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)	Lifetime, Five-Year and Past-Year Prevalence of Homelessness in Europe: a cross-national survey in eight European nations
AUTHORS	Loubiere, Sandrine; Taylor, Owen; Tinland, Aurelie; Vargas-Moniz, Maria; Spinnewijn, Freek; Manning, Rachel; Gaboardi, Marta; Wolf, Judith; Bokszczanin, Ana; Bernad, Roberto; Kallmen, Hakan; Toro, Paul; Ornelas, Jose; AUQUIER, Pascal

VERSION 1 – REVIEW

REVIEWER	Matthew Morton Chapin Hall at the University of Chicago
REVIEW RETURNED	12-Sep-2019

GENERAL COMMENTS	This paper makes a very important contribution to the literature and policy discussions on homelessness by using a representative survey approach to estimate homelessness prevalence and analyze its correlates across multiple European countries. Overall, I think the paper and underlying methodology are strong and key limitations are transparently reported. Points that could, in my view, be better addressed or reported include the following: METHODS – Please include the actual measures, especially for homelessness. How exactly were the homelessness questions asked? The paper says that "literal homelessness" was defined as "sleeping in the street, in a car, or living in an emergency or temporary shelter." Is this exactly how the question was asked? If so, it seems like it might omit forms of literal homelessness, such as sleeping in an abandoned building, 24-hour laundry or eating facility, outside/in parks, or bus/train station, which the respondent might not read into "sleeping in the street." This could also bias estimates toward urban homelessness experiences. METHODS – Please explain why non-"literal" homelessness was excluded – e.g., couch surfing/hopping or doubling up due to a lack of safe and stable place to stay. It makes sense to disaggregate literal vs. other forms of homelessness in the analysis, but I don't understand the rationale for excluding other forms of homelessness altogether. METHODS – Were rural, semi-urban, urban self-defined or based on using geocodes to determine standardized classifications? METHODS – Given the stark disproportionalities reported elsewhere in homelessness by race, ethnicity, sexual orientation, and gender identity, the authors should explain why such demographic information was not collected and analyzed in this study. METHODS – Response rates are described as low in the 30-40% range, but these actually appear to be quite high compared to typical response rates for phone-based surveys, at least in the

	U.S., which continue to decline: https://www.pewresearch.org/methods/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/. It would be helpful for the authors to explain why the expected the response rates that they did. Are these trends different in Europe than for the U.S.? METHODS – Please explain how weights were constructed. To what extend did the timing of data used to create weights align with timing of the survey (2017)? METHODS – It would seem more appropriate to use one-year prevalence, vs. lifetime prevalence, as the dependent variable for the GLM analysis, especially for age as the independent variable. Lifetime prevalence would inherently involve bias toward older age due to cumulative experience. RESULTS – (Just a note that the tables were cut off in my PDF so I couldn't review all information.) RESULTS – What percentage of respondents answered landline vs mobile phones? Were either of these deliberately oversampled (METHODS)? RESULTS – Table IV – Looks like this presents compositions of those reporting homelessness. It may be more useful to give homelessness prevalence rates for each subpopulation (also or instead of). DISCUSSION – Important points on discrepancies between official estimates and the study's estimates of homelessness. This could also be a good opportunity to explain why annual/lifetime prevalence is different and important vs point-in-time estimates. These discrepancies in scale are related to at least two issues: (1) sampling (i.e., representative phone-based survey vs street/shelter counts/outreach) and (2) timeframe (annual/lifetime vs point-in-time). DISCUSSION – Between-country variation is highlighted, but not the variations in policy contexts (e.g., social safety nets) that might help explain these variations. In some ways, the between-country variations offer potential hope that policies matter. While the study wasn't designed to assess the impact of those policy differences, it would be useful to discuss the potential for how policy differences might help explain homelessness prevalence variations and warrant further research. DISCUSSION – The discussion of between-country variations in duration of homelessness experiences might be a little misleading. Yes, Spain has lower share of *its homelessness population* experiencing longer periods of homelessness, but it also had the highest homelessness population. So it could still have a greater share of *its overall population* experiencing longer periods of homelessness than the other countries. I think it'd be more useful to present, compare, and discuss the share of countries' overall populations that experience different durations of homelessness.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer:

METHODS – Please include the actual measures, especially for homelessness.
Authors:

