Reviewer’s report

Title: Utilisation of rehabilitation services for non-migrant and migrant groups of higher working age in Germany – results of the lidA cohort study

Version: 0 Date: 08 Oct 2019

Reviewer: Joana Joana alves

Reviewer's report:

The article entitled "Utilization services for non-migrants and migrant groups of higher working age in Germany - results of the lidA-cohort study" aimed to analyze whether the use of rehabilitation services differed by migration status (non-migrants versus first- and second-generation migrants). The authors concluded that differences in use were not fully explained by the variables considered.

The findings are important to those authors with similar research interests.

However, I propose some changes to clarify the manuscript.

Abstract:
1 The background and the objectives of the study are not very clear.
2 In the methods section is not possible to distinguish the dependent from the independent variables.
3 The word "chance" should be replaced by "likelihood.
4 The conclusions are confusing and need some clarification.

Introduction:

5 Page 2, line 44: I could not understand the differences between DRV and DGUV. Why some people are entitled to one and not to the other? Are they both included in the lidA-cohort?
6 Page 3, line 67: these results were found in which countries?
7 Page 4, Line 87: differ in terms of what? The objectives need to become clearer.
Methods

8 I suggest a reformulation of the section study design and participants. First to describe the methods from l1dA; then to describe the criteria for inclusion and exclusion of participants. Now they are somewhat mixed.

9 Page 4, line 107: exclude the sentence "note the paragraph below about migration background". It is clear from the objective that this study wants to analyze differences according to MB.

10 It is not clear to me what happened to those participants who were not born in Germany, but had German citizenship and both parents were born in Germany. Imagine that they were born in other country because parents were spending a period abroad… They were included in G2 MB?

11 Covariates section: The reason for including those covariates should be explained. E.g.: why should we expect that physical and psychosocial work exposure to be associated with the use of rehabilitation services? Also, the occupational status is very important per se, and not only because it represents education. It reflects social influences, networks, working conditions, etc.

12 Page 5, line 142: why including variables (as physical and psychosocial work exposure) that explain poor health, since the variable for poor health is also included? Also, why did the authors choose to include two psychosocial variables?

13 Page 6, line 150: why the need for dichotomization of the variable for physical and psychosocial work exposure? The cut off values differ according to each variable: the authors chose tertiles for some variables, high vs low for other variables, and ¾ for other. The reasoning should be explained.

14 Page 6, line 153: The methods for the ERI questionnaire need to be clarified. These methods are taken from the ERI, or they were created by the authors?
15 Page 6, line 155: when the ERI ratio had up to two missing items, the missing observations were replaced by the individual mean. However, these missing values might not be missing at random… imagine that the participants with higher stress deliberately choose to omit some answers. It is not clear for me why using multiple imputation approaches for some variables and this particular method for the ERI variable.

16 Page 6, line 175: language spoken at home was included as a "non-work related factor". The authors justify this inclusion since lacking knowledge of German might be a barrier to access (page 16, line 371). But not using German as a first language at home does not mean that they do not understand the language… Since LidA questionnaire was applied in German (page 16, line 382), the participants had to have at least a minimum knowledge of the language. Furthermore, other studies use the language spoke at home as a definition for the migrant status. See for example: Bhopal (2004) "Glossary of terms relating to ethnicity and race: for reflection and debate" or Lorant (2016) "A social network analysis of substance use among immigrant adolescents in six European cities".

17 Page 7, line 182: "hierarchical logistic regression modelling" can be confused with "Hierarchical linear modelling", a regression-based analysis that takes the hierarchical structure of the data into account (e.g. students clustered within classrooms within school). Therefore, I suggest to call it only logistic regressions, mentioning that they were performed in a sequential way, adjusting block-wise for the group of variables.

Discussion:

18 Page 13, lines 298-301: it is not clear to me what the authors wanted to say.

19 Page 14, line 317: Can we observe a lower use of outpatient services among participants with MB because they do not have enough conditions at home to be treated in outpatient care?

20 Page 15, lines 346-354: this was not the focus of your study, and thus this is not relevant for the discussion. I suggest to comment only when the group of variables included decreased the coefficients for MB.
Page 16, line 371: Again, I have some doubts about the language mainly spoken at home, since we are talking about migrants.

Page 16, lines 377-379: this is not clear for a non-German reader.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

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

Not suitable for publication unless extensively edited

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