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. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Previous sickness absence and current low perceived social support at work among employees in the general population: a historical cohort study
AUTHORS	Knapstad, Marit; Holmgren, Kristina; Hensing, Gunnel; Øverland, Simon

VERSION 1 - REVIEW

REVIEWER	Jenni Ervasti Finnish Institute of Occuppational Health, Finland
REVIEW RETURNED	08-Jul-2014

GENERAL COMMENTS	My major concern is related to your study design. Based on your data, you cannot draw causal conclusions. Since you don't know the baseline level of social support, you can't say that you examine whether various patterns of previous long-term sickness absences PREDICT current low social support, even if you have data on previous sick leaves. You should simply talk about associations. Your main finding would thus be that "Sick leave history is associated with current low support". The study aim (research question), discussion and conlusions should be modified accordingly. If a person has immigrated or emigrated or out of working life, that person is not at risk of sick leave. That data should not be imputed. Instead, these persons should be excluded from the analyses. The grouping of "patterns of sick leaves" (five groups) was somewhat unclear. If you used some kind of statistical methods, those should be described in the "Statistical Analyses" section. The grouping is also rather crude. Maybe you could try analyzing different trajectories of sick leaves in relation to social support? If you decide to stick with this grouping, a more detailed description about the analyses is needed. The statement about ethical consent/ethics committee approval is missing.

REVIEWER	Merete Labriola Rehabilitation Center Marselisborg, Department of Public Health, Section of social medicine and rehabilitation, Aarhus University, and CFK, Central Denmark Region. Denmark
REVIEW RETURNED	29-Jul-2014

GENERAL COMMENTS	Interesting paper, especially the section strengths and limitations plus the interpretation

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

My major concern is related to your study design. Based on your data, you cannot draw causal conclusions. Since you don't know the baseline level of social support, you can't say that you examine whether various patterns of previous long-term sickness absences PREDICT current low social support, even if you have data on previous sick leaves. You should simply talk about associations. Your main finding would thus be that "Sick leave history is associated with current low support". The study aim (research question), discussion and conlusions should be modified accordingly.

If a person has immigrated or emigrated or out of working life, that person is not at risk of sick leave. That data should not be imputated. Instead, these persons should be excluded from the analyses.

The grouping of "patterns of sick leaves" (five groups) was somewhat unclear. If you used some kind of statistical methods, those should be described in the "Statistical Analyses" section. The grouping is also rather crude. Maybe you could try analyzing different trajectories of sick leaves in relation to social support? If you decide to stick with this grouping, a more detailed description about the analyses is needed.

The statement about ethical consent/ethics committee approval is missing.

Response to reviewer 1:

We thank the reviewer for the thoughtful suggestions for further improvement of the manuscript. In the following we will clarify and discuss solutions regarding the issues raised by the reviewer about a) study design, b) missing data on sickness absence, c) categorization of the sickness absence groups and d) ethical content.

a) Study design:
Regarding study design, we fully agree that only having measured social support at one time point is a key limitation of our study, and that the design precludes making causal inferences. However, a large amount of the previous literature on the association between social support and sickness absence has used cross-sectional design, and results are often interpreted in a uni-directional mode where low social support (as other negative factors at work) leads to sickness absence. We therefore think our study is important as it after all is longitudinal, and sheds light on that these relationships might also be bidirectional and that sickness absence also can influence social support. As highlighted in the manuscript, we recommend that further studies use repeated measurements and open up for the possible bi-directional relationships between psychosocial work factors and sickness absence.
To avoid any confusion regarding our aims and what conclusions can be drawn from our study, we have modified the manuscript, as suggested, and only use the term associations.

b) Missing data on previous sickness absence
The choice of how to treat missing data on sickness absence is complex. Thank you for pointing out that some of those included might not be at risk and should be excluded from the analyses. There are

n=65 with missing data one or more of the seven follow-up years (2001-2007). The choice about inclusion/exclusion of all or some of these cases should, as we see it, be informed by the answer to the following three interrelated questions; 1) what is the cause of missing for the n=65, 2) does missing mean not being at risk for sickness absence and 3) should cases be excluded though being "not at risk" only some of the follow up years (2001-2007).

We know from a report from the Statistics Sweden from 2011 on the LISA register 1990-2009 (SCB report 2011:4) that missing data could be due to births and death, migration, and as well as registration errors. The two first causes are not applicable to our study sample. The latter can be regarded random and arguably be treated by multiple imputation. The case of migration is somewhat more complex:

Migration:
Some of the missing is most likely due to migration. Firstly, as noted in the report from the Statistics Sweden (SCB report 2011:4), only those having Swedish citizenship the year of registration are included in LISA. Secondly, registry data on the HAP participants shows that there are significantly fewer being registered as Swedish citizenship in 2008 (46.2% vs 97.6%) and having Sweden as their country of birth (41.5% vs 90.8%) among those with missing than among those without missing data. Some might have immigrated during the follow-up period (2001-2007). Others might have emigrated (but must have returned to Sweden by 2008 to be included in the HAP study).

