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) | The relation between occupation, gender dominance in the occupation and workplace and suicide in Sweden: a longitudinal study. |
|---------------------|-----------------------------------------------------------------------------------------------------------------|
| AUTHORS             | Matilla Santander, Nuria; Blazevska, Bianca; Carli, Vladimir; Hadlaczy, Gergő; Linnersjö, Anette; Bodin, Theo; Johansson, Gun |

VERSION 1 – REVIEW

| REVIEWER           | Assari, Shervin  |
|--------------------|------------------|
|                    | University of Michigan, Psychiatry |
| REVIEW RETURNED    | 04-Feb-2022      |
| GENERAL COMMENTS   | This retrospective cohort investigates how sex/gender, sex/gender representation in the job, male and female-dominated workplaces contribute to incidence of suicide. This is an epidemiological paper with odds ratio, and a DAG is provided to control for bias. Here are my concerns with the paper: |
|                    | 1- When the data has time and event, then why not to report Hazard ratio. The authors have mainly reported Odds ratio which ignores if the incident was in 2006 or 2010. |
|                    | 2- Why only 2006-2010 are included? |
|                    | 3- Missing data can be a real threat to this study. Sensitivity analysis without and with imputation should be better utilized and discussed. |
|                    | 4- In an ideal study, multi-level determinants could be investigated. People are nested to jobs, and those people have similar characteristics other than sex, education, etc. If you were thinking about this study as a multi-level, then you could test cross-level interactions like: |
|                    | a- males vs females in female-dominated workplaces and male-dominated workplaces. |
|                    | b- low vs high education in female-dominated workplaces and male-dominated. |
|                    | c- etc. |
|                    | 5- Other conditions not including stress and unknown role of employment income, wealth, and benefits |
|                    | 6- In the DAG, education could also impact past suicide, and past mental health problems. |
|                    | 7- Thus, overall, data may require re-analysis to make it more multi-level and comprehensive. |

| REVIEWER           | Mathieu, Sharna |
|--------------------|------------------|
|                    | Griffith University, Australian Institute for Suicide Research and Prevention |
| REVIEW RETURNED    | 08-Feb-2022      |
| GENERAL COMMENTS   | This observational study utilizes a population register in Sweden to
examine the association between occupation and suicide, and to determine the effect of gender composition at the occupational and workplace level on suicide, in a large sample of Swedish adults (N=3,318,727; 2006-2010). This manuscript has many strengths, most notably, adjusting for employment characteristics (e.g., precariousness). Several limitations are discussed.

Introduction
The references appear out of order, jumping in number from 9 to 13. Furthermore, reference 11 in the final reference list appears incomplete. I suggest the authors review their references and reference list for accuracy.

Materials and Methods
Could the authors please clarify in the study criteria what is meant by point (v) in the exclusion criteria. Currently it reads as ‘occupational code from more than 5 years ago’ from which point?

Could the authors please provide an example reference for the line ‘The reason for including undetermined intents is because several studies have concluded that some deaths classified as undetermined intent are in fact suicides.’ Currently there are none. It would be beneficial to readers if the authors could define precarious employment briefly (as measured by the SWE-ROPE).

Three possible categories of unemployment are described (0 months unemployed, 1-6 months unemployed, more than 6 months unemployed) yet in the results only one percentage is provided. Were these groups collapsed?

I am not clear on how the authors have distinguished workplace from occupation. Given this is seemingly an important distinction based upon the aims of the study I suggest the authors clarify this for readers. How is workplace defined or assigned, how is this different from the occupational codes assigned using the SSYK-96 classification system at the 2-digit level (and later the 4-digit level)?

Results
The remainder of the results section is succinct and clear.

