PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Socioeconomic differences in recruitment and sickness absence in a large NHS health organisation – a cross-sectional study
AUTHORS	Daras, Konstantinos; Baker, Wesam; Rafferty, Joe; Oates, Amanda; Edwards, Louise; Wyatt, Steven; Barr, Ben

VERSION 1 – REVIEW

REVIEWER	Hashmi, Rubayyat University of Southern Queensland, School of Commerce
REVIEW RETURNED	08-Apr-2021

GENERAL COMMENTS	Thank you for the opportunity to review the article titled "Patterns of employment and inequalities in sickness absence in a large community and mental health organisation - an observational study". Using the Mersey Care NHS foundation trust's Electronic Staff record, this study examined the association between the sickness absence days and sociodemographic factors (deprivation indices of areas, wage band, occupational group and age-sex demographics) in North-West of England. I have enjoyed reading the article and some of the results seem interesting. However, I have some concerns and confusions regarding how the article is presented. I hope the authors find the following comments and suggestions helpful: Major comments: The major concern for this study is the use of the word "employment". Employment as in being recruited/hired/employed (other states: being unemployed/ not in the labour force) should not be interchangeable with terms like occupational groups or productivity. Thus, part of the title "Patterns of employment" is confusing. Since all the data in Mersey Care NHS foundation trust's Electronic Staff record are already employed individuals, there are no patterns of employment in this study by definition. Thus, if there is variation in NHS recruitment by geography, in my opinion, that is not employment pattern rather geographic variation in the supply of NHS staff. Given this background, any discussion on employment should be carefully presented. For example, in some part of the article, it has been suggested that increasing the recruitment from the disadvantaged communities can reduce the health inequalities (see abstract objective, introduction, the implication for policy and practice, conclusion). Although this hypothesis could be true, this article, however, does not provide such evidence. The article only found that high deprivation areas are more associated with sick days (if controlled for with occupation groups this was not significant) which has different interpretations. I hope the authors will clarify these issues. Other comments:

	Introduction: The introduction should more adequately introduce the variables considered in the study and set the purpose of the research. The contribution of the research is not clear. The aims are poorly outlined and hypotheses are completely missing. Methods: It would be good to have a separate descriptive statistics table (detail) of all the variables used in the study. Besides, justification is required as to why such variables were chosen for the study. A conceptual framework for the study linking the hypotheses will be useful. In the models, interaction effects among factors could be explored. Results: All results should report 95% confidence intervals (For example in the abstract section or in figure 2. Instead of reporting in median days in figure 2b, it would be good to have to mean sick days since it will use all the ESR data). The information in the first two sentences of the abstracts results section is not available in the detail results section. There should be a justification on which of the model (model 1 or model 2) should be reported in the abstract. Discussion: In my opinion, the discussion section is very small and not detail. The discussion section could be improved by comparing and contrasting results from other studies. The detailed contribution of this study compared with other studies should be discussed. Conclusion: The conclusion needs to be revised according to the study findings.

REVIEWER	Teixeira, Liliane Centro de Estudos da Saúde do Trabalhador e Ecologia Humana, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Sala 55 Rua Leopoldo Bulhões, 1480, Manguinhos, Rio de Janeiro, RJ, CEP: 21040-360, Brasil, National School of Public Health
REVIEW RETURNED	09-Jun-2021

GENERAL COMMENTS	Dear Authors, The topic is important for public health. However, I do not feel comfortable analyzing the article. I thought in the introduction the authors should make a comparison between "NHS" and other national health systems, considering the differences between developed countries and developing countries. In addition, the references used for social determinants of health and psychosocial and health effects of the workplace are old and again questioned me, since we know that low income-related occupations lead to psychosocial complaints when presenteeism and absenteeism. So I think it important to send the article to experts from the field of planning and management of health systems. And, in the Methods - Data and measures: I think it's important to explain the theoretical basis for defining the categories of age groups. And reallocate the phrase "To avoid small number effects Introduced in our analysis, we combined pay bands 1 and 2 along with pay bands 8 and 9." To close the wage band.

REVIEWER	Chuang, Hung-Yi Kaohsiung Medical University Hospital, Department of Occupational and Environmental Medicine
REVIEW RETURNED	14-Jun-2021

GENERAL COMMENTS	The study used data from Mersey Care Electronic Staff Record to analyse the distribution of the workforce by socioeconomic deprivation and how sickness absence rates vary based on wage

level, occupational group and level of deprivation. The authors used negative binomial regression models to estimate the number of sick days per employee and the % of staff with any sickness absence in the year for each deprivation quintile, the wage band, and the occupational group. They found that the sickness absence rate for the most deprived quintile was 1.41 (95% CI 1.16 to 1.70) times higher than that for the least deprived quintile. In addition, the model revealed the higher sickness rate in the lower wage bands; furtherly, adjusted sickness absence rates for the staff in the groups of care assistants and the nursing and midwifery registered group were 1.72 and 1.84 time higher than the administrative and clerical group. Finally, they concluded "NHS organisations potentially have large share of their workforce living in disadvantaged areas, however, these groups are likely to experience higher level of sickness absence. By increasing recruitment from these communities and developing effective policies for improving health and working conditions for these groups, the NHS can contribute to reducing health inequalities through its workforce policies."

