PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cumulative incidence of SARS-CoV-2 infection within the homeless population: insights from a citywide longitudinal study
AUTHORS	MOSNIER, Emilie; loubiere, sandrine; Monfardini, Elisabetta; Alibert, Agathe; Landier, Jordi; Ninoves, Laeticia; Bosetti, Thomas; AUQUIER, Pascal; Mosnier, Marine; Wakap, Stephanie Nguengang; Warszawski, Josiane; Tinland, Aurelie

VERSION 1 – REVIEW

REVIEWER	Sagar , Manish Boston University School of Medicine, Medicine
REVIEW RETURNED	28-Jun-2022

GENERAL COMMENTS	This study investigates SARS-CoV-2 prevalence in Marseille (mostly homeless population) at 2 different periods. Investigators did a serological evaluation of participants that are deemed to be experiencing homelessness. They find significant increase in SARS-CoV-2 prevalence. They find that SARS-CoV-2 population is higher among those without housing as compared to the general population. 1) Investigators should indicate the number of individuals that declined to participate in the blood draw and survey. This will indicate if the surveyed population is representative or selective. 2) Please indicate the number of participants that were resampled between the first and second survey. How did they insure that same individuals was not sampled multiple times. If there was resampling, were there instances where people with previous positive serology had subsequent negative serology. 3) They state that a total of 180 had positive serology. There was 74 and 136 positive in the first and second sampling. These numbers are incongruent unless there was resampling. 4) Table 2 states that total participants was 1241. Thus, this only incorporates data from the first sampling? 5) I am not sure a Kaplan Meier analysis is appropriate. Using serology as a marker for infection acquisition is not appropriate because serology indicates current or past infection. 6) It is unclear if the participants were asked about symptoms at the time of sampling or prior. Symptoms at time of sampling would be inappropriate because serology could reflect past infection. Furthermore, past symptoms would be liable to recall bias. 7) Figure 1 should be removed. It contains data that is not part of this study. 8) Investigators don't explain why physical distancing and inability to use hygiene materials should be associated with greater infection incidence. 9) There are multiple studies that have shown that lack of stable housing engenders greater risk for SARS-CoV-2 infection. Authors

	have highlighted some of the previous literature. I would include, Bean et. al., Clinical Infectious Diseases

REVIEWER	Racape, Judith ULB École de Santé Publique, Research Centre in "Epidemiology, Biostatistics and Clinical Research"
REVIEW RETURNED	06-Sep-2022

GENERAL COMMENTS	The manuscript entitled “Cumulative incidence of SARS-CoV-2 infection within the homeless population: insights from a citywide longitudinal study” is an interesting study to determine risk factors associated with SARS-CoV-2 in a cohort of homeless people. However, the analysis and the methodology of the paper need to be reviewed and clarified. Methods 1. The major point concerns the methodology and the use of survival analysis for your data. For longitudinal studies as survival analysis, you need a follow-up of the individual people during the time of your study. It is not clear if you have constructed your study for this methodology? You need to clarify this point. - How the people have been tested? They have been followed at different time period? - A person tested positive at a time X can have been positive before. Or a person tested negative at a time Y can have been positive later. How did you manage those points during the follow-up? - You mentioned two testing sessions 1/ June 5 to august 5 (n=1241) and 2/ September 11 to December 18 (n=721) . Your analyses should concern the people tested for both periods (n=721)? Why did you present the population characteristics for all the cohort (table 1) ? 2. You mentioned that “We fitted a multivariate Cox model by considering as eligible variables those that were significant in a univariate analysis at the 10% level ….” - Why did you not include "age" (p=0.05) and "hand washing" (p=0.06) in your model? (linked to preventing measures as mentionned in your discussion) - Did you check PH assumption of your model? 3. Why did you not separate the category Africa in north Africa and sub-Saharan Africa for country of birth? The risks factors are different in the literature and the comparison in your discussion (p9 l54) of “black ethnic group” must be nuanced. Results: - Table1: as mentioned before, the population characteristics must be presented for the population analyzed in both periods. - Table 2: Why did you not present the results concerning ETHOS accommodations? - p15-17: I am confused with the presentation of the graphs which are not refered into the text? - p22-51: What is the link between the report (mortality COVID-19) attached and the manuscript?