We assume that the measures referred to are those we set out to asses in the study – although we did mention these in the section on 'measures', we note that they are not quite foregrounded as much as we would like, and we have adjusted the beginning of this section accordingly.
Text amended as follows (lines 156-7):
The measures for this study were the prevalence rates for lifetime, five-year and past-year homelessness; total time spent homeless; and respondents socio-demographic data.
METHODS – How exactly were the homelessness questions asked? The paper says that "literal homelessness" was defined as "sleeping in the street, in a car, or living in an emergency or temporary shelter." Is this exactly how the question was asked? If so, it seems like it might omit forms of literal homelessness, such as sleeping in an abandoned building, 24-hour laundry or eating facility, outside/in parks, or bus/train station, which the respondent might not read into "sleeping in the street." This could also bias estimates toward urban homelessness experiences.
Authors:
We agree with the reviewer that the exact phrasing of our definition is somewhat restrictive, and that this does give rise to the concern, raised by the reviewer, that respondents might feel that certain forms of homelessness fall outside of this definition. Our reasoning behind the particular phrasing was an aim to capture ETHOS 1&2, with a focus on the most visible forms of homelessness. The phrase 'in the street' was intended to be interpreted broadly, but the questionnaire was also delivered in 8 different languages – the English language version only being delivered in Ireland. In each case it is possible that certain experiences of homelessness might appear to be excluded. However, the interviewers were trained in the general type of homelessness we were aiming to cover, and given the context of a telephone discussion we can also hope that a respondent who wasn't sure if their experience qualified would mention it to the interviewer who, if it was literal homelessness in a rural environment, or in the reviewers examples, that the interviewer would explain that yes, this constituted homelessness for our study. That said, we do appreciate the possible limitations raised by the reviewer.
With regards to our definition of homelessness and the way the question was asked, we have clarified this in the text with the following amendment (lines 158-67):
Firstly, respondents were informed that, in this study, we defined homelessness as 'sleeping in the street, in a car, or living in an emergency or temporary shelter', encompassing ETHOS 1 and 2 from the European Typology of Homelessness and Housing Exclusion, namely the most visible forms of literal homelessness. The respondents were then posed the open question, 'have you ever been homeless?'
And to the Limitations (lines 473-477):
It is also possible that the phrasing of our definition of homelessness referring to sleeping in the street or in a car will create a bias towards urban forms of homelessness. However, we hope this will be mitigated by the open discussion of the telephone interview with the possibility for respondents to clarify any doubts with the interviewer.

METHODS – Please explain why non-"literal" homelessness was excluded – e.g., couch surfing/hopping or doubling up due to a lack of safe and stable place to stay. It makes sense to disaggregate literal vs. other forms of homelessness in the analysis, but I don't understand the rationale for excluding other forms of homelessness altogether.
Authors:
Whilst we sympathise with the reviewer in that non-literal forms of homelessness are of deep concern and are worth measuring using the methods we have employed, the prevalence questions formed part of a broader survey including opinions about homelessness. It was necessary to make the interview as concise as possible, and therefore we decided to restrict our focus to literal homelessness. Hopefully further studies in this vein will produce more data for non-literal homelessness.
The following text has been added to the limitations section to clarify our approach (lines 470-473):

In addition, as part of a broader project examining long-term homelessness, the prevalence questions in our survey focused on literal homelessness, as it was necessary to keep the interview as concise as possible. This does, however, restrict the scope of this study to the most visible forms of homelessness.

METHODS – Were rural, semi-urban, urban self-defined or based on using geocodes to determine standardized classifications?
Authors:
Respondents were asked whether they lived in an urban, semi-urban, or rural environment, therefore these categories were defined according to the respondent's own conception of where they lived. We agree with the reviewer that this does lack standardisation, but it has the upside of capturing an individual's own perception or experience.
We have added the following qualification to the text to ensure clarity for the reader (lines 179-180):
Data on socio-demographic characteristics of respondents were collected, namely: gender, age, educational level, professional status, annual household income, marital status and living area as self-reported by respondents (with the question 'would you say that you live in' followed by the options of urban, semi-urban or rural).
And in the limitations (lines 498-501):
In addition, Living area was limited by being defined by the respondents own perception of where they lived, and therefore these data are not standardized. For legal reasons it was not possible for our study to use geocodes or other methods of locating respondents.

METHODS – Given the stark disproportionalities reported elsewhere in homelessness by race, ethnicity, sexual orientation, and gender identity, the authors should explain why such demographic information was not collected and analysed in this study.
Authors:
The reviewer raises an important issue, which we very much appreciate. The primary reason for not collecting this information is the fact that the survey was conducted in eight different countries with different ethical restrictions on data collection which restricted the uniform collection of this kind of data. For example, in France permission to collect this kind of data will only be given if the information is clearly necessary to address the primary outcomes of the study. Because our survey was assessing a variety of elements alongside the prevalence of homelessness, this was not the case. Furthermore, our sample size of just 700 per country limits the capacity of this data set to shed light on the reasons for respondents' experiences of homelessness, particularly in relation to potential discrimination.
Nevertheless, this is an important limitation of the study, and we have added the following text to the limitations section (lines 506-511):
.. data on ethnicity, sexual orientation and gender identity were not investigated due to legal restrictions on data collection in some of the participating countries limiting our ability to investigate these well-explored correlates of homelessness.

