for understanding the dynamics of social support among methadone patients?

Measures:

Provide a more detailed explanation of the recruitment and sampling procedures. Clarify how participants were approached and calculate the percentage of individuals approached (556) relative to the total number of patients within the selected clinics. Describe the data collection methods, specifying whether interviews or self-administered surveys were employed, and whether these were computer-assisted or paper-based. Additionally, address any instances of missing data.

Discussion:

Given the minimal representation of female participants (approximately 1%), please refrain from overemphasizing gender differences.

Please discuss the study results within the context of the COVID-19 pandemic. For example, explore how the findings compare to those from other countries, both during and outside the pandemic. Furthermore, examine any disparities in the COVID-19 situation (such as epidemic status, policies, etc.) among the three provinces during the data collection period.

---

**Reviewer**                    **3**  
**Name**                         **Bastos, Francisco**  
**Affiliation**                 **Fundacao Oswaldo Cruz, Health Information**  
**Date**                         **30-Aug-2024**  
**COI**                         **None**

---

Please, fix the sentence as follows:

“R software version 4.3.1”

R is NOT a software, as carefully explained in the link below. It may be integrated with different softwares, but it’s a computational language and environment. Please. See the correct citation as follows: <https://ropensci.org/blog/2021/11/16/how-to-cite-r-and-r-packages/>

The proper definition is as follows: R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. (verbatim)

The whole abstract does not contain a single word about the N of the study. One should remember that sample size should be defined based on statistical power or precision (<https://pubmed.ncbi.nlm.nih.gov/29912015/>).

Without any of them, it’s impossible to avert type II errors and/or to get a less than optimal precision.

Everything suggests that the variables as follows have some degree of collinearity:

“patient’s occupation ( $p < 0.001$ ), occupation type ( $p = 0.001$ ), family’s monthly income ( $p < 0.001$ )”

Please, include in the analysis a comprehensive assessment of their putative collinearity.

Please, see the plain definition by Voss (D. Stephen Voss, in Encyclopedia of Social Measurement, 2005):

When two or more explanatory variables overlap completely, with one a perfect linear function of the others, such that the method of analysis cannot distinguish them from one another. This condition prevents a multiple regression from estimating coefficients and the equation becomes unsolvable.

The limitation does exist, but the explanation as follows is misleading:

“Using a cut-off point of 0.001 to identify statistically significant variables can help increase the reproducibility probability of findings”

Actually a strict and meaningful cut-off complains with the ASA statement on p-values (<https://www.tandfonline.com/doi/full/10.1080/00031305.2016.1154108>), trying to minimize spurious associations and other errors. But this cannot be described as “reproducibility probability”.

The sentence as follows is partially true:

“Causal relationships between independent factors and patients’ social support cannot be determined because this is only a cross-sectional study.”

For sure, cross-sectional studies are ALWAYS associated with limitations, but the authors could have made some attempts to create DAGs. They are nothing but the graphic

expression of Bayesian networks. Please, see the tutorial as follows:

[https://causalnex.readthedocs.io/en/stable/04\\_user\\_guide/04\\_user\\_guide.html](https://causalnex.readthedocs.io/en/stable/04_user_guide/04_user_guide.html).

The sentence as follows is partially true:

“Employing a convenience sampling method for recruiting patients can give rise to selection bias.”

Convenience samples ALWAYS challenge statistical inference. There are means to improve statistical inference for non-probability samples, but they are NOT discussed by the authors. Selection bias is just ONE (among several other problems) of statistical inference based on non-probability samples). Please, see the seminal paper as follows:

<https://projecteuclid.org/journals/statistical-science/volume-32/issue-2/Inference-for-Nonprobability-Samples/10.1214/16-STS598.full>.

“In recent years, online sales have made access to illicit drugs simpler than ever.”

The sentence above is correct, but there is no reference, especially one that should demonstrate this phenomenon is taking place in Vietnam.

The sentence as follows has no clear meaning:

“Participants were recruited using a convenience sampling method”

There is not a given method that could be called A “convenience sampling method”.

This could be accomplished by recruiting consecutive patients attending a service, using snowballing, using a first come-first recruited basis, etc. Please, clarify.