	Discussion: - P9 l35 : The comparison with the general population is tricky because of the heterogeneity of the spreading in different regions/cities in France (see introduction p4 l17) and the different testing modalities during the pandemic. - P9 l43 : The adjustment with age should be useful ? - P9 l 50 : A discussion on the characteristics of the shelter should be adequate - How have been managed the homeless issue during COVID-19 pandemic? What were the public health measures: concerning the isolation of homeless people ? hostel have been request ? References missing: - P12 l34/35 - P13 l45/46 General comments The main issue for this paper is the methodology: is the survival analysis appropriate? If you want to compare both testing session, the analysis is different. The authors should clarify those points and present the results in line of those analysis. I would also suggest a discussion of the results with those from a previous study of the authors (same cohort apparently): Loubiere S, Monfardini E, Allaria C, Mosnier M, Allibert A, et al. (2021) Seroprevalence of SARS-CoV-2 antibodies among homeless people living rough, in shelters and squats: A large population-based study in France. PLOS ONE 16(9): e0255498. https://doi.org/10.1371/journal.pone.0255498

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Manish Sagar , Boston University School of Medicine
Comments to the Author:
This study investigates SARS-CoV-2 prevalence in Marseille (mostly homeless population) at 2 different periods. Investigators did a serological evaluation of participants that are deemed to be experiencing homelessness. They find significant increase in SARS-CoV-2 prevalence. They find that SARS-CoV-2 population is higher among those without housing as compared to the general population.

1) Investigators should indicate the number of individuals that declined to participate in the blood draw and survey. This will indicate if the surveyed population is representative or selective.
We proposed to all eligible people to participate in the study in Marseille city. Refusals were not reported because it was not possible to conduct a complete census in slums or squats due to feasibility and security concerns at the time of inclusion. However, an estimate of the total population size is given in the table, in supl file, with the most recent data and shows a relatively good representativeness of our study.

For more precision on the representativeness we have added the following sentence in the results section « Approximately 37% of eligible ETHOS 2,3 and 8 participants were included in the study (Supplementary file 1). » We added also the last (and only) result of census of homeless people living in street.

Supplementary Table 1: Representativeness of the sample

Type of accommodation	Estimation of the target population*	Source	Number of effectively enrolled people
Squats or slums (Ethos 8)	619 to 817	NGO (MDM - Doctors of the World) – April 2020	363
Emergency shelters (Ethos 2)	795	Official administrative data (SIAO) – April 2020	358
Collective transitional shelters (Ethos 3)	634	Official administrative data (SIAO) – April 2020	196
Hostels mobilized during the covid crisis (Ethos 2)	893	Official administrative data (Service+) – July 2020	197
Street (Ethos 1)	455	First census of homeless people living in streets – January 2022 - No usable source of data in 2020	98

* *Adult homeless population living in the city of Marseille*. MDM (Médecins du Monde): Doctors of the World NGO; SIAO (Services intégrés de l'accueil et de l'orientation): Integrated reception and guidance services

2) Please indicate the number of participants that were resampled between the first and second survey. How did they insure that same individuals was not sampled multiple times. If there was resampling, were there instances where people with previous positive serology had subsequent negative serology.

A total of 1241 participants were included and 721 could be followed up and had a new serology 3 months later. In order to clarify this point we added in methodology following sentences "Each participant receives individualized follow-up and repeated testing at the inclusion and 3 month later. There was no resampling." Each included participant received a unique anonymous identification number linked to his or her identity verified by an ID card. Same participants were follow up in a cohort. We have tried to be as exhaustive as possible (Supl Table 1).

Thank you for your suggestion, for more clarity about false negative results, we have added in the results part the following sentences.

« Of the 74 participants with positive serology at the start of the study, 43 were able to be followed up and have a new serology 3 months later. A total of 69.8% (n=30) still had positive serology. Thus, in 30.2% of cases (n=13) there was a rapid negativation of serology. »

3) They state that a total of 180 had positive serology. There was 74 and 136 positive in the first and second sampling. These numbers are incongruent unless there was resampling.