Discussion
In their interpretation the authors do well to describe some of the prominent explanations for why certain occupations may be at higher risk of suicide (e.g., access to lethal means). However, given the aims of the study more interpretation around gender may be warranted. For example, there is some good discussion of social status and men working in attendant nursing professions, however, it is merely mentioned in passing why access to means for women working as Veterinarians in particular may be more relevant (than for men working as veterinarians). Do the authors have any ideas as to why this might be, and how it relates to other professions with access to other lethal means? Or any ideas for future research to find this out? Overall more gender based discussion is warranted – perhaps linking to theories mentioned in the introduction.

The authors state ‘Our results add to the literature of the effect of occupation gender composition on mental health, and for the first time, to the best of our knowledge, suicide is explored as an outcome.’ However, the authors cite elsewhere in the manuscript, a paper by Milner & King (2019) which examined the rate ratios of
Reviewer: 1  
Dr. Shervin Assari, University of Michigan

Comments to the Author:
This retrospective cohort investigates how sex/gender, sex/gender representation in the job, male and female-dominated workplaces contribute to incidence of suicide. This is an epidemiological paper with odds ratio, and a DAG is provided to control for bias. Here are my concerns with the paper:

1- When the data has time and event, then why not to report Hazard ratio. The authors have mainly reported Odds ratio which ignores if the incident was in 2006 or 2010.  
Thanks for your suggestion. The exposure variables we use in our analysis (occupation type, gender dominance in the occupation and gender dominance in the workplace) are not incident exposures (meaning that individuals may have been exposed previously to this exposure). In this regard, we are measuring the exposure at a single-point time (2005). Therefore, time until the event is not an outcome of interest as the exposure is not incident. Further, we are using cumulative incidence of suicide between 2006-2010 for increasing the number of cases in our study, as suicide is such a rare outcome (in our study, there are a total of 607 suicide cases happening in 5 years).

2- Why only 2006-2010 are included?  
Thanks for your comment. We are using 5 years for increasing the number of cases in the study population, and we are not including more years because the likelihood of individuals changing their occupational titles during the follow-up period would increase. Therefore, it would be difficult to establish the association between occupational title and suicide risk. We have clarified this choice in the methods as follows:  
“Cumulative incidence for the period of 2006-2010 was used for increasing the number of suicides and therefore, statistical power.”

3- Missing data can be a real threat to this study. Sensitivity analysis without and with imputation should be better utilized and discussed.  
We agree with the reviewer that missing data in occupational codes may have an effect in our estimates, this is especially true for self-employed individuals as we discuss in the limitations of the study. Following the reviewer recommendations, we have done a comparison of the cumulated incidence of suicide per 100000 persons among included and excluded participants due to missing values in the occupational code.  
We have added this table to the supplementary material (Table S3).

Included  
(n= 3318727)  
Excluded  
(Workers excluded due to missing values in occupational codes)  
(n= 752,196)  
Sex (% female) 48.5% 46.2%  
Level of education (% elementary education) 64.2% 67.6%  
Country of birth (% foreign born) 9.9% 15.8%
Unemployment (% any spell of unemployment) 5.6% 24.6%
Mean age 43.6 35.6
Suicide cumulative incidence per 100,000 persons (2006-2010) 18 26

Further, we have added an explanation in the statistical analysis:
“We also compared the suicide rates among the included study population and the excluded due to missing values in their occupational codes.”
In the results section:
“Excluded individuals due to missing values in occupational codes were more frequently foreign born (15.8% vs 9.9%), younger (mean age 35.6 vs 43.6) and had spells of unemployment (24.6% vs 5.6%) compared to the included population. Also, the cumulative incidence of suicide per 100,000 persons was higher among the excluded individuals (26 vs 18) (Supplementary Table S3).”
And in the discussion section:
“In this regard, the comparison of the characteristics of excluded and included individuals in our study showed that individuals working without information in their occupational codes were more likely to be young and foreign born. Also, they had more frequently any spells of unemployment, which may also explain why they also had a higher incidence of suicide.”

4- In an ideal study, multi-level determinants could be investigated. People are nested to jobs, and those people have similar characteristics other than sex, education, etc. If you were thinking about this study as a multi-level, then you could test cross-level interactions like:
a- males vs females in female-dominated workplaces and male-dominated workplaces.
b- low vs high education in female-dominated workplaces and male-dominated workplaces.
c- etc.