Again, why 180 individuals? Based on what? Convenience samples do NOT exclude the need to calculate the statistical power and/or precision of a sample. There is not a single word about it.

The purpose of descriptive statistics is to summarize and reduce data. This is NOT what the authors have described as follows:

“Frequencies and percentages were used to DESCRIBE categorical variables (such as sex and place of residence), while means (standard deviations - SD) and medians (25th-75th/min-max) were employed to REPORT numeric variables” (verbatim, my emphasis).

Please, see the tutorial as follows, which has been originally implemented for the analysis of genetic databases: <https://www.nature.com/articles/s41467-021-27438-7>.

LASSO is an excellent method to select variables

(<https://spectdata.com/index.php/2019/08/08/variable-selection-using-lasso/>), but DOES not address the key gaps: i.e. there is no analysis of putative collinearity, no assessment of putative interactions, and not a single word about the fit of the model.

As a famous statistician says: a model without any kind of diagnosis is “an act of faith”, but not science. It has zero replicability.

Please, provide a careful assessment of the fit of the model.

There are infinite ways to do it, varying from the old classics to the most modern alternatives.

Just two basic guidelines:

Please, browse chapter five (Assessing the Fit of the Model) of the book as follows:

[https://www.amazon.co.uk/Applied-Logistic-Regression-Probability-Statistics/dp/0471356328/ref=sr\\_1\\_1?crid=1G09O2258L4IG&dib=eyJ2ljojMSJ9.wRZdgJfBl0BRNA5m7WKfy4Rkvbrwukqfave6u3kkxchVMbk\\_9VO\\_4ji99iJ6Fy6mNddvWQTX2W2zuA5twl1tDFdbhJWqNCPJ26Qr-oWfKBqtFnxmZmq04Eez0ye4DFilG9-Mu7bTVO-4AWZ2fTthmw.Ev8m50x5jeeBrYC8eqzj-CvKTq1LZBRy22CXYYvjIMM&dib\\_tag=se&keywords=hosmer+and+lemeshow&qid=1725025030&s=books&sprefix=hosmer+and+lemeshow%2Cstripbooks%2C253&sr=1-1-catcorr](https://www.amazon.co.uk/Applied-Logistic-Regression-Probability-Statistics/dp/0471356328/ref=sr_1_1?crid=1G09O2258L4IG&dib=eyJ2ljojMSJ9.wRZdgJfBl0BRNA5m7WKfy4Rkvbrwukqfave6u3kkxchVMbk_9VO_4ji99iJ6Fy6mNddvWQTX2W2zuA5twl1tDFdbhJWqNCPJ26Qr-oWfKBqtFnxmZmq04Eez0ye4DFilG9-Mu7bTVO-4AWZ2fTthmw.Ev8m50x5jeeBrYC8eqzj-CvKTq1LZBRy22CXYYvjIMM&dib_tag=se&keywords=hosmer+and+lemeshow&qid=1725025030&s=books&sprefix=hosmer+and+lemeshow%2Cstripbooks%2C253&sr=1-1-catcorr)

Contemporary alternatives are comprehensively described as follows:

<https://royalsocietypublishing.org/doi/10.1098/rsbl.2019.0174>

By now, it's impossible to discern whether results are accurate or secondary to sampling errors and/or less than optimal analyses.

Please, try to follow the very basics of statistics "by the book" and then present and discuss findings on a sound basis.

---

## VERSION 1 - AUTHOR RESPONSE

### Reviewer: 1

1. In the literature review section, I would encourage the authors to expand on this section by discussing the COVID-19 context of Vietnam in more detail. Furthermore, I would encourage the authors to ground their research in a theoretical framework.

Regarding the COVID-19 context, we added some information involving this pandemic in Vietnam in the Introduction.

Some previous studies used EFA/CFA to analyze the validity of the MOS-SSS questionnaire. We used this questionnaire to calculate the social support scores. In this study, we did not want to use SEM to investigate factors associated with the social support of methadone maintenance patients because it is relatively complex. However, in the Methods section, we also described the construct of the MOS-SSS questionnaire, which was employed to measure patients' social support (with four domains).