In total of the 1241 participants were included in the cohort and followed at different time period. A total of 721 participants have both serological resultats ta inclusion and month 3 follow-up. We report here the results for all participants with a positive serology of SARS CoV2. In order to clarify this point we added in the result part « At inclusion (n=1241) or as part of the cohort follow-up (n=721) ». A total of 180 had positive serology at any point of the study.

4) Table 2 states that total participants was 1241. Thus, this only incorporates data from the first sampling?
Data on Table 2 incorporate data at inclusion (n=74 positive serology on a total of 1241 participants) or as part of the cohort follow-up at 3 month (n=136 positive serology on a total 721 participants). We specified this in the table legend as « Februaryand December 2020". This includes 30 prevalent cases and 106 incident cases.

5) I am not sure a Kaplan Meier analysis is appropriate. Using serology as a marker for infection acquisition is not appropriate because serology indicates current or past infection.
We understand your point of view. However, in this case we were at the beginning of the emergence of COVID in Marseille and the serologies of our participants who positive between inclusion in the cohort and 3 month later were representative of a recent and new contact with the virus.

6) It is unclear if the participants were asked about symptoms at the time of sampling or prior. Symptoms at time of sampling would be inappropriate because serology could reflect past infection. Furthermore, past symptoms would be liable to recall bias.
Table 3 reported symptoms declared by participants according their serological status before the last 3 months of the sampling. There was obviously a memory bias. To clarify this point we added « reported in the last 3 months prior the serological test" in the legend of the table 3".

7) Figure 1 should be removed. It contains data that is not part of this study.
Thank you for your suggestion. The figure was removed. For your information the purpose of this figure was to contextualize the dates of sampling and explain the local profile of the epidemic.

8) Investigators don't explain why physical distancing and inability to use hygiene materials should be associated with greater infection incidence.
These two criteria are significant in bivariate but not in multivariate. In fact, those who declare to have problems accessing hygiene products are mostly participants living in squats, on the street or in slums and they have lower prevalence SARS-CoV-2 infection than those who benefit from hygiene products but who live in emergency accommodation. We can hypothesize a confounding factor, with the association of the use of hygiene materials and greater infection incidence being due to the link between the type of precarious housing and access to hygiene products.
We did not also keep "physical distancing" in the final model because it was highly correlated to ETHOS.

9) There are multiple studies that have shown that lack of stable housing engenders greater risk for SARS-CoV-2 infection. Authors have highlighted some of the previous literature. I would include, Bean et. al., Clinical Infectious Diseases
Thank for your suggestion, this interesting article was not published at the time we submitted. We added the reference in the introduction.

Reviewer: 2
Ms. Judith Racape, ULB École de Santé Publique
Comments to the Author:
The manuscript entitled "Cumulaive incidence of SARS-CoV-2 infection within the homeless population: insights from a citywide longitudinal study" is an interesting study to determine risk factors associated with SARS-CoV-2 in a cohort of homeless people. However, the analysis and the methodology of the paper need to be reviewed and clarified.

Methods

1. The major point concerns the methodology and the use of survival analysis for your data. For longitudinal studies as survival analysis, you need a follow-up of the individual people during the time of your study. It is not clear if you have constructed your study for this methodology? You need to clarify this point.

- How the people have been tested? They have been followed at different time period?
- A person tested positive at a time X can have been positive before. Or a person tested negative at a time Y can have been positive later. How did you manage those points during the follow-up?
- You mentioned two testing sessions 1/ June 5 to august 5 (n=1241) and 2/ September 11 to December 18 (n=721) . Your analyses should concern the people tested for both periods (n=721)? Why did you present the population characteristics for all the cohort (table 1) ?

Our study is a cohort study with an individualized follow-up of the participants which allows a survival analysis. We have clarified this point in the manuscript in the methods section and the study design as follows: « Each participant receives individualized follow-up and repeated testing at the inclusion and 3 months later".

All individuals received a serological test at inclusion and 3 months later. Given the epidemiological situation (arrival of the first cases of infection in March 2020 in Marseille), we considered that they were previously seronegative. This point is described in data analysis section of methodology. All participants were monitored at the same time.