1. This is a cross-sectional study, which cannot mention causation. The part of discussion was too short to discuss why increasing recruitment workforce living in disadvantaged areas would reduce health inequalities.

2. Uses of sickness absence as inequalities or inequalities in sickness absence need further elaboration otherwise differential discussion.

3. The model 2 in the table 2 may consider carefully, since the IMD is a composite indicator based on seven domains: Income, Employment, Education, Skills and Training, Health, Crime, Housing and the Environment, I suggest IMD cannot be an independent variable in the same model with wage and occupational groups, or require specialist statistical review.

VERSION 1 – AUTHOR RESPONSE

Reviewer name/ref	Comment No.	Reviewer suggestion/comment (exactly as it came)	Our response	Done
Editor	1	Please revise the title to indicate the research question, setting, and study design. This is the preferred format for the journal.	Title revised to indicate the research question, setting and study design.	☑
Editor	2	Please ensure that all acronyms are defined on first mention, including those in the abstract.	All acronyms are defined on first mention as requested.	☑
Editor	3	Please reformat the abstract so that it follows the structured abstract recommended in the journal's instructions for authors for research articles. See: https://bmjopen.bmj.com/pages/authors/#research	Abstact reformated to follow the structure recommended in the journal's instructions.	☑
Editor	4	Please include any relevant statistical results in the results section of the Abstract.	Results section of the Abstract have been updated reporting 95% confidence intervals.	☑

| Editor | 5 | Along with your revised manuscript, please include a copy of the STROBE checklist indicating the page/line numbers of your manuscript where the relevant information can be found (https://strobe-statement.org/index.php?id=strobe-home). | STROBE checklist completed indicating the page/line numbers of our manuscript where the relevant information can be found, | ☑ |
| Reviewer: 1 | 6 | The major concern for this study is the use of the word "employment". Employment as in being recruited/hired/employed (other states: being unemployed/ not in the labour force) should not be interchangeable with terms like occupational groups or productivity. Thus, part of the title "Patterns of employment" is confusing. Since all the data in Mersey Care NHS foundation trust's Electronic Staff record are already employed individuals, there are no patterns of employment in this study by definition. Thus, if there is variation in NHS recruitment by geography, in my opinion, that is not employment pattern rather geographic variation in the supply of NHS staff. Given this background, any discussion on employment should be carefully presented. For example, in some part of the article, it has been suggested that increasing the recruitment from the disadvantaged communities can reduce the health inequalities (see abstract objective, introduction, the implication for policy and practice, conclusion). Although this hypothesis could be true, this article, however, does not provide such evidence. The article only found that high deprivation areas are more associated with sick days (if controlled for with occupation groups this was not significant) which has different interpretations. I hope the authors will clarify these issues. | Thank you, we agree that employment was not quite the right term here - as it is potentially confused with other labour market concepts such as employment rates, unemployment rates etc. Our aim was to understand the geographical distribution of workers at this organisation in relation to area deprivation, and how sickness absence varied by area deprivation of residence. This is important because, by recruiting from more disadvantaged populations NHS organisations can help improve employment in those communities and by improving the health of these employees help reduce health inequalities. We have therefore changed the wording - using the term recruitment of employees, and the geographical distribution of employees/ staff / workforce- instead of employment per se.

We are not claiming to test the hypothesis as to whether increasing recruitment from the disadvantaged communities can reduce the health inequalities. We highlight that this is potentially true, given the existing evidence for the health benefits of employment. But a necessary condition for that to be true would be that NHS organisations recruit from those communities and targets workplace health programmes proportional to differences in their needs. The purpose of the study is therefore descriptive and aims to understand:
1. The distribution of the workforce of this NHS trust by socioeconomic deprivation and,
2. How sickness absence rates vary by level of deprivation , and the extent to which this was | ☑ |

explained by levels of wages and occupation.