Thanks for your suggestion. Our data may be seen as having several levels as the reviewer mentions: worker (level 2) and workplace or occupation (level 1).
The minimum required units per level necessary for running multilevel analyses are usually 20, and in the case of our study, this would only apply at the occupational level (where we have 20 categories) but not for the variables gender dominance in the occupation or in the workplace, for which we only have three units per level: male-dominated, female-dominated and gender-balanced. Further, due to the rare outcome that we are exploring in this study, we did not think that multilevel analysis would be suitable for our outcome of study, as there would be levels without the outcome or few cases. Therefore, parameters of additional levels would be harder to estimate. We believe that applying multilevel analysis would be a better fit for studying more prevalent outcomes, such as common mental disorders or musculoskeletal disorders. We will consider this approach in our future studies.
We have included the following information in the limitations and strengths of the study as follows:
“This study is exploring the effects of occupation and workplace in the individual incidence of suicide. This may have been approached using mixed-method models, where workplace and/or occupation are included as random effects in the model. Due to the low incidence of our outcome and the small number of levels in the variables gender dominance in the occupation and workplace, this study is not considering the clustering of individuals at the occupation or workplace level.”

5- Other conditions not including stress and unknown role of employment income, wealth, and benefits
We apologize, but we are not sure if we understood this comment.

6- In the DAG, education could also impact past suicide, and past mental health problems.
Thanks for this suggestion. We did not include socio-economic variables happening before baseline in the DAG because they are predictors of the socio-economic variables at baseline. Therefore, to control for the confounding effect of socio-economic variables happening before baseline, it is enough to adjust for socio-economic baseline characteristics. Also, we wanted to simplify the variables and connections shown in the DAG.
Thus, overall, data may require re-analysis to make it more multi-level and comprehensive. Thank you so much for suggestion. Please see the response to comment 4.

Reviewer: 2
Dr. Sharna Mathieu, Griffith University
Comments to the Author:
This observational study utilizes a population register in Sweden to examine the association between occupation and suicide, and to determine the effect of gender composition at the occupational and workplace level on suicide, in a large sample of Swedish adults (N=3,318,727; 2006-2010). This manuscript has many strengths, most notably, adjusting for employment characteristics (e.g., precariousness). Several limitations are discussed.

Introduction
The references appear out of order, jumping in number from 9 to 13. Furthermore, reference 11 in the final reference list appears incomplete. I suggest the authors review their references and reference list for accuracy.
Thanks for spotting this. We have gone through the references and fixed this.

Materials and Methods
Could the authors please clarify in the study criteria what is meant by point (v) in the exclusion criteria. Currently it reads as ‘occupational code from more than 5 years ago’ from which point?
We have clarified the exclusion criteria in the study design as follows:
“registered occupational code reported from more than 5 years ago.”

Could the authors please provide an example reference for the line ‘The reason for including undetermined intents is because several studies have concluded that some deaths classified as undetermined intent are in fact suicides.’ Currently there are none.
Absolutely. We have added two references for this:
33 Tøllefsen IM, Helweg-Larsen K, Thiblin I, et al. Are suicide deaths under-reported? Nationwide re-evaluations of 1800 deaths in Scandinavia. BMJ Open 2015;5:e009120. doi:10.1136/bmjopen-2015-009120
34 Tøllefsen IM, Thiblin I, Helweg-Larsen K, et al. Accidents and undetermined deaths: re-evaluation of nationwide samples from the Scandinavian countries. BMC Public Health 2016;16:449. doi:10.1186/s12889-016-3135-5

It would be beneficial to readers if the authors could define precarious employment briefly (as measured by the SWE-ROPE).
We apologize for the lack of clarity. We have expanded the definition and measurement of precarious employment in the methods as follows:
“SWE-ROPE consists of five components (contractual employment insecurity, temporariness, multiple job-holding, income level, coverage under collective bargaining agreements) covering the three dimensions (employment insecurity, income inadequacy and lack of rights and protection) of precarious employment as identified by Kreshpaj et al. Individuals were assigned a score according to the 5 components that ranged from –10 to +2 (being –10 very precarious), and then they were classified as precarious if their score was lower than –3.”