A person tested serologically positive was considered positive on the date of the first positive test. His or her status was considered positive regardless of the results of subsequent serological tests. We added « A survival analysis was carried out to address the spread of COVID 19 among the targeted population. The time (in months) was defined as follows: the starting date was the date of February 01 2020, date at which none positive cases were registered in Marseille, that is a when all participants could be considered to have a COVID19-negative status. The event was a positive SARS-CoV-2 status, whatever a positive SARS-CoV-2 PCR or a positive serological test informed the diagnosis. His or her status was considered positive regardless of the results of subsequent serological tests. For those with a PCR test achieved, the date of the event corresponded to the PCR date, corrected with the date of the first symptoms when reported. For positive participant with a rapid serological test, the reported date of the first symptoms was considered. For participant with a positive serological test but with no history of symptoms, we considered the date of the testing strategy performed by the research team. No additional corrections were made in absence of any informative data. Participants tested negative at the first testing wave but being lost to follow up at the second testing wave were censored at the date of the last collection data. The cut-off date was December 18 2020, precisely 11.2 months after the starting date. » in the methodology section in order to clarify this point.

We have presented in the table the results for all included participants (n=1241) because some were already positive at inclusion in the first round of testing. We presented in table 1 the socio-demographic presentation of all participants of the study.

2. You mentioned that "We fitted a multivariate Cox model by considering as eligible variables those that were significant in a univariate analysis at the 10% level …."

- Why did you not include "age" (p=0.05) and "hand washing" (p=0.06) in your model? (linked to preventing measures as mentionned in your discussion)

There are many variables in this study. We have therefore kept only the significant ones in univariate at the 5% threshold.

- Did you check PH assumption of your model?

We tested the assumption of proportional hazards using Schoenfeld residuals. We clarified this point in methodology section.

3. Why did you not separate the category Africa in north Africa and sub-Saharan Africa for country

of birth? The risks factors are different in the literature and the comparison in your discussion (p9 l54) of “black ethnic group” must be nuanced.
Thanks for your suggestion. Unfortunately the question was asked like this but and includes the 2 groups: migrants from North Africa and sub-Saharan Africa. We are agree with you, but even if we include migrants from North Africa and sub-Saharan Africa, the risk is significantly higher in our study. According to your recommendation we have nuanced this point in our discussion. We added in the discussion « This is despite the fact that North and Sub-Saharan Africans are grouped together for analysis in our study » in order to clarify this point.

Results:

- Table1: as mentioned before, the population characteristics must be presented for the population analyzed in both periods.
Following your remarks, we hope to have clarified our approach to longitudinal follow-up of a cohort. In the context of a cohort, it is the characteristics of the baseline cohort that we feel are relevant to describe here.

- Table 2: Why did you not present the results concerning ETHOS accommodations?
We chose to keep the notion of time spent in an emergency shelter. This variable was highly correlated with the ETHOS variable, so we had to choose one or the other.

- p15-17: I am confused with the presentation of the graphs which are not refered into the text?
Figure 1 has been removed (taking up the argument of the 1st reviewer). There are 3 figures left. All 3 are now correctly numbered and referenced in the text.

- p22-51: What is the link between the report (mortality COVID-19) attached and the manuscript?
If the seroprevalence is higher in our study population, it seems important to look at whether this is related to an increase in mortality. This is that we discuss in the discussion section.

Discussion:

- P9 l35 : The comparison with the general population is tricky because of the heterogeneity of the spreading in different regions/cities in France (see introduction p4 l17) and the different testing modalities during the pandemic.
Thank you for raising this point, which has engaged our attention. The comparison with the general population was made on the same region, as described in the methodology section: « Seroprevalence data of our study were compared with data from a representative sample of the general population living in Marseille, which were derived from a national seroprevalence survey (EpiCov) [19]. Results of serprevalence in the general population were obtained from home self-samples of dried blood spots, in order to detect IgG antibodies (Euroimmun ELISA-S) [19].”
In order to remove the confusion, we propose to use the term “EpiCoV-Marseille” instead of “EpiCov”, which implies that we use the whole national database.

- P9 l43 : The adjustment with age should be useful ?
The objective of the study was to know what was the impact of the epidemic in the precarious population compared to the general population. The aim of the study was not to compare specifically the 2 populations but rather what were the risk factors associated to a possible over risk in precarious population in order to adapt public health prevention strategies.
If you suggest adjusting our cohort data with the EpiCoV data, unfortunately the adjustment is impossible because the data have been processed in EpiCoV and we do not have access to the raw data.