METHODS – Response rates are described as low in the 30-40% range, but these actually appear to be quite high compared to typical response rates for phone-based surveys, at least in the U.S., which continue to decline: https://www.pewresearch.org/methods/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/. It would be helpful for the authors to explain why the expected the response rates that they did. Are these trends different in Europe than for the U.S.?
Authors:
We did not intend to describe the response rates as low, but rather in a range that we expected based upon prior experience in France and the assumption that the topic of homelessness might attract lower response rates than some of the Eurobarometer studies (Health Barometer 2016: 52% landline and 48% mobile phone response rates). In fact, some recent telephone surveys in France have achieved high response rates (Richard, Méthode d'enquête du Baromètre santé, 2014: 57% and

Santé Publique France, Baromètre Santé 2017: 48.5%) but it does seem the there is a downward trend here, although with much higher response rates than those described for the US in the reference you kindly linked. That said, we have clarified the text to explain why we expected the range we did, and rephrased those sections to avoid the impression that we considered our response rate to be low:
Text amended as follows:
At line 147-8:
Based on European telephone survey response rates and our survey topic of homelessness, we expected response rates…
Lines 220-1:
Response rates were slightly better than expected, and ranged from 30.4% to 33.5%, resulting in a total number of 5,631 respondents producing 5,295 valid questionnaires.

METHODS – Please explain how weights were constructed.
Authors: Following this advice, we have added the explanation of how weights were calculated to the Methods section:
Text amended (lines 190-196):
The 2017 census data on distribution of sex, education and age variables were obtained through the World Bank and Eurostat. Our weight variable was based on the known distribution of the entire population according to age, sex, and education. In our case, the weight variable was post-stratified to match the entire population in each country to ensure that calculated estimates were truly representative of the surveyed European populations. Each participant in the survey dataset was therefore assigned a unique weight.
METHODS – To what extent did the timing of data used to create weights align with timing of the survey (2017)?
Authors: As with the previous question, we have added this information to the main text in the Methods section:
Lines 204-209:
Between September 2017 and December 2017, a control of the quality of the database was undertaken and percentages of missing data for each variable were calculated. Full descriptive analyses were conducted of all variables from the database; specifically the distribution of gender, age and education was assessed in each national sample to address representativeness. A weighted sample, as described above, was built in January 2018. Missing data for the overall study was less than 0.5% and missing variables were not included in the analysis.

METHODS – It would seem more appropriate to use one-year prevalence, vs. lifetime prevalence, as the dependent variable for the GLM analysis, especially for age as the independent variable. Lifetime prevalence would inherently involve bias toward older age due to cumulative experience.
Authors:
Although we agree with the reviewer, unfortunately the sample size for past-year prevalence was too low to conduct effective statistical analysis – out of the total sample of 5,261, overall there were 32 respondents reporting past-year homelessness, with only 5 people in Sweden and 2 in Ireland. As a result, we used lifetime and five-year prevalence for the statistical analysis. Note that we do mention this problem when we introduce the GLM models, at lines 294-5.
In order to add clarity, we have added the following line to the limitations section (lines 503-505):
… our sample size prevented effective GLM analysis of past-year prevalence. For certain socio-demographic data, particularly age, this may induce bias (it is likely that cumulative experience will bias lifetime prevalence toward older age groups).

RESULTS – (Just a note that the tables were cut off in my PDF so I couldn't review all information.)
RESULTS – What percentage of respondents answered landline vs mobile phones? Were either of these deliberately oversampled (METHODS)?

Authors:
The reviewer raises another good point. Unfortunately these data were not collected by the telephone survey professionals hired to conduct the survey. One of the reasons for this was that many of the interviews were conducted in two or more stages, or with the respondent requesting a call back. In these instances the different calls often took place on different numbers, at the request of the respondent. As a result, although these data existed initially, they swiftly became too intermingled to be able to collect it usefully.
We have added the following text to the limitations section (464-467):
Unfortunately in this study we were not able to collect relative percentages of mobile phone or landline responses, in part due to the repeated call back procedure and respondents' requests to receive a call on a different line leading to mixed data. We are therefore not able to report the relative proportions of responses by landline and mobile phone.

RESULTS – Table IV – Looks like this presents compositions of those reporting homelessness. It may be more useful to give homelessness prevalence rates for each subpopulation (also or instead of).
Authors:
We agree with the referee, and we have altered the table to give prevalence rates for each subpopulation.