2. The methodology is adequately described. However, how were clinics in each province selected?

We selected these clinics on purpose. It is difficult to approach healthcare facilities if we do not have relationships with their leaders. We had to obtain the permission of clinic directors before interviewing patients. As a result, we selected several clinics where the first author (Ms. Huong) has established connections with the leadership.

3. The Discussion section is comprehensive and the authors have done well in comparing and contrasting their findings with the existing literature. However, the differences between the three provinces should be combined with their COVID-19 pandemic and the difference between pandemic time and normal time situation should be discussed.

We added some information in the Discussion part to discuss the difficulty involving social support for MMT patients in the context of the COVID-19 pandemic.

4. The implications are too general.

We revised this part.

If there is anything inappropriate or inadequate, please let us know. Thank you so much for your valuable comments.

Best regards.

**Reviewer: 2**

Introduction:

The rationale for investigating social support within the framework of the COVID-19 pandemic needs to be clearly articulated. Why is this context essential for understanding the dynamics of social support among methadone patients?

We added some information involving the COVID-19 pandemic and its potential effects on MMT patients. However, to avoid repeating the information, we discussed these topics in more detail in the Discussion part.

Measures:

Provide a more detailed explanation of the recruitment and sampling procedures. Clarify how participants were approached and calculate the percentage of individuals approached (556) relative to the total number of patients within the selected clinics. Describe the data collection methods, specifying whether interviews or self-administered surveys were employed, and whether these were computer-assisted or paper-based. Additionally, address any instances of missing data.

We added some information involving the data collection process in the Methods. We hope that the necessary information has been provided.

Discussion:

Given the minimal representation of female participants (approximately 1%), please refrain from overemphasizing gender differences.

Please discuss the study results within the context of the COVID-19 pandemic. For example, explore how the findings compare to those from other countries, both during and outside the pandemic. Furthermore, examine any disparities in the COVID-19 situation (such as epidemic status, policies, etc.) among the three provinces during the data collection period.

Regarding the variable "sex", we moved this paragraph to the end of the Discussion part and shortened it. Thank you for this meaningful comment.

As for issues involving the COVID-19 pandemic, we added some information in this part. However, it is difficult to analyze in detail the differences between the three provinces in

terms of policies because anti-epidemic policies and measures often change quite quickly during a pandemic.

If there is anything inappropriate or inadequate, please let us know. Thank you so much for your valuable comments.

Best regards.

**Reviewer: 3**

Please, fix the sentence as follows: “R software version 4.3.1”

R is NOT a software, as carefully explained in the link below. It may be integrated with different softwares, but it’s a computational language and environment. Please. See the correct citation as follows: <https://ropensci.org/blog/2021/11/16/how-to-cite-r-and-rpackages/>

The proper definition is as follows: R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. (verbatim)

We revised this issue as recommended.

The whole abstract does not contain a single word about the N of the study. One should remember that sample size should be defined based on statistical power or precision (<https://pubmed.ncbi.nlm.nih.gov/29912015/>). Without any of them, it’s impossible to avert type II errors and/or to get a less than optimal precision.

In the Abstract, we mentioned the number of patients in our study (N=540 patients) as per the guideline of BMJ Open for the Abstract. Regarding the total number of MMT patients in Vietnam, we mentioned this data in the Introduction: “According to the statistics of the Ministry of Health, as of 2023, roughly 50,353 patients used methadone to treat opioid dependence in 343 methadone clinics in Vietnam”.

“In recent years, online sales have made access to illicit drugs simpler than ever.”

The sentence above is correct, but there is no reference, especially one that should demonstrate this phenomenon is taking place in Vietnam.

We added a citation for this information. Thank you for this helpful comment.

The limitation does exist, but the explanation as follows is misleading:

“Using a cut-off point of 0.001 to identify statistically significant variables can help increase the reproducibility probability of findings”

Actually a strict and meaningful cut-off complains with the ASA statement on p-values (<https://www.tandfonline.com/doi/full/10.1080/00031305.2016.1154108>), trying to minimize spurious associations and other errors. But this cannot be described as “reproducibility

probability”.