- P9 l 50 : A discussion on the characteristics of the shelter should be adequate

Thank for your suggestion we added the following sentence in order to clarify this point « The emergency shelters are short term shelters that can accommodate several hundred people in Marseille. Collective transitional shelters are longer term, smaller facilities offering more consistent social work. »

- How have been managed the homeless issue during COVID-19 pandemic? What were the public health measures: concerning the isolation of homeless people ? hostel have been request ?

The first French lockdown was ordered on March 17, 2020, and it mobilised emergency social action to ensure that a maximum of homeless people were given shelter: hotels and holiday resorts were used in addition to existing shelters that were already full.
We clarify this point in the discussion « The first French lockdown was ordered on March 17, 2020, as emergency shelters are already full in normal times, hotels were required. In these hotels, people did not have kitchen facilities and often found themselves in high-density grouping areas, especially at meal times or in the few outdoor spaces available. »

References missing:

- P12 l34/35
- P13 l45/46

We are sorry we don't see what missing references you are referring to.

General comments

The main issue for this paper is the methodology: is the survival analysis appropriate? If you want to compare both testing session, the analysis is different. The authors should clarify those points and present the results in line of those analysis.
It seems to us that with the different follow-up points a survival analysis is the appropriate method. We have made some clarifications in this document and manuscript and remain at the disposal of the reviewers and the editor if some points remain unclear.

I would also suggest a discussion of the results with those from a previous study of the authors (same cohort apparently):
Loubiere S, Monfardini E, Allaria C, Mosnier M, Allibert A, et al. (2021) Seroprevalence of SARS-CoV-2 antibodies among homeless people living rough, in shelters and squats: A large population-based study in France. PLOS ONE 16(9): e0255498. https://doi.org/10.1371/journal.pone.0255498
Thank for your suggestion we added the following sentence in Strengths and limitation « and confirm the result of the first cross-sectional study in the same population (Loubiere S and al).”

Thank you for your consideration.

Yours faithfully,

Dr Emilie Mosnier on behalf of all the authors

VERSION 2 – REVIEW

REVIEWER	Sagar , Manish Boston University School of Medicine, Medicine
REVIEW RETURNED	26-Oct-2022

GENERAL COMMENTS	In this revised manuscript, the authors have responded well to the previous critiques. They have done a better job at explaining the methodology, which clarifies the subsequent results. I am not able to adequately review the manuscript because of the reasons stipulated below. Furthermore, the authors could clarify some points that may strengthen the manuscript. 1) Please relate how the data was handled when a demographic variable changed between 1st the 2nd sampling. For instance, how was the individual characterized if housing status changed between 1st and 2nd sampling. This would affect the Kaplan Meier analysis. This stipulation also applies to the examination of factors associated with SARS-CoV-2 positivity. 2) In addition, the survival analysis does not account for demographic differences among the people with different form of housing. 3) The protocol is entirely in French. An English version should be provided. 4) There are no attached figures in the review material.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Dr. Manish Sagar , Boston University School of Medicine
Comments to the Author:
In this revised manuscript, the authors have responded well to the previous critiques. They have done a better job at explaining the methodology, which clarifies the subsequent results. I am not able to adequately review the manuscript because of the reasons stipulated below. Furthermore, the authors could clarify some points that may strengthen the manuscript.

1) Please relate how the data was handled when a demographic variable changed between 1st the 2nd sampling. For instance, how was the individual characterized if housing status changed between 1st and 2nd sampling. This would affect the Kaplan Meier analysis. This stipulation also applies to the examination of factors associated with SARS-CoV-2 positivity.

A total of 30% of the participants in the cohort changed their place of residence during follow-up. For these participants, the data considered in the analyses relate to the target period. That is data collected at the time of the positive test for SARS CoV2 participants and for negative patients throughout the follow-up, at the time of the last test. Concerning the socio-demographic variables, such as the type of work or resources, as indicated, the values at the time of the positive test or the last test were taken into account. We also chose to take into account a variable that defined the time spent in a type of accommodation (i.e. emergency accommodation) to take into account a possible change that occurred very recently.

In order to clarify this point we have added in the methods part in the following sentences.