Reviewer: 1	7	Introduction: The introduction should more adequately introduce the variables considered in the study and set the purpose of the research. The contribution of the research is not clear. The aims are poorly outlined and hypotheses are completely missing.	We agree that the introduction had not adequately set up the purpose and contribution of the research and the study aims. We have now extensively revised the introduction in light of these comments and clarified the aims. We have also more clearly described the variables in the methods.	☑
Reviewer: 1	8	Methods: It would be good to have a separate descriptive statistics table (detail) of all the variables used in the study. Besides, justification is required as to why such variables were chosen for the study. A conceptual framework for the study linking the hypotheses will be useful. In the models, interaction effects among factors could be explored.	Thank you we have now added a descriptive statistics table as sugested to the supplamentary file (Table 1). We have added text to introduction and discussion to better explain the rationale of this study.	☑

Reviewer: 1	9	Results: All results should report 95% confidence intervals (For example in the abstract section or in figure 2. Instead of reporting in median days in figure 2b, it would be good to have to mean sick days since it will use all the ESR data). The information in the first two sentences of the abstracts results section is not available in the detail results section. There should be a justification on which of the model (model 1 or model 2) should be reported in the abstract.	All results report 95% CIs and mean sick days used instead of median in figure 2b. Results mentioned only in abstract have been added to the results section too. We have added a rationale for the two models - the first is to explore the age and sex adjusted sickness absence ratios between deprivation groups and the second is to investigate whether this is explained by difference in sickness absence between wage bands and occupational group. We have also highlighted in discussion the implication for these findings - in that targeting workplace interventions to reduce differences in sickness absence across wage levels and occupational groups would largely address the differences in sickness absence between deprivation groups.	☑
Reviewer: 1	10	Discussion: In my opinion, the discussion section is very small and not detail. The discussion section could be improved by comparing and contrasting results from other studies. The detailed contribution of this study compared with other studies should be discussed.	We agree that the discussion section could be improved by comparing results to other studies and explainig what this study contiburtes. We have added a paragraph in this section discussing our findings.	☑
Reviewer: 1	11	Conclusion: The conclusion needs to be revised according to the study findings.	Conclusion section have been revised according to the study findings.	☑
Reviewer: 2	12	The topic is important for public health. However, I do not feel comfortable analyzing the article. I thought in the introduction the authors should make a comparison between "NHS" and other national health systems, considering the differences between developed countries and developing countries. In addition, the references used for social determinants of health and psychosocial and health effects of the workplace are old and again questioned me, since we know that low income-related occupations lead to psychosocial complaints when presenteeism and absenteeism. So I think it important to send the article to experts from the field of planning and management of health systems.	We have added text to introduction and discussion to better explain the rationale of this study. We also investigated further for more up to date references related to sickness absence in health services but there is limited evidence specific to the NHS or other health services. An international comparison between NHS and other national health systems is beyond the scope of this paper.	☑

Reviewer: 2	13	in the Methods - Data and measures: I think it's important to explain the theoretical basis for defining the categories of age groups. And reallocate the phrase "To avoid small number effects Introduced in our analysis, we combined pay bands 1 and 2 along with pay bands 8 and 9." To close the wage band.	The age groups are according to the way the Mersey Care ESR records staff information. We also added the numbers for each band affected by small numbers (band1 and 9). We have explained further the structure of age groups and wage band in the methods section.	☑
Reviewer: 3	14	This is a cross-sectional study, which cannot mention causation. The part of discussion was too short to discuss why increasing recruitment workforce living in disadvantaged areas would reduce health inequalities.	We agree with the reviewer that we are not trying to ascertain causation. This study tries to describe patterns of recruitment and Sickness Abscence. We have added to introduction and discussion sections to better explain the rationale and relationship to health inequalities.	☑
Reviewer: 3	15	Uses of sickness absence as inequalities or inequalities in sickness absence need further elaboration otherwise differential discussion.	We have removed the use of term "inequalities in sickness absence".and replaced it with "differences in sickness abcence rates between socioeconomic groups".	☑
Reviewer: 3	16	The model 2 in the table 2 may consider carefully, since the IMD is a composite indicator based on seven domains: Income, Employment, Education, Skills and Training, Health, Crime, Housing and the Environment, I suggest IMD cannot be an independent variable in the same model with wage and occupational groups, or require specialist statistical review.	We further clarified that the IMD domains mainly consisted of deprivation indicators. The IMD is an area based measure whilst the others are individual based measures. Firstly we are interested in how patterns relate to area based measures - as these relate to recruitment to policies - i.e to address health inequalities, we need to recruit from those communities and improve their health, but then we want to understand if just using wages and occupation to target workplace policies would address the differences in sickness abcence by IMD. So these are measuring different things at different levels. We also examined the multicollinearity between the variables in both models using Variance Inflation Factors (VIF). VIF results for both models were less than 2.0 when considering the number of coefficients (Df) in the variable, suggesting that multicollinearity was not a cause for concern.	☑

VERSION 2 – REVIEW

REVIEWER	Chuang, Hung-Yi Kaohsiung Medical University Hospital, Department of Occupational and Environmental Medicine
REVIEW RETURNED	18-Oct-2021

GENERAL COMMENTS	I have no major questions, since they revised the manuscript with the previous comments. If the other reviewer had no more comment, the manuscript could be accepted.