We removed this sentence to avoid misunderstanding.

The sentence as follows is partially true:

“Causal relationships between independent factors and patients’ social support cannot be determined because this is only a cross-sectional study.”

For sure, cross-sectional studies are ALWAYS associated with limitations, but the authors could have made some attempts to create DAGs. They are nothing but the graphic expression of Baysean networks. Please, see the tutorial as follows: [https://causalnex.readthedocs.io/en/stable/04\\_user\\_guide/04\\_user\\_guide.html](https://causalnex.readthedocs.io/en/stable/04_user_guide/04_user_guide.html).

Thank you for your meaningful suggestion. However, in this study, we only used a questionnaire to calculate patients’ social support scores and employed linear regression models to identify some associated factors (following the protocol registered). We do not aim to use SEM/DAGs in this study.

The sentence as follows is partially true: “Employing a convenience sampling method for recruiting patients can give rise to selection bias.” Convenience samples ALWAYS challenge statistical inference. There are means to improve statistical inference for non-probability samples, but they are NOT discussed by the authors. Selection bias is just ONE (among several other problems) of statistical inference based on non-probability samples). Please, see the seminal paper as follows: <https://projecteuclid.org/journals/statisticalscience/volume-32/issue-2/Inference-for-Nonprobability-Samples/10.1214/16-STS598.full>.

The sentence as follows has no clear meaning:

“Participants were recruited using a convenience sampling method”

There is not a given method that could be called A “convenience sampling method”.

This could be accomplished by recruiting consecutive patients attending a service, using snowballing, using a first come-first recruited basis, etc. Please, clarify.

Again, why 180 individuals? Based on what? Convenience samples do NOT exclude the need to calculate the statistical power and/or precision of a sample. There is not a single word about it.

We used the term “convenience sampling” as per many previous documents/articles. For example, <https://www.sciencedirect.com/topics/computer-science/nonprobability-sample>

[https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2\\_41](https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2_41)

<https://uta.pressbooks.pub/foundationsofsocialworkresearch/chapter/6-2-nonprobabilitysampling/>

We did not ask the participants to introduce their friends to participate in our study. Therefore, snowballing may be inappropriate. As we can see, in one MMT clinic, MMT patients have to visit this clinic to take methadone. All patients will visit this clinic in one day. By reason of the large number of patients, we could not approach all of them in only one day. "Consecutive patients attending a service" and "using a first come-first recruited basis" are also inappropriate. In a day, we can only interview roughly 20 patients and other patients must be approached on other days.

We added the formula to compute the minimum sample size. We strived to approach as many patients as possible to increase the number of patients participating in our research.

The purpose of descriptive statistics is to summarize and reduce data. This is NOT what the authors have described as follows: "Frequencies and percentages were used to DESCRIBE categorical variables (such as sex and place of residence), while means (standard 129 deviations - SD) and medians (25th-75th/min-max) were employed to REPORT numeric variables" (verbatim, my emphasis).

Please, see the tutorial as follows, which has been originally implemented for the analysis of genetic databases: <https://www.nature.com/articles/s41467-021-27438-7>.

We revised this issue. Thank you for this meaningful comment.

Everything suggests that the variables as follows have some degree of collinearity: "patient's occupation ( $p < 0.001$ ), occupation type ( $p = 0.001$ ), family's monthly income ( $p < 0.001$ )". Please, include in the analysis a comprehensive assessment of their putative collinearity. Please, see the plain definition by Voss (D. Stephen Voss, in Encyclopedia of Social Measurement, 2005): When two or more explanatory variables overlap completely, with one a perfect linear function of the others, such that the method of analysis cannot distinguish them from one another. This condition prevents a multiple regression from estimating coefficients and the equation becomes unsolvable.

LASSO is an excellent method to select variables (<https://spectdata.com/index.php/2019/08/08/variable-selection-using-lasso/>), but DOES not address the key gaps: i.e. there is no analysis of putative collinearity, no assessment of putative interactions, and not a single word about the fit of the model.