As noted by the reviewer those immigrated or emigrated are in general not at risk (that is, when not being registered in Sweden). Of note, some of those emigrating for a shorter or longer period, may nonetheless be entitled to sick pay and could hence be at risk. Unfortunately, we cannot discriminate with accuracy between those at risk and those not at risk among the n=65.

"Being out of work":
In Sweden, unemployed and students have access to sick pay, and could hence be at risk in relation to compensation though not working. It could though be other reasons for excluding this group, e.g. we have excluded all those not reporting working as an employee in 2008. The information about this group was a bit unclear in the manuscript and is updated in the revised version

In sum, some of those n=65 are probably not at risk one or more of the follow-up years. Though having some information about this group, we cannot know with certainty whom were not at risk. We first chose to employ multiple imputation rather than to exclude these cases, because many of them only had missing data a few of the seven years (e.g. did n=19 of the n=65 cases have missing data only at one data point, and thus were "at risk" 6 of 7 of the follow-up years). We hence reasoned that we could make an estimate of these cases' overall sickness absence pattern based on data the years they reliably were at risk. The main advantage of this approach is to keep as many cases with data as possible. It could especially be of value to keep as many as possible of the non-Swedish subgroup, as these already are underrepresented in the study sample. The main drawback of this solution is the "over inclusion" of cases not at risk, which may cause bias or noise in our analyses. One alternative could have been to require a certain amount of data (e.g. thee data points) for inclusion. It could however be theoretically difficult to choose how many data points required.
To be sure we do not introduce bias to our analyses, we thus agree with the reviewer that it is more sensible to exclude all the n=65 cases with missing one or more of the follow-up years. As can be seen if comparing the results reported in the original and revised manuscript, the results are robust whether or not including these cases.

c) Categorization of the sickness absence groups
Thank you for pointing out that the categorization of "sickness absence groups" was not sufficiently

transparent. Further, we agree that examining different trajectories of sick leave could have been an interesting way to explore our research aim. As we had sufficient data points initially performed latent class analyses (LCA), which could help us find meaningful subgroups without constrains regarding linearity. There were two main reasons that we did not use this approach for the grouping employed in the manuscript: 1) We needed to accommodate information from a subgroup (n=86) that were granted the more a long-term sickness compensation part- or full-time one or more of the follow-up years, and this not adequately feasible for LCA. 2) We got difficulties regarding statistical power due to small groups if excluding the mentioned subgroup from the LCA. We thus chose to construct categories based on the widely applied median splits, informed by the results of the latent class analyses. Theoretically, it gives meaning to split the sample into highs and lows, and the split between the "former" and "latter" period was informed by results of the LCA. The main drawback of the "median split" alternative is, as noted by the reviewer, that it gives a rather crude measure. We nonetheless believe that this way of grouping the sickness absence patterns are more transparent and accessible for readers not familiar with mixture modelling techniques. Important to note, the groupings where rather similar and main results robust across statistical method employed. In addition, the groupings were similar to those emerged from a previous published trajectory analysis (Haukka E, Kaila-Kangas L, Ojajärvi A, et al. Pain in multiple sites and sickness absence trajectories: A prospective study among Finns. Pain. 2013;154(2):306-12).

Based on knowledge from the various attempts in grouping the "sick leave patterns" and discussions in the author group, we have landed on sticking to the grouping employed in the original submitted manuscript. In the revised manuscript we have expanded on and sharpened the description about how we performed the analyses. We have also added a table that describe the group allocation. We have not described the LCA procedures in detail. It only was used as an initial exploratory analysis, and since the results from this analysis are not shown in the manuscript, we believe a description would confuse more than inform the reader. We may add more information on the LCA if the reviewers or editor find this more appropriate.

d) Ethical consent
The HAP study was approved by the Ethics Committee, University of Gothenburg, Sweden. We have moved and expanded the statement on ethical consent to the method section.

Reviewer 2:
Reviewer Name Merete Labriola
Institution and Country Rehabilitation Center Marselisborg, Department of Public Health, Section of social medicine and rehabilitation, Aarhus University, and CFK, Central Denmark Region. Denmark
Please state any competing interests or state 'None declared': None declared

Interesting paper, especially the section strengths and limitations plus the interpretation

Response to reviewer 2:
We thank the reviewer for the positive response to our manuscript. Based on comments from reviewer 1 and the editor, we have made some changes in the manuscript to improve its quality and readability.

VERSION 2 – REVIEW

REVIEWER	Jenni Ervasti
	Finnish Institute of Occupational Health, Finland
REVIEW RETURNED	22-Aug-2014

| GENERAL COMMENTS | In the revised manuscript, the authors have addressed my concerns. Some minor comments: on p. 10 you say that age variable was mean? I think you mean that it was continuous (not classified, as in previous version). On Table 3, you still have a footnote about missing responses handled with imputations, which is not the case anymore (?) That should be deleted. |