DISCUSSION – Important points on discrepancies between official estimates and the study's estimates of homelessness. This could also be a good opportunity to explain why annual/lifetime prevalence is different and important vs point-in-time estimates. These discrepancies in scale are related to at least two issues: (1) sampling (i.e., representative phone-based survey vs street/shelter counts/outreach) and (2) timeframe (annual/lifetime vs point-in-time).
Authors:
We are very much in agreement with the reviewer on this point. Although we do discuss this briefly in the introduction (from line 107), in the interests of space we did not develop much depth in this paper. We have amended the discussion section as follows (lines 342-351):
These results highlight the importance of prevalence studies in comparison to existing methods (point-in-time and service-user estimates) in order to more effectively assess the extent of homelessness in Europe. The large discrepancies between our prevalence data and official figures relate to at least two issues – firstly the sampling process itself, using representative phone-based surveys rather than street or shelter counts and other outreach methods. The latter methods are well-known to underestimate the number of people experiencing homeless as they will inevitably miss a proportion of the population. Secondly, prevalence counts assess the number of people experiencing homelessness over time rather than in a single night (or week or month counts), and therefore provide a much better estimate of the transient portion of people experiencing homelessness.

DISCUSSION – Between-country variation is highlighted, but not the variations in policy contexts (e.g., social safety nets) that might help explain these variations. In some ways, the between-country variations offer potential hope that policies matter. While the study wasn't designed to assess the impact of those policy differences, it would be useful to discuss the potential for how policy differences might help explain homelessness prevalence variations and warrant further research.
Authors:
Again we agree with the reviewer's insight on this point. As the reviewer notes, we do address this very briefly in the discussion, at lines 341-344, but only with a reference to national policy. The challenges that confront a cross-national analysis of policy contexts that influence pathways into homelessness are formidable. Policy contexts are complicated and require historical context, extensive study of available legislation and committee meetings, and in most cases will involve a mix of national and local governmental interventions. Internally our team has conducted some research just on the French context, and effective treatment of this topic would make an extensive paper just on its own. This topic is being led by one of the HOME_EU consortium partners, with a cross-national

policy study looking at the period 2012-2017, and the results of their investigation are still pending. Even if we had access to this material, we feel that it would only really provide context for past-year (and to a lesser extent to five-year) prevalence rates in our study. For lifetime prevalence we would require in-depth, long-term (more in the region of twenty years) policy studies across all eight countries, and then to subsequently make a comparative study of these data to get a sufficiently detailed picture. For these reasons we have had to keep our observations limited on this front. We propose the following additional text in the manuscript to draw attention to this:
AMENDED TEXT: at 344
It is certain that the different policy contexts between the different countries in this study will have an effect on the prevalence rates, however to date the variations and definitions in approach between European states put such analyses beyond the scope of this study. Research in this area is ongoing.

DISCUSSION – The discussion of between-country variations in duration of homelessness experiences might be a little misleading. Yes, Spain has lower share of *its homelessness population* experiencing longer periods of homelessness, but it also had the highest homelessness population. So it could still have a greater share of *its overall population* experiencing longer periods of homelessness than the other countries. I think it'd be more useful to present, compare, and discuss the share of countries' overall populations that experience different durations of homelessness.
Authors:
This is a good point. Initially we opted to discuss the proportions of the population experiencing homelessness as we felt this profile, in general, provided interesting comparisons between countries, irrespective of the actual numbers and their relationship to national populations. In particular, we felt this might provide some insight to national governments concerned with their individual homeless population profiles. That said, the reviewer makes an astute observation. We have amended this discussion section to take this point into account.
Lines 377-384:
The rate of longer experiences of homelessness reported in Poland, Italy, the Netherlands, and Sweden is particularly striking. However, if we account for lifetime prevalence rates, as proportions of their populations (Eurostat, 2017) exposed to experiences of homelessness lasting more than a year Sweden, Italy and Spain have the highest rates with approximately 1.9%, 1.6% and 1.2% respectively, with the Netherlands around 1% and Poland at 0.6%. For Italy this is particularly concerning, as with the largest population among these countries, 1.6% amounts to around 960,000 people having experienced homelessness lasting more than a year (with approximately 560,000 in Spain and 190,000 in Sweden)…
Lines 391-395:
It is notable that these results appear to cut across certain aspects of homelessness policy, as Poland and Italy have only recently begun to formulate national approaches to homelessness, Spain introduced a national strategy in 2015, and Sweden and the Netherlands have had national strategies in place since 2007 and 2006 respectively.

VERSION 2 – REVIEW

REVIEWER	Matthew Morton Chapin Hall at the University of Chicago, USA
REVIEW RETURNED	11-Nov-2019

GENERAL COMMENTS	All of my comments were adequately addressed. This manuscript will make an important contribution the homelessness prevention literature.