As a famous statistician says: a model without any kind of diagnosis is “an act of faith”, but not science. It has zero replicability. Please, provide a careful assessment of the fit of the model. There are infinite ways to do it, varying from the old classics to the most modern alternatives. Just two basic guidelines:

Please, browse chapter five (Assessing the Fit of the Model) of the book as follows:

[https://www.amazon.co.uk/Applied-Logistic-Regression-Probability-Statistics/dp/0471356328/ref=sr\\_1\\_1?crid=1G09O2258L4IG&dib=eyJ2IjojMSJ9.wRZdqJfBI0BRNA5m7WKfy4Rkvbrwukqfave6u3kkxchVMbk\\_9VO\\_4ji99iJ6Fy6mNddvWQTX2W2zuA5twl1tDFdbhJWqNCPJ26Qr-oWfKBqtFnxmZmq04Eez0ye4DFilG9-Mu7bTVO-4AWZ2fTthmw.Ev8m50x5jeeBrYC8eqzj-CvKTq1LZBRy22CXYYvjIMM&dib\\_tag=se&keywords=hosmer+and+lemeshow&qid=1725025030&s=books&sprefix=hosmer+and+lemeshow%2Cstripbooks%2C253&sr=1-1-catcorr](https://www.amazon.co.uk/Applied-Logistic-Regression-Probability-Statistics/dp/0471356328/ref=sr_1_1?crid=1G09O2258L4IG&dib=eyJ2IjojMSJ9.wRZdqJfBI0BRNA5m7WKfy4Rkvbrwukqfave6u3kkxchVMbk_9VO_4ji99iJ6Fy6mNddvWQTX2W2zuA5twl1tDFdbhJWqNCPJ26Qr-oWfKBqtFnxmZmq04Eez0ye4DFilG9-Mu7bTVO-4AWZ2fTthmw.Ev8m50x5jeeBrYC8eqzj-CvKTq1LZBRy22CXYYvjIMM&dib_tag=se&keywords=hosmer+and+lemeshow&qid=1725025030&s=books&sprefix=hosmer+and+lemeshow%2Cstripbooks%2C253&sr=1-1-catcorr)

Contemporary alternatives are comprehensively described as follows: <https://royalsocietypublishing.org/doi/10.1098/rsbl.2019.0174>

By now, it’s impossible to discern whether results are accurate or secondary to sampling errors and/or less than optimal analyses.

Please, try to follow the very basics of statistics “by the book” and then present and discuss findings on a sound basis.

Thank you for these meaningful comments. We calculated VIFs to assess the collinearity in multivariate linear regression models.

> vif(model1) #LASSO

	GVIF	Df	GVIF^(1/(2*Df))
Location	1.903652	2	1.174619
Occupation	3.763648	4	1.180188
Worktype	2.291390	1	1.513734
Familymembers	1.288074	1	1.134933
FamilysIncome	1.507066	1	1.227626
Closefriend	1.397772	1	1.182274

>vif(model2) #BMA

	GVIF	Df	GVIF^(1/(2*Df))
Location	1.670085	2	1.136801
Worktype	1.269198	2	1.061408
Familymembers	1.237229	1	1.112308
FamilysIncome	1.402983	1	1.184476

Closefriend	1.396453	1	1.181716
-------------	----------	---	----------

VIFs of all independent variables in two multivariate linear models are lower than 3.8.

Regarding the goodness of fit of multivariate linear regression models, we reported multiple R-squared and adjusted R-squared as per the recommendations of this document (<http://www.medicine.mcgill.ca/epidemiology/joseph/courses/EPIB-621/fit.pdf>).

Due to the paucity of funding, we cannot purchase and read some of the closed-access documents above. However, we truly appreciate and thank Reviewer 3 for the many documents and scientific articles that you suggested.

If there is anything inappropriate or inadequate, please let us know. Thank you so much for your valuable comments.

Best regards.

---

## VERSION 2 - REVIEW

---

<b>Reviewer</b>	<b>1</b>
<b>Name</b>	<b>Van, Hoang Thi Hai</b>
<b>Affiliation</b>	<b>Hanoi Medical University, Global Health</b>
<b>Date</b>	<b>03-Dec-2024</b>
<b>COI</b>	

---

The author's revision meets the requirements of reviewers.