“In order to take into account in the analysis, in regards to the mobility of the participants, in terms of the place of residence and the possibility of changes in socio-demographic characteristics, we took the data at the time of the positive test for SARS CoV2 participants and at the time of the last test for negative patients throughout the follow-up”

In order to clarify this point we have added in the results part in the following sentences.

“A total of 30% of the participants in the cohort changed their place of residence during follow-up. »

2) In addition, the survival analysis does not account for demographic differences among the people with different form of housing.

Concerning the adjustment, the variables such as Age, sex were not retained in the analysis. Indeed we had advanced in the method that in the multivariate model only the variables significant at the threshold of 0.1 would be kept in analyses. However, we propose in the complementary analysis seen in the appendix (see supplement). Model remains unchanged even by adjusting on age and sex.

In order to clarify this point we have shown this in the supplement file below.

Supplement file: Cox multivariable logistic regression analysis adjusted by age and sex

of risk factors of SARS-CoV-2 seroprevalence in homeless people in Marseille.

Having_Comorb_PsyAddic: Psychiatric or addictive comorbidities

Variable		N	Hazard ratio	p
Sex		1048	1.12 (0.75, 1.68)	0.576
Age		1048	1.21 (0.69, 2.13)	0.503
Household_status	Isolated adults	589	Reference	
	Isolated parent	121	1.62 (1.05, 2.49)	0.028
	Family	338	0.75 (0.48, 1.17)	0.207
Percent_time_shelter	<=33	764	Reference	
	>=66	145	1.90 (1.13, 3.18)	0.015
	33 to 66	139	1.73 (1.11, 2.68)	0.015
Difficultyaccess_hygiene	No,never,rarely	685	Reference	
	Yes,always,often	363	0.74 (0.50, 1.12)	0.154
Have_FinancialResource	No	459	Reference	
	Yes	589	0.71 (0.51, 0.98)	0.037
Tobacco	Non smoker	452	Reference	
	Smoker	596	0.47 (0.32, 0.67)	<0.001
Have_Comorb_PsyAddic	No	784	Reference	
	Yes	264	0.52 (0.32, 0.84)	0.008
Number_Daily_contact	<=5	634	Reference	
	>15	95	1.47 (0.70, 3.09)	0.311
	5-15	319	1.86 (1.28, 2.70)	0.001

3) The protocol is entirely in French. An English version should be provided.

We are very sorry but we do not have the budget for the translation of the protocol. However, we would like to inform you that we already have 3 papers published in this study that informs about the study protocol.

1 "Locked down outside": Perception of hazard and health resources in COVID-19 epidemic context among homeless people.

Allaria C, Loubière S, Mosnier E, Monfardini E, Auquier P, Tinland A. SSM Popul Health. 2021 Sep;15:100829. doi: 10.1016/j.ssmph.2021.100829. Epub 2021 May 28. PMID: 34079855

2 Seroprevalence of SARS-CoV-2 antibodies among homeless people living rough, in shelters and squats: A large population-based study in France.

Loubiere S, Monfardini E, Allaria C, Mosnier M, Allibert A, Ninove L, Bosetti T, Farnarier C, Hamouda I, Auquier P, Mosnier E, Tinland A. PLoS One. 2021 Sep 15;16(9):e0255498. doi: 10.1371/journal.pone.0255498. eCollection 2021. PMID: 34525096 Free PMC article. Clinical Trial.

3 Residential Mobility of a Cohort of Homeless People in Times of Crisis: COVID-19 Pandemic in a European Metropolis.

Allibert A, Tinland A, Landier J, Loubière S, Gaudart J, Mosnier M, Farnarier C, Auquier P, Mosnier E. Int J Environ Res Public Health. 2022 Mar 7;19(5):3129. doi: 10.3390/ijerph19053129. PMID: 35270823

4) There are no attached figures in the review material.

We will make sure that the figures are with the revised manuscript

Thank you for your consideration.

Yours faithfully,

Dr Emilie Mosnier on behalf of all the authors

VERSION 3 – REVIEW

REVIEWER	Sagar , Manish Boston University School of Medicine, Medicine
REVIEW RETURNED	29-Dec-2022

GENERAL COMMENTS	Authors have responded to previous comments. I have no further comments