Three possible categories of unemployment are described (0 months unemployed, 1-6 months unemployed, more than 6 months unemployed) yet in the results only one percentage is provided.
Were these groups collapsed?
Yes, these groups were collapsed. We have clarified this in the methods as follows:
“unemployment spells in 2005 (categorized as none/any)”

I am not clear on how the authors have distinguished workplace from occupation. Given this is seemingly an important distinction based upon the aims of the study I suggest the authors clarify this for readers. How is workplace defined or assigned, how is this different from the occupational codes assigned using the SSYK-96 classification system at the 2-digit level (and later the 4-digit level)? Thanks for your suggestion. We have clarified the measurement of workplace and occupation in the methods as follows:
“Workplace was measured through the workplace identity number (Statistics Sweden's Business Database establishes an eight-digit, unique, serial number for each workplace). Everyone is linked to a workplace identity number. A workplace is any address, property or group of nearby properties where a company conducts business.”

Results
The remainder of the results section is succinct and clear. Thanks for your comment.

Discussion
In their interpretation the authors do well to describe some of the prominent explanations for why certain occupations may be at higher risk of suicide (e.g., access to lethal means). However, given the aims of the study more interpretation around gender may be warranted. For example, there is some good discussion of social status and men working in attendant nursing professions, however, it is merely mentioned in passing why access to means for women working as Veterinarians in particular may be more relevant (than for men working as veterinarians). Do the authors have any ideas as to why this might be, and how it relates to other professions with access to other lethal means? Or any ideas for future research to find this out? Overall more gender-based discussion is warranted — perhaps linking to theories mentioned in the introduction.
We have further developed the possible explanation why access to lethal means could be more relevant for women in certain occupations compared to men and added a reference for this as follows:
“This could be in part, due to the different preferences for suicide method among men and women, where men more often use more lethal and violent suicide methods (i.e., hanging, guns) compared to women (i.e., drugs and carbon monoxide poisoning) [44].”

Reference:
44. Mergl R, Koburger N, Heinrichs K, et al. What Are Reasons for the Large Gender Differences in the Lethality of Suicidal Acts? An Epidemiological Analysis in Four European Countries. PLoS One 2015;10:e0129062. doi:10.1371/journal.pone.0129062

The authors state ‘Our results add to the literature of the effect of occupation gender composition on mental health, and for the first time, to the best of our knowledge, suicide is explored as an outcome.’ However, the authors cite elsewhere in the manuscript, a paper by Milner & King (2019) which examined the rate ratios of suicide for men working in male or female dominated occupations and for women working in male or female dominated occupations (with various adjustments). I suggest the authors clarify
We apologize for the lack of clarity. We have modified the statement as follows:
“Our results add to the literature of the effect of occupation gender composition on mental health.”
Reviewer 2, Dr. Sharna Mathieu, Griffith University
I commend the authors on their revision and addressing reviewer comments. It is my view that the authors have responded appropriately to the feedback I provided. I would note that there could be some minor English language editing, for example, the switching of past and present tense throughout.

While the authors decided against suggestions made by the first reviewer regarding multi-level analyses, I believe there was effort made to provide thoughtful rebuttal for this and has been acknowledged in the limitations. Nevertheless, the authors do not provide references for their assertions and rule of thumb which makes it hard to confirm.

We have provided two references that discuss the sufficient sample size for multilevel modeling in order to provide valid estimates.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Moineddin, R., Matheson, F.I. & Glazier, R.H. A simulation study of sample size for multilevel logistic regression models. BMC Med Res Methodol 7, 34 (2007). https://doi.org/10.1186/1471-2288-7-34