Reviewer’s report

Title: Counting the non-existing: causes of death of undocumented migrants in Brussels-Capital Region (Belgium), 2005 -2010

Version: 0 Date: 10 May 2019

Reviewer: Jean Tafforeau

Reviewer’s report:

The article is interesting particularly because it is dealing with statistics about hard to reach population. Results are well described. Discussion is a bit long. My main problems are with the methods (see below).

Background

In the last paragraph of the introduction, where you are explaining the aim of this publication you are in my opinion repeating 2-3 times the same thing so that at the end we are quite confuse about the real aim. Please be clear and concise.

Methods

That's the part of the article where have most difficulties.

In "how were the data gathered" …

1. the sentence "death certificates bearing no link to population register" is quite unclear to me : why should those certificates have a link with the register? How do you verify this? Why do you conclude that this means that they don't belong to official population?

2. for those who can be "linked" … how are you able to divided them in the 2 groups "foreign" and "Belgian"?

3. why did you limit your study to the period 2006-2010 (attention, further below, in the "study population" you say 2005-10!)

4. you speak in the next paragraph on the link between census and death certificates … is this different from the link already mentioned? How is this link performed? Why? By whom?

5. about the link with census, you mention "cross-validation" as one of the main reason … but which validation are you peaking about? For which information/variable and how?
5. still about the link with census, you mention that it is providing "extra information" … which information? And has it been used in the current study?

6. don't grasp why you decided to limit your study to Brussels; the fact that most of the undocumented migrants live there is not a reason, at the opposite: you could have increased your numbers by including all cases from the whole country!

In "study population" …

1. selection of undoc migrants … you remove those without national identification number but as far as I know the death certificates are anonymous!!

2. in the second step you are excluding the "tourists" … based on criterium??

3. 3nd step is based on WHO group A … why is it in function of child / adult mortality? what are the thresholds used? what is the rationale behind? migrants can come from rich countries with low mortality but for political reasons … and you are actually excluding them!

In "variables" …

1. why did you decide to work with the underlying cause of death and not other causes?

2. why age at death: don't you have birth date on the certificate?

3. cause of death is coded in ICD 10: how? by whom?
4. due to small numbers you aggregated causes to larger groups: are those groups equivalent to the "chapters" of the ICD10 classification?

5. for external causes you decided to keep smaller categories such as suicide, homicide, … was such a decision made a priori or after a first analysis: changing the study design "ad hoc" based on first results is indeed not very sound and may lead to abusive interpretations!

In "statistical analysis"

1. you are using logistic regression … my first remark when reading your text is that this is not the best choice as mortality is a rare event and poisson regression is best suited in that case. But it is only in the results part that we see that you are not estimating mortality rates (where poisson regression should be used) but rather (only) proportion of causes of specific causes of death compared to all causes. You should thus explain this first in this statistical analysis part!

2. age at death is included in your logistic regression model while this is rather a continuous variable … how did you deal with this in the model?

3. length of stay … be careful with this concept: in the health domain this term is normally used to describe the duration of stay when someone is admitted to the hospital, in number of days. explain more clearly what you are speaking about.
4. you have no information on the population of undoc migrants … meaning that you also don't have a denominator. A consequence of that is as you say that you cannot perform a survival analysis but in my opinion the first consequence is that you cannot estimate mortality rates!! … reason why you are limited in the analysis of proportion of specific causes of death!

In "results"

1. comparisons of the results with Chauvin's and Myria's publications should not figure in the presentation of the results but rather in the discussion section.

2. when presenting results with OR you are showing the confidence interval to verify their statistice significance … why not add p value as well?

In "strengths and limitations"

1. undoc migrant's population is unknown … reason why mortality rates could not be estimated

2. migrants could leave Belgium (and dye somewhere else) for reason that could be related to some specific diseases (and causes of death) which could introduce a bias. instance a migrant with a chronic disease may prefer to go back home (and will thus not dye in Belgium) while others (healthy) while die in Belgium from an external cause!

3. you decided to exclude migrants from Flanders and Wallonia, and by doing so you may have introduced a potential bias as we may think that migrants in Brussels could be different from those in the other regions. At least check in a sensitivity analysis that when including the whole country you get the same results